

ISSN 2078-4481

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХЕРСОНСЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ**

ВІСНИК

**ХЕРСОНСЬКОГО НАЦІОНАЛЬНОГО
ТЕХНІЧНОГО УНІВЕРСИТЕТУ**

4(95)

Частина 3

Рекомендовано до друку Вченою радою
Херсонського національного технічного університету
(протокол № 6 від 11.12.2025 року)

Журнал включено до Переліку наукових фахових видань України
категорії «Б» зі спеціальностей: C1, D1, D2, D3, D5, D7, J3
(Наказ МОН України від 17.03.2020 № 409);
D4 (Наказ МОН України від 29.06.2021 № 735);
J5, J6, J7, J8, F2, F3, F5, F6, F7, G7, G8, G9, G11
(Наказ МОН України від 02.07.2020 № 886);
G1, G3, G4, G15 (Наказ МОН України від 24.09.2020 № 1188)

Журнал включено до наукометричних баз, електронних бібліотек та репозитаріїв:
GoogleScholar, Crossref, National Library of Ukraine (Vernadsky)



Видавничий дім
«Гельветика»
2025

Редакційна колегія

Головний редактор

Шерстюк В. Г.

д.т.н., професор, проректор з навчальної роботи

Заступник головного редактора

Сарібекова Ю. Г.

д.т.н., професор

Відповідальний секретар

Лур'є І. А.

к.т.н., доцент, начальник навчально-методичного відділу, доцент кафедри інформатики і комп'ютерних наук

Члени редакційної колегії

Баганов Є. О.	к.т.н., доцент
Березовський Ю. В.	д.т.н., доцент
Бойко Г. А.	к.т.н., доцент
Вороненко М. О.	к.т.н., доцент
Гончар О. І.	д.е.н., професор
Горбачов П. Ф.	д.т.н., професор
Дімітрова В.	д.н., доцент
Джерелюк Ю. О.	д.е.н., професор
Євтушенко В. В.	к.т.н., доцент
Жарікова М. В.	д.т.н., професор
Зубкова К. В.	к.т.н., доцент
Корчевська Л. О.	д.е.н., професор
Кузьміна Т. О.	д.т.н., професор
Куник О. М.	к.т.н., доцент
Луб'яний П. В.	к.т.н., доцент
Наумов О. Б.	д.е.н., професор
Олійник Н. М.	к.т.н., доцент
Плющ Р. М.	д.держ.упр., професор
Половцев О. В.	д.держ.упр., к.т.н., професор
Рудакова Г. В.	д.т.н., професор
Русанов С. А.	к.т.н., доцент
Савін С. Ю.	д.е.н., доцент
Салєба Л. В.	к.т.н., доцент
Семешко О. Я.	д.т.н., професор
Сідельникова Л. П.	д.е.н., професор
Smolarz A. (Польща)	dr.hab.inz.
Стоянова О. В.	к.т.н., доцент
Топалова Е. Х.	к.держ.упр., доцент
Філіппова В. Д.	д.держ.упр., професор
Хрущ Н. А.	д.е.н., професор
Чепелюк О. В.	д.т.н., професор
Шандова Н. В.	д.е.н., професор
Шарко О. В.	д.т.н., професор
Шарко М. В.	д.е.н., професор
Шевченко І. І.	д.т.н., професор

ISSN 2078-4481

**MINISTRY OF EDUCATION AND SCIENCE OF UKRAINE
KHERSON NATIONAL TECHNICAL UNIVERSITY**

VISNYK

**OF KHERSON NATIONAL
TECHNICAL UNIVERSITY**

4(95)

Part 3

Recommended for publication by the Academic Council
of Kherson National Technical University
(Minutes № 6 on 11th December 2025)

The journal is included in the List of Scientific Professional Editions of Ukraine Category “B”
in specialties: C1, D1, D2, D3, D5, D7, J3

(Ukraine Education and Science Ministry Order dated 17.03.2020, № 409);

D4 (Ukraine Education and Science Ministry Order dated 29.06.2021, № 735);

J5, J6, J7, J8, F2, F3, F5, F6, F7, G7, G8, G9, G11

(Ukraine Education and Science Ministry Order dated 02.07.2020, № 886);

G1, G3, G4, G15

(Ukraine Education and Science Ministry Order dated 24.09.2020, № 1188)

The journal is included in the scientometric bases, electronic libraries and repositories:

Google Scholar, Crossref, National Library of Ukraine (Vernadsky)



Publishing House
“Helvetica”
2025

Editorial Board

Editor-in-Chief

Sherstiuk V. H.
DSc (Engineering), Professor, Vice-Rector for Academic Affairs

Deputy Editor-in-Chief

Saribekova Yu. G.
Doctor of Engineering Science, Professor

Executive Secretary

Lurie I. Yu.
PhD (Engineering), Associate Professor, Head of the Educational and Methodical Department, Associate Professor of the Department of Informatics and Computer Sciences

Members of Editorial Board

Baganov Ye. O.	Ph.D., Associate Professor
Berezovsky Yu. V.	Doctor of Engineering Science, Associate Professor
Boiko H. A.	Ph.D., Associate Professor
Voronenko M. O.	Ph.D., Associate Professor
Honchar O. I.	Doctor of Economic Sciences, Professor
Horbachov P. F.	Doctor of Economics, Professor
Dimitrova V. Ya. (Bulgaria)	Ph.D., Associate Professor
Dzhereliuk Yu. A.	Doctor of Economics, Professor
Yevtushenko V. V.	Ph.D., Associate Professor
Zharikova M. V.	Doctor of Engineering Science, Professor
Zubkova K. V.	Ph.D., Associate Professor
Korchevska L. A.	Doctor of Economics, Professor
Kuzmina T. O.	Doctor of Technical Sciences, Professor
Kunyk O. N.	Ph.D., Associate Professor
Lubianyi P. V.	Ph.D., Associate Professor
Naumov O. B.	Doctor of Economics, Professor
Oliinyk N. M.	Ph.D., Associate Professor
Pliushch R. M.	Doctor of Sciences in Public Administration, Professor
Polovtsev O. V.	Doctor of Sciences in Public Administration, Professor
Rudakova H. V.	Doctor of Engineering Science, Professor
Rusanov S. A.	Ph.D., Associate Professor
Savin S. Yu.	Doctor of Economics, Associate Professor
Saleba L. V.	Ph.D., Associate Professor
Semeshko O. Ya.	Doctor of Engineering Science, Professor
Sidelnikova L. P.	Doctor of Economics, Professor
Smolarz A. (Poland)	Ph.D., Associate Professor
Stoianova O. V.	Ph.D., Associate Professor
Topalova E. K.	Ph.D., Associate Professor
Filippova V. D.	Doctor of Science in Public Administration, Professor
Khrushch N. A.	Doctor of Economics, Professor
Chepeliuk O. V.	Doctor of Technical Sciences, Professor
Shandova N. V.	Doctor of Economics, Professor
Sharko M. V.	Doctor of Economics, Professor
Sharko O. V.	Doctor of Engineering Science, Professor
Shevchenko I. I.	Doctor of Technical Sciences, Professor

ЗМІСТ

ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

С. Г. Антошук, О. М. Іванова. Швидка локалізація осередків програмного коду FPGA, придатних для стеганографічного зберігання додаткових даних.....	9
О. Yu. Barkovska. Formal description of interaction and data flows in multimodal assistive systems for user autonomy support.....	15
М. Е. Бондаренко, Г. С. Іващенко. Адаптивний розподіл обчислювального навантаження в системах голосової ідентифікації.....	21
М. О. Вернік, Л. М. Олещенко. Програмні методи та архітектурні рішення для обробки великих даних через квантові нейронні мережі та метаевристичні оптимізатори.....	29
І. В. Гавриш, А. Г. Кривоухата. Розробка архітектури адаптивного інтерфейсу редактора веб-застосунку для створення віртуальних турів.....	37
Т. О. Грінченко, Д. М. Малярова, О. П. Нарсжний. Прототип веб-сервісу квантового генератора випадкових чисел для захищеного обміну повідомленнями.....	47
К. І. Дорофєєва, О. В. Сєверінов, З. М. Сидоренко, В. М. Сухотеппий. Застосування інструмента аналізу безпеки для виявлення критичних вразливостей у веб-додатках.....	62
Л. О. Древич. Сегментація смуг для оптимізації трафік-потоків.....	69
М. С. Дутчак, М. В. Пікуляк, Р. О. Тичковський, І. Б. Полатайко. Моделювання життєвого циклу розробки програмного забезпечення інтелектуальних навчальних систем.....	81
Я. Р. Д'яков. Модель ієрархічного семантичного графу знань для багаторівневого представлення знань.....	88
Г. А. Завгородня, В. В. Завгородній. Розробка масштабованої розподіленої архітектури для масових багатокористувацьких онлайн-систем.....	99
Ю. Ю. Іляш, В. О. Горєлов, В. А. Ровінський. Алгоритми та методи процедурної генерації та адаптивної складності рівнів в іграх-платформерах.....	107
М. Р. Калашніков, В. С. Харченко. Інтелектуальний підхід до підвищення стабільності систем автоматизованого тестування на базі навчання з підкріпленням.....	117
А. О. Каргін, О. І. Проніна, О. Є. П'ятикоп, Є. А. Перцев. Технологія проектування штучного інтелекту, що підтримує автономні розподілені системи.....	126
О. В. Крайнюк, Ю. В. Буц, М. М. Пікєасов, Н. В. Діденко, Б. О. Походенко. Функціональне моделювання та алгоритмізація процесів забезпечення кіберстійкості автоматизованих систем видачі ЗІЗ.....	137
Я. А. Мартиненко, О. В. Сєверінов, А. М. Євгенєєв, М. С. Скибенко. Дослідження SOAR-платформ та обґрунтування вибору Splunk Phantom.....	145
Д. М. Нехороших, П. В. Буслов, А. М. Євгенєєв, М. С. Скибенко. Аналіз фреймворків та технічних засобів захисту для інфраструктури інтернету роботизованих речей.....	153
А. М. Онищенко, О. П. Остапенко, А. А. Лойшин. Інтегровані інформаційні моделі управління сталим розвитком на засадах еколого-економічного аналізу.....	161
О. С. Пархоменко, В. В. Коваль. Еволюція шахрайського контенту в умовах розвитку інформаційних технологій та ШІ.....	171
В. Передеренко. Методологія міграції Android-проекту на основі перегляду проекту MVVM до архітектури Jetpack Compose та MVI.....	175
Б. Р. Попович, Р. Б. Попович. Дискретна задача побудови елементів великого порядку в мультиплікативній групі скінченного поля.....	182
Д. М. Ратушний, Р. А. Руденський. Архітектурні підходи до вибору стеку технологій для зберігання IoT-даних в задачах управління логістикою.....	187
І. В. Рубан, В. М. Ткачов. Метод робастної варіаційної оптимізації траєкторій функціонування інформаційної системи на мобільній платформі в умовах деструктивних впливів.....	195
V. I. Ruzhentsev. Analysis of cryptographic resistance of iterative algorithms based on consideration of iterative differential characteristics.....	213
S. A. Sgadov. Analysis of STM32 family support in the .NET eco-system and perspectives for community targets development.....	220

І. П. Сініцин, С. В. Задворний, О. П. Ільїна. Моделі та методи підтримки користувацької центрованості інтерфейсу програмних систем оборонного призначення з використанням машинного навчання.....	228
О. А. Сніюсек, О. П. Нарєжний. Модель постквантової інфраструктури відкритих ключів з використанням квантового генератора випадкових чисел для автономної системи диференціальної корекції глобальних навігаційних супутникових систем.....	235
Т. А. Tkachenko, V. O. Martovitsky, N. M. Bolohova. Methods of optimising queries in distributed databases.....	245
Р. О. Ткачук, О. В. Муляревич, С. В. Василюк-Зайцева. Адаптивні алгоритми обробки природної мови для багатомовного інтерфейсу користувача.....	252
О. І. Федюшин, І. К. Ткачук, Р. Ю. Гвоздьов, В. В. Стаднік. Динамічне керування правами доступу на основі поведінкових моделей користувачів і сервісів.....	261
О. І. Федюшин, П. В. Шулік, О. М. Товма, В. П. Коцюба. Виявлення аномалій у мережевому трафіку веб-додатків з використанням алгоритму Random Forest.....	269
І. С. Чепурна. Модель збору метрик обчислювальних ресурсів блокчейн-орієнтованої інформаційної системи у хмарному середовищі	276
В. М. Чижик, М. В. Харченко, А. О. Твердохліб. Використання блокчейн-технологій для прозорості ланцюгів постачання в транспортній логістиці.....	285
Я. В. Чумаченко. Розробка комп'ютерних алгоритмів в проєктах дослідження хвильових процесів у дискретних середовищах.....	294
В. П. Шаповал, Я. В. Тарасенко. Метод інтелектуального вимірювання психолінгвістичних показників для динамічних систем безперервного моніторингу психологічного стану.....	301
О. М. Шушура, Б. О. Сухенко, С. О. Соломаха, Л. А. Асєєва. Інтеграція програмних реєстраторів розрахункових операцій з ERP-системами на прикладі ODOO.....	309

СФЕРА ОБСЛУГОВУВАННЯ

О. О. Долинська, І. М. Шоробура, Т. Г. Гільберг. QR-коди як інструмент цифрового маркетингу та інформаційно-навігаційної підтримки в сучасному туризмі.....	317
В. М. Подоляк. Проєктування інноваційного комплексу устаткування мобільної кав'ярні з урахуванням ресурсної автономності.....	323

СОЦІАЛЬНІ ТА ПОВЕДІНКОВІ НАУКИ

Т. Ю. Осадча. Макроекономічні детермінанти розвитку підприємництва в Україні: аналіз поточного стану та прогностичні орієнтири.....	330
С. І. Пастушенко, Г. Д. Берегова. Університет як синтез добірних практик вищої професійної та гуманітарної освіти.....	335
Н. В. Шандова, Н. В. Мєшкова-Кравченко. Оцінка соціально-економічної безпеки підприємства.....	345

CONTENTS

INFORMATION TECHNOLOGIES

S. G. Antoschuk, O. M. Ivanova. Fast localization of FPGA program code elements suitable for steganographic storage of additional data.....	9
O. Yu. Barkovska. Formal description of interaction and data flows in multimodal assistive systems for user autonomy support.....	15
M. E. Bondarenko, H. S. Ivashchenko. Adaptive load distribution in voice identification systems.....	21
M. O. Vernik, L. M. Oleshchenko. Software methods and architectural solutions for big data processing via quantum neural networks and metaheuristic optimizers.....	29
I. V. Havrysh, A. G. Kryvokhata. Development of the architecture of an adaptive interface of the editor of a web application for creating virtual tours.....	37
T. O. Hrinenko, D. M. Maliarova, O. P. Nariezhnii. Prototype of a quantum random number generator web service for secure message exchange.....	47
K. I. Dorofieieva, O. V. Sievierinov, Z. M. Sydorenko, V. M. Sukhoteplyi. Using a security analysis tool to identify critical vulnerabilities in web applications.....	62
L. O. Dretych. Lane segmentation to optimize traffic flow.....	69
M. S. Dutchak, M. V. Pikulyak, R. O. Tychkovskiy, I. B. Polatayko. Modeling the software development life cycle of intelligent learning systems.....	81
Ya. R. Diakov. A hierarchical semantic knowledge graph model for multilevel knowledge representation.....	88
G. A. Zavhorodnia, V. V. Zavhorodnii. Development of a scalable distributed architecture for massively multiplayer online systems.....	99
Yu. Yu. Iliash, V. O. Horielov, V. A. Rovinsky. Algorithms and methods of procedural generation and adaptive difficulty in platformer games.....	107
M. R. Kalashnikov, V. S. Kharchenko. Intelligent approach to enhancing the stability of automated testing systems based on reinforcement learning.....	117
A. O. Kargin, O. I. Pronina, O. Ye. Piatyko, Ye. A. Pertsev. Design technology of the artificial intelligence that supports autonomous distributed systems.....	126
O. V. Krainiuk, Yu. V. Buts, M. M. Piksasov, N. M. Didenko, B. O. Pohodenko. Functional modeling and algorithmization of processes for ensuring cyber resilience of automated dispensing systems for personal protective equipment.....	137
Ya. A. Martynenko, O. V. Sievierinov, A. M. Yevheniev, M. S. Skybenko. Research on SOAR platforms and justification for choosing Splunk Phantom.....	145
D. M. Nekhoroshykh, P. V. Buslov, A. M. Yevheniev, M. S. Skybenko. Analysis of frameworks and technical security measures for the infrastructure of the internet of robotic things.....	153
A. M. Onyshchenko, O. P. Ostapenko, A. A. Loishyn. Integrated information models of sustainable development management based on environmental and economic analysis.....	161
O. S. Parkhomenko, V. V. Koval. The evolution of fraudulent content in the context of the development of information technology and AI.....	171
V. Perederenko. Android project migration methodology based on MVVM project review to Jetpack Compose and MVI architecture.....	175
B. R. Popovych, R. B. Popovych. Discrete problem of constructing high order elements in multiplicative group of finite field.....	182
D. M. Ratushnyi, R. A. Rudenskiy. Architectural approaches to selecting a technology stack for IoT data storage in logistics management tasks.....	187
I. V. Ruban, V. M. Tkachov. A method of robust variational optimization of operational trajectories of an information system on a mobile platform under destructive impacts.....	195
V. I. Ruzhentsev. Analysis of cryptographic resistance of iterative algorithms based on consideration of iterative differential characteristics.....	213
S. A. Sgadov. Analysis of STM32 family support in the .NET eco-system and perspectives for community targets development.....	220

I. P. Sinitsyn, S. V. Zadvornyi, O. P. Ilina. Models and methods for supporting user-centered interface design of defense information systems using machine learning	228
O. A. Sniosek, O. P. Nariezhnii. Model of a post-quantum public key infrastructure using a quantum random number generator for an autonomous differential correction system of global navigation satellite systems.....	235
T. A. Tkachenko, V. O. Martovitsky, N. M. Bolohova. Methods of optimising queries in distributed databases.....	245
R. O. Tkachuk, O. V. Muliarevych, S. V. Vasylyuk-Zaitseva. Adaptive natural language processing algorithms for multilingual user interfaces.....	252
O. I. Fediushyn, I. K. Tkachuk, R. Yu. Gvozdoz, V. V. Stadnik. Dynamic access control management based on behavioral models of users and services.....	261
O. I. Fediushyn, P. V. Shulik, O. M. Tovma, V. P. Kotsiuba. Anomaly detection in web application network traffic using the Random Forest algorithm.....	269
I. S. Chepurna. Model of computational resource metric collection for a blockchain-oriented information system in a cloud environment.....	276
V. M. Chyzhyk, M. V. Kharchenko, A. O. Tverdokhlib. Using blockchain technologies to ensure transparency in supply chains within transportation logistics.....	285
Ya. V. Chumachenko. Development of computational algorithms for the study of wave processes in discrete media.....	294
V. P. Shapoval, Ya. V. Tarasenko. Method for intelligent measurement of psycholinguistic indicators for dynamic systems of continuous psychological state monitoring.....	301
O. M. Shushura, B. O. Sukhenko, S. A. Solomakha, L. A. Asieieva. Integration of software fiscal registrars with ERP systems on the example of ODOO	309

SERVICE SECTOR

O. O. Dolynska, I. M. Shorobura, T. H. Hilberh. QR-codes as a tool of digital marketing and information-navigational support in modern tourism.....	317
V. M. Podoliak. Designing an innovative equipment complex for a mobile coffee shop considering resource autonomy.....	323

SOCIAL AND BEHAVIORAL SCIENCES

T. Yu. Osadcha. Macroeconomic determinants of entrepreneurship development in Ukraine: analysis of the current situation and forecast guidelines.....	330
S. I. Pastushenko, H. D. Berehova. University as a synthesis of selected practices of higher professional and humanitarian education.....	335
N. V. Shandova, N. V. Mieshkova-Kravchenko. Assessment of socio-economic security of an enterprise.....	345

ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

УДК 004.056.53

DOI <https://doi.org/10.35546/kntu2078-4481.2025.4.3.1>

С. Г. АНТОЩУК

доктор технічних наук, професор, директор
Інститут комп'ютерних систем
Національного університету «Одеська політехніка»
ORCID: 0000-0002-9346-145X

О. М. ІВАНОВА

старший викладач кафедри комп'ютерних систем
Національний університет «Одеська політехніка»
ORCID: 0000-0002-4743-6931

ШВИДКА ЛОКАЛІЗАЦІЯ ОСЕРЕДКІВ ПРОГРАМНОГО КОДУ FPGA, ПРИДАТНИХ ДЛЯ СТЕГАНОГРАФІЧНОГО ЗБЕРІГАННЯ ДОДАТКОВИХ ДАНИХ

Статтю присвячено використанню стеганографічного підходу до зберігання контрольних даних, які використовуються в процесі прихованого моніторингу характеристик безпеки програмного коду мікросхем FPGA. За такого підходу контрольні дані утворюють з інформаційним об'єктом програмного коду єдине ціле, а сам факт присутності цих даних в програмному коді не є очевидним. Традиційно для стеганографічного вбудовування додаткових даних в програмний код FPGA застосовують методи, що базуються на еквівалентних перетвореннях елементів програмного коду. Однак обсяг зберігання даних, забезпечений зазначеними методами часто є недостатнім для одночасного зберігання контрольних даних декількох видів моніторингу. Інша група методів стеганографічного зберігання даних базується на нееквівалентних перетвореннях програмного коду. Ці методи застосовуються для обчислювачів, реалізованих в середовищі FPGA, які виконують наближену обробку даних. В структурі зазначених обчислювачів мають місце елементарні блоки FPGA, які обчислюють тільки несуттєві розряди результатів наближених операцій. Зміна в нееквівалентний спосіб програмного коду таких блоків не призводить до некоректного функціонування пристрою. Відомі методи даного класу ґрунтуються на моделюванні, яке знаходять всі несуттєві блоки, але має значну часову складність. В даній роботі пропонується підхід, який дозволяє знайти неповну множину несуттєвих блоків, але зі значним скороченням часової складності. Експериментальне дослідження запропонованого підходу показує, що його використання дозволяє забезпечити додатковий до методів еквівалентних перетворень обсяг стеганографічного зберігання контрольних даних в середовищі програмного коду FPGA. Пропозиції роботи можуть бути застосовані до обчислювачів, розмірність яких не дозволяє виконати локалізацію несуттєвих існуючими методами.

Ключові слова: стеганографічне зберігання даних, FPGA, наближена обробка даних, операції над числами з плаваючою крапкою, блоки LUT.

S. G. ANTOSCHUK

Doctor of Technical Sciences, Professor, Director
Institute of Computer Systems
of Odesa Polytechnic National University
ORCID: 0000-0002-9346-145X

O. M. IVANOVA

Senior Lecturer at the Department of Computer Systems
Odesa Polytechnic National University
ORCID: 0000-0002-4743-6931

FAST LOCALIZATION OF FPGA PROGRAM CODE ELEMENTS SUITABLE FOR STEGANOGRAPHIC STORAGE OF ADDITIONAL DATA

This article is devoted to the use of a steganographic approach to storing control data used in the process of hidden monitoring of the security characteristics of FPGA program code. With this approach, the control data forms a single whole with the information object of the program code, and the very fact of the presence of this data in the program code is not obvious. Traditionally, methods based on equivalent transformations of software code elements are used for steganographic embedding of additional data into FPGA program code. However, the volume of data storage provided by such methods is often insufficient for the simultaneous storage of control data for several types of monitoring. Another group of steganographic data storage methods is based on non-equivalent transformations of program code. These methods are used for computing elements implemented in an FPGA environment that perform approximate data processing. The structure of these computing elements includes elementary FPGA units that calculate only insignificant digits of the results of approximate operations. The program code of such units can be changed in a non-equivalent way without causing the device to malfunction. Known methods of this class are based on simulation, which finds all insignificant units, but they are significantly time complexity. This paper proposes an approach that allows finding an incomplete set of insignificant blocks, but with a significant reduction in time complexity. An experimental study of the proposed approach shows that its use allows for an increase in the volume of steganographic storage of control data in the FPGA software code environment. The proposals of this work are applicable to computing elements whose complexity does not allow for the localization of insignificant elementary units using existing methods.

Key words: steganographic data storage, FPGA, approximate data processing, operations on floating-point numbers, LUT units.

Постановка проблеми

Мікросхеми FPGA є програмно-керованими цифровими пристроями. Подібно мікропроцесорам та мікроконтролерам їх поведінка може змінюватися під впливом програмного коду. На відміну від мікропроцесорів, програмний код змінює структуру зав'язків та функції елементарних блоків FPGA тим самим налаштовуючи поведінку цих пристроїв. Паралельна структура FPGA створює можливість для отримання значно більшої продуктивності обчислень ніж продуктивність, доступна для мікропроцесорів.

FPGA є типовою елементною базою для високопродуктивних обчислень в галузях критичного застосування: енергетика, швидкісний транспорт, комунікації [1]. Критичність цих галузей підвищує ризики, пов'язані з зловмисним втручанням в функціонування комп'ютерних систем на базі FPGA. Втручання на апаратному рівні може бути реалізоване при наявності фізичного доступу до комп'ютерної систем. В той же час для втручання на програмному рівні достатньо обмеженого фізичного або віддаленого доступу. Зазначені фактори обумовлюють важливість процесів забезпечення цілісності та автентичності програмного коду FPGA для систем підвищеного ризику.

Одним з ефективних підходів до запобігання втручання в функціонування систем на базі FPGA є оперативний моніторинг цілісності та автентичності програмного коду цих мікросхем [2]. Однак проблемним моментом процедур моніторингу є місце та спосіб зберігання контрольних даних, які використовуються в процесі моніторингу. В даній роботі розглядається підхід до зберігання контрольних даних, який полягає в прихованому їх вбудовуванні в програмний код FPGA в стеганографічний спосіб.

Аналіз останніх досліджень і публікацій

Одним із ефективних підходів до зберігання контрольних даних є стеганографічний підхід. Найбільшого свого розвитку цей підхід отримав в застосуванні до мультимедійних стеганографічних інформаційних контейнерів, таких як растрові зображення, цифрове відео та цифровий звук [3]. Це пояснюється тим, що зазначені контейнери мають аналогову природу походження, а їх елементарних одиниць (пікселі, семпли) мають наближене подання. В результаті незначні спотворення цих даних не тільки не призводять до деградації контейнера, але й зазвичай не реєструється.

На відміну від мультимедійних інформаційних контейнерів, програмний код (і зокрема програмний код мікросхем FPGA) є інформаційним контейнером з точно поданими даними. Спотворення бітів програмного коду призводить до спотворення функціонування пристрою, що програмується цим кодом. Тому для стеганографічного вбудовування додаткових даних у програмний код FPGA зазвичай використовують методи еквівалентних перетворень [4]. Ці методи не змінюють функціонування мікросхеми FPGA та обсяг програмного коду, проте дозволяють приховано вбудувати в цей код додаткові дані.

Методи стеганографічного вбудовування даних у програмний код FPGA, які використовують еквівалентні перетворення є практичними, проте забезпечують малий, потенційно доступний для зберігання обсяг додаткових даних. Це часто призводить до дефіциту обсягу для зберігання контрольних даних у задачах оперативного моніторингу програмного коду FPGA, а також зменшує кількість різновидів моніторингу, що можуть бути застосовані до програмного коду.

Але для точно поданого інформаційного контейнера, яким є програмний код FPGA, також можливе стенографічне вбудовування додаткових даних, подібне до того, що використовується в інформаційних контейнерах з наближено поданими елементарними одиницями. Зазначена можливість полягає в тому, що при реалізації наближеної обробки даних на FPGA, ця наближеність опосередковано переноситься на точно поданий програмний код. При виконанні наближених обчислень частина елементарних блоків FPGA беруть участь у обчисленні лише несуттєвих розрядів результатів обчислень [5]. Програмний код таких елементарних блоків може бути змінений в нееквівалентний спосіб для вбудовування додаткових даних. Причому таке нееквівалентне вбудовування не відбивається на коректності виконання функції пристрою.

Основними елементарними блоками структури мікросхем FPGA є блоки LUT (Look Up Table). У сучасних серіях FPGA кількість таких блоків становить мільйони та десятки мільйонів. Блок LUT має m входів та виконує обчислення однієї логічної функції від m змінних. Блоки LUT є програмованими блоками та налаштовуються на реалізацію конкретної логічної функції 2^m бітним двійковим програмним кодом. При реалізації зазначених вище наближених арифметичних операцій на FPGA в її структурі можливо виділити елементарні обчислювальні блоки LUT, які беруть участь у формуванні тільки несуттєвих розрядів результату. В подальшому такі блоки LUT будуть навіватися несуттєвими блоками. В силу цього програмний код таких блоків може бути модифікований і така модифікація не буде мати впливу на результат роботи пристрою та його параметри.

В роботах [6, 7] запропоновано підхід до пошуку несуттєвих блоків LUT в структурі FPGA-проектів, які мають в своєму складі модулі наближеної обробки даних. Цей підхід дозволяє достатньо швидко знаходити несуттєві блоки для обчислювачів, в яких сукупна розрядність операндів не перевищує 34-36 розрядів. Для пошуку несуттєвих блоків в обчислювачах з більшою розрядністю пошук може потребувати значних часових ресурсів. Враховуючи цей недолік існуючого підходу, в даній статті пропонується не базований на переборі варіантів спосіб пошуку несуттєвих блоків, який має меншу часову складність та здатність виявити більшу частину несуттєвих блоків.

Формулювання мети статті

Розробити підхід до не базованого на переборі варіантів пошуку несуттєвих блоків LUT в структурі FPGA-проектів, що реалізують наближену обробку даних. На основі цього підходу оцінити обсяг даних, які можливо вбудувати в програмний код FPGA в стеанографічний спосіб, використовуючи несуттєві блоки LUT.

Викладення основного матеріалу дослідження

В роботах [6, 7] запропоновано для локалізації несуттєвих блоків LUT виконувати пошук суттєвих блоків, інвертування значень на виході яких дає помилки в суттєвих розрядах результату обчислень. Однак цей підхід дозволяє швидко локалізувати потрібні блоки тільки в складі обчислювачів сукупна розрядність операндів яких не перевищує 34-36 розрядів. В даній роботі пропонується підхід, що не має зазначеного обмеження. Підхід базується на врахуванні структурних та арифметичних особливостей побудови обчислювачів, в складі яких виконується пошук несуттєвих блоків.

За умови вимоги про відповідність формату результату форматам операндів, обчислювач при наближеній обробці даних в частині обробки мантис за n -розрядними операндами отримує повний $2n$ -розрядний результат з послідовним округленням та відкиданням молодших n -розрядів. Відомим є метод скорочення структури матриці помноження мантис для описаного випадку [8]. Відповідно до цього методу матриця помноження розбивається на старшу та молодшу частини. Молодша частина матриці складається з k -стовбців. Метод скорочення структури матриці помноження обумовлює виключення молодшої частини матриці зі схеми. При цьому старша частина матриці обчислює усічений $2n-k$ -розрядний результат $V(2n \dots k+1)$, старші n -розрядів якого складають кінцевий результат $V(2n \dots n+1)$, формат якого відповідає формату операндів.

В зазначеному методі [8] величина k визначається за умови відсутності впливу молодшої частини схеми матриці на округлений кінцевий результат $V(2n \dots n+1)$. Значення k розраховуються виходячи з максимального значення складової повного результату, яку створює молодша частина матриці. Таке максимальне значення утворюється коли вся молодша частина матриці кон'юнкцій добутку заповнена одиницями. Відповідно до положень методу [8] значення k приближено оцінюється за формулою:

$$k = n - \log_2 n;$$

де n – розмірність операндів та кінцевого результату.

Зазначені положення методу скорочення матриці помноження мантис пропонується використати для встановлення нижньої оцінки кількості несуттєвих блоків LUT та їх локалізації в структурі обчислювача. Блоки LUT, які відповідають молодшим k -стовбцям матриці в структурі обчислювача є несуттєвими через відсутність впливу значень на виходах цих блоків на округлений кінцевий результат. Будь-які спотворення значень на виходах зазначених блоків не приводять до змін в суттєвих розрядах результату та розрядах кінцевого результату обчислення. Аналіз структури матриці обчислювача у вигляді списків блоків LUT, а також списків зв'язків цих блоків між собою і з входами та виходами обчислювача дає можливість виявити сукупність блоків, які відносяться до молодшої частини матриці обчислювача.

Для визначення множини суттєвих блоків LUT та оцінки кількості несуттєвих блоків в роботах [6, 7] було запропоновано та розроблено два програмних додатки:

1. Додаток LUTListExtractor, призначення якого полягає в добуванні детальної інформації про структуру FPGA-проєкту з внутрішньої бази даних САПР Intel/Altera Quartus Prime. Ця детальна інформація містить список блоків LUT FPGA-проєкту, значення програмних кодів цих блоків і список зв'язків блоків LUT між собою.

2. Програмний додаток EsLUTGetter, який приймає від додатка LUTListExtractor інформацію про сукупність блоків LUT, їх зв'язки в FPGA-проєкті та виконує моделювання LUT-схеми з метою визначенні кількості суттєвих блоків LUT, що дає змогу оцінити кількість несуттєвих блоків LUT.

Для задач даної роботи було розроблено програмний додаток LMatrixLUTGetter, який пропонується долучити до набору додатків з локалізації несуттєвих блоків LUT (рис. 1). Цей додаток приймає від додатка LUTListExtractor інформацію про сукупність блоків LUT та їх зв'язки в FPGA-проєкті, аналізує структуру матриці обчислювача, визначає розмір молодшої частини матриці в структурі обчислювача та виконує пошук блоків LUT, які відносяться до молодшої частини матриці. За результатом функціонування додатка LMatrixLUTGetter отримується список несуттєвих блоків LUT, які розміщені в молодшій частині матриці обчислювача для кожного FPGA-проєкта. На основі аналізу цього списку виконується нижня оцінка кількості несуттєвих блоків LUT.

В середовищі вказаних програмних додатків було виконано експериментальне дослідження з локалізації несуттєвих блоків LUT в молодшій частині матриці обчислювача. Вихідні дані експерименту становили п'ять FPGA-проєктів, які складаються з модулів помноження мантис та мають розрядність кожного з операндів відповідно 4, 6, 8, 12 та 16 розрядів та розрядність до округлення відповідно 8, 12, 16, 24 та 32 розряди. Після округлення розрядність результату збігається з розрядністю операндів. Синтез, розміщення і трасування FPGA-проєктів виконувалися в САПР Intel Quartus Prime 20.1 Lite Edition для цільових мікросхем FPGA Intel Cyclone IV EP4CE15F23A7.

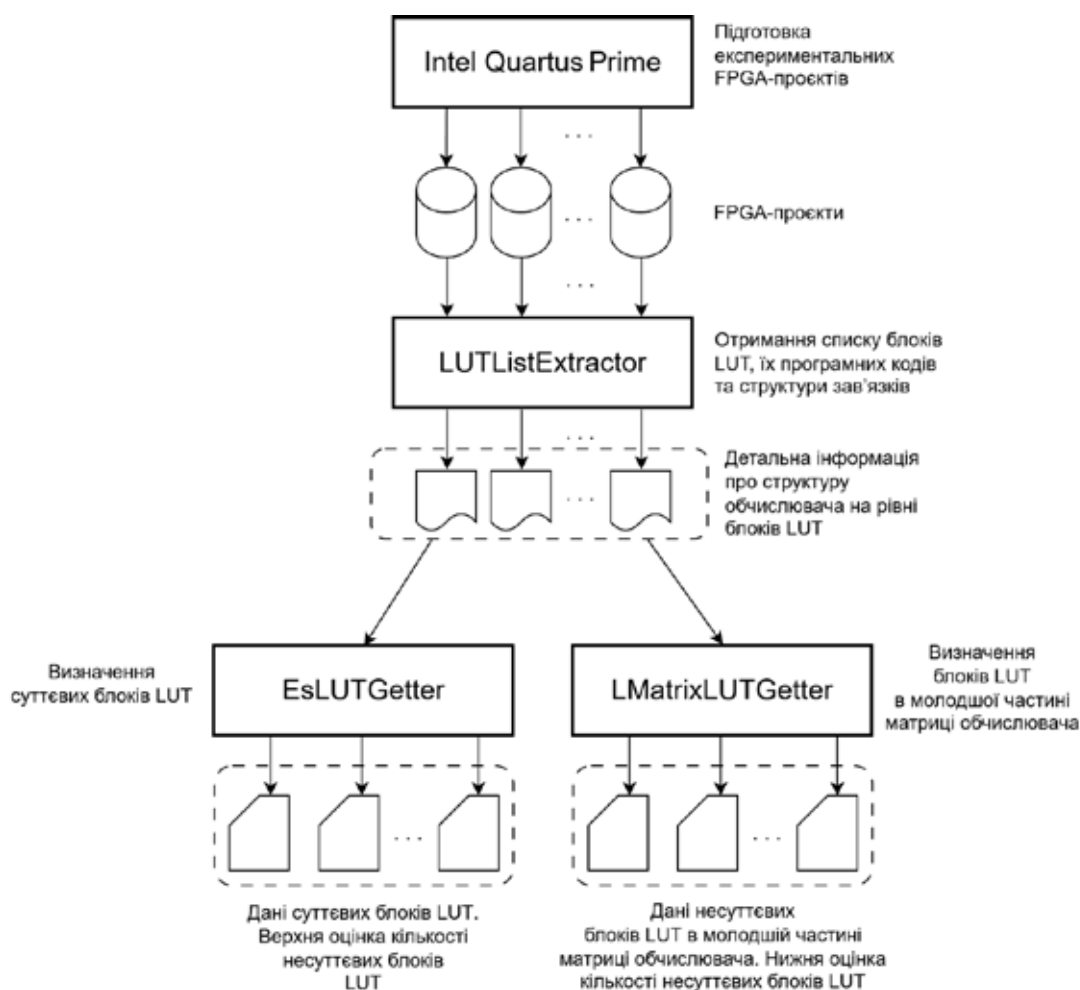


Рис. 1. Зв'язки між програмними додатками, які становлять середовище проведення експерименту з локалізації несуттєвих блоків LUT

Результати експериментального дослідження з визначення нижньої оцінки кількості несуттєвих блоків LUT та їх локалізації представлені в табл. 1. В експериментальних FPGA-проектах нижня оцінка частки несуттєвих блоків LUT лежить в діапазоні 16,67% ... 29,91% від загальної кількості блоків LUT проектів. Обсяг, доступний для зберігання даних в стеганографічний спосіб, забезпечений несуттєвими блоками LUT, за отриманою нижньою оцінкою змінюється від 5 до 102 бітів та від 80 до 1632 бітів при використанні для вбудовування стего-даних відповідно тільки одного або всіх розрядів програмного коду несуттєвих блоків LUT. В ході проведення експериментального дослідження на обчислювальній системі на базі 8-ядерного процесора AMD Ryzen 7 при обсязі оперативної пам'яті 16 ГБ часова складність пошуку блоків LUT в молодшій частині матриці обчислювачів становила від $2.1 \cdot 10^{-5}$ с до $2.6 \cdot 10^{-4}$ с. В той час, як для цих же проектів з застосуванням переборних методів в роботах [6, 7] час моделювання склав від $0.5 \cdot 10^{-4}$ с до 821.75 с.

Таблиця 1

Результати експериментального визначення нижньої оцінки кількості несуттєвих блоків LUT

Розрядність операндів та кінцевого результату	4	6	8	12	16
Розрядність результату до округлення	8	12	16	24	32
Загальна кількість блоків LUT	30	61	101	208	341
Нижня оцінка кількості несуттєвих блоків LUT	5	10	22	51	102
Нижня оцінка частки несуттєвих блоків LUT серед усіх блоків	16,67%	16,39%	21,78%	24,52%	29,91%
Мінімальний обсяг, доступний для зберігання даних (біт)	5	10	22	51	102
Максимальний обсяг (біт), доступний для зберігання даних при кількості входів блоків LUT – 4	80	160	352	816	1632

На комплексній діаграмі (рис. 2) з результатами експериментів з дослідження інформаційних стеганографічних ресурсів, забезпечених несуттєвими блоками LUT, показано порівняння абсолютних, а також відносних значень загальної кількості блоків LUT, кількість експериментально визначених в роботах [6, 7] суттєвих блоків LUT, верхньої та визначеної в даній роботі нижньої оцінки кількості несуттєвих блоків LUT для кожного з експериментальних FPGA-проектів.

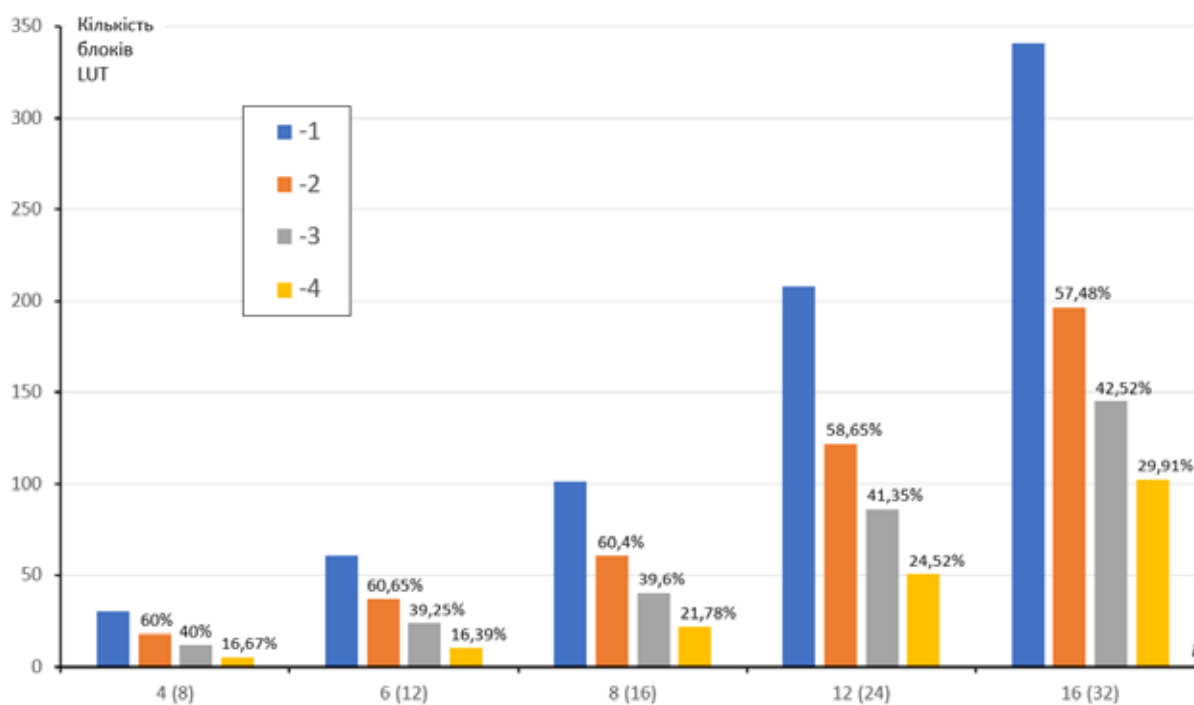


Рис. 2. Результати експериментального дослідження з урахуванням нижньої оцінки кількості несуттєвих блоків LUT

1 – загальна кількість блоків LUT в реалізації FPGA-проекту; 2 – з них кількість експериментально визначених суттєвих блоків LUT; 3 – верхня оцінка кількості несуттєвих блоків LUT; 4 – нижня оцінка кількості несуттєвих блоків LUT

Висновки. В роботі запропоновано підхід швидкої локалізації несуттєвих блоків LUT в структурі обчислювачів, які виконують обробку чисел з плаваючою крапкою та реалізовані на елементній базі FPGA. На відміну

від відомого підходу до локалізації несуттєвих блоків LUT, підхід що запропоновано, має малу обчислювальну складність, але дає змогу визначити тільки нижню оцінку кількості цільових блоків LUT та виконати їх локалізацію. В сукупності з відомими підходами, пропозиції даної роботи дають можливість обрати пошук повної множини несуттєвих блоків LUT при високій обчислювальній складності відомих підходів або пошук множини блоків, кількість елементів якої обмежено нижньою оцінкою кількості при низькій обчислювальній складності.

Виконане експериментальне дослідження показало, що навіть кількість несуттєвих блоків LUT, оцінена за нижньою оцінкою дає обсяг для прихованого збереження додаткових даних в програмному коді FPGA від 5 до 102 бітів при використанні для вбудовування стего-даних тільки одного розряду та від 80 до 1632 бітів при використанні для вбудовування стего-даних всіх розрядів програмного коду несуттєвих блоків LUT. Такі обсяги є співставними з розмірами хеш-сум та похідних кодів, що використовуються в якості контрольних в задачах моніторингу програмного коду. Зазначені обсяги суттєво доповнюють простір для прихованого зберігання контрольних даних та дають можливість забезпечити стеганографічне зберігання контрольних даних більшої кількості видів моніторингу.

Список використаної літератури

1. Tu K., Tang S., Yu X., Josipovic L., Chu Z. *FPGA EDA: Design Principles and Implementation*. Singapore: Springer, 2024. 248 p.
2. Ashutosh M., Goswami M., Manoj K., Rajput N. *Hardware Security: Challenges and Solutions*. Singapore: Springer, 2025. 274 p.
3. Shih F. *Digital Watermarking and Steganography: Fundamentals and Techniques*. 2nd ed. USA, Boca Raton: CRC Press. 2017. 320 p.
4. Drozd A., Antoshchuk S., Drozd J., Zashcholkin K., Drozd M., Kuznietsov N., Al-Dhabi M., Nikul V. Checkable FPGA Design: Energy Consumption, Throughput and Trustworthiness. *Studies in Systems, Decision and Control: Green IT Engineering: Social, Business and Industrial Applications*. Switzerland, Cham: Springer International Publishing, 2019. Vol. 171. P. 73-94.
5. Anderson A., Muralidharan S., Gregg D. Efficient Multibyte Floating Point Data Formats Using Vectorization, *IEEE Transactions on Computers*, vol. 66, no. 12 (2017) 2081-2096. DOI: 10.1109/TC.2017.2716355.
6. Zashcholkin K., Drozd O., Antoshchuk S., Ivanova O., Sachenko O. Steganographic Resources of FPGA-based Systems for Approximate Data Processing. *CEUR-WS*. 2021. Vol. 2864. P. 324-333.
7. Zashcholkin K., Drozd O., Ivanova O., Shaporin R., Kuznietsov M. An Approach to Stego-Container Organization in FPGA Systems for Approximate Data Processing. *CEUR-WS*. 2021. Vol. 2853. P. 527-536.
8. Дрозд О.В. Теоретичні основи, методи та засоби функціонального діагностування вузлів обчислювальних пристроїв з використанням природної надмірності при виконанні приблизних обчислень: дис. ... д-ра техн. наук: 05.13.05. Одеса, 2003. 327 с.

References

1. Tu, K., Tang, S., Yu, X., Josipovic, L., Chu, Z. (2024). *FPGA EDA: Design principles and implementation*. Springer.
2. Ashutosh, M., Goswami, M., Manoj, K., Rajput, N. (2025). *Hardware security: Challenges and solutions*. Springer.
3. Shih, F. (2017). *Digital watermarking and steganography: Fundamentals and techniques* (2nd ed.). CRC Press.
4. Drozd, A., Antoshchuk, S., Drozd, J., Zashcholkin, K., Drozd, M., Kuznietsov, N., Al-Dhabi, M., Nikul, V. (2019). Checkable FPGA design: Energy consumption, throughput and trustworthiness. In *Studies in Systems, Decision and Control: Green IT Engineering: Social, Business and Industrial Applications* (Vol. 171, pp. 73-94). Springer International Publishing.
5. Anderson, A., Muralidharan, S., Gregg, D. (2017). Efficient multibyte floating point data formats using vectorization. *IEEE Transactions on Computers*, 66(12), 2081-2096. <https://doi.org/10.1109/TC.2017.2716355>
6. Zashcholkin, K., Drozd, O., Antoshchuk, S., Ivanova, O., Sachenko, O. (2021). Steganographic resources of FPGA-based systems for approximate data processing. *CEUR Workshop Proceedings*, 2864, 324-333.
7. Zashcholkin, K., Drozd, O., Ivanova, O., Shaporin, R., Kuznietsov, M. (2021). An approach to stego-container organization in FPGA systems for approximate data processing. *CEUR Workshop Proceedings*, 2853, 527-536.
8. Drozd, O.V. (2003). *Theoretic Bases, Methods and Means for On-line Testing of Computing Devices Units with Use of Natural Redundancy in Approximated Calculations Execution* (Doctoral dissertation).

Дата першого надходження рукопису до видання: 10.11.2025
Дата прийнятого до друку рукопису після рецензування: 08.12.2025
Дата публікації: 31.12.2025

O. YU. BARKOVSKA

Candidate of Technical Sciences, Associate Professor,
Associate Professor at the Department of Electronic Computers
Kharkiv National University of Radio Electronics
ORCID: 0000-0001-7496-4353

FORMAL DESCRIPTION OF INTERACTION AND DATA FLOWS IN MULTIMODAL ASSISTIVE SYSTEMS FOR USER AUTONOMY SUPPORT

The growing number of people with disabilities, particularly among war veterans, generates a pressing demand for technological solutions capable of providing support in everyday activities, mobility, education, and information access. This paper presents a formalized description of multimodal data flows within context-aware assistive interaction systems designed for users with special needs.

A complete data processing cycle is described – from the initial input (including voice commands, visual cues, spatial parameters, and physiological indicators) to the generation of an adaptive system response, tailored to the current context such as environment, user condition, and interaction history. Special attention is given to the synchronization of input channels and prioritization strategies in cases of incomplete or noisy data.

Representative use-case contexts are explored (e.g., spatial navigation, information requests, task execution), in which the system must adapt its output in the form of voice prompts, visual guidance, or haptic feedback. A functional model is proposed, outlining the key stages of processing: identification of the user's state, dynamic update of their ergonomic profile, and selection of the most relevant interaction scenario.

The paper also addresses technological constraints, including sensor data quality, influence of external conditions (lighting, noise, spatial obstacles), computational demands, and the need for user-specific personalization.

The proposed description provides a foundation for further development and implementation in multimodal systems for indoor navigation, independent living support, educational platforms, and information access services for individuals with disabilities.

Key words: multimodal systems, assistive technologies, context-aware interaction, people with disabilities, personalized interfaces.

О. Ю. БАРКОВСЬКА

кандидат технічних наук, доцент,
доцент кафедри електронних обчислювальних машин
Харківський національний університет радіоелектроніки
ORCID: 0000-0001-7496-4353

ФОРМАЛЬНИЙ ОПИС ВЗАЄМОДІЇ ТА ПОТОКІВ ДАНИХ У МУЛЬТИМОДАЛЬНИХ АСИСТИВНИХ СИСТЕМАХ ДЛЯ ПІДТРИМКИ АВТОНОМНОСТІ КОРИСТУВАЧІВ

Зростаюча кількість людей з інвалідністю, особливо серед ветеранів війни, формує запит на технологічні рішення, здатні забезпечити підтримку у побутовій діяльності, мобільності, навчанні та доступі до інформації. У статті формалізовано потік мультимодальних даних у контекстно-залежних системах асистивної взаємодії для користувачів з особливими потребами.

Описано повний цикл обробки даних – від моменту отримання вхідної інформації (включаючи голосові команди, візуальні сигнали, просторові параметри, фізіологічні індикатори) до формування адаптивної відповіді, що враховує поточний контекст: середовище, стан користувача, історію взаємодій. Основну увагу приділено синхронізації каналів та визначенню пріоритетів у випадку неповних або зашумлених даних.

Розглянуто приклади контекстів (рух у просторі, запит на інформацію, виконання завдань), у яких система має адаптувати свій вихід – голосовий супровід, візуальні підказки, тактильний фідбек. Запропоновано функціональну модель, що описує етапи обробки: ідентифікація стану користувача, динамічне оновлення його ергономічного профілю, вибір найбільш релевантного сценарію взаємодії.

Особливу увагу приділено технологічним обмеженням: залежності від якості сенсорних сигналів, впливу зовнішніх факторів (освітлення, шум, просторові бар'єри), обчислювальним вимогам та необхідності персоналізації під конкретного користувача.

Запропонований опис може бути основою для подальшої розробки та впровадження в мультимодальні системи навігації, підтримки побутової незалежності, навчальних платформ та сервісів доступу до інформації для людей з інвалідністю.

Ключові слова: мультимодальні системи, асистивні технології, контекстно-залежна взаємодія, люди з інвалідністю, персоналізовані інтерфейси.

Introduction

Before the full-scale invasion, approximately 2.7 million people with disabilities were officially registered in Ukraine (2021). As of 2024–2025, the Ministry of Social Policy and humanitarian agencies report over 3.0–3.4 million registered; estimates of the actual number exceed 6 million (taking into account new injuries, trauma, age, and unregistered cases). Separately, the Ministry of Veterans reports over 130,000 veterans with disabilities, increasing the need for rehabilitation, employment, and technological support. This situation exacerbates an already difficult labor market, where the employment rate of people with disabilities before the war did not exceed 2–3%, making assistive technologies not just tools of support but a key factor in socio-economic recovery and inclusion under martial law and subsequent reconstruction.

Inclusion is the creation of conditions under which people with different abilities have equal access to resources, services, and participation in social life. In the technical (IT) context, inclusion means designing and implementing information technologies that reduce barriers in interaction with the environment, information, and other people.

The key idea of this work is the use of two types of multimodal data to build a context-adaptive support system for people with disabilities:

- human-centered data
- environmental data.

The relevance of assistive technology for people with disabilities in the modern world is driven by a number of powerful factors that affect all areas of social life.

Analysis of the latest research and publications

Analyzing the scenarios, barriers, characteristics of technical solutions, technical requirements, and user features typical for ensuring a comfortable and safe life and activity for people with disabilities and special needs [1], this work formulates the main scenarios that require the development and implementation of scientific and methodological foundations for creating an assistive information technology based on multimodal data analysis (Figure 1).

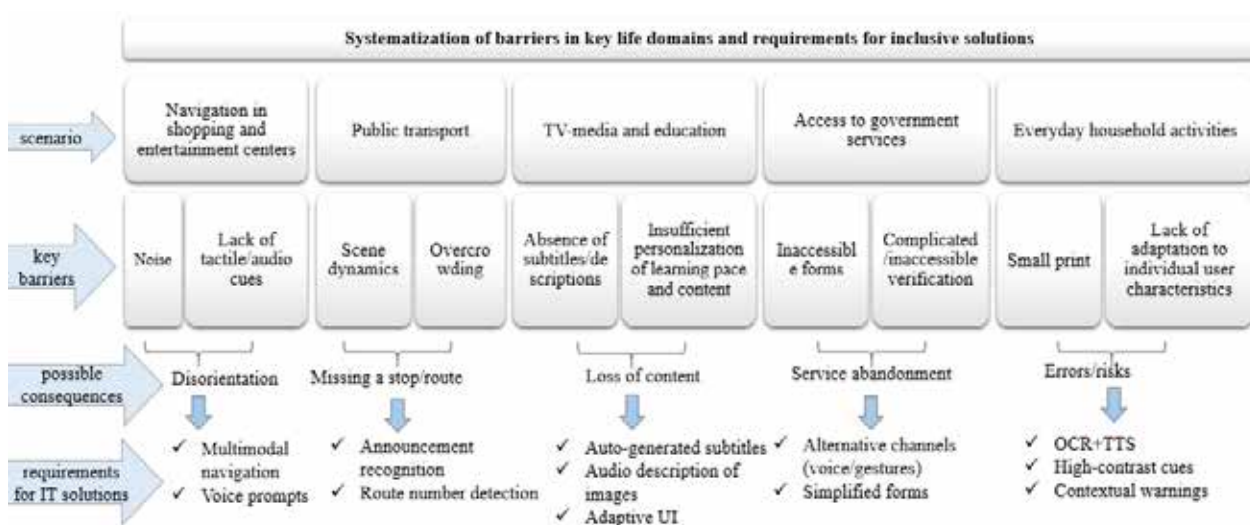


Fig. 1. Ontological map of problems and requirements for assistive and support technologies for people with disabilities

Navigation in shopping and entertainment centers, hospitals, and universities is complicated by the user moving through a complex multi-zone space while searching for stores, exits, and elevators [2,3]. Critical barriers include noise and the absence of clear tactile or audio cues, which lead to disorientation. This can be mitigated by multimodal navigation with an audio guide and/or tactile feedback, complemented by indoor localization for stable positioning within the building.

When using public transport, a person needs to reliably identify the route, avoid missing their stop, and make connections amid noise and crowds; the dynamics of the scene and overcrowding can easily cause confusion, creating the risk of boarding the wrong vehicle or missing a stop. This can be addressed by combining automatic recognition of

announcements and visual route cues (displays, numbers, directions) with geolocation, as well as short audio and vibration prompts with low latency that guide the user during movement and work even offline. However, solving this task is accompanied by challenges due to variable environmental conditions [4].

In telemedicine and education, during online lectures and instructional videos, it is crucial not to lose content due to the absence of subtitles and text descriptions of images; when these formats are missing, part of the material simply “drops out” [5]. Accessibility is restored through real-time auto-generated subtitles, captioning for graphics and slides, as well as an adaptive interface with adjustable fonts, contrast, and keyboard navigation.

In access to digital services, complex or inaccessible verification procedures stop the user midway and often result in abandoning the service. The problem is addressed by alternative interaction channels (voice, gestures, simplified text mode), simplified forms with clear step-by-step guidance, and accessible methods of identity verification without barrier-based CAPTCHAs [6].

In daily household activities, reading small labels or instructions causes errors and risks, especially under low lighting. Assistance is provided by instant text-to-speech via OCR+TTS, high-contrast on-screen cues, and contextual warnings about dangerous actions or incorrect choices.

Formulation of the purpose of the research

The aim of this work is to formalize the architecture and information flow in multimodal assistive systems for context-dependent interaction, supporting autonomy and access to information for people with disabilities, particularly war veterans, under variable environmental conditions and individual user limitations.

To achieve this goal, the following tasks must be addressed:

- analyze the key types of input data (audio, video, spatial, physiological) and usage scenarios for multimodal systems supporting users with special needs;
- describe the functional stages of input data processing in context-dependent interaction systems, from collection to adaptive response;
- identify the main technological constraints and requirements for implementing assistive systems in the areas of daily living, mobility, education, and access to digital services.

Materials and methods

The operation of the developed information technology is aimed at maximizing ergonomic efficiency and functional independence for users with disabilities through the sequential implementation of four logical stages: data collection, preprocessing, analytical processing, and direct user interaction. The diagram below (Figure 2) provides a detailed illustration of the architecture and the sequence of information flow between the key components of the technology, revealing the logic of data transformation from event to action.

Each block in the diagram performs a specific data transformation function. A key advantage is the integration of the context-aware core with a dynamic ergonomic profile, which provides personalized system adaptation to the individual needs of the user and to changing environmental conditions.

At the sensor stage, the technology converts physical phenomena from the environment and the user’s state into digital representations. Hardware sensors (cameras, microphones, LiDAR/ultrasound, light and noise sensors, GPS/BLE, physiological sensors) digitize light waves, sound vibrations, radio signals, and physiological changes. This creates two state sources at time t : the environment C_t and the user U_t . Raw data streams are heterogeneous, redundant, and noisy; they are unsuitable for direct semantic analysis, so the next step is unification and quality enhancement.

At the multisensor preprocessing stage, data are transformed from low-level signal representations into high-level feature representations. The operator $\varphi(\cdot)$ performs noise suppression, exposure normalization, time-series synchronization, channel reliability assessment, and feature extraction (key points, spectral descriptors, depth maps, etc.). The result is a structured feature vector

$$Z_t = \varphi(C_t, U_t),$$

which radically reduces the entropy of the input streams, unifies heterogeneous sensors, and prepares the data for semantic interpretation. Importantly, Z_t preserves the temporal synchronization of different modalities-this is critical for subsequent fusion.

Next, the transition from data to information occurs. The context-aware core $C(\cdot)$ performs high-level fusion and semantic interpretation of features, forming the context vector

$$Ctx_t = C(Z_t),$$

which contains situation descriptors (indoor/outdoor, room zone, illumination level L_{lux} , background noise N_{dB} , estimated distance to screens/objects $\bar{d}_t$ [7], presence of obstacles, scene density, channel trustworthiness, etc.). At this stage, the interaction risk/criticality level is also calculated:

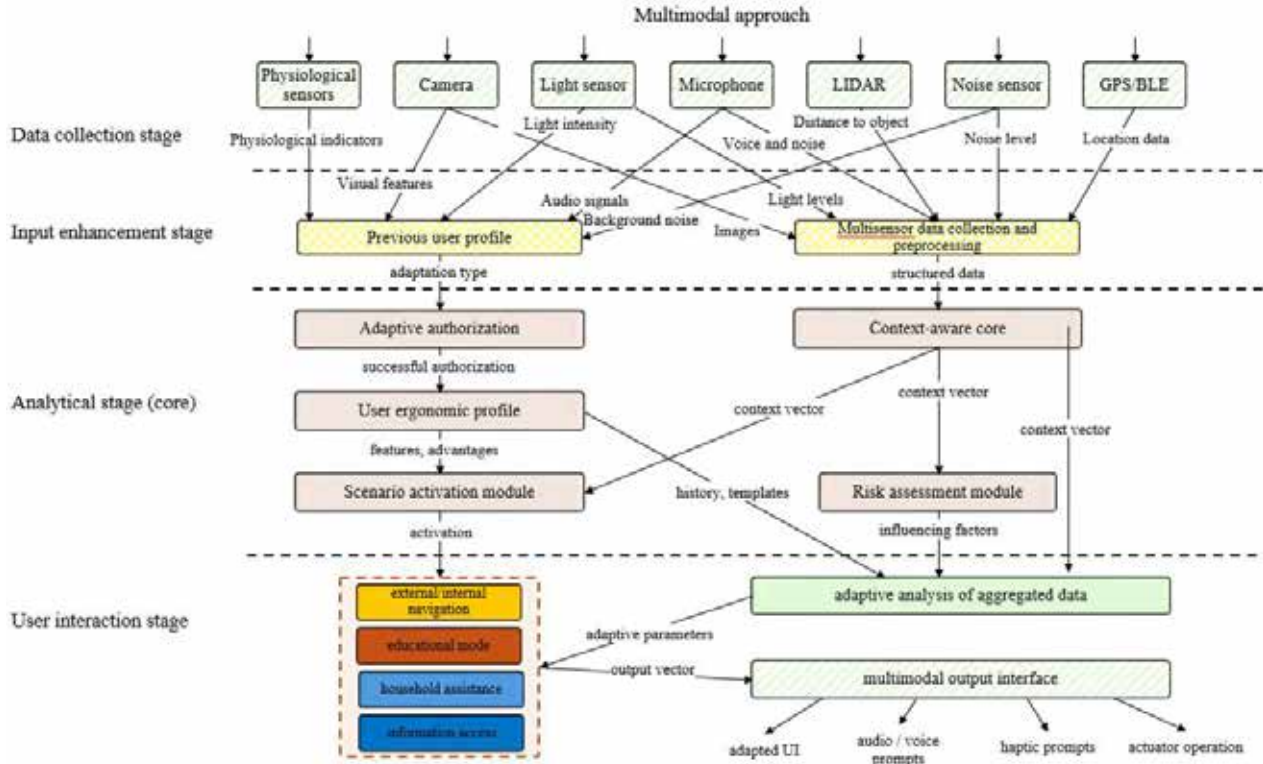


Fig. 2. Activity diagram of data flows in the human-machine interaction technology for people with disabilities.

$$\rho_i = R(Ctx_i),$$

which subsequently affects the output duplication policy (audio + haptics, etc.) to ensure guaranteed user perception.

With the context established, the system proceeds to personalized interaction planning. For each available modality $m \in M$, the utility is calculated:

$$J_m(Ctx_i, E_i) = \alpha q_m(Ctx_i) + \beta Pref_m(E_i) - \gamma Effort_m(Ctx_i, E_i),$$

where E_i – is the user's ergonomic profile (history H_i , priority matrix W_i , biometric templates). The component $q_m(Ctx_i)$ reflects the channel quality in the given context (SNR for voice; illumination/glare for vision; availability of vibrofeedback for haptics), $Pref_m(E_i)$ – represents the individual familiarity/success of the channel based on history $Effort_m$ – denotes the costs (fine motor demands, cognitive complexity, etc.). The selection of the priority modality:

$$m_i^* = \arg \max_{m \in M} J_m(Ctx_i, E_i).$$

This formalizes the principle: “select the interaction method that is simultaneously reliable under current conditions and least demanding for this particular user”.

Once the optimal channel is determined, the policy π_i generates the control (command) vector Cmd_i with parameters for display/audio/haptic output and commands for the “smart home” actuators:

$$Cmd_i = \pi_i(Ctx_i, E_i) = \begin{cases} \{m_i^*, parameters\}, \rho_i < \tau_p, \\ \{m_i^*, m_i', parameters\}, \rho_i \geq \tau_p \text{ (duplication } y \geq 2 \text{ channels)} \end{cases}$$

where τ_p – is the criticality threshold, m_i' – is the backup channel (for example, haptics in addition to audio). For the visual modality a distance-aware UI is applied: to ensure angular legibility $\theta \in [24, 30] \text{ arcmin}$ (recommended for x-height), the size of elements is scaled according to distance $\bar{d}_i$:

$$size_{px} = \frac{2\bar{d}_i \tan(\theta/2)}{mm/px},$$

and is also adjusted by the contrast-sensitivity factor $k(L_{lux}, glare)$. For interactive targets that are pressed or fixed by gaze, the parameters are additionally aligned with Fitts' law (a compromise between target size and distance).

At the output stage, the multimodal interface decodes the command into physically perceivable signals:

$$y_t = O(Cmd_t),$$

which materialize as voice/spatial prompts (TTS), visual changes (adaptive rendering, ROI highlighting), haptic patterns, as well as commands to actuators (lighting, locks, navigation beacons). If $\rho_t \geq \tau_p$, the message is mandatorily duplicated in at least two channels, which increases reliability of perception under noise/glare/stress.

After executing an action, the system collects operational metrics (time, number of repetitions, errors, need for confirmation) and self-learns by updating the profile:

$$E_{t+1} = U(E_t, Ctx_t, y_t, metrics(y_t)).$$

The update $U(\cdot)$ can be implemented through exponential smoothing of the priorities W_t , context-dependent multi-armed bandits (for selecting channels based on reward), as well as incremental learning of biometric templates (voice/gaze/gestures) with drift control. As a result, with each session the cognitive and motor load decreases, and the decisions π_t become increasingly accurate for the specific person in specific contexts.

In summary, the system implements a reproducible cycle:

$$\underbrace{(C_t, U_t)}_{\text{sensors}} \xrightarrow{\varphi} \underbrace{Z_t}_{\text{sensors}} \xrightarrow{C} \underbrace{Ctx_t, \rho_t}_{\text{context, risk}} \xrightarrow{\pi_t} \underbrace{Cmd_t}_{\text{action plan}} \xrightarrow{O} \underbrace{y_t}_{\text{multimodal output}} \xrightarrow{U} \underbrace{E_{t+1}}_{\text{self-learning}}$$

Each module has a clear function and defined parameters, with transparent dependencies between levels: context and risk depend on the quality of sensor data and preprocessing; policy depends on context and the personal profile; output depends on policy and the criticality threshold; the profile depends on the history of successes under various conditions. This multilayered structure ensures context-aware, personalized, and reliable interaction for people with disabilities.

Conclusions

The article systematizes the operating principles of multimodal context-aware systems for assistive support of people with disabilities. It formalizes the data flow from sensor input to the generation of adaptive responses that take into account the environment, user state, and task. It is noted that successful functioning of such systems requires synchronization of different channels, flexible updating of the user profile, and robustness to partially lost or noisy data.

The practical significance of the study lies in creating a unified approach to designing multimodal assistive solutions in the areas of rehabilitation, daily navigation, education, and social integration. In the future, the proposed approaches can be applied to the design of interfaces and services supporting voice, visual, and tactile interaction scenarios, adapted to the specific limitations of the user.

Bibliography

1. Cook, A. M., Polgar, J. M., & Encarnação, P. *Assistive technologies: Principles & practice*. Elsevier. 2019.
2. Cowan, R. E., Fregly, B. J., Boninger, M. L., Chan, L., Rodgers, M. M., & Reinkensmeyer, D. J. Recent trends in assistive technology for mobility. *Journal of neuroengineering and rehabilitation*, 2012. 9(1), 20. <https://doi.org/10.1186/1743-0003-9-20>
3. Barkovska, O., Holovchenko, O., Storchai, D., Kostin, A., & Lehezin, N. Investigation of computer vision techniques for indoor navigation systems. *Сучасний стан наукових досліджень та технологій в промисловості*, 2025. (2(32), 5–15. <https://doi.org/10.30837/2522-9818.2025.2.005>
4. Serdechnyi, V., Barkovska, O., Kovalenko, A., Havrashenko, A., & Martovytskyi, V. Research on machine learning methods for detecting objects in difficult shooting conditions. *Radioelectronic and Computer Systems*, 2025, (2), 64-77. <https://doi.org/10.32620/reks.2025.2.04>
5. McNicholl, A., Casey, H., Desmond, D., & Gallagher, P. The impact of assistive technology use for students with disabilities in higher education: a systematic review. *Disability and rehabilitation: assistive Technology*, 2021. 16(2), 130-143. <https://doi.org/10.1080/17483107.2019.1642395>
6. Barkovska, O., Ruban, I., Tymoshenko, D., Holovchenko, O., & Yankovskyi, O. Research on mobile machine learning platforms for human gesture recognition in human-machine interaction systems. *Technology Audit and Production Reserves*, 2025. 2(82), 6–14. <https://doi.org/10.15587/2706-5448.2025.325423>
7. Barkovska, O., Shapiro, A., Mavrynskyi, O., & Zhebin, P. Research on the sensitivity of the distance measurement method based on the facemesh algorithm. *Системи управління, навігації та зв'язку*, 2025. 2(80), 76-82. <https://doi.org/10.26906/SUNZ.2025.2.076>

References

1. Cook, A. M., Polgar, J. M., & Encarnação, P. (2019). *Assistive technologies: Principles & practice*. Elsevier.
2. Cowan, R. E., Fregly, B. J., Boninger, M. L., Chan, L., Rodgers, M. M., & Reinkensmeyer, D. J. (2012). Recent trends in assistive technology for mobility. *Journal of neuroengineering and rehabilitation*, 9(1), 20. <https://doi.org/10.1186/1743-0003-9-20>
3. Barkovska, O., Holovchenko, O., Storchai, D., Kostin, A., & Lehezin, N. (2025). Investigation of computer vision techniques for indoor navigation systems. *Innovative technologies and scientific solutions for industries*, (2(32), 5–15. <https://doi.org/10.30837/2522-9818.2025.2.005>
4. Serdechnyi, V., Barkovska, O., Kovalenko, A., Havrashenko, A., & Martovytskyi, V. (2025). Research on machine learning methods for detecting objects in difficult shooting conditions. *Radioelectronic and Computer Systems*, (2), 64-77. <https://doi.org/10.32620/reks.2025.2.04>
5. McNicholl, A., Casey, H., Desmond, D., & Gallagher, P. (2021). The impact of assistive technology use for students with disabilities in higher education: a systematic review. *Disability and rehabilitation: assistive Technology*, 16(2), 130-143. <https://doi.org/10.1080/17483107.2019.1642395>
6. Barkovska, O., Ruban, I., Tymoshenko, D., Holovchenko, O., & Yankovskyi, O. (2025). Research on mobile machine learning platforms for human gesture recognition in human-machine interaction systems. *Technology Audit and Production Reserves*, 2(82), 6–14. <https://doi.org/10.15587/2706-5448.2025.325423>
7. Barkovska, O., Shapiro, A., Mavrynskyi, O., & Zhebin, P. (2025). Research on the sensitivity of the distance measurement method based on the facemesh algorithm. *Control, Navigation and Communication Systems*, 2(80), 76-82. <https://doi.org/10.26906/SUNZ.2025.2.076>

Дата першого надходження рукопису до видання: 20.11.2025

Дата прийнятого до друку рукопису після рецензування: 17.12.2025

Дата публікації: 31.12.2025

М. Е. БОНДАРЕНКО

аспірант кафедри електронних обчислювальних машин
Харківський національний університет радіоелектроніки
ORCID: 0000-0002-2500-7626

Г. С. ІВАЩЕНКО

кандидат технічних наук, доцент,
доцент кафедри електронних обчислювальних машин
Харківський національний університет радіоелектроніки
ORCID: 0000-0003-1027-5262

АДАПТИВНИЙ РОЗПОДІЛ ОБЧИСЛЮВАЛЬНОГО НАВАНТАЖЕННЯ В СИСТЕМАХ ГОЛОСОВОЇ ІДЕНТИФІКАЦІЇ

У роботі запропоновано підхід до підвищення ефективності систем голосової ідентифікації шляхом розподілу обчислювального навантаження між центральним (CPU) та графічним (GPU) процесорами. Існуючі реалізації подібних систем ґрунтуються на фіксованих схемах розподілу, у межах яких окремі етапи обробки апріорі асоційовані з CPU або GPU. Така статична організація обчислень не враховує динамічної варіативності параметрів мовного сигналу, поточного стану обчислювальних ресурсів та істотних відмінностей у обчислювальній складності окремих операцій. За таких умов система позбавлена механізмів адаптивного перерозподілу навантаження, що зумовлює асиметричне використання CPU та GPU і, відповідно, призводить до зниження інтегральної продуктивності.

Запропоновано адаптивний алгоритм, який реалізує керування розподілом навантаження на основі аналізу комплексу характеристик мовних записів, що включає довжину кадру, ступінь перекриття, енергетичний вміст та спектральну насиченість сигналу. Використання цих параметрів дозволяє кількісно оцінювати обчислювальну складність поточного сегмента оброблюваного сигналу та динамічно визначати співвідношення обчислювальних операцій, виконуваних на CPU і GPU. Це забезпечує узгоджену взаємодію процесорів, мінімізацію простой і підвищення загальної швидкодії системи.

Проведений порівняльний аналіз показав, що використання адаптивного алгоритму дозволяє суттєво зменшити середній час обробки мовних фрагментів порівняно з підходами, у яких розподіл здійснюється лише за обсягом даних. Запропонований блок адаптивного керування та розподілу навантаження підвищує загальну продуктивність систем голосової ідентифікації, особливо під час обробки великих та структурно складних наборів сигналів, і може бути інтегрований у сучасні багатопроцесорні архітектури.

Ключові слова: системи голосової ідентифікації, MFCC, спектральне віднімання, вейвлет-фільтрація, CPU, GPU, адаптивний розподіл навантаження, керування навантаженням, розподіл задач, динамічне балансування ресурсів, продуктивність у реальному часі, багатопроцесорні системи.

М. Е. BONDARENKO

Postgraduate Student at the Department of Electronic Computers
Kharkiv National University of Radio Electronics
ORCID: 0000-0002-2500-7626

H. S. IVASHCHENKO

Candidate of Technical Sciences, Associate Professor,
Associate Professor at the Department of Electronic Computers
Kharkiv National University of Radio Electronics
ORCID: 0000-0003-1027-5262

ADAPTIVE LOAD DISTRIBUTION IN VOICE IDENTIFICATION SYSTEMS

This paper examines an approach to improving the efficiency of voice identification systems by distributing the computational load between the central processing unit (CPU) and the graphics processing unit (GPU). Existing implementations of such systems are based on fixed distribution schemes, within which individual processing stages are a priori associated with either the CPU or the GPU. Such a static organisation of computations does not take into account

the dynamic variability of speech signal parameters, the current state of computing resources, and significant differences in the computational complexity of individual operations. Under such conditions, the system lacks mechanisms for adaptive load redistribution, resulting in asymmetric use of the CPU and GPU, and consequently, a decrease in overall performance.

The proposed approach involves an adaptive algorithm that implements load distribution control based on the analysis of a set of speech recording characteristics, including frame length, overlap degree, energy content, and spectral saturation of the signal. The use of these parameters enables a quantitative assessment of the computational complexity of the current segment of the processed signal and the dynamic determination of the ratio of computational operations performed on the CPU and GPU. This ensures coordinated interaction between processors, minimises downtime, and increases the overall performance of the system.

The comparative analysis revealed that the use of an adaptive algorithm significantly reduces the average processing time of speech fragments compared to approaches that rely solely on data volume. The proposed adaptive control and load distribution block increases the overall performance of voice identification systems, especially when processing large and structurally complex sets of signals, and can be integrated into modern multiprocessor architectures.

Key words: voice identification systems, MFCC, spectral subtraction, wavelet filtering, CPU, GPU, adaptive load balancing, load management, task distribution, dynamic resource balancing, real-time performance, multiprocessor systems.

Постановка проблеми

Голосова ідентифікація, як різновид біометрії, аналізує індивідуальні акустичні параметри мовного сигналу: частоту основного тону, формантну структуру, амплітудно-частотні характеристики та тембр, що відображають унікальні властивості мовного апарату людини: будову голосових зв'язок, форму резонаторних порожнин і особливості артикуляційних рухів. Голосова біометрія вже активно використовується у банківській сфері, службах підтримки, системах «розумного дому» та мобільних пристроях, забезпечуючи надійну та зручну перевірку особи [1-3].

Складність алгоритмів вилучення, аналізу та розпізнавання біометричних ознак потребує значних обчислювальних ресурсів [4]. Сучасні системи біометричної ідентифікації реалізуються на основі багатоядерних центральних процесорів (CPU), графічних процесорів (GPU) або їх комбінації у вигляді гібридних обчислювальних архітектур [5-6]. GPU забезпечує ефективну паралельну обробку великої кількості фреймів сигналу, що дозволяє скоротити час навчання та розпізнавання, тоді як CPU виконує керування потоками даних, попередню нормалізацію мовного сигналу та синхронізацію між етапами обчислень [7].

Одним із перспективних напрямів підвищення ефективності біометричних систем є керування розподілом обчислювального навантаження між CPU та GPU з урахуванням характеру операцій, ступіню їх паралельності, обсягів даних та завантаження апаратних ресурсів [8].

Аналіз останніх досліджень і публікацій

Розподілення навантаження між CPU та GPU з метою підвищення продуктивності гібридних обчислювальних систем висвітлюється у низці досліджень [9-11].

У [9] представлено підхід до спільної обробки аудіосигналів на центральному та графічному процесорах із використанням середовища OpenCL. Балансування навантаження здійснюється шляхом адаптивного розподілу фреймів або каналів сигналу між процесорами залежно від поточного завантаження CPU. Недоліком описаного підходу є ускладнена оцінка стану GPU, що знижує ефективність розподілу навантаження.

Дослідження [10] присвячено забезпеченню роботи в реальному часі при відеокодуванні H.264/AVC на гібридних системах, що поєднують багатоядерні CPU та GPU. CPU виконує керування, ініціалізацію та передачу даних через шину PCIe, тоді як GPU виконує частини алгоритмів у масовопаралельному режимі. Реалізація базується на технологіях CUDA, OpenMP та інструментах виробників апаратного забезпечення. Розподіл навантаження між CPU і GPU здійснюється адаптивно за допомогою ітераційного алгоритму балансування, який оцінює фактичний час обробки макроблоків, оновлює параметри продуктивності після обробки кожного кадру і корегує розподіл, враховуючи продуктивність пристроїв та пропускну здатність каналів. Основним недоліком є часові витрати на передавання даних через шину PCIe.

Дослідження [11] присвячене розробці методу адаптивного балансування обчислювального навантаження між CPU та GPU у системах розпізнавання мовлення. Обчислення виконуються паралельно: CPU обробляє етапи з низькою паралельністю або складною логікою (керування потоками, оцінка ймовірнісних моделей), а GPU прискорює масово паралельні обчислення, зокрема спектральні перетворення, фільтрацію та обчислення MFCC. Розподіл навантаження між процесорами здійснюється адаптивно, на основі аналізу часу виконання попередніх завдань та розміру оброблюваних блоків даних. Недоліком запропонованого підходу є підвищені витрати на обмін даними між процесорами, оскільки в процесі адаптивного балансування одні й ті самі фрейми можуть помилково передаватися між пристроями кілька разів, що спричиняє додаткові затримки та зменшує загальну ефективність обчислювального процесу.

Розглянуті дослідження підтверджують доцільність адаптивного керування розподілом навантаження між CPU та GPU в системах паралельної обробки даних, зокрема голосових сигналів.

Формулювання мети дослідження

Метою роботи є розробка та дослідження ефективності системи обробки голосових сигналів із алгоритмом адаптивного балансування навантаження між CPU та GPU. Підхід повинен передбачати динамічну оцінку складності кожного кадру сигналу на основі статистичних та спектральних характеристик і відповідне розподілення обчислювальних ресурсів при роботі з сигналами різної складності.

Викладення основного матеріалу дослідження

Обробка голосових сигналів у системах ідентифікації мовця включає підготовку сигналу (нормалізація, фільтрація, фреймування, кадрування за допомогою віконної функції), виділення ознак, класифікацію та прийняття рішення [7]. Перед обробкою кожен кадр сигналу проходить швидкий попередній аналіз у окремому потоці CPU. Це дозволяє оцінити характеристики сигналу та визначити його обчислювальну складність до застосування ресурсоємних алгоритмів.

Одним із параметрів, що визначаються при попередньому аналізі, є середньоквадратичне відхилення амплітуди кадру, що відображає розсіяння значень сигналу відносно середнього рівня та характеризує його коливальну активність:

$$\sigma = \sqrt{\frac{1}{N} \sum_{n=0}^{N-1} (x[n] - \bar{x})^2}, \quad (1)$$

$$\bar{x} = \frac{1}{N} \sum_{n=0}^{N-1} x[n],$$

де $x[n]$ – відліки сигналу у кадрі, а N – кількість відліків у кадрі. Зростання σ вказує про високу амплітудну варіабельність і, відповідно, більшу обчислювальну складність операцій нормалізації та фільтрації.

Також враховується середня енергія кадру, яка відображає активність сигналу та використовується для визначення впливу амплітуди на обробку:

$$E = \frac{1}{N} \sum_{n=0}^{N-1} x^2[n] \quad (2)$$

Високі значення енергії вказують на наявність активних кадрів сигналу, що потребують інтенсивної обробки.

Для оцінки спектральної структури кадру розраховується коефіцієнт варіації спектра, який характеризує відносний розкид амплітуд спектральних компонентів та дозволяє визначити рівень деталізації сигналу:

$$CV_{\text{spec}} = \frac{\sigma_{\text{spec}}}{\mu_{\text{spec}}}, \quad (3)$$

$$\mu_{\text{spec}} = \frac{1}{N/2} \sum_{k=0}^{N/2-1} |X[k]|,$$

$$\sigma_{\text{spec}} = \sqrt{\frac{1}{N/2} \sum_{k=0}^{N/2-1} (|X[k]| - \mu_{\text{spec}})^2},$$

де $X[k]$ – дискретне Фур'є-перетворення кадру, μ_{spec} – середнє значення амплітуд спектра, σ_{spec} – середньоквадратичне відхилення амплітуд спектра. Значення CV_{spec} зростає зі збільшенням спектральної варіативності, що прямо впливає на обчислювальне навантаження при порівнянні з шаблонами.

Додатково розраховується спектральна ентропія, яка є мірою інформаційної насиченості та складності сигналу:

$$H = - \sum_{k=0}^{N/2-1} P[k] \log P[k], \quad (4)$$

$$P[k] = \frac{|X[k]|^2}{\sum_{j=0}^{N/2-1} |X[j]|^2},$$

де $P[k]$ – нормалізована спектральна потужність у точці частотного індексу k .

Високі значення спектральної ентропії вказують на наявність численних частотних компонентів, що ускладнює розпізнавання сигналу та підвищує потребу в обчислювальних ресурсах.

На основі розрахованих статистичних параметрів формується вектор ознак $F = [f_1, f_2, \dots, f_m]$, який передається модулю оцінювання складності. Попередньо всі параметри нормалізуються за виразом:

$$\rho_i^n = \frac{\rho_i - \rho_i^{\min}}{\rho_i^{\max} - \rho_i^{\min}}, \quad (5)$$

де $\rho_i^{\min}$ та $\rho_i^{\max}$ – мінімальне та максимальне значення параметра у поточній вибірці.

Модуль адаптивного керування розподілом навантаження обчислює зведений індекс складності кадру шляхом зваженого підсумовування всіх описаних нормалізованих параметрів із урахуванням вагових коефіцієнтів, що визначаються в процесі адаптації системи:

$$C_f = \sum_{i=1}^m \omega_i \rho_i^n, \quad (6)$$

де ω_i – вагові коефіцієнти, що характеризують відносну важливість кожного параметра у загальній оцінці. Вагові коефіцієнти визначаються таким чином, щоб сумарна вага всіх параметрів дорівнювала одиниці, що забезпечує нормування внеску кожного параметра у підсумкову оцінку C_f та робить індекс незалежним від кількості використовуваних характеристик:

$$\sum_{i=1}^m \omega_i = 1, \quad (7)$$

Оцінювання вагових коефіцієнтів відбувається динамічно: система безперервно відстежує статистику змін параметрів у часовому вікні, обчислюючи ковзні середні значення та дисперсії для кожного кадру. Це дозволяє адаптувати ваги в реальному часі до умов сигналу та підвищувати значущість енергетичного параметра при наявності шумів або спектрального при багатокomпонентній структурі мовлення.

Аналіз динаміки ознак голосового сигналу дає змогу системі реалізувати адаптивні стратегії розподілу навантаження між CPU та GPU. Для скорочення обчислень та підвищення ефективності роботи системи реалізовано механізм адаптивного пропускання кадрів, який приймає рішення на основі зваженої сумарної оцінки кожного кадру. Цей механізм дає змогу динамічно регулювати частоту виконання обчислень, уникаючи зайвого аналізу кадрів із низькою інформативністю.

Алгоритм прийняття рішення вибору виконавчого пристрою передбачає, що система обчислює вектор ознак для кожного кадру. Якщо виявлено суттєві зміни статистичних або спектральних характеристик (наприклад, різке зростання енергетичного рівня, зміна частотної структури або поява шумових компонентів), це свідчить про появу нового стану сигналу, що потребує оновлення параметрів для точної оцінки його складності. Якщо змін не виявлено, виконується розрахунок зваженої сумарної характеристики кадру, що дозволяє оцінити інтегральну складність без повторного повного аналізу:

$$S_i = \sum_{k=1}^m \omega_k f_{ik}, \quad (8)$$

де w_k – вагові коефіцієнти, визначені під час адаптації системи та відображають значущість окремих ознак.

Якщо обчислене значення S_i виявляється меншим за адаптивний поріг $S_{\text{пор}}$, кадр пропускається. Пропущений кадр вважається малозначущим для подальшої обробки, що дозволяє скоротити кількість операцій без втрати інформативності. Порогове значення $S_{\text{пор}}$ визначається статистично на основі середнього рівня складності попередніх кадрів із врахуванням їх дисперсії або ковзного стандартного відхилення:

$$T_{\text{адмт}} = \mu_{\text{поперед}} + k + \sigma_{\text{поперед}}, \quad (9)$$

де $\mu_{\text{поперед}}$ – середнє значення складності попередніх кадрів, $\sigma_{\text{поперед}}$ – стандартне відхилення, k – коефіцієнт, що визначає чутливість алгоритму до змін сигналу.

У результаті система динамічно регулює кількість необхідних обчислень: при появі нових або складних фрагментів сигналу активується повний розрахунок параметрів, у стабільних фрагментах використовується лише оцінка складності, а в неінформативних інтервалах вихідного сигналу кадри пропускаються.

Система реалізує механізм повторного перенаправлення обчислювальних задач, у межах якого кожен вузол повідомляє блок керування про стан виконання кадру. Якщо час обробки перевищує встановлений поріг, кадр автоматично перенаправляється на процесор іншого типу. Такий підхід дозволяє адаптивно перерозподіляти навантаження між процесорами, уникати вузьких місць і підтримувати стабільну продуктивність навіть за значних коливань характеристик сигналів або завантаженості обчислювальних пристроїв, забезпечуючи ефективну роботу системи в режимі реального часу без простоїв та втрат швидкодії.

Для кожного кадру визначається поріг часу обробки $T_{\text{поріз}}$, перевищення якого або виявлення перевантаження вузла ініціює автоматичне перенаправлення кадру на інший вузол, що гарантує своєчасну обробку. Поріг часу обчислюється на основі середнього часу обробки останніх кадрів $T_{\text{сеп}}$ та стандартного відхилення σ_T :

$$T_{\text{поріз}} = T_{\text{сеп}} + k \cdot \sigma_T, \quad (10)$$

де $T_{сер}$ – середній час обробки попередніх кадрів, σ_T – стандартне відхилення часу обробки, k – емпіричний коефіцієнт, що визначає допустиме відхилення.

З урахуванням поточного завантаження вузла $L_{вузол}$ поріг коригується:

$$T_{поріг}^{скор} = T_{поріг} \cdot (1 - \alpha L_{вузол}), \quad (11)$$

де α – коефіцієнт адаптації порога до навантаження.

Для оцінки продуктивності системи голосової ідентифікації з адаптивним розподілом завдань між CPU та GPU метрикою ефективності обрано час виконання обробки кадру (у секундах) та ступінь завантаженості обчислювальних вузлів (у відсотках), що дозволяє оцінити, наскільки рівномірно розподіляються обчислювальні ресурси між CPU та GPU.

Експериментальна модель системи голосової ідентифікації була розгорнута у середовищі Google Colab, із підтримкою багатопотокових обчислень на CPU за допомогою бібліотек OpenMP та масово-паралельних обчислень на GPU. Такий підхід дозволяє оцінити ефективність адаптивного розподілу навантаження між CPU Intel Core i7-13700H та дискретним графічним прискорювачем NVIDIA архітектури Ada Lovelace (GTX 4060), забезпечуючи дослідження взаємодії між процесором і графічним модулем.

Для оцінки ефективності обробки голосових сигналів використано датасет SEWA, який містить понад 2000 хвилин аудіо-записів 398 учасників із шести культур. Емоційний стан учасників описується двома параметрами: позитивністю/негативністю та інтенсивністю прояву емоцій [12]. На основі цього датасету сформовано набори різного обсягу та складності, що включають короткі нейтральні фрагменти з низькою спектральною складністю і мінімальним шумом, середні за тривалістю сигнали з різною інтенсивністю емоцій та фоновим шумом, а також довгі діалогові сегменти з динамічними коливаннями енергетичної інтенсивності, паузами та періодичними шумовими перешкодами (Таблиця 1).

Таблиця 1

Тестові набори даних

№	Кількість файлів	Середня тривалість, с	Опис сигналу
1	100	0,8	Короткі фрази
2	200	1,5	Фрази середньої довжини
3	500	3,0	Запис із помірним шумом
4	1000	4,2	Фонові перешкоди
5	2000	5,0	Сильний шум середовища
6	5000	8,0	Розмовна мова у приміщенні
7	10000	12,0	Фонові голоси
8	20000	20,0	Уривки лекційних занять
9	50000	35,0	Складні акустичні умови
10	100000	60,0	Діалоги з високим фоновим шумом

Кадри організовано у багатовимірні тензори, де перша вісь відповідає номеру кадру, а наступні – часовим і частотним вимірам, що забезпечує паралельну обробку кількох фреймів і повніше використання ресурсів обчислювальних вузлів (CPU та GPU).

Для ускладнення роботи блоку адаптивного розподілу обчислювального навантаження в наборах тестових даних передбачено короткі імпульсні сегменти з високими піковими коливаннями амплітуди та фоновим шумом, а також чергування коротких, середніх та довгих кадрів у межах одного пакету, імітуючи реальні умови потокового мовлення з різкою зміною складності спектра та енергії.

Розглянуто чотири сценарії виконання обробки голосових сигналів, залежно від задіяних обчислювальних вузлів: лише на CPU, лише на GPU, спрощений розподіл за розміром сигналу (Size-based Scheduler) та повна адаптивна система (Full Adaptive Scheduler), яка аналізує характеристики сигналу для оптимального розподілу завдань.

Результати виконання методів попередньої обробки (спектральний аналіз, вейвлет-фільтрація), наведені в таблиці 2, свідчать, що час роботи адаптивної системи є найменшим серед розглянутих варіантів, що підтверджує її ефективність.

Розрахунок мел-частотних кепстральних коефіцієнтів MFCC в системі з застосуванням алгоритму адаптивного розподілу потребує найменший час (Таблиця 3), що підтверджує ефективність запропонованого підходу у задачах різного рівня навантаження. Водночас для окремих наборів даних спостерігається незначне зменшення часу виконання порівняно з CPU- та GPU-орієнтованими системами, що свідчить про коректність вибору обчислювального вузла відповідно до характеристик вхідних даних та їхньої обчислювальної складності.

Таблиця 2

Час виконання етапів попередньої обробки

№ тестового набору	CPU, с	GPU, с	Size-based Scheduler, с	Full Adaptive Scheduler, с
1	0,42	0,61	0,4	0,38
2	0,85	0,73	0,69	0,62
3	2,10	1,20	1,23	0,91
4	4,75	2,45	1,98	1,72
5	9,40	4,60	3,92	3,35
6	21,2	10,3	18,5	7,24
7	43,7	20,4	19,9	15,5
8	87,9	39,5	37,8	29,6
9	193	82,1	90,2	62,8
10	391	164	195	127

Таблиця 3

Час обчислення MFCC

№ тестового набору	CPU, с	GPU, с	Size-based Scheduler, с	Full Adaptive Scheduler, с
1	0,58	0,66	0,55	0,51
2	1,12	0,84	0,75	0,69
3	2,84	1,49	1,35	1,18
4	6,10	3,10	7,76	2,41
5	12,5	6,25	5,32	4,64
6	228,3	13,8	11,9	10,1
7	55,4	27,1	173,3	20,2
8	112	54,3	47,5	40,9
9	2238	113	2981,1	84,5
10	486	223	198	167

Проведене експериментальне дослідження дозволило оцінити навантаженість обчислювальних вузлів під час обробки різнорідних наборів мовних сигналів при використанні розподілу за обсягом та з урахуванням численних характеристик сигналу.

Таблиця 4

Результати навантаження вузлів системи під час роботи системи

№ тестового набору	Size-based Scheduler		Full Adaptive Scheduler	
	CPU, %	GPU, %	CPU, %	GPU, %
1	92	5	66	81
2	31	44	72	63
3	96	13	68	42
4	24	83	74	63
5	86	12	62	84
6	77	15	53	42
7	34	36	81	32
8	23	91	61	46
9	56	26	74	44
10	88	20	36	22

Як показано в таблиці 4, обсягоорієнтований підхід до розподілу навантаження (Size-based Scheduler) спричиняє дисбаланс: один вузол працює на межі обчислювальних можливостей, тоді як інший простоє, що знижує сумарну продуктивність і збільшує загальний час обробки. Тоді як запропонований адаптивний алгоритм забезпечує динамічний розподіл задач між CPU та GPU, враховуючи складність кадрів, поточне завантаження вузлів і варіативність характеристик сигналів.

Висновки

Запропоновано адаптивний розподіл обчислювального навантаження у системах обробки голосових сигналів, який використовує комплекс характеристик мовних записів, зокрема довжину кадру, рівень енергетичної активності та спектральну насиченість.

Експериментальні дослідження виконувалися на основі наборів даних різного обсягу та рівня складності сигналів, з урахуванням наявності викривлень, що дозволило всебічно перевірити ефективність адаптивного

механізму розподілу навантаження, оцінити його здатність визначати пріоритетність обробки кадрів та використовувати ресурси CPU та GPU незалежно від тривалості, спектральної складності, емоційного навантаження та наявності шумів у вхідних сигналах. Проведено порівняльний аналіз роботи системи без адаптивного розподілу, з його спрощеною версією (що враховує лише обсяг даних), та запропонованого рішення, яке дозволяє розподіляти обчислення між вузлами, підвищуючи ефективність використання ресурсів та загальну продуктивність системи.

Список використаної літератури

1. Samonte M. J. C., Callejo J. K., Lumbea D. C. N., Ocaya J. C. B. Mitigating Vishing in Digital Banking Through Caller Authentication and Verification Technologies. *2024 14th International Conference on Software Technology and Engineering ICSTE*, Macau, China. 2024. Pp. 102–108. DOI: 10.1109/ICSTE63875.2024.00025
2. Kambampati P., Rane S., Shoeb A., Dhannawat R. PAYV Payment Voice A Platform using Voice Recognition to Enable Payment Transactions. *2024 Asia Pacific Conference on Innovation in Technology APCIT*, Mysore, India. 2024. Pp. 1–6. DOI: 10.1109/APCIT62007.2024.10673442
3. Li Y., Gao X., Song Q., Wang Y., Lyu P., Zhang H. BoneAuth A Bone-Conduction-Based Voice Liveness Authentication for Voice Assistants. *IEEE Internet of Things Journal*. 2025. Vol. 12, no. 6. Pp. 6997–7009. DOI: 10.1109/IJOT.2024.3494024
4. Bao L., Zuo Y. Speaker Identification based on MFSC Voice Feature Extraction using Transformer. *2023 IEEE International Conference on Data Mining Workshops ICDMW*, Shanghai, China. 2023. Pp. 1–7. DOI: 10.1109/ICDMW60847.2023.00008
5. Chen Q., Gu Z., Lu L., Xu X., Ba Z., Lin F., Liu Z., Ren K. Conan's Bow Tie A Streaming Voice Conversion for Real-Time VTuber Livestreaming. *Proceedings of the 29th International Conference on Intelligent User Interfaces*. 2024. Pp. 35–50. DOI: 10.1145/3640543.3645146
6. Mykhailichenko I., Ivashchenko H., Barkovska O., Liashenko O. Application of Deep Neural Network for Real-Time Voice Command Recognition. *IEEE 3rd KhPI Week on Advanced Technology KhPIWeek*, Kharkiv, Ukraine. 2022. Pp. 1–4. DOI: 10.1109/KhPIWeek57572.2022.9916473
7. Бондаренко М. Е., Іващенко Г. С. Використання послідовності методів попередньої обробки в системах голосової ідентифікації. *Системи управління, навігації та зв'язку*. Полтава: ПНТУ. 2025. № 2 (80). С. 90–96. DOI: 10.26906/SUNZ.2025.2.090
8. Бондаренко М. Е., Іващенко Г. С. Організація паралельного виконання методів обробки голосових сигналів на багатоядерних CPU та GPU. *Системи управління, навігації та зв'язку*. Полтава: ПНТУ. 2025. № 4 (82). С. 39–44. DOI: 10.26906/SUNZ.2025.4.39-44
9. Gjermundsen A. CPU and GPU Co-processing for Sound. MA thesis. *Norwegian University of Science and Technology*. 2010. 173 p.
10. Momcilovic S., Ilic A., Roma N., Sousa L. Dynamic Load Balancing for Real-Time Video Encoding on Heterogeneous CPU+GPU Systems. *IEEE Transactions on Multimedia*. 2014. Vol. 16, no. 1. Pp. 108–121. DOI: 10.1109/TMM.2013.2284892
11. Kim J., Lane I. Accelerating Large Vocabulary Continuous Speech Recognition on Heterogeneous CPU-GPU Platforms. *2014 IEEE International Conference on Acoustics, Speech and Signal Processing ICASSP*, Florence, Italy. 2014. Pp. 3291–3295. DOI: 10.1109/ICASSP.2014.6854209
12. Kossaifi J., Walecki R., Panagakis Y., Shen J., Schuller B., Pantic M. SEWA DB A Rich Database for Audio-Visual Emotion and Sentiment Research in the Wild. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2021. Vol. 43, no. 3. Pp. 1022–1040. DOI: 10.1109/TPAMI.2019.2944808

References

1. Samonte, M. J., C. Callejo, J. K., Lumbea, D. C. N., Ocaya, J. C. B. (2024). Mitigating Vishing in Digital Banking Through Caller Authentication and Verification Technologies. *2024 14th International Conference on Software Technology and Engineering (ICSTE)*. Macau, China, pp. 102–108. DOI: 10.1109/ICSTE63875.2024.00025
2. Kambampati, P., Rane, S., Shoeb, A., Dhannawat, R. (2024). PAYV – Payment Voice: A Platform using Voice Recognition to Enable Payment Transactions. *2024 Asia Pacific Conference on Innovation in Technology (APCIT)*. Mysore, India, pp. 1–6. DOI: 10.1109/APCIT62007.2024.10673442
3. Li, Y., Gao, X., Song, Q., Wang, Y., Lyu, P., Zhang, H. (2025). BoneAuth: A Bone-Conduction-Based Voice Liveness Authentication for Voice Assistants. *IEEE Internet of Things Journal*. Vol. 12, No. 6, pp. 6997–7009. DOI: 10.1109/IJOT.2024.3494024
4. Bao, L., Zuo, Y. (2023). Speaker Identification based on MFSC Voice Feature Extraction using Transformer. *2023 IEEE International Conference on Data Mining Workshops (ICDMW)*. Shanghai, China, pp. 1–7. DOI: 10.1109/ICDMW60847.2023.00008

5. Chen, Q., Gu, Z., Lu, L., Xu, X., Ba, Z., Lin, F., Liu, Z., Ren, K. (2024). Conan's Bow Tie: A Streaming Voice Conversion for Real-Time VTuber Livestreaming. *Proceedings of the 29th International Conference on Intelligent User Interfaces*, pp. 35–50. DOI: 10.1145/3640543.3645146
6. Mykhailichenko, I., Ivashchenko, H., Barkovska, O., Liashenko, O. (2022). Application of Deep Neural Network for Real-Time Voice Command Recognition. *IEEE 3rd KhPI Week on Advanced Technology (KhPIWeek)*. Kharkiv, Ukraine, pp. 1–4. DOI: 10.1109/KhPIWeek57572.2022.9916473
7. Bondarenko, M. E., Ivashchenko, H. S. (2025). Using a sequence of preprocessing methods in voice identification systems. *Control, navigation and communication systems*. Poltava: NUPP. Vol. 2 (80), pp. 90–96. DOI: 10.26906/SUNZ.2025.2.090
8. Bondarenko, M. E., Ivashchenko, H. S. (2025). Parallel implementation of voice signal processing methods on multicore CPU and GPU. *Control, navigation and communication systems*. Poltava: NUPP. Vol. 4 (82), pp. 39–44. DOI: 10.26906/SUNZ.2025.4.39-44
9. Gjermundsen, A. (2010). CPU and GPU Co-processing for Sound. MA thesis. *Department of Computer and Information Science, Norwegian University of Science and Technology*. Norwegian University of Science and Technology. 173 p.
10. Momcilovic, S., Ilic, A., Roma, N., Sousa, L. (2014). Dynamic Load Balancing for Real-Time Video Encoding on Heterogeneous CPU+GPU Systems. *IEEE Transactions on Multimedia*. Vol. 16, no. 1, pp. 108–121. DOI: 10.1109/TMM.2013.2284892
11. Kim, J., Lane, I. (2014). Accelerating Large Vocabulary Continuous Speech Recognition on Heterogeneous CPU-GPU Platforms. *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Florence, Italy, pp. 3291–3295. DOI: 10.1109/ICASSP.2014.6854209
12. Kossaifi, J., Walecki, R., Panagakis, Y., Shen, J., Schuller, B., Pantic, M. (2021). SEWA DB: A Rich Database for Audio-Visual Emotion and Sentiment Research in the Wild. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Vol. 43, no. 3, pp. 1022–1040. DOI: 10.1109/TPAMI.2019.2944808

Дата першого надходження рукопису до видання: 13.11.2025
Дата прийнятого до друку рукопису після рецензування: 10.12.2025
Дата публікації: 31.12.2025

М. О. ВЕРНІК

аспірант кафедри програмного забезпечення комп'ютерних систем
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
ORCID: 0009-0008-6156-1051

Л. М. ОЛЕЩЕНКО

кандидат технічних наук,
доцент кафедри програмного забезпечення комп'ютерних систем
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
ORCID: 0000-0001-9908-7422

ПРОГРАМНІ МЕТОДИ ТА АРХІТЕКТУРНІ РІШЕННЯ ДЛЯ ОБРОБКИ ВЕЛИКИХ ДАНИХ ЧЕРЕЗ КВАНТОВІ НЕЙРОННІ МЕРЕЖІ ТА МЕТАЕВРИСТИЧНІ ОПТИМІЗАТОРИ

У статті подано запропоновані програмні методи та гібридну класично-квантову архітектуру для обробки великих даних з використанням квантово-ядрового методу *QSVC* та варіаційного квантового класифікатора *VQC/QNN* для розв'язання задачі обчислення ментальної енергії на основі *ЕЕГ*-сигналів. Запропонований фреймворк працює з даними, отриманими від носимого *ЕЕГ*-пристрою, що реєструє діапазони δ , θ , α та β на чотирьох каналах (4-канальний *ЕЕГ*-пристрій *Muse-S Athena*). На основі цих сигналів формуються вектори ознак, які відображаються у квантові стани за допомогою квантової відображальної функції (*quantum feature map*) і надалі використовуються або в квантовому ядровому методі *SVM*, або у варіаційній квантовій нейронній мережі для задач класифікації чи регресії показників, пов'язаних з ментальною енергією. Для оптимізації побудованих квантових схем («квантових нейронних мереж») застосовано метаевристичні алгоритми, зокрема, *Particle Swarm Optimization (PSO)*, *Genetic Algorithm (GA)*, *Grey Wolf Optimizer (GWO)*, а також багатокритеріальні модифікації, такі як *NSGA*. Ці оптимізатори спільно налаштовують параметри квантових кіл, масштаби кодування ознак та елементи структури ансамблю з урахуванням реалістичних обмежень *NISQ*-архітектур. Основне завдання полягає у зведенні вихідної багатокритеріальної постановки до придатної для розв'язання однокритеріальної форми шляхом побудови функції вартості, що мінімізує зважену суму ключових параметрів: кількості вимірювань (*shots*) на квантовому комп'ютері, глибини квантової схеми та загального часу виконання при одночасному збереженні прийнятної якості класифікації ментальних станів. Також запропоновано програмну архітектуру, побудовану на основі квантових мікросервісів для обчислення квантових ядер та виконання *VQC* у поєднанні з оркестратором метаевристичної оптимізації та компонентами моніторингу ресурсів. Запропонована архітектура орієнтована на інтеграцію з моделями ментальної енергії, побудованими на *ЕЕГ*-даних, що дає змогу виконувати масштабовані експерименти з квантовими моделями на великих наборах *ЕЕГ*-сигналів. Отримані результати формують методологічну основу для використання квантових нейронних мереж та метаевристичних оптимізаторів у когнітивній інженерії та нейроінженерних застосуваннях, зокрема, для кількісного оцінювання ментальної енергії в реальних умовах.

Ключові слова: програмне забезпечення, великі дані, машинне навчання, метаевристичні алгоритми, квантові нейронні мережі, *QML*, оптимізація, квантова оптимізація, ментальна енергія.

М. О. VERNIK

Postgraduate Student at the Department of Computer Systems Software
National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute”
ORCID: 0009-0008-6156-1051

L. M. OLESHCHENKO

Candidate of Technical Sciences,
Associate Professor at the Department of Computer Systems Software
National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute”
ORCID: 0000-0001-9908-7422

SOFTWARE METHODS AND ARCHITECTURAL SOLUTIONS FOR BIG DATA PROCESSING VIA QUANTUM NEURAL NETWORKS AND METAHEURISTIC OPTIMIZERS

The article presents proposed software methods and a hybrid classical–quantum architecture for big data processing using a quantum kernel method QSVC and a variational quantum classifier VQC/QNN to address the problem of computing mental energy from EEG signals. The proposed framework operates on data acquired from a wearable EEG device that records δ , θ , α , and β bands across four channels (the 4-channel Muse-S Athena EEG device). From these signals, feature vectors are constructed and mapped into quantum states via a quantum feature map, which are then used either in a quantum kernel SVM or in a variational quantum neural network for classification or regression of mental energy-related indicators. To optimize the resulting quantum circuits (“quantum neural networks”), the work employs metaheuristic algorithms such as Particle Swarm Optimization (PSO), Genetic Algorithm (GA), Grey Wolf Optimizer (GWO), and multi-objective variants like NSGA. These optimizers jointly tune circuit parameters, feature-encoding scales and elements of the circuit structure under realistic NISQ constraints. The main task is to reduce an inherently multi-objective problem to a tractable single-objective formulation by defining a cost function that minimizes a weighted sum of key parameters: the number of measurement shots on the quantum computer, the depth of the quantum circuit, and the total execution time, while preserving an acceptable classification quality of mental states. The paper also proposes a software architecture based on quantum microservices (for quantum kernel evaluation and VQC execution) combined with a metaheuristic optimization orchestrator and resource-tracking components. This architecture is designed to integrate with EEG-driven mental energy models, enabling scalable experimentation with quantum models on large EEG datasets. The results obtained provide a methodological basis for using quantum neural networks and metaheuristic optimizers in cognitive and neuroengineering applications, with a particular focus on quantifying mental energy in real-world conditions.

Key words: software, big data, machine learning, metaheuristic algorithms, quantum neural networks, QML, optimization, quantum optimization, metaheuristics, EEG, mental energy.

Problem statement

The relevance of the research lies in bridging the gap between theoretical quantum machine learning and its practical application by designing a scalable hybrid framework, building mathematical foundations for quantum model training, enabling multi-objective optimization of quantum circuits, and demonstrating applicability on complex datasets such as EEG signals – thereby advancing both quantum computing and intelligent neuroinformatics.

The proposed research focuses on the development of a hybrid computational framework that integrates classical and quantum processing components to enable scalable data analysis and efficient model training. This framework leverages metaheuristic optimization methods to overcome the inherent limitations of current quantum hardware – such as noise, restricted qubit counts, and limited circuit depth – which typically constrain the applicability of quantum machine learning to large datasets. By combining classical optimization pipelines with quantum subroutines, the system aims to facilitate the training of quantum kernel-based models (e.g., Quantum Support Vector Classifiers, QSVC) as well as variational quantum classifiers and regressors (VQC/QNN) in a resource-aware and performance-optimized manner.

Within this research, a comprehensive set of mathematical models will be formulated to describe the encoding of input data into quantum states, the structure and behavior of quantum kernel methods and variational circuits, the quality metrics for evaluating these models, and the operational constraints imposed by quantum hardware. Building on these foundations, the study will define a multi-objective optimization problem that captures the trade-offs between model accuracy, quantum circuit depth, the required number of measurement shots, and system latency. Metaheuristic algorithms – such as evolutionary strategies, particle swarm optimization, simulated annealing, differential evolution, and CMA-ES – will be explored and adapted to search for optimal parameter configurations and circuit architectures under these competing objectives.

A modular software architecture will be designed to support quantum microservices, including a Q-Kernel service dedicated to quantum kernel computation and a Q-VQC service responsible for executing variational circuits. These services will operate under a centralized metaheuristic orchestration layer that manages optimization workflows, tracks quantum resource consumption, and incorporates mechanisms for load balancing and fault tolerance to ensure robust system performance. The resulting hybrid infrastructure will provide a unified platform for experimentation, evaluation, and deployment of quantum-enhanced machine learning pipelines. To validate the proposed methodology, extensive experiments will be conducted using public datasets as well as proprietary data sources, including EEG signals collected from Muse S Athena devices. The performance of quantum models will be benchmarked against classical machine learning baselines, with particular attention to assessing their generalization ability and resource efficiency. Additionally, the research will examine and validate the Countable Mental Energy (CME) indicator as an engineering metric for quantifying cognitive state based on EEG data, providing further insight into the potential of hybrid quantum–classical approaches for neurophysiological signal analysis.

Related research

In modern studies of quantum machine learning for EEG analysis, the possibility of using variational quantum classifiers (VQC) and quantum kernel methods (QSVC) for brain signal classification tasks has been demonstrated. In particular,

in the work devoted to the classification of P300-EEG in the brain–computer interface paradigm, it has been shown that quantum classifiers are able to successfully learn on EEG data: the quantum-augmented support vector machine (QSVC) achieves a balanced accuracy of about 83% on the training sample, exceeding classical baseline models [1-3]. At the same time, the limited generalization ability of the models on the test data and the dependence of the results on the configuration of quantum circuits and the amount of training data are emphasized, which indicates the need for further optimization of architectures and hyperparameters [4-7]. In study [2], quantum methods for the classification of motor imagery from EEG were proposed, including a quantum genetic algorithm (QGA) for feature selection on NISQ equipment; it has been demonstrated that the hybrid quantum approach outperforms a number of classical classifiers on reference datasets, which confirms the potential of quantum kernels in working with high-dimensional EEG spaces. In parallel, the direction of using metaheuristic algorithms to optimize features and model parameters in EEG analysis tasks is actively developing. One of the works proposed the GALoRIS model, which combines a genetic algorithm with logistic regression for selecting EEG features in the task of assessing the driver's cognitive load, where it is shown that the method automatically reduces the feature space by more than 50%, while maintaining the informativeness of the signal [3]. As a result, classifiers based on selected features achieve an average accuracy of about 90.4%, and in combination with RBF-SVM – up to ~96%, which significantly exceeds the performance of traditional feature selection methods (in particular, approaches based on mutual information or PCA), which provide an accuracy of about 84% [3]. These results demonstrate that evolutionary search is capable of forming more discriminative subspaces of EEG features compared to linear methods, however, this approach has so far been used only in conjunction with classical models [4]. A separate direction directly related to the subject of this work concerns the development of software frameworks for the quantitative assessment of cognitive load and "mental energy" in everyday conditions. Within the framework of one such framework, the use of a portable EEG device with time-aligned audio segmentation for continuous assessment of the indicator of countable mental energy throughout the day is proposed: the system combines EEG recordings from a Muse-type headset with an audio lifelogger, where the user's speech activity is transcribed and segmented, and for each segment, EEG-conditioned effort levels are calculated $L(t)$ and energy expenditure rate w_i [8]. Analytical expressions are proposed that relate the spectral characteristics of the EEG (frontal theta, occipital alpha, beta activation) to energy metrics, and a "flow" coefficient is introduced to take into account task involvement and an energy reserve multiplier associated with fatigue, which together form an integral indicator of daily cognitive/mental energy. An intelligent software framework for processing large audio data sets and assessing cognitive energy is presented, focused on scalable processing of speech streams and integration with cognitive load models, which allows combining big data audio processing with engineering metrics of mental energy [9].

Proposed software methods

Quantum encoding model

Let's say we have a sample of EEG data:

$$\mathcal{D} = \{(x_t, y_t, a_t)\}_{t=1}^N, \quad (1)$$

$x_t \in \mathbb{R}^d$ is vector of features calculated from the raw EEG signal in the time interval (window) t . The features used are, for example, the relative powers of the δ , θ , α , β ranges on 4 channels, asymmetry indicators (Fp1–Fp2, AF7–AF8), the θ/α ratio, etc.

$y_t \in \{0,1\}$ is target variable (mental state class, e.g. "in flow / not in flow"). If the problem is regression, $y_t \in \mathbb{R}$.

$a_t \in [0,1]$ is an indicator of the complexity or type of activity at a given point in time t (estimated using additional sensors or contextual data, e.g. based on audio or the user's task).

Preprocessing and feature calculation

For each EEG window of ~5 s duration (with ~50% overlap between consecutive windows), signal filtering (range 1 – 40 Hz), feature normalization (componentwise) and feature vector calculation x_t using the function $\phi: \text{EEG (raw)} \rightarrow x \in \mathbb{R}^d$. Thus, x_t compactly describes the state of brain activity in the interval, and a_t characterizes context (complexity task, availability distracting factors).

Quantum encoding of features

To map a classical vector $x \in \mathbb{R}^d$ into a quantum state, a parametric *quantum feature map* $U_\alpha(x)$ is used. In particular, angle encoding of the components x on the amplitudes or phases of qubits with a scale parameter is used α , as well as layers of local entanglement of qubits to take into account correlations between features. This mapping $U_\alpha(x)$ prepares a multidimensional quantum state $|\psi(x)\rangle = U_\alpha(x)|0\rangle^{\otimes N}$ corresponding to the input data [6].

Quantum kernel method (QSVC)

A quantum support vector classifier uses a quantum kernel defined by the overlap of two quantum states of the data. For two samples, x_i, x_j the kernel function is defined as the probability of obtaining a zero state when encoded sequentially x_i and x_j on the same quantum register:

$$K_\alpha(x_i, x_j) = |\langle 0 | U_\alpha(x_i) U_\alpha(x_j) | 0 \rangle|^2, \quad (2)$$

which is the square of the inner product of the quantum states $|\psi(x_i)\rangle$ and $|\psi(x_j)\rangle$ [1]. Having calculated the kernel matrix $K = [K(x_i, x_j)]_{i,j=1}^m$ on the training data, we solve the classification problem using the support vector method:

$$f(x) = \text{sign}\left(\sum_{i=1}^m \alpha_i y_i K_\alpha(x_i, x) + b\right), \quad (3)$$

where α_i, b are the parameters found in the process of training SVM (solving the dual problem).

Thus, a quantum computer is used to efficiently calculate the kernel in a spatiotemporal high-dimensional Hilbert space, which can provide advantages when analytical calculation of the kernel by classical methods is impossible.

Variational quantum classifier (VQC/QNN) is a quantum model of the “quantum neural network” type, which is tuned by optimizing the parameters of the *ansatz* – a parameterized quantum circuit.

Data encoding – a scheme is applied $U_\alpha(x)$ that encodes the input vector x into a quantum state.

Ansatz V_θ is a set of quantum gates with parameters θ (several layers of single-qubit rotations and entanglements) is applied to the qubits, which act as an analogue of the layers of a neural network. The *ansatz* has a small depth (according to NISQ constraints) and can include only local entanglements.

After applying $U_\alpha(x)$ and V_θ some quantum observables (e.g., observables Z or ZZ on individual qubits) are measured to obtain a classical output. In particular, for each qubit, the mathematical expectation of the measurement can be calculated (e.g., $\langle Z_i \rangle$), which gives a set of values from which the probability of belonging to a certain class is calculated $p_\theta(y|x)$. Thus, VQC estimates the class probability based on the measured mean values after the *ansatz*. Obtaining these expected values on a real quantum device requires multiple iterations of the schemes (performing a certain number of *shots*) and averaging the results. The parameters are trained θ by a classical optimizer, reducing the problem to a variational algorithm (VQA). Note that *metaheuristic algorithms* (gradient methods) can be used to optimize the parameters of variational quantum schemes.

QUBO-statements for subproblems

Some combinatorial subproblems within our problem (e.g., feature subset selection, representative data window selection, *ansatz* topology optimization) can be formulated as a quadratic unconditional binary optimization (QUBO) problem. General form: $\min_{z \in \{0,1\}^n} z^T Q z$, where z is a binary vector of decision variables, and Q is a symmetric matrix of coefficients. QUBO is equivalent to the Ising-type Hamiltonian minimization problem and, therefore, is a “native” format for both QAOA-type algorithms on gate quantum computers and D-Wave quantum annealing machines. Such QUBO problems can be solved by quantum annealing or QAOA algorithms, obtaining optimal or near-optimal combinations (e.g., the most informative features).

Multi-criteria metaheuristic optimization

Training quantum models under resource constraints can be viewed as a multi-criteria optimization. When searching for the optimal configuration of a quantum model, we have several competing goals:

- maximizing model quality (minimizing the loss function $\mathcal{L}$ on the validation sample or maximizing accuracy/quality metrics);
- minimizing the use of quantum resources: quantum circuit depth (Depth), number of measurements (Shots), and total execution time (Latency).

Denote the set of all parameters and hyperparameters of the model as

$$\Theta = (\alpha, \theta, \tau), \quad (4)$$

where α is the coding scale parameter, θ is the vector of variational parameters, and τ is the structure parameters of the quantum *ansatz* (e.g., the entanglement pattern, the number of layers, the selection of observables for measurement). Let us define the vector of objective functions:

$$J(\Theta) = (\mathcal{L}(\Theta), \text{Shots}(\Theta), \text{Depth}(\Theta), \text{Latency}(\Theta)). \quad (5)$$

Thus, we look for such Θ that improve all components of J . In practice, an approximation of the Pareto-optimal front is typically constructed – the set solutions where no objective can be improved without worsening another. Metaheuristic approaches are used to find the Pareto front (e.g., multi-criteria versions of evolutionary algorithms such as NSGA-II). Alternatively, the problems can be reduced to a single-criteria problem by scalarization – minimizing the weighted sum:

$$\mathcal{F}(\Theta) = \mathcal{L}(\Theta) + \lambda_s \text{Shots}(\Theta) + \lambda_d \text{Depth}(\Theta) + \lambda_l \text{Latency}(\Theta), \quad (6)$$

where $\lambda_s, \lambda_d, \lambda_l$ – weighting factors reflecting the “cost” of using the relevant resources. This approach allows the use of single-criteria optimizers (e.g. GWO or PSO) to find a compromise solution. It is worth noting that recent studies also propose quantum algorithms for multi-criteria optimization – for example, the formation of a quantum state containing Pareto-optimal solutions in a superposition, with subsequent selection solutions from this state. However, the main emphasis is on classical metaheuristics for tuning quantum models.

During the optimization process, it is necessary to additionally take into account the strict limitations of hardware resources:

1. number of qubits $\leq q_{\max}$ (determined by the specific quantum device);
2. circle depth $\leq D_{\max}$ (due to decoherence time);
3. number of measurements (post-selection shots) per iteration $\leq S_{\max}$ (due to runtime limitations on a real QPU or quantum simulator).

These constraints are either introduced as penalties into the optimization function or are controlled by solution generators (for example, evolutionary mutation operators do not create schemes deeper than $D_{\max}$). The goal is to avoid solutions that are not suitable for execution on the available hardware.

Countable Mental Energy (CME) is a user-defined indicator that integrates information about the energetics of brain activity and the context of the task to estimate the “mental energy expended” per session. It is proposed as an aggregate measure of the cognitive load and the user’s “flow” state. The calculation of CME is based on EEG data and a focus state classification model.

For each window, t the relative power of brain waves in different ranges is calculated:

$$P_{\text{relative}}(b, t) \text{ for } b \in \{\delta, \theta, \alpha, \beta\}. \quad (7)$$

For one electrode (and one window t):

$$P_{\text{relative}}(b, t) = \frac{P_{\text{abs}}(b, t)}{P_{\text{abs}}(\delta, t) + P_{\text{abs}}(\theta, t) + P_{\text{abs}}(\alpha, t) + P_{\text{abs}}(\beta, t)}. \quad (8)$$

These values are obtained by normalizing the absolute spectral powers of the window by the total power. The energy of the window is calculated as the weighted sum of the powers:

$$E_{\text{band}}(t) = \sum_b w_b P_{\text{relative}}(b, t), \quad (9)$$

where the coefficients w_b specify the contribution of each band. For example, if β -waves are considered more “energy-intensive” for the brain, w_β may be larger.

Modification by context and focus

The CME value for a window t is defined as:

$$\text{CME}(t) = \kappa E_{\text{band}}(t) \cdot g(c(t), f_\Theta(x_t)) \cdot \Delta t, \quad (10)$$

where $c(t) = a_t \in [0, 1]$ – complexity/activity index, $f_\Theta(x_t) = p_\Theta(\text{flow} | x_t) \in [0, 1]$ – probability of the “flow” state (or high concentration) according to the model with parameters Θ , $g(c, f)$ – a monotonic function that combines the effects of complexity and focus.

Total score per session

$$\text{CME}_{\text{session}} = \sum_{t \in \text{session}} \text{CME}(t). \quad (11)$$

Reflects the total “mental energy” expended during a work/training session. Typically calculated per day as a component of the total “mental energy” for a day of activity.

Research results

The conducted research resulted in the development of a hybrid quantum–classical framework that enables scalable training and evaluation of quantum machine-learning models under realistic hardware constraints. A set of formal mathematical models was constructed for quantum data encoding, quantum kernel estimators (QSVC), and variational quantum classifiers and regressors (VQC/QNN). These models incorporate explicit resource-aware constraints, including circuit depth, number of qubits, and measurement complexity, thereby allowing their systematic optimization. A multi-objective formulation for training quantum models was proposed, balancing model accuracy, circuit expressiveness, and computational latency. For this purpose, a family of metaheuristic optimization strategies –including evolutionary algorithms, particle swarm optimization, simulated annealing, differential evolution, and CMA-ES was adapted to the quantum domain.

The evaluation demonstrated that metaheuristic search enables stable parameter discovery even under noisy intermediate-scale quantum (NISQ) conditions. A modular microservice-based software architecture was implemented to operationalize these theoretical contributions.

The developed software includes dedicated quantum microservices such as Q-Kernel for computing quantum kernels and Q-VQC for executing variational circuits, coordinated by an orchestration layer responsible for metaheuristic optimization, resource tracking, and fault tolerance. Experimental validation was carried out using both publicly available datasets and EEG signals from the Muse S Athena device. The hybrid approach consistently outperformed purely classical or purely quantum baselines in scenarios constrained by qubit count and circuit depth. Additionally, preliminary analysis supported the feasibility of using quantum-enhanced models for cognitive-state estimation and provided empirical

evidence toward validating the Countable Mental Energy (CME) indicator derived from EEG data. For software under development, automata models (finite automata/UML state diagrams) for describing behavior, Petri nets, or process calculus for analyzing concurrency and coordinating microservices may be important.

Petri models can be used to model the process of executing tasks in a queue of quantum experiments: where there are positions "task in queue", "task is being executed on the backend", "completed successfully"/"failed" and transitions between them. Analysis of such a network will confirm that there will be no situation where a task is "stuck" between states or lost. No "hangs", correct transitions, boundary conditions (timeout $\rightarrow$ retry $\leq k$, completion logging), finite automata are also suitable for performing the current task. In general, in the work, you need to focus on methods that bring direct benefit – they check the correctness of the system's behavior. The Fig. 1 shows a streamlined sequence of interactions in the hybrid quantum–classical framework, tracing how EEG data is processed from preprocessing to quantum evaluation and final inference.

After the researcher initiates an experiment, the API gateway sends prepared data to the metaheuristic orchestrator, which manages iterative optimization, calling the Q-Kernel Service for quantum kernel matrices and the Q-VQC Service for evaluating variational circuits. Both services run on IBM QPU Runtime to execute circuits and collect measurements.

The orchestrator analyzes intermediate results (kernel quality, VQC accuracy, circuit depth, shot count), adjusts parameters, and repeats the loop until convergence. Once optimized, the system aggregates final metrics AUC/ROC, accuracy, depth, shots, latency and returns them to the researcher.

The workflow integrates quantum computation with classical orchestration to automate training and evaluation of quantum machine-learning models (Fig. 2).

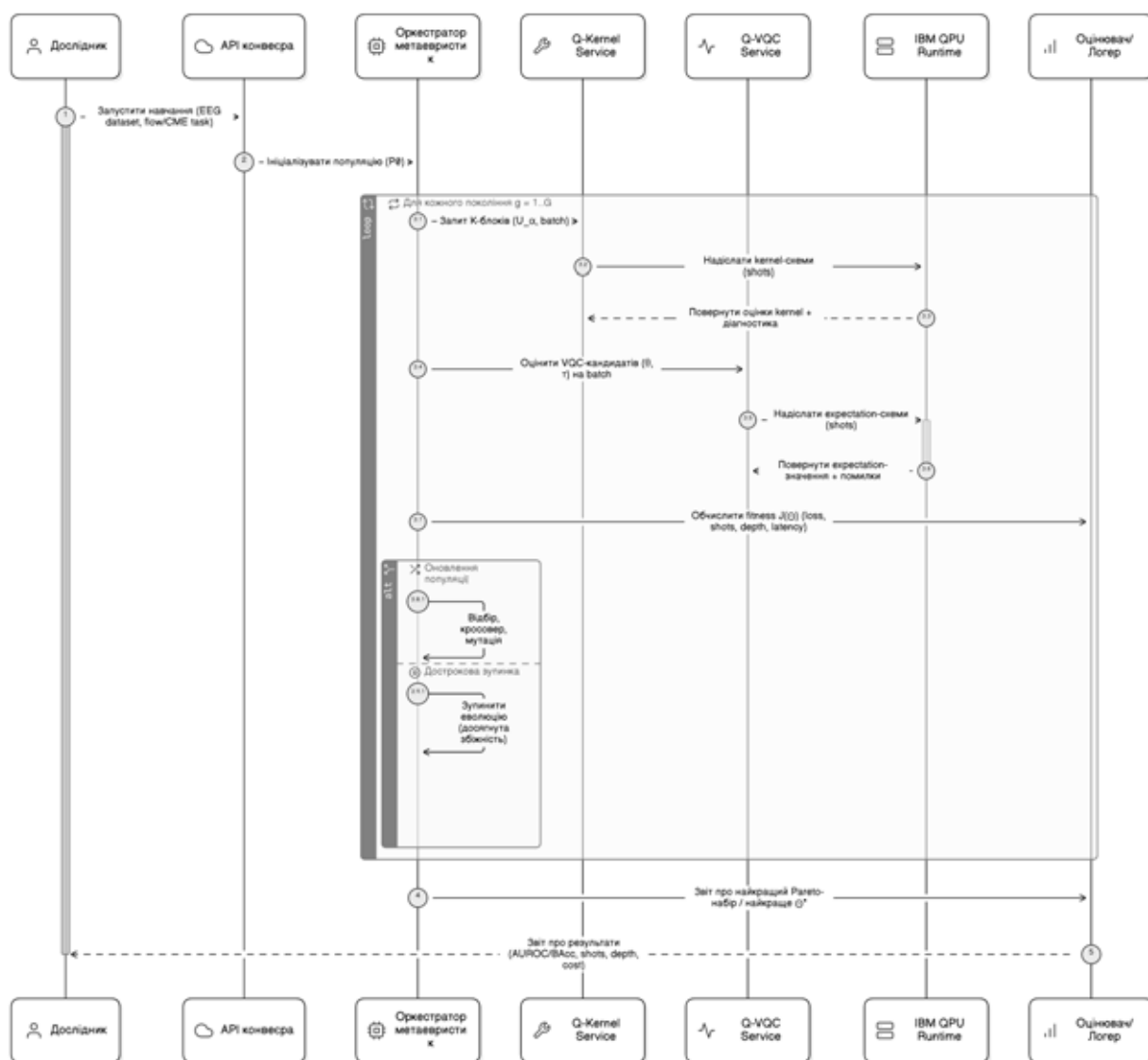


Fig. 1. Software sequence diagram "Training a model with metaheuristics"

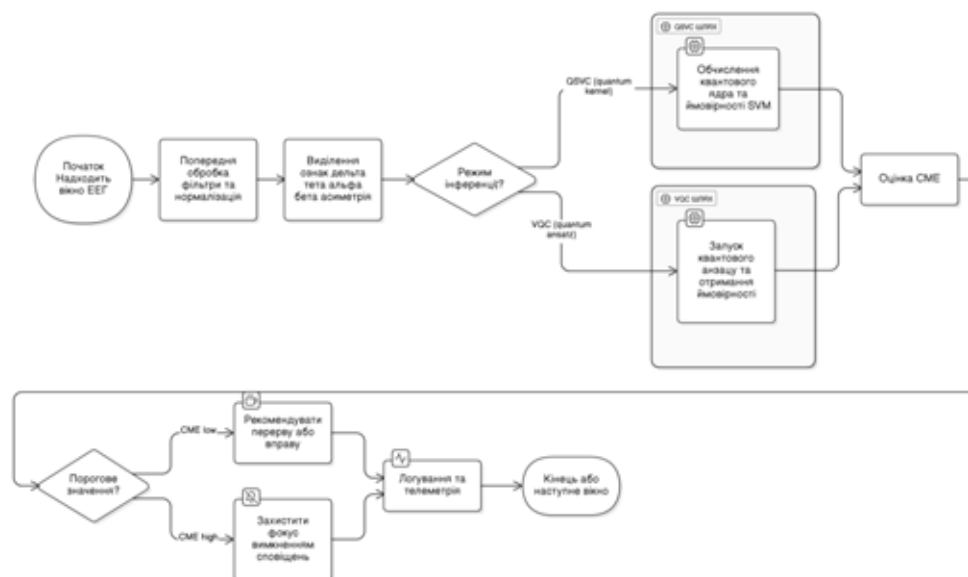


Fig. 2. Activity diagram “CME system and real-time recommendations”

Conclusions and future work

This research demonstrates the practicality and effectiveness of hybrid quantum–classical approaches for training machine-learning models on contemporary NISQ hardware. By establishing a rigorous mathematical foundation for quantum data encoding, kernel-based classification, and variational modeling, and by embedding these formulations into a multi-objective optimization framework, the study provides a systematic methodology for improving quantum model performance in resource-constrained environments. The integration of advanced metaheuristic algorithms proved essential for navigating the highly non-convex parameter landscapes of quantum circuits and mitigating the detrimental effects of noise and limited circuit depth. The developed quantum microservice architecture further confirms that scalable, modular, and fault-tolerant quantum computation workflows can be achieved through software-centric engineering methods. The experimental results, including the application to EEG-based cognitive-state analysis, illustrate the potential of quantum-enhanced analytics for real-world data processing tasks. The validation of the CME indicator within the hybrid framework suggests new opportunities for quantum approaches in neuroengineering, human–computer interaction, and biomedical signal analysis. Overall, the outcomes of the study highlight a viable technological pathway for advancing quantum machine learning prior to the emergence of fully fault-tolerant quantum processors, and they open multiple directions for future research in optimization, circuit design, and quantum software engineering. Future work will focus on extending the hybrid quantum–classical framework in several directions to further enhance its robustness, scalability, and applicability across domains.

Bibliography

1. Cattani G.H., Quemy A. Case-Based and Quantum Classification for ERP-Based Brain–Computer Interfaces. *Brain Sciences*, 2023. 13(2), 303. <https://doi.org/10.3390/brainsci13020303> (access date: 05.11.2025).
2. Olvera C., Ross O.M., Rubio Y. EEG-Based Motor Imagery Classification with Quantum Algorithms. *Expert Systems with Applications*, 2024. 247, 123354. <https://doi.org/10.1016/j.eswa.2024.123354> (access date: 05.11.2025).
3. Becerra-Sánchez P., del Campo Á., Bustos A. and Hernández-González S. Feature Selection Model Based on EEG Signals to Assess the Cognitive Workload in Drivers (GALoRIS). *Sensors*, 2020. 20(20), 5881. <https://doi.org/10.3390/s20205881> (access date: 05.11.2025).
4. IBM Quantum. Quantum kernel methods. IBM Quantum Learning. URL: <https://quantum.cloud.ibm.com/learning/courses/quantum-machine-learning/quantum-kernel-methods> (access date: 05.11.2025).
5. PennyLane Community. Variational classifier circuit. *PennyLane Discussion Forum*, 2021. 12 July. URL: <https://discuss.pennylane.ai/t/variational-classifier-circuit/1187> (access date: 05.11.2025).
6. Chen S.Y.-C. Evolutionary Optimization for Designing Variational Quantum Circuits with High Model Capacity. *arXiv preprint arXiv:2412.12484v1*. 2024. URL: <https://arxiv.org/html/2412.12484v1> (access date: 05.11.2025).
7. Masood A. Quantum Sundays [28] QUBO, Inc.: Selecting the Right Solver in a NISQ World. 2025. URL: <https://medium.com/@adnanmasood/quantum-sundays-27-qubo-inc-selecting-the-right-solver-in-a-nisq-world-0ee797d3f2c3> (access date: 05.11.2025).

8. Vernik M. and Oleshchenko L.M. Wearable EEG Framework with Time-Aligned Audio Segmentation for Quantifying Countable Mental Energy. *Trends, Issues, and Challenges in Modern Science: Proceedings of the II International Scientific Conference*, Cambridge, UK, 5 September 2025, pp. 90–92. <https://doi.org/10.64076/iedc250905.17> (access date: 06.11.2025).

9. Vernik M., Oleshchenko L.M. Intelligent Software Framework for Big Data Audio Processing and Cognitive Energy Estimation. *Вчені записки ТНУ імені В.І. Вернадського. Серія: Технічні науки*, 2025. 36(75), 3, С. 125–131. <https://doi.org/10.32782/2663-5941/2025.3.2/17> (access date: 06.11.2025).

References

1. Cattani G.H., Quemy A. (2023). Case-Based and Quantum Classification for ERP-Based Brain–Computer Interfaces. *Brain Sciences*, 13(2), 303. <https://doi.org/10.3390/brainsci13020303> (access date: 05.11.2025).

2. Olvera C., Ross O.M., Rubio Y. (2024). EEG-Based Motor Imagery Classification with Quantum Algorithms. *Expert Systems with Applications*, 247, 123354. <https://doi.org/10.1016/j.eswa.2024.123354> (access date: 05.11.2025).

3. Becerra-Sánchez P., del Campo Á., Bustos A. and Hernández-González S. (2020). Feature Selection Model Based on EEG Signals to Assess the Cognitive Workload in Drivers (GALoRIS). *Sensors*, 20(20), 5881. <https://doi.org/10.3390/s20205881> (access date: 05.11.2025).

4. IBM Quantum. Quantum kernel methods. IBM Quantum Learning. URL: <https://quantum.cloud.ibm.com/learning/courses/quantum-machine-learning/quantum-kernel-methods> (access date: 05.11.2025).

5. PennyLane Community (2021). Variational classifier circuit. *PennyLane Discussion Forum*, 12 July. URL: <https://discuss.pennylane.ai/t/variational-classifier-circuit/1187> (access date: 05.11.2025).

6. Chen S.Y.-C. (2024). Evolutionary Optimization for Designing Variational Quantum Circuits with High Model Capacity. *arXiv preprint arXiv:2412.12484v1*. URL: <https://arxiv.org/html/2412.12484v1> (access date: 05.11.2025).

7. Masood A. (2025). Quantum Sundays |28| QUBO, Inc.: Selecting the Right Solver in a NISQ World. URL: <https://medium.com/@adnanmasood/quantum-sundays-27-qubo-inc-selecting-the-right-solver-in-a-nisq-world-0ee797d3f2c3> (access date: 05.11.2025).

8. Vernik M. and Oleshchenko L.M. (2025). Wearable EEG Framework with Time-Aligned Audio Segmentation for Quantifying Countable Mental Energy. *Trends, Issues, and Challenges in Modern Science: Proceedings of the II International Scientific Conference*, Cambridge, UK, 5 September 2025, pp. 90–92. <https://doi.org/10.64076/iedc250905.17> (access date: 06.11.2025).

9. Vernik M., Oleshchenko L.M. (2025). Intelligent Software Framework for Big Data Audio Processing and Cognitive Energy Estimation. *Scientific Notes of V.I. Vernadsky Taurida National University. Series: Technical Sciences*, 36(75), 3, pp. 125–131. <https://doi.org/10.32782/2663-5941/2025.3.2/17> (access date: 06.11.2025).

Дата першого надходження рукопису до видання: 21.11.2025
Дата прийнятого до друку рукопису після рецензування: 19.12.2025
Дата публікації: 31.12.2025

І. В. ГАВРИШ

студент II курсу магістратури
за спеціальністю 121 Інженерія програмного забезпечення
Запорізький національний університет
ORCID: 0009-0007-5294-2146

А. Г. КРИВОХАТА

кандидат фізико-математичних наук,
доцент кафедри програмної інженерії
Запорізький національний університет
ORCID: 0000-0001-9664-0659

РОЗРОБКА АРХІТЕКТУРИ АДАПТИВНОГО ІНТЕРФЕЙСУ РЕДАКТОРА ВЕБ-ЗАСТОСУНКУ ДЛЯ СТВОРЕННЯ ВІРТУАЛЬНИХ ТУРІВ

Стрімкий розвиток та поширення технологій віртуальної реальності сприяє їх активному використанню у сферах туризму, нерухомості та освіти. Одним із способів представлення інформації у віртуальному середовищі є інтерактивні VR-тури. У роботі проаналізовано існуючі системи для створення та редагування віртуальних екскурсій, досліджено сучасні підходи до побудови мобільних інтерфейсів і виявлено низку суттєвих недоліків їх поточних реалізацій. Зокрема, встановлено, що більшість сервісів мають обмежений функціонал на мобільних пристроях, особливо щодо роботи з меню інструментів, керування сценою та стабільності відображення дрібних елементів на малих екранах.

Метою роботи є розробка та тестування архітектури адаптивного інтерфейсу редактора веб-застосунку віртуальних екскурсій, здатної повністю відтворювати функціональність настільної версії в умовах обмеженого екранного простору. Запропонована архітектура ґрунтується на стратегії Mobile-First, системі модульних компонентів і принципах розподілу робочих зон інтерфейсу. Вона включає структурні елементи, такі як верхня панель навігації, робочий простір 3D-сцени, bottom sheet та бічне меню інструментів. Такий підхід забезпечує раціональне використання наявної площі екрана, чіткий поділ між навігацією, інструментарієм та середовищем перегляду віртуальної сцени, а також передбачувану поведінку інтерфейсу під час зміни розміру вікна перегляду.

Для практичної перевірки архітектури було створено макет інтерфейсу з використанням фреймворку Vue.js. Запропонований підхід протестовано за допомогою набору автоматизованих UI-тестів з використанням Playwright, що охоплюють перевірку розташування, масштабування та доступності елементів під час роботи з різними розмірами вікна перегляду (viewport). Результати підтвердили відсутність конфліктів верстки, коректну роботу компонентів навігації і стабільність відображення навіть на малих екранах (наприклад, iPhone 5). Таким чином, архітектура демонструє ефективність і практичну придатність для створення інтуїтивних та зручних інтерфейсів для редакторів віртуальних екскурсій, орієнтованих на мобільні пристрої.

Ключові слова: віртуальні екскурсії, адаптивний інтерфейс, мобільний дизайн, UI/UX дизайн, автоматизоване тестування, архітектура інтерфейсу.

I. V. HAVRYSH

2nd year Master's Student
of Specialty 121 Software Engineering
Zaporizhzhia National University
ORCID: 0009-0007-5294-2146

A. G. KRYVOKHATA

Candidate of Physical and Mathematical Sciences,
Associate Professor at the Department of Software Engineering
Zaporizhzhia National University
ORCID: 0000-0001-9664-0659

DEVELOPMENT OF THE ARCHITECTURE OF AN ADAPTIVE INTERFACE OF THE EDITOR OF A WEB APPLICATION FOR CREATING VIRTUAL TOURS

The rapid development and spread of virtual reality technologies facilitates their active use in the fields of tourism, real estate, and education. VR tours are one of the ways to present information in a virtual environment. This paper analyzes existing systems for creating and editing virtual tours, examines modern approaches to building mobile interfaces, and identifies several significant limitations in their current implementations. In particular, it has been found that most services have limited functionality on mobile devices, particularly those related to tool menus, scene control, and the stability of displaying tiny elements on small screens.

The goal of this work is to develop and test the architecture of an adaptive interface for an editor of a virtual tour web application that is capable of fully replicating the functionality of the desktop version in a limited screen space. The proposed architecture is based on the Mobile-First strategy, a system of modular components, and the principles of distributing interface workspaces. It includes structural elements such as the top navigation bar, 3D scene workspace, bottom sheet, and navigation drawer. This approach ensures rational use of the available screen space, a clear separation between navigation, tools, and the virtual scene working environment, as well as predictable interface behavior when the viewport size changes.

An interface mockup was created using the Vue.js framework to test the architecture in practice. The proposed approach was tested using a set of automated UI tests with Playwright, covering the verification of the location, scaling, and accessibility of elements when working with different viewport sizes. The results confirmed the absence of layout conflicts, the correct operation of navigation components, and stable display even on small screens (e.g., iPhone 5). Thus, the architecture demonstrates its effectiveness and practical applicability for creating intuitive and convenient interfaces for virtual tour editors for mobile devices.

Key words: virtual tours, adaptive interface, mobile design, UI/UX design, automated testing, interface architecture.

Постановка проблеми

Упродовж останнього десятиліття технології віртуальної (англ. VR – Virtual Reality) та доповненої реальності (англ. AR – Augmented Reality) зазнали стрімкого розвитку й набули широкого застосування у засобах масової комунікації, навчання й цифрового маркетингу. Інтеграція цих технологій у сферу інтернет-ресурсів, освіти, архітектури та ігрової індустрії змінила способи взаємодії з контентом завдяки впровадженню нових підходів, що ґрунтуються на принципах повної або часткової присутності користувача у віртуальному просторі. VR і AR дозволяють користувачам не лише споглядати інформацію, а й взаємодіяти з нею у тривимірному середовищі, що підвищує рівень залучення, мотивації та ефективності засвоєння даних. Актуальність цих технологій підтверджується їх активним впровадженням у освітні процеси [1, 2], професійну підготовку [3, 4], медичну практику [5] та бізнес-сектор [6], де створення якісного користувацького досвіду є важливим завданням.

Значну роль у поширенні технологій VR та AR відіграє WebXR – відкритий стандарт, розроблений W3C, який забезпечує єдиний програмний інтерфейс для створення та відтворення XR-контенту безпосередньо у веб-браузері. Практична реалізація програмного забезпечення з підтримкою WebXR здійснюється за допомогою спеціалізованих фреймворків, зокрема A-Frame, Three.js та Babylon.js, які надають високий рівень абстракції для створення інтерактивних 3D-середовищ [7]. Ці інструменти дозволяють швидко прототипувати та розгортати складні тривимірні сцени безпосередньо у браузері, а також адаптувати їх під різноманітні бізнес-завдання. Одним із поширених варіантів застосування даних технологій стали VR-тури 360 – інтерактивні віртуальні екскурсії, що поєднують панорамні фотографії, відео та інколи тривимірну графіку з елементами навігації для створення справжнього імерсивного досвіду, достовірного відтворення реальних локацій та пам'яток культури.

Водночас активне зміщення медіаспоживання в бік мобільних пристроїв підвищує вимоги до гнучкості та зручності користувацьких інтерфейсів. Сучасні підходи до UX/UI-дизайну наголошують, що мобільний UI не може бути простою адаптацією десктопної версії – він потребує окремого проєктування за принципом Mobile-First, який ставить у пріоритет простоту, контекстність та інтуїтивну взаємодію [8]. У VR-застосунках зазначене завдання ускладнюється додатковими чинниками: тривимірним середовищем, використанням жестового та сенсорного керування, обмеженою продуктивністю мобільних пристроїв, а також відсутністю звичних фізичних інструментів введення.

Попри наявний попит на створення віртуальних екскурсій, аналіз ринку демонструє, що лише невелика кількість існуючих застосунків пропонує повнофункціональні мобільні інтерфейси для створення VR-контенту [9, 10]. Інші сервіси або обмежують функціональні можливості роботи на мобільних пристроях, або взагалі не надають їх. Ключова проблема полягає у складності адаптації багатофункціонального інтерфейсу для роботи з 3D-об'єктами до умов обмеженого екранного простору та сенсорного керування. Дослідження у галузі перевірки використовуваності (англ. usability) VR-застосунків для туризму свідчать, що навіть базові інтерфейси потребують ретельного тестування для забезпечення позитивного користувацького досвіду [11, 12]. Тому завдання проєктування повноцінного адаптивного інтерфейсу редактора для мобільних платформ залишається актуальною науково-практичною проблемою.

Аналіз останніх досліджень і публікацій

Автори дослідження [13] зазначають, що адаптація веб-інтерфейсів досягається за допомогою принципів Responsive Web Design (RWD), які передбачають використання гнучких сіток, адаптивних зображень, відносних одиниць вимірювання та CSS media queries для динамічної зміни компоновки залежно від розміру та типу екрана. Даний підхід, запропонований в [14], став стандартом для більшості сучасних веб-застосунків, оскільки забезпечує єдину кодову базу для різних типів пристроїв та мінімізує потребу у створенні окремих версій сайту. Однак під час створення складних інтерактивних систем, зокрема веб-редакторів, проста зміна компоновки елементів інтерфейсу не гарантує забезпечення оптимального користувацького досвіду, оскільки сценарії використання суттєво відрізняються між десктопними та мобільними середовищами.

На відміну від RWD, який динамічно адаптує єдину HTML-сторінку до розміру вікна перегляду (англ. viewport), Adaptive Web Design (AWD) використовує заздалегідь підготовлені версії макета сайту, щоб забезпечити його оптимальний вигляд і роботу на різних пристроях. У роботі [15] вказано, що такий підхід забезпечує точний контроль над користувацьким досвідом, але вимагає додаткових ресурсів для підтримки та оновлення декількох версій одночасно. Підхід AWD використовує серверну або клієнтську логіку для визначення характеристик пристрою і надає відповідний варіант інтерфейсу, оптимізованого під конкретний варіант взаємодії та роздільну здатність екрана. Це є особливо ефективним у складних інтерактивних системах, де якість взаємодії залежить не лише від розміру екрана, а й від способу введення, обчислювальної потужності пристрою та швидкості Інтернет з'єднання.

Підхід Desktop-First, характерний для звичних веб-застосунків, може призводити до зниження зручності, інтуїтивності та продуктивності при використанні на мобільних пристроях. Зокрема, багатоклонкові макети та статичні панелі інструментів, ефективні в десктопному середовищі, на малих екранах можуть створювати надмірне когнітивне навантаження та ускладнювати процес навігації. У дослідженні [16] зазначається, що використання архітектурних шаблонів для десктопних систем без врахування контексту мобільних пристроїв призводить до втрати інтуїтивності та ефективності користувацького інтерфейсу. Він стверджує, що сучасне проєктування має базуватися на принципах Mobile-First, які передбачають створення дизайну насамперед для мобільного середовища з подальшою адаптацією до більших екранів. Такий підхід забезпечує логічну послідовність елементів, робить навігацію веб-застосунком більш інтуїтивною та зручною, і покращує доступність контенту.

Дизайн ефективних мобільних макетів, описаних у роботі [17], ґрунтується не лише на адаптивній верстці, а й на застосуванні розповсюджених практик до взаємодії з користувачем за допомогою таких елементів графічного інтерфейсу, як панелі навігації, меню та модальні компоненти, що забезпечують логічну структуру інтерфейсу та зручну взаємодію з застосунком. Подібні елементи, згідно з дослідженням [18], допомагають зосередити увагу користувача на основних діях, що сприяє пришвидшенню адаптації до роботи з додатком та покращує ефективність виконання цільових задач.

Водночас успішна адаптація інтерфейсу до мобільних платформ неможлива без урахування архітектурних особливостей побудови застосунків. Як зауважено у роботі [19], недостатньо забезпечити лише візуальну адаптивність елементів інтерфейсу. Архітектура мобільного UI має комплексно враховувати вимоги до продуктивності, масштабованості та доступності, оскільки саме ці характеристики безпосередньо визначають якість користувацького досвіду та стабільність роботи системи.

Отже, з огляду літератури можна зробити висновок, що сучасна адаптація веб-інтерфейсів виходить за межі класичного RWD і трансформується у комплексний підхід, що поєднує принципи дизайну Mobile-First, оптимізації структури макету в залежності від контексту взаємодії та використання типових елементів керування та навігації. Такий підхід дозволяє створювати інтерфейси, які забезпечують ефективну взаємодію з користувачем, високу продуктивність та зручність використання саме в умовах мобільних платформ.

Формулювання мети дослідження

Метою дослідження є розробка та тестування архітектури адаптивного інтерфейсу для редактора сцен VR-турів. Запропонована концепція має забезпечити ефективну взаємодію з інтерактивними та мультимедійними елементами віртуального середовища в умовах обмеженості ресурсів мобільних пристроїв. Особливу увагу приділено відображенню складних багаторівневих меню та панелей налаштувань без втрати функціональності.

Викладення основного матеріалу дослідження

Адаптація веб-інтерфейсів під мобільні пристрої вимагає не лише гнучкої верстки чи масштабування елементів, а передусім – архітектурного переосмислення взаємодії користувача з інтерфейсом. Сучасні підходи ґрунтуються на принципах Mobile-First, контекстної взаємодії та зменшення використання зайвих елементів, що є ключовими складовими нативного UX-дизайну. Відповідно до цих принципів, інтерфейс має забезпечувати інтуїтивну навігацію, зручне керування контентом із застосування сенсорного введення та адаптацію контекстних елементів до обмеженого простору екрана. Ключовим архітектурним рішенням є ієрархічна організація інформації, коли основний функціонал займає більшу частину вільного простору екрану, а вторинні дії приховуються у модальних вікнах, висувних панелях та контекстних меню. Ергономіка сенсорної взаємодії вимагає розміщення

ключових елементів керування переважно в нижній частині екрана. Окрім цього, важливу роль відіграє керування станами інтерфейсу, коли набір інструментів повністю змінюється залежно від обраного режиму (наприклад, режим перегляду проти режиму редагування). Ефективна мобільна архітектура також повинна пропагувати використання сенсорних жестів (наприклад, стискання для масштабування або свайпи) для виконання дій, що раніше вимагали окремих кнопок.

Архітектура адаптивного інтерфейсу, зображеного на рисунку 1, для мобільних пристроїв веб-застосунку редактора віртуальних турів побудована за модульно-компонентним принципом і складається з чотирьох основних структурних компонентів:

- верхня панель керування (рис. 1.а) – компонент із фіксованим розташуванням у верхній частині екрану, який відображає вміст та дії, пов'язані з поточною сценою. Панель може використовуватись для розміщення брендингу, заголовків екранів, елементів навігації та основних контекстних дій. Головне призначення панелі – забезпечення швидкого доступу до глобальних дій без перевантаження робочого простору;

- тіло, або робочий простір сцени (рис. 1.б) – центральний компонент архітектури, який відображає вміст тривимірного середовища та дозволяє користувачу взаємодіяти з 3D-об'єктами за допомогою сенсорних жестів. Компонент безпосередньо використовується як контейнер для WebXR-сумісних фреймворків (A-Frame, Three.js, Babylon.js), що дозволяє відокремити логіку побудови 3D-графіки від решти користувацького інтерфейсу. Додаткові елементи керування можна реалізувати з використанням плаваючих кнопок для основних дій або контекстних елементів, що з'являються лише при взаємодії з об'єктами сцени;

- bottom sheet (рис. 1.в) – універсальний компонент призначений для розміщення додаткового контенту, елементів меню або дій, закріплений у нижній частині екрана. Може підтримувати два стани: згорнутий – виступає в ролі елемента “footer” та відображає допоміжну інформацію, і розгорнутий – займає частковий або весь простір екрану. У згорнутому виді можна розмістити, наприклад, горизонтальний список всіх об'єктів сцени для забезпечення більш зручної та швидкої навігації у 3D-середовищі. У розгорнутому стані, bottom sheet може мати свій макет інтерфейсу, що дозволяє ізолювати додатковий або контекстно-залежний функціонал та запобігти перевантаженню основного робочого простору. Наприклад, такий підхід можна використати для відображення панелі налаштувань об'єктів сцени, реалізованої з використанням вкладок (англ. tabs) для логічного групування елементів та використання текстових полів, слайдерів та випадаючих меню для редагування параметрів налаштувань;

- бічне меню (англ. navigation drawer) (рис. 1.г) – компонент інтерфейсу, який використовується для забезпечення як основної, так і додаткової навігації застосунком та може відкриватися/закриватися за допомогою іконки меню навігації. Він є типовим для мобільних застосунків і може надавати як постійний доступ до глобальних елементів навігації, так і мати власну реалізацію в залежності від екрану. В контексті редактора віртуальних турів, бічне меню може використовуватись для відображення списку всіх сцен туру для забезпечення швидких переходів між ними.

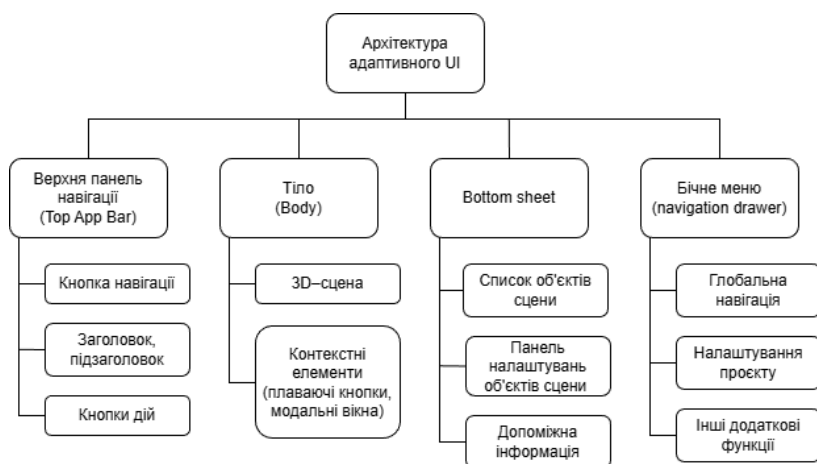


Рис. 1. Архітектура адаптивного UI: а) верхня панель навігації; б) тіло; в) bottom sheet; г) бічне меню

Таким чином, в основі архітектури лежить поєднання адаптивного та компонентного підходів. Структура інтерфейсу дотримується ключових принципів мобільного дизайну: чіткої ієрархії, модульності та відображення інформації в залежності від контексту. Особливу увагу приділено використанню звичних принципів взаємодії – основні елементи керування розміщені в нижній частині екрана для зручного доступу, а діалогові вікна реагують на жести, знайомі користувачам пристроїв з операційними системами Android та iOS. Така організація забезпечує природну взаємодію та наближає використання веб-застосунку до досвіду роботи з мобільними додатками.

Для перевірки працездатності архітектури та оцінки коректності поведінки компонентів інтерфейсу у різних умовах було створено практичний приклад реалізації редактора, який відтворює ключові сценарії взаємодії користувача. Реалізація побудована за принципом Mobile-First і відповідає вимогам до продуктивності, інтуїтивності та стабільної роботи на мобільних пристроях. Загальна структура (рис. 2) передбачає три основні складові: систему навігації, що забезпечує перехід між сценами та збереження результатів; робочий простір 3D-сцени, де користувач взаємодіє з елементами віртуальної екскурсії (наприклад, на базі A-Frame); та панель налаштувань, яка містить контекстно-залежні вкладки для керування параметрами вибраних елементів. Практичну реалізацію створено з використанням фреймворку Vue.js, однак архітектура залишається універсальною й може бути впроваджена за допомогою інших бібліотек або класичного JavaScript у поєднанні з CSS.



Рис. 2. Загальна структура інтерфейсу

Система навігації є базовим елементом інтерфейсу редактора та забезпечує доступ до ключових функцій застосунку – переходу між сценами, збереженням результатів роботи та поверненню до попереднього екрану. Одними з основних компонентів для навігації на мобільних пристроях є верхня панель керування та бічне меню. Відповідно до прототипу, зображеному на рисунку 3, панель керування складається з кнопки “Назад”, назви поточної сцени та кнопки збереження змін. Таке компонування елементів відповідає звичним практикам побудови інтерфейсів та сприяє більш швидкій адаптації користувачів до роботи з застосунком.

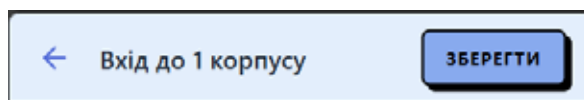


Рис. 3. Реалізація верхньої панелі керування

Бічне меню містить список всіх сцен (рис. 4), з яких складається віртуальна екскурсія, що забезпечує зручну та ефективну навігацію між ними. Кожен елемент списку відображає мініатюру сцени та її назву, що дозволяє користувачу швидко ідентифікувати потрібну локацію. Використання висувної панелі замість постійно видимого списку дозволяє оптимально використати обмежений простір мобільного екрану.

Компонент робочого простору 3D-сцени (рис. 5) є основним місцем взаємодії користувача з віртуальним середовищем для додавання, редагування та розміщення елементів сцени. За реалізацію цього компоненту відповідають фреймворки, що забезпечують створення віртуальних 3D-середовищ з підтримкою стандарту WebXR, наприклад, A-Frame. Фреймворк відповідає за такі ключові механізми взаємодії з VR-контентом, як рендеринг панорам 360 та примітивів, відображення інтерактивних елементів, зображень тощо. Користувач може повертати сцену за допомогою жестів (свайп, перетягування), змінювати масштаб та взаємодіяти з об'єктами з використанням сенсорного керування.

За додавання нових елементів до сцени відповідає плаваюча кнопка з меню (англ. floating action button menu), яка є одним з компонентів дизайн-системи Material Design 3 від Google. Такий тип кнопки активно використовується в застосунках Google, наприклад, Google Drive або Keep Notes, для відображення декількох варіантів додавання нового ресурсу. В розробленому макеті, при натисканні на кнопку розгортається список доступних

типів об'єктів: хотспоти для переходу між сценами, мультимедійні елементи (аудіо, відео, зображення) та текстові анотації. Після вибору потрібного типу, об'єкт розміщується відповідно до координат розташування віртуального курсора, який потім користувач може вільно перемістити в межах сцени.



Рис. 4. Реалізація бічного меню



Рис. 5. Макет компоненту 3D-сцени

Панель налаштувань, зображена на рисунку 6, виконує функцію контекстного інструменту редагування, який динамічно змінює вміст в залежності від обраного об'єкта сцени. Для відображення такого складного елементу інтерфейсу використовується компонент bottom sheet. У згорнутому стані (рис. 6.а) він відображає горизонтальний список всіх об'єктів сцени, що дозволяє швидко переключатись між ними. У розгорнутому (рис. 6.б) – надає доступ до параметрів налаштувань обраного об'єкта. Всі доступні налаштування логічно згруповані та розділені на окремі вкладки. Це сприяє структурованості інтерфейсу та зменшує когнітивне навантаження при роботі з великою кількістю параметрів. Для кожного типу параметрів використовуються оптимальні елементи керування, що забезпечують баланс між точністю налаштування та зручністю взаємодії. Зокрема, позиціонування об'єктів у сцені реалізовано з використанням комбінації слайдера та числового поля введення для координат X, Y та Z. Для параметрів, що мають обмежений набір варіантів, таких як колір тексту, або цільова сцена для переходу, використано випадаючі списки. Всі зміни автоматично застосовуються до обраного об'єкта та відображаються у вікні сцени. Користувач може змінювати стан панелі за допомогою свайпу вгору/вниз або натискання на відповідний маркер.

Важливим аспектом реалізації є оптимізація продуктивності при роботі з панеллю налаштувань на мобільних пристроях з обмеженими ресурсами. Вміст вкладок має завантажуватись динамічно лише при їх активації, щоб зменшити початкове навантаження на систему та прискорити відкриття панелі. Також можливе застосування

техніки дебаунсингу (англ. debouncing) для обробки змін параметрів, що запобігає надмірному оновленню сцени при швидкому оновленні значень за допомогою слайдерів чи текстових полів.

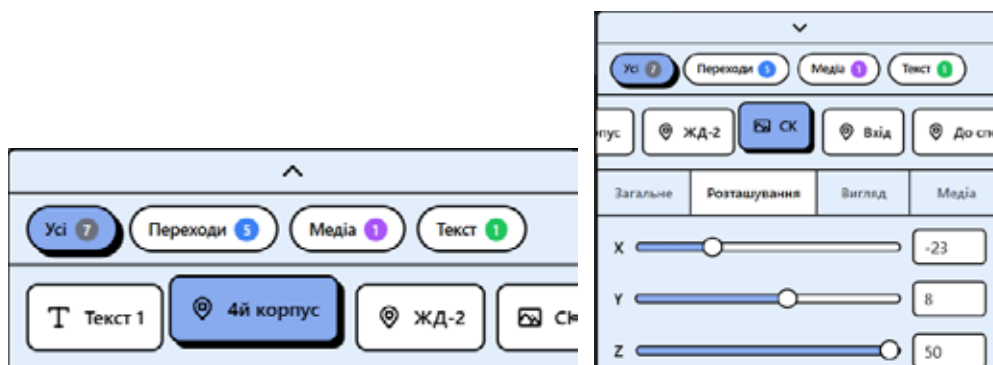


Рис. 6. Макет панелі налаштувань об'єктів сцени: а) згорнутий вид; б) розгорнутий вид

Тестування є одним з етапів циклу розробки програмного забезпечення та необхідне для забезпечення стабільності, передбачуваності роботи та якості кінцевого продукту. Тестування дозволяє виявити помилки на стадії реалізації застосунку, зменшити ризики помилок та перевірити відповідність системи функціональним та нефункціональним вимогам. Автоматизоване тестування забезпечує швидку та систематичну перевірку функціональності застосунку шляхом автоматичного виконання попередньо визначених тестових сценаріїв, що зменшує потребу в ручному тестуванні. Автори дослідження [20] відзначають, що автоматизація суттєво знижує витрати на підтримку програмного забезпечення та розширює тестове покриття, що є критичним для комплексних та масштабних проєктів. Існують різні види автоматизованого тестування, зокрема модульне тестування, інтеграційні тести, UI-тести та тестування продуктивності, кожен з яких призначений для перевірки окремих компонентів системи або їх взаємодії. Для тестування фронтенд-застосунків та мобільних веб-інтерфейсів широко застосовуються фреймворки Jest, Vitest, Mocha, Playwright, Cypress, Selenium WebDriver, Appium тощо, які дозволяють автоматизувати як перевірку логіки, так і змодельовати взаємодію користувача з системою. Використання зазначених інструментів для тестування адаптивних інтерфейсів дозволяє перевіряти коректність роботи UI при різних розмірах екранів, конфігураціях пристроїв та сценаріях взаємодії.

Для перевірки запропонованої архітектури адаптивного інтерфейсу було розроблено набір автоматизованих тестів з використанням Playwright [21]. Playwright – це фреймворк з відкритим кодом для автоматизації комплексного та кроссбраузерного тестування веб-застосунків. Серед особливостей цього фреймворку можна виділити підтримку емуляції мобільних браузерів Google Chrome та Safari, необхідних для перевірки роботи адаптивного інтерфейсу. Метою тестування є оцінка ефективності обраного архітектурного підходу для створення мобільного інтерфейсу редактора VR-турів. Тести охоплюють валідацію таких елементів UI, як панель навігації, бічне меню, сцена та компоненти bottom sheet, і спрямовані на перевірку роботи інтерфейсу, його адаптивності для типових сценаріїв використання користувачем на різних мобільних пристроях. Тестування дозволить виявити такі проблеми у структурі інтерфейсу, як накладання елементів один на одного, неправильне масштабування компонентів або порушення їхнього розташування. Крім того, перевірка переходів між екранами та зміни відображуваних даних спрямована на оцінку роботи системи навігації застосунку та взаємодії його компонентів.

Для тестування коректної роботи адаптивності інтерфейсу було підібрано набір поширених значень розмірів вікна перегляду, які відповідають розмірам екранів існуючих мобільних пристроїв (табл. 1) – від компактних смартфонів (iPhone 5/SE, iPhone 8) до сучасних фаблетів (iPhone 12 Pro Max). Увагу приділено критичним точкам, де найчастіше виникають проблеми з розташуванням елементів при адаптації веб-сторінки до поточної ширини екрану. Такий підхід дозволяє перевірити поведінку інтерфейсу в умовах різної щільності пікселів, співвідношення сторін і фізичних розмірів екрану, що забезпечує комплексну оцінку адаптивності інтерфейсу.

Таблиця 1

Використанні значення розміру вікна перегляду для тестування адаптивності [22]

Пристрій	Розмір екрану в пікселях	Розмір вікна перегляду
iPhone 5/SE	640 x 1136	320 x 568
iPhone 8	750 x 1334	375 x 667
iPhone 12 Pro Max	2778 x 1284	428 x 926

На рисунку 7 представлено приклад автоматизованого тесту, який перевіряє адаптивність бічного меню при зміні розміру вікна перегляду. Тест має забезпечити, що всі пункти меню залишаються видимими, не перекривають одне одного та не виходять за межі екрану. Така перевірка гарантує, що навігація залишається стабільною та доступною незалежно від розміру пристрою.

```
test('Drawer Menu adapts to different screen sizes', async ({ page }) => {
  const drawerElements = ['mobile-drawer-overlay', 'mobile-drawer-header', 'mobile-drawer-scenes-list'];

  for (const viewport of viewports) {
    await page.setViewportSize(viewport);

    // Open the drawer
    await page.getByTestId('drawer-trigger').click();

    const drawerMenu = page.getByTestId('mobile-drawer-menu');
    await expect(drawerMenu).toBeVisible();

    await Promise.all(
      drawerElements.map(elementTestId =>
        expect(page.getByTestId(elementTestId)).toBeVisible()
      )
    );

    console.log('Drawer menu adapts correctly to viewport ${viewport.width}x${viewport.height}');

    // Close the drawer before next iteration
    await page.getByTestId('drawer-close-button').click();
    await expect(drawerMenu).toBeHidden();
  }
});
```

Рис. 7. Приклад автоматизованого тесту

Результати виконання повного набору автоматизованих тестів підтвердили стабільність роботи мобільного інтерфейсу редактора (рис. 8). Валідацію компонентів – верхньої панелі навігації, бічного меню, робочої області сцени та панелі bottom sheet – на трьох цільових розмірах viewport було виконано без виявлення конфліктів верстки чи втрати функціональності. Інтерфейс демонстрував коректну поведінку при зміні станів, зокрема під час навігації, відкриття меню та трансформації панелей. Отримані дані свідчать про ефективність розробленої архітектури адаптивного інтерфейсу у забезпеченні передбачуваного користувацького досвіду на мобільних пристроях.

Simulator Mobile UI - Adaptive UI		
✓ Bottom Sheet adapts to different screen sizes	Mobile Chrome	24s
✓ Drawer Menu adapts to different screen sizes	Mobile Chrome	25s
✓ Header adapts to different screen sizes	Mobile Chrome	16s
✓ Scene Body adapts to different screen sizes	Mobile Chrome	11s
Simulator Mobile UI - Bottom Sheet		
✓ Opens and closes to fixed heights correctly	Mobile Chrome	11s
✓ Contains expected content	Mobile Chrome	19s
✓ Shows important UI elements when closed	Mobile Chrome	7s
✓ Covers less important UI elements when opened	Mobile Chrome	6s
Simulator Mobile UI - Drawer Menu		
✓ Contents are scrollable when expanded	Mobile Chrome	12s
✓ Contents fit within bottom sheet width	Mobile Chrome	16s
Simulator Mobile UI - Drawer Menu		
✓ Toggles correctly	Mobile Chrome	14s
✓ Clicking on the overlay area closes the drawer	Mobile Chrome	16s
✓ Contains header, scenes list, and close button	Mobile Chrome	14s
✓ Displays over other content	Mobile Chrome	15s
✓ Clicking on a scene navigates to that scene	Mobile Chrome	15s

Рис. 8. Результати виконання повного набору тестів

Висновки

У роботі проаналізовано проблему відсутності єдиної архітектури користувацького інтерфейсу для створення та редагування VR-турів 360 на мобільних пристроях. Виявлено, що доступні рішення обмежені у використанні через складність адаптації комплексних інтерфейсів до екранів смартфонів та особливостей сенсорного керування, пропонують лише часткову функціональність або взагалі не підтримують адаптивний дизайн. Для вирішення цієї задачі запропоновано та обґрунтовано архітектуру адаптивного інтерфейсу, що базується на підході Mobile-First та використанні контекстно-залежних елементів, таких як компонент bottom sheet, модальні вікна та меню. Розроблена архітектурна модель спрямована на забезпечення повної підтримки функціональності редактора віртуальних екскурсій без перевантаження доступного екранного простору. Перевірку запропонованого

рішення було здійснено шляхом реалізації прототипу та проведення серії автоматизованих тестів на різних розмірах екрану. Результати тестування підтвердили коректність розташування та масштабування елементів, а також надійність механізмів навігації на мобільних пристроях. Таким чином, розроблене рішення забезпечує надійну основу для створення сучасного та масштабованого адаптивного інтерфейсу для редактора віртуальних екскурсій.

Перспективи подальших досліджень пов'язані з оптимізацією продуктивності рендерингу 3D-сцен та компонентів інтерфейсу на мобільних пристроях із низькою обчислювальною потужністю, що залишається важливим технічним обмеженням. Крім того, перспективним напрямком є дослідження можливостей масштабування запропонованої архітектури для планшетів або пристроїв зі складними екранами. Також доцільно провести розширені емпіричні дослідження адаптивного інтерфейсу за участі реальних користувачів для оцінки його використовуваності під час тривалої роботи з редактором. Актуальним напрямком є й дослідження розробки інтерфейсу редактору віртуальних екскурсій для використання у режимі повного занурення безпосередньо у VR-гарнітурах.

Список використаної літератури

1. Burke D., Crompton H., Nickel C. The Use of Extended Reality (XR) in Higher Education: A Systematic Review. *TechTrends*. 2025. № 69. P. 998-1011. URL: <https://doi.org/10.1007/s11528-025-01092-y>
2. Thangavel S., Sharmila K., Sufina K. Revolutionizing Education Through Augmented Reality (AR) and Virtual Reality (VR): Innovations, Challenges and Future Prospects. *Asian Journal of Interdisciplinary Research*. 2025. Vol. 8, Iss. 1. P. 1-28. URL: <https://doi.org/10.54392/ajir2511>
3. Wang K. Research and Application of VR in Training and Learning. *Highlights in Science, Engineering and Technology*. 2024. Vol. 85. P. 1306–1313. URL: <https://doi.org/10.54097/anhqyy16>
4. Shahu A., Kinzer K., Michahelles F. Enhancing Professional Training with Single-User Virtual Reality: Unveiling Challenges, Opportunities, and Design Guidelines. *MUM '23: Proceedings of the 22nd International Conference on Mobile and Ubiquitous Multimedia (Vienna, 3-6 December 2023)*. Vienna, 2024. P. 224-256. URL: <https://doi.org/10.1145/3626705.3627791>
5. Virtual reality and augmented reality in medical education: an umbrella review / T. Tene, D. F. V. Lopez, P. E. V. Aguerre et al. *Frontiers in Digital Health*. 2024. Vol. 6:1365345. DOI: 10.3389/fdgh.2024.1365345
6. What is XR? Towards a Framework for Augmented and Virtual Reality / P. A. Rauschnabel, R. Felix, C. Hinsch et al. *Computers in Human Behavior*. 2022. Vol. 133. P. 107289. URL: <https://doi.org/10.1016/j.chb.2022.107289>
7. Macario G. WebXR, A-Frame and Networked-Aframe as a Basis for an Open Metaverse: A Conceptual Architecture. 2024. URL: <https://arxiv.org/abs/2404.05317>
8. Govindaraju M. The Impact Of Mobile-First Development Approaches On User Experience And Development Efficiency. *International Journal of Creative Research Thoughts*. 2025. Vol. 13, no. 3. URL: <https://www.ijcrt.org/papers/IJCRT25A3445.pdf>
9. Kuula: #1 Virtual Tour Software for any business. *Kuula*. URL: <https://kuula.co/> (дата звернення: 20.09.2025).
10. Virtual Tour Easy. *360° Virtual Tour software*. URL: <https://virtualtoureasy.com/> (дата звернення: 20.09.2025).
11. Fang Y.-M., Lin C. The Usability Testing of VR Interface for Tourism Apps. *Applied Sciences*. 2019. Vol. 9(16). P. 3215. URL: <https://doi.org/10.3390/app9163215>
12. Kamińska D., Zwoliński G., Laska-Leśniewicz A. Usability Testing of Virtual Reality Applications—The Pilot Study. *Sensors*. 2022. Vol. 22(4). P. 1342. URL: <https://doi.org/10.3390/s22041342>
13. Bhanarkar N., Paul A., Mehta A. Responsive web design and its impact on user experience. *International Journal of Advanced Research in Science, Communication and Technology*. 2023. Vol. 3, Is. 4. P. 50-55. DOI: 10.48175/ijarsct-9259
14. Marcotte E. *Responsive Web Design*. 2nd ed. New York : A Book Apart, 2014. 153 p.
15. Responsive web design: principles and implementation / R. Rupali, D. P. Mankame, B. Patil et al. *International Journal of Research Publication and Reviews*. 2024. Vol. 5, Is. 7. P. 4291-4293. URL: <https://ijrpr.com/uploads/V5ISSUE7/IJRPR31795.pdf>
16. Prakash B. Retrofitting Mobile First Design, Responsive Design: Driving Factors, Approach, Best Practices and Design Considerations. *Current Trends in Computer Sciences & Applications*. 2020. Vol. 2, Is. 2. DOI: 10.32474/CTCSA.2020.02.000131
17. Kim H., Moritz D., Hullman J. Design Patterns and Trade-Offs in Responsive Visualization for Communication. 2021. URL: <https://arxiv.org/abs/2104.07724>
18. Understanding Visual Saliency in Mobile User Interfaces / L. A. Leiva et al. 2021. URL: <https://arxiv.org/abs/2101.09176>
19. Bilohub D., Skrypchenko M., Tytenko S. Quality attributes and architectural patterns of modern mobile apps. *Modern Engineering And Innovative Technologies*. 2023. № 1(29-01). P. 33-38. URL: <https://doi.org/10.30890/2567-5273.2023-29-01-056>
20. Gu R., Rojas J. M., Shin D. Software testing for extended reality applications: a systematic mapping study. *Automated Software Engineering*. 2025. № 32(56). URL: <https://doi.org/10.1007/s10515-025-00523-7>

21. Fast and reliable end-to-end testing for modern web apps. *Playwright*. URL: <https://playwright.dev/> (date of access: 07.11.2025).
22. Common Viewport Sizes for Mobile. *Award Winning Web Design Newport – Digital Marketing*. URL: <https://www.icwebdesign.co.uk/common-viewport-sizes> (date of access: 07.11.2025)

References

1. Burke, D., Crompton, H., & Nickel, C. (2025). The Use of Extended Reality (XR) in Higher Education: A Systematic Review. *TechTrends*, (69), 998-1011. <https://doi.org/10.1007/s11528-025-01092-y>
2. Thangavel, S., Sharmila, K., & Sufina, K. (2025). Revolutionizing Education Through Augmented Reality (AR) and Virtual Reality (VR): Innovations, Challenges and Future Prospects. *Asian Journal of Interdisciplinary Research*, 8(1), 1-28. <https://dx.doi.org/10.2139/ssrn.5195697>
3. Wang, K. (2024). Research and Application of VR in Training and Learning. *Highlights in Science, Engineering and Technology*, 85, 1306-1313. <https://doi.org/10.54097/anhqyy16>
4. Shahu, A., Kinzer, K., & Michahelles, F. (2024). Enhancing Professional Training with Single-User Virtual Reality: Unveiling Challenges, Opportunities, and Design Guidelines. *Proceedings of the 22nd International Conference on Mobile and Ubiquitous Multimedia*, 224-256. <https://doi.org/10.1145/3626705.3627791>
5. Tene, T., Vique López, D. F., Valverde Aguirre, P. E., Orna Puente, L. M., & Vacacela Gomez, C. (2024). Virtual reality and augmented reality in medical education: an umbrella review. *Frontiers in Digital Health*, 6:1365345. <https://doi.org/10.3389/fdgth.2024.1365345>
6. Rauschnabel, P. A., Felix, R., Hinsch, C., Shahab, H., & Alt, F. (2022). What is XR? Towards a framework for augmented and virtual reality. *Computers in human behavior*, 133, 107289. <https://doi.org/10.1016/j.chb.2022.107289>
7. Macario, G. (2024). WebXR, A-Frame and Networked-Aframe as a Basis for an Open Metaverse: A Conceptual Architecture. <https://arxiv.org/abs/2404.05317>
8. Govindaraju, M. (2025). The Impact Of Mobile-First Development Approaches On User Experience And Development Efficiency. *International Journal of Creative Research Thoughts*, 13(3). <https://www.ijcrt.org/papers/IJCRT25A3445.pdf>
9. Kuula: #1 Virtual Tour Software for any business. (n.d.). Kuula. <https://kuula.co/>
10. Virtual Tour Easy. (n.d.). *360° Virtual Tour software*. <https://virtualtoureasy.com/>
11. Fang, Y.-M., & Lin, C. (2019). The Usability Testing of VR Interface for Tourism Apps. *Applied Sciences*, 9(16), 3215. <https://doi.org/10.3390/app9163215>
12. Kamińska, D., Zwoliński, G., & Laska-Leśniewicz, A. (2022). Usability Testing of Virtual Reality Applications – The Pilot Study. *Sensors*, 22(4), 1342. <https://doi.org/10.3390/s22041342>
13. Bhanarkar, N., Paul, A., & Mehta, A. (2023). Responsive Web Design and Its Impact on User Experience. *International Journal of Advanced Research in Science, Communication and Technology*, 3(4), 50-55. <https://doi.org/10.48175/ijarsct-9259>
14. Marcotte, E. (2014). *Responsive web design* (2nd ed.). New York : A Book Apart.
15. R, R., P Mankame, D., Patil, B., & Dhavalgi, V. (2024). Responsive Web Design: Principles and Implementation. *International Journal of Creative Research Thoughts*, 5(7), 4291-4293. <https://ijrpr.com/uploads/V5ISSUE7/IJRPR31795.pdf>
16. Prakash, B. (2020). Retrofitting Mobile First Design, Responsive Design: Driving Factors, Approach, Best Practices and Design Considerations. *Current Trends in Computer Sciences & Applications*, 2(2). <https://doi.org/10.32474/ctcsa.2020.02.000131>
17. Kim, H., Moritz, D., & Hullman, J. (n.d.). Design Patterns and Trade-Offs in Responsive Visualization for Communication. <https://arxiv.org/abs/2104.07724>
18. Leiva, L. A., Xue, Y., Bansal, A., Tavakoli, H. R., Köroğlu, T., Dayama, N. R., & Oulasvirta, A. (n.d.). Understanding Visual Saliency in Mobile User Interfaces. <https://arxiv.org/abs/2101.09176>
19. Bilohub, D., Skrypchenko, M., & Tytenko, S. (2023). QUALITY ATTRIBUTES AND ARCHITECTURAL PATTERNS OF MODERN MOBILE APPS. *Modern engineering and innovative technologies*, 1(29-01), 33-38. <https://doi.org/10.30890/2567-5273.2023-29-01-056>
20. Gu, R., Rojas, J. M., & Shin, D. (2025). Software testing for extended reality applications: a systematic mapping study. *Automated Software Engineering*, 32(56). <https://doi.org/10.1007/s10515-025-00523-7>
21. Fast and reliable end-to-end testing for modern web apps. (n.d.). Playwright. <https://playwright.dev/>
22. Common Viewport Sizes for Mobile. (n.d.). Award Winning Web Design Newport – Digital Marketing. <https://www.icwebdesign.co.uk/common-viewport-sizes>

Дата першого надходження рукопису до видання: 25.11.2025
 Дата прийнятого до друку рукопису після рецензування: 12.12.2025
 Дата публікації: 31.12.2025

Т. О. ГРІНЕНКО

кандидат технічних наук, доцент,
доцент кафедри безпеки інформаційних технологій
Харківський національний університет радіоелектроніки
ORCID: 0000-0002-8251-8991

Д. М. МАЛЯРОВА

здобувач магістерського рівня кафедри безпеки інформаційних технологій
Харківський національний університет радіоелектроніки
ORCID: 0009-0003-2506-6970

О. П. НАРЕЖНИЙ

кандидат технічних наук,
доцент кафедри кібербезпеки інформаційних систем, мереж і технологій
Харківський національний університет імені В. Н. Каразіна
ORCID: 0000-0003-4321-0510

ПРОТОТИП ВЕБ-СЕРВІСУ КВАНТОВОГО ГЕНЕРАТОРА ВИПАДКОВИХ ЧИСЕЛ ДЛЯ ЗАХИЩЕНОГО ОБМІНУ ПОВІДОМЛЕННЯМИ

У даній роботі особливу увагу приділено криптографічному механізму HMAC як одному з найбільш ефективних засобів забезпечення автентичності та цілісності повідомлень. Досліджено принципи функціонування HMAC, описано його модифікацію HMAC-SHA-256, який визнано сучасним стандартом у безпечному обміні даними. Порівняння різних реалізацій HMAC виявило баланс між стійкістю та продуктивністю, що робить його придатним для використання у веб-сервісах з високими вимогами до продуктивності та безпеки, а саме, у веб-сервісі QRNG.

Обґрунтований вибір інструментів та середовища розробки прототипу веб-сервісу QRNG, наданий опис його основних компонентів та побудовані блок-схеми логіки автентифікації користувачів та обміну приватними повідомленнями. Надані результати тестування ключових функцій системи захисту за різними сценаріями, включаючи захист від підміни, помилкової автентифікації та обробку виняткових ситуацій.

Ефективне забезпечення захищеного обміну повідомленнями у веб-сервісі QRNG потребує комплексного підходу, що включає впровадження надійних криптографічних методів, надійне управління сесіями та ключами, а також застосування сучасних систем автентифікації. Такий підхід дозволяє не лише забезпечити високий рівень безпеки, але й підтримувати високу продуктивність веб-сервісу при високих навантаженнях, що є критично важливим для його стабільної роботи в сучасному цифровому середовищі.

Комплексний підхід щодо побудови захищеного обміну повідомленнями у веб-сервісі є необхідним для створення надійного і безпечного веб-сервісу QRNG, який буде здатним протистояти сучасним кіберзагрозам і забезпечувати безпеку даних.

Ключові слова: QRNG, HMAC, SHA-256, TLS, WEB, автентифікація, веб-сервіс, захищений обмін, квантовий генератор випадкових чисел, клієнт-серверна архітектура, код автентифікації повідомлень, парольна автентифікація.

Т. О. HRINENKO

PhD, Associate Professor,
Associate Professor at the Department of Information Technologies Security
Kharkiv National University of Radio Electronics
ORCID: 0000-0002-8251-8991

D. M. MALIAROVA

Master's Degree Student at the Department of Information Technology Security
Kharkiv National University of Radio Electronics
ORCID: 0009-0003-2506-6970

O. P. NARIEZHNI

PhD, Associate Professor at the Department of Cybersecurity
of Information Systems, Networks and Technologies
V. N. Karazin Kharkiv National University
ORCID: 0000-0003-4321-0510

PROTOTYPE OF A QUANTUM RANDOM NUMBER GENERATOR WEB SERVICE FOR SECURE MESSAGE EXCHANGE

This work pays special attention to the HMAC cryptographic mechanism as one of the most effective means of ensuring message authenticity and integrity. The principles of HMAC operation are examined, and its modification HMAC-SHA-256 – recognized as a modern standard in secure data exchange – is described. A comparison of different HMAC implementations revealed a balance between security and performance, which makes it suitable for use in web services with high performance and security requirements, including the QRNG web service.

The choice of tools and the development environment for the QRNG web service prototype is substantiated, and its main components are described. Block diagrams of the user authentication logic and private messaging processes are presented. The results of testing the system's key security functions under various scenarios are provided, including protection against spoofing, false authentication, and the handling of exceptional situations.

Effective secure message exchange in the QRNG web service requires a comprehensive approach that includes the implementation of reliable cryptographic methods, secure and reliable session and key management, as well as the use of modern authentication systems. Such an approach not only ensures a high level of security but also maintains the high performance of the web service under heavy loads, which is critically important for its stable operation in the modern digital environment.

A comprehensive approach to building secure message exchange within the web service is essential for creating a reliable and secure QRNG platform capable of withstanding modern cyber threats and ensuring data security.

Key words: QRNG, HMAC, SHA-256, TLS, web, authentication, web service, secure communication, client-server architecture, message authentication code, password authentication.

Постановка проблеми

Веб-сервіси стали ключовими інструментами обміну даними, організації бізнес-процесів і комунікації. Сучасні веб-сервіси обробляють значний обсяг персональних, фінансових та корпоративних даних, що робить їх привабливою мішенню для кіберзлочинців. Найпоширенішими атаками залишаються SQL-ін'єкції, міжсайтовий скриптинг (Cross-Site Scripting, XSS), підробка міжсайтових запитів (Cross-Site Request Forgery, CSRF), атака типу «людина посередині» (Man-In-The-Middle, MITM), атаки грубої сили, атаки на відмову в обслуговуванні (Distributed Denial of Service, DDoS-атаки) й інші [1].

Застосування веб-сервісів квантового генератора випадкових чисел (Quantum Random Number Generator, QRNG) стало критично важливим для забезпечення безпеки банківських систем, інтернету речей та інших високочутливих платформ. Наприклад, при шифруванні даних QRNG виступають як надійний інструмент для створення криптографічних ключів. Однак їх інтеграція в реальні системи породжує нові ризики, пов'язані з потенційними атаками на апаратному та програмному рівнях [2].

Ключовими проблемами, з якими стикаються веб-сервіси QRNG в контексті безпеки, є захист від атак типу перехоплення повідомлень (MITM-атаки) та повторного відтворення повідомлень (replay attacks). MITM-атаки можуть призвести до серйозних наслідків, оскільки перехоплені повідомлення можуть бути підроблені або використані для несанкціонованого доступу до конфіденційних даних. Повторне відтворення повідомлень (replay attacks), при якому зловмисники намагаються повторно використати перехоплені дані, є ще однією важливою загрозою, що вимагає ефективних заходів захисту.

Також серйозну загрозу для веб-сервісів QRNG становлять атаки на геш-функції, особливо при використанні алгоритмів, що можуть бути вразливими до колізій. Зловмисники можуть модифікувати геш-значення для зміни переданих даних, що ускладнює процес забезпечення цілісності інформації. Зокрема, такі атаки можуть бути спрямовані на порушення надійності системи автентифікації та перевірки цілісності даних.

Також важливою загрозою є слабкість традиційних методів автентифікації, таких як парольна автентифікація, що не здатні забезпечити належний рівень безпеки у випадку атаки зловмисників. Простота використання паролів і ймовірність їх компрометації або витоку робить цей метод уразливим до атак типу brute-force, фішинг, а також до атак на системи зберігання паролів. Сучасні зловмисники можуть також застосовувати атаки на системи зберігання паролів, що дозволяє їм отримувати доступ до захищених ресурсів без необхідності обходити інші механізми безпеки. У зв'язку з цим, виникає необхідність у впровадженні більш складних методів автентифікації, таких як багатофакторна автентифікація або коди автентифікації повідомлень, які гарантують більш високий рівень захисту.

Однак, навіть при використанні сучасних криптографічних алгоритмів, існує ще одна важлива проблема – управління секретними ключами. Ключі, що використовуються для криптографічного захисту даних, повинні бути надійно захищені від витоку або компрометації, оскільки компрометація ключа може призвести до втрати всієї безпеки системи. У зв'язку з цим, важливим є впровадження надійних механізмів для зберігання і обміну ключами, а також створення засобів їхнього автоматичного оновлення, щоб мінімізувати ризики, пов'язані з їх компрометацією.

Іншим важливим аспектом є забезпечення продуктивності веб-сервісу QRNG у високонавантажених середовищах. Використання складних криптографічних операцій може негативно впливати на швидкість роботи сервісу, що є критичним в умовах високих навантажень і масштабованих рішень (наприклад, у веб-сервісах з великою кількістю одночасних сесій). Тому важливо знаходити баланс між високим рівнем захисту і необхідною продуктивністю системи, що дозволяє підтримувати її ефективність навіть при значних навантаженнях.

Враховуючі вказані загрози, основними напрямками, які потребують вирішення, для забезпечення захищеного обміну повідомленнями у веб-сервісі QRNG, є:

- вибір криптографічних алгоритмів, що є стійкими до атак на колізії та перехоплення даних. Вони повинні бути надійними та ефективними для використання у веб-сервісах, де критично важлива продуктивність;
- безпечне управління ключами, яке охоплює зберігання, обмін і оновлення ключів. Надійні механізми захисту ключів є важливою умовою забезпечення високого рівня безпеки;
- забезпечення продуктивності при високих навантаженнях, де необхідно балансувати між швидкістю виконання криптографічних операцій та рівнем захисту;
- розробка заходів протидії MITM-атакам і підробці повідомлень, які включають використання сучасних методів автентифікації, криптографічних механізмів перевірки цілісності і збереження конфіденційності при передачі даних.

Вирішення цих проблем потребує комплексного підходу, що поєднує технічні та організаційні заходи для забезпечення надійного захисту даних у веб-сервісі QRNG.

Аналіз останніх досліджень і публікацій

Аналіз наукових публікацій [1-5] та стандартів [6-10] показує, що проблема захищеного обміну повідомленнями у веб-сервісах залишається надзвичайно актуальною, особливо через зростаючу кількість кібератак. Аналіз відповідних стандартів дозволяє сформулювати архітектуру сервісу QRNG з урахуванням актуальних вимог до захисту даних, зокрема щодо конфіденційності, цілісності та автентичності.

Протокол WebSocket, описаний у RFC 6455 [6] є основою для реалізації двостороннього зв'язку між клієнтом і сервером. Він підтримує постійне TCP-з'єднання, завдяки чому обмін повідомленнями відбувається без повторного встановлення HTTP-запитів, що суттєво знижує затримки під час передачі повідомлень. Оскільки базова версія WebSocket не має вбудованого шифрування чи автентифікації, у веб-сервісі передбачається використання захищеного з'єднання wss://, яке базується на протоколі TLS. Це дозволяє забезпечити конфіденційність і цілісність даних при їх передачі мережею, що є обов'язковою умовою для протидії атакам типу «людина посередині».

Міжнародний стандарт ISO/IEC 27002:2022 [7] надає рекомендації щодо впровадження контролю інформаційної безпеки, включаючи використання криптографічних методів для забезпечення цілісності та автентичності даних. У межах захищеного обміну повідомленнями актуальним є застосування механізму HMAC, що дозволяє перевірити, чи було повідомлення змінено під час передачі. Крім того, можливе збереження унікального ідентифікатора користувача у локальному сховищі, що дозволяє підтримувати сесійну взаємодію з сервером.

НД ТЗІ 2.5-010-03 [8] встановлює вимоги до технічних та організаційних заходів захисту інформації WEB-сторінки в мережі Інтернет. Нормативний документ містить вимоги до захищеності інформації та визначає необхідність використання криптографічних засобів при обміні даними та впровадження механізмів автентифікації, що повністю відповідає вимогам до безпечного обміну повідомленнями.

Федеральний стандарт FIPS 198-1 [9] визначає механізм HMAC як надійний спосіб автентифікації повідомлень на основі криптографічної геш-функції та секретного ключа. Додатково, спеціальна публікація NIST SP 800-224 [10] конкретизує технічні вимоги до реалізації HMAC, включаючи вимоги до довжини ключа (не менше 128 біт), використання сучасних геш-функцій, зокрема SHA-256, та неприпустимість обрізання тегу автентифікації.

Аналіз наукових джерел і нормативних документів підтверджує необхідність впровадження комплексних заходів захисту у веб-сервісі QRNG. Зокрема, застосування HMAC-SHA-256 у поєднанні з рекомендаціями ISO/IEC, FIPS та NIST дозволяє досягти високого рівня безпеки при обміні повідомленнями.

Формулювання мети дослідження

Метою дослідження є обґрунтування вибору алгоритму HMAC та його модифікації HMAC-SHA-256 для автентифікації повідомлень у відкритому мережевому середовищі; розробка прототипу веб-сервісу QRNG для захищеного обміну повідомленнями з використанням механізмів автентифікації та перевірки цілісності; тестування ключових компонентів системи захисту за різними сценаріями, включаючи захист від підміни, помилкової автентифікації та обробку виняткових ситуацій.

Викладення основного матеріалу дослідження

1. Прототип веб-сервісу QRNG

Схема апаратно-програмного комплексу QRNG наведена на рисунку 1. Основною складовою комплексу є QRNG, інтерфейсний модуль IEEE 488 та сервер, роль якого виконує Raspberry Pi.

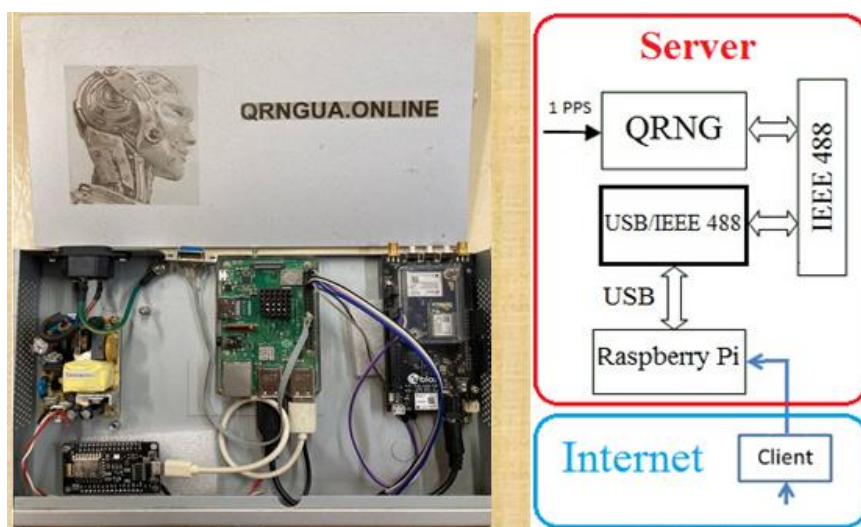


Рис. 1. Схема прототипу апаратно-програмного комплексу QRNG

Користувачі можуть отримати доступ до QRNG як джерела генерування випадкових послідовностей через сайт, де послідовність надається у вигляді рядка цифр розміром послідовності до 1024 чисел (рис. 2). Або через розроблений прототип веб-сервісу QRNG, який за допомогою запитів надає послідовності випадкових чисел у декількох форматах (рис. 3).



Рис. 2. Приклад сторінки сайту з QRNG

Варіант роботи у вигляді веб-сервісу розроблений за стандартом SOAP, а всі результати повертаються у вигляді JSON повідомлень. Доступ до веб-сервісу отримують тільки зареєстровані користувачі, які отримали особистий ключ доступу (api key). Веб-сервіс надає користувачу послугу – генерація випадкової послідовності чисел. Послідовність випадкових чисел надається у вигляді масиву, розмір масиву та формат елементів масиву задається користувачем при запиті. Розмір масиву може бути довільним, а формат елементів має два варіанти:

десятькові числа (uint8) або шістнадцятькові (hex8). Особистий ключ доступу (api key) обов'язково передається при кожному запиті.

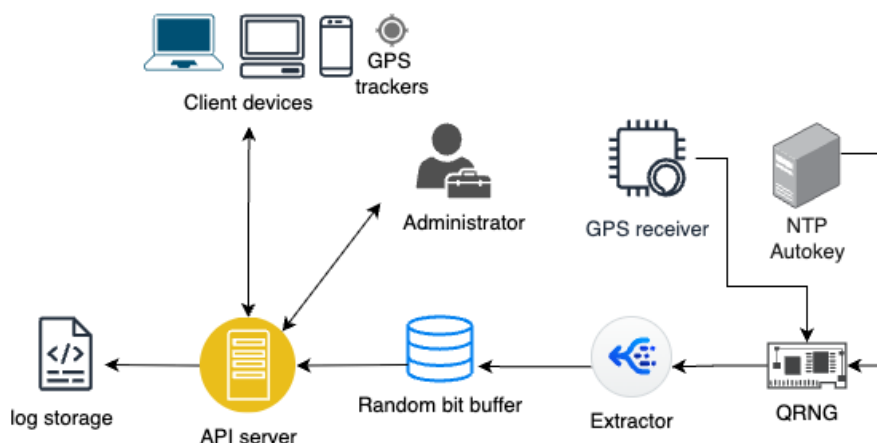


Рис. 3. Схема прототипу сервісу QRNG

2. Комплексний підхід щодо забезпечення безпеки обміну повідомленнями у веб-сервісі QRNG

Побудова моделі загроз та моделі порушника є одним із ключових етапів аналізу інформаційної безпеки, що дозволяє своєчасно виявити критичні вразливості веб-сервісу QRNG, оцінити потенційні ризики та обґрунтувати доцільність впровадження конкретних засобів захисту. Для забезпечення належної безпеки в веб-сервісі QRNG необхідно не тільки запобігати конкретним типам атак, але ще і виявляти слабкі місця в архітектурі та інфраструктурі сервісу. У роботі [11] досліджені типові архітектури таких сервісів, визначені потенційні вектори атак та профіль порушника, а також обґрунтовані та надані рекомендації щодо посилення інформаційної безпеки даного класу систем.

Ефективне забезпечення захищеного обміну повідомленнями у веб-сервісі QRNG потребує комплексного підходу, що включає впровадження надійних криптографічних методів, надійне управління сесіями та ключами, а також застосування сучасних систем автентифікації та захисту від атак типу MITM. Такий підхід дозволяє не лише забезпечити високий рівень безпеки, але й підтримувати високу продуктивність веб-сервісу при високих навантаженнях, що є критично важливим для його стабільної роботи в сучасному цифровому середовищі.

Ключовим елементом захисту є застосування криптографії – основного засобу забезпечення конфіденційності та цілісності даних. Одним із найпоширеніших механізмів захисту є TLS (Transport Layer Security), який забезпечує шифрування всіх переданих повідомлень, знижуючи ризик їх перехоплення або модифікації при передачі через незахищену мережу. Використання протоколу TLS дозволяє реалізувати захищене з'єднання між клієнтом і сервером, захищаючи дані від атак типу «людина посередині» (MITM). Однак для досягнення високого рівня безпеки важливо застосовувати сучасні версії TLS з підтримкою складніших механізмів автентифікації та шифрування, що додатково знижує ризики підслуховування та маніпуляцій з даними.

Також важливим інструментом є HMAC (Hash-based Message Authentication Code) – криптографічний механізм для перевірки автентичності і цілісності повідомлень. HMAC використовує секретний ключ у поєднанні з геш-функцією для створення спеціального коду автентифікації повідомлень. Це дозволяє точно перевірити, чи не було змінено передане повідомлення під час його доставки, забезпечуючи високу надійність захисту. Для сучасних веб-сервісів рекомендованим є використання HMAC-SHA-256 – модифікації HMAC, що використовує одну з найстійкіших геш-функцій сімейства SHA-2. Цей алгоритм забезпечує високу стійкість до атак на колізії та є ідеальним рішенням для збереження цілісності даних в умовах сучасних кіберзагроз.

Однак навіть використання найсучасніших криптографічних методів не може гарантувати повної безпеки, якщо управління сесіями та криптографічними ключами некоректно налаштовано. Адже компрометація секретних ключів або сесійних токенів може призвести до значних витоків даних і втрати контролю над системою. У зв'язку з цим важливим є використання апаратних модулів захисту для надійного зберігання і захисту криптографічних ключів, що забезпечує максимальний рівень безпеки. Крім того, важливим є впровадження процедур для регулярного оновлення ключів, що дозволяє знизити ризики від тривалого використання одного і того ж ключа.

Продуктивність при високих навантаженнях є важливим аспектом, оскільки криптографічні операції, зокрема шифрування та створення кодів автентифікації, можуть суттєво впливати на швидкість роботи веб-сервісу. Для підтримки високої продуктивності в умовах великих обсягів даних і значної кількості одночасних сесій важливо

оптимізувати криптографічні операції, застосовуючи паралельні обчислення або апаратне прискорення криптографії, що дозволяє мінімізувати затримки при забезпеченні високого рівня захисту.

Для запобігання несанкціонованому доступу до ресурсів веб-сервісу необхідно також застосовувати механізми автентифікації користувачів. Найпоширенішим методом є пароліна автентифікація, однак цей метод має свої обмеження та вразливості, зокрема, він є схильним до атак типу brute-force, фішинг, а також витоків облікових даних. Тому багато веб-сервісів використовують багатофакторну автентифікацію (Multifactor Authentication, MFA), яка поєднує паролі з додатковими методами перевірки, такими як одноразові паролі, що генеруються за допомогою мобільних застосунків, або біометричні дані.

Іншим важливим аспектом є управління сесіями. Для запобігання несанкціонованому доступу до облікових записів веб-сервіси використовують сесійні токени, що зберігаються у захищених місцях, таких як cookie, доступні лише через захищене з'єднання. Крім того, важливо впроваджувати механізми автоматичного завершення сесії після певного часу неактивності, що допомагає знизити ризик несанкціонованого доступу до облікових записів.

Для захисту від атак MITM та забезпечення автентичності переданих даних використовуються SSL/TLS сертифікати. Вони підтверджують автентичність серверів і забезпечують шифрування переданої інформації, запобігаючи її перехопленню або модифікації. Використання сертифікатів дозволяє впевнитися в тому, що сервер, з яким взаємодіє клієнт, є дійсно тим, за кого себе видає, і що з'єднання між клієнтом і сервером є захищеним від зовнішніх атак.

Комплексний підхід щодо побудови захищеного обміну повідомленнями у веб-сервісі є необхідним для створення надійного і безпечного веб-сервісу QRNG, який буде здатним протистояти сучасним кіберзагрозам і забезпечувати безпеку даних.

3. Коди автентифікації повідомлень

У захищених веб-сервісах, для забезпечення автентичності повідомлень використовуються коди автентифікації повідомлень (Message Authentication Code, MAC), які дозволяють перевірити, що повідомлення дійсно надійшло від заявленого джерела та не було змінено в процесі передавання. Одним із ключових елементів побудови MAC-кодів є криптографічно стійка геш-функція. Саме криптографічно стійка геш-функція є необхідною умовою для побудови MAC-кодів, оскільки вона забезпечує стійкість до базових криптоатак, зокрема підбору прообразу та пошуку колізій.

На основі геш-функцій будуються ключові геш-коди, які на вхід приймають повідомлення та секретний ключ. Ці механізми дають змогу гарантувати цілісність і автентичність даних без залучення додаткових протоколів, базуючись на принципах симетричної криптографії [12].

MAC-функція h , повинна задовольняти таким умовам [12]:

- результат $h(K;X)$ має фіксовану довжину n біт, X може бути довільної довжини, K – це секретний ключ;
- при наявності даних h і X (але з невідомим K), повинно бути складно визначити $h(K;X)$ з імовірністю успіху значно більшою $1/2^n$.

Окремо можна виділити три підходи побудови MAC-кодів:

- MAC-коди, побудовані на основі БСШ (CBC-MAC);
- MAC-коди, побудовані на основі безключових геш-функцій (HMAC);
- MAC-коди, побудовані на основі сімейства універсальних гешфункцій (UMAC).

4. Коди автентифікації повідомлень на основі геш-функцій

Серед багатьох методів гешування та способів реалізації кодів автентифікації повідомлень, особливе місце займає HMAC (Hash-based Message Authentication Code, HMAC) – вкладає конструкція, яка обчислює MAC для основної функції геша h , повідомлення X і секретний ключ K . Альтернативою для HMAC є конструкції MDx-MAC, які можуть бути засновані на MD5, SHA, RIPEMD або аналогічних геш-функціях [13].

Код автентифікації повідомлень HMAC широко використовується у сучасних веб-сервісах і протоколах обміну даними. Його застосування забезпечує високий рівень захисту від підробки повідомлень, втручання та інших загроз. HMAC визначено як обов'язковий (mandatory to implement) механізм у межах IPsec, а також він застосовується в ключових протоколах TLS (Transport Layer Security, TLS), який замінив SSL, і SET (Secure Electronic Transaction, SET). На відміну від SSL, де для створення MAC використовувався застарілий MD5, TLS реалізує HMAC із сучасними геш-функціями, такими як SHA-256, що значно підвищує рівень криптографічного захисту.

HMAC став одним із ключових інструментів побудови безпечних веб-сервісів. Його використання забезпечує баланс між високим рівнем захисту, продуктивністю та простотою реалізації, що робить його оптимальним вибором для побудови захищених систем обміну повідомленнями.

Конструкція HMAC має такий загальний вигляд [9]:

$$MAC(text) = HMAC(K, text) = H((K_0 \oplus opad) \| H((K_0 \oplus ipad) \| text)),$$

де B – розмір блоку (у байтах) вхідних даних для затвердженої геш-функції, L – розмір блоку (у байтах) вихідних даних затвердженої геш-функції, H – затверджена геш-функція, K – секретний ключ, спільний між

відправником та цільовим одержувачем, $text$ – дані, для яких обчислюється $HMAC$, K_0 – ключ K після попередньої обробки для формування B -байтового ключа, $opad$ – зовнішній блок байту $0x5c$ повторений B разів, $ipad$ – внутрішній блок байту $0x36$ повторений B разів, $\parallel$ – операція конкатенації, $\oplus$ – операція виключне АБО.

5. Порівняльний аналіз швидкодії HMAC алгоритмів

Швидкодія алгоритмів HMAC значною мірою залежить від обраної геш-функції, її реалізації та обчислювального середовища.

У таблиці 1 наведені значення швидкодії різних реалізацій HMAC для декількох типів процесорів. Швидкість обчислень визначається кількістю циклів процесора, затрачених на обробку одного байту повідомлення [13].

Геш-функції сімейства SHA-1 та SHA-2 (SHA-256, SHA-384, SHA-512) демонструють гарний баланс між криптографічною стійкістю та швидкодією [14, 15]. Хоча HMAC-MD5 і HMAC-MD4 демонструють вищу швидкодію, вони не рекомендуються до використання через вразливості. Зокрема, MD5 був зламаний на практиці, що дозволило підробляти цифрові підписи та сертифікати.

Таблиця 1

Порівняння швидкодії HMAC алгоритмів, Мбіт/с

Алгоритм	Довжина MAC-коду (біт)	Довжина ключа (біт)	Тип ПЕОМ				
			Pentium2	PIII/Linux	Pentium4	Xeon	AMD
HMAC-Whirlpool	512	512	86	72	98	103	100
HMAC-MD4	128	512	4.7	4.7	6.4	6.4	4.7
HMAC-MD5	128	512	7.2	7.3	9.4	9.4	7.4
HMAC-RIPEMD	160	512	23	18	27	26	21
HMAC-SHA-0	160	512	16	15	23	23	13
HMAC-SHA-1	160	512	16	15	25	24	12
HMAC-SHA-2	256	512	40	39	40	39	33
	384		84	84	124	132	72
	512		84	84	124	132	72
HMAC-Tiger	192	512	24	21	28	26	20

Таким чином, HMAC є універсальним та надійним механізмом автентифікації повідомлень. Серед усіх розглянутих варіантів HMAC-SHA-256 є найбільш рекомендованим для впровадження у веб-сервіси QRNG та інших відкритих середовищах передачі даних, де актуальні загрози MITM-атак, підробки або втрати цілісності інформації.

6. Основні компоненти захищеного веб-сервісу QRNG

На етапі практичної реалізації веб-сервісу QRNG для захищеного обміну повідомленнями основна увага приділялася забезпеченню автентифікації користувачів, перевірці цілісності та достовірності повідомлень, а також захисту даних від несанкціонованого доступу. Впровадження механізму HMAC на основі SHA-256 дозволяє гарантувати як цілісність переданих даних, так і їхню автентичність.

Архітектура веб-сервісу QRNG (рис. 3) побудована на взаємодії двох основних компонентів – клієнта та сервера – через захищене WebSocket-з'єднання. Кожен компонент має чітко визначені обов'язки:

- клієнтська частина відповідає за введення та валідацію даних, формування структур запитів, створення HMAC-кодів для перевірки цілісності повідомлень, підтримку сесії за допомогою cookie, візуалізацію повідомлень у чаті та обробку відповідей від сервера;
- серверна частина здійснює обробку вхідних підключень, автентифікацію користувачів, перевірку формату та достовірності повідомлень, передавання повідомлень між клієнтами, обробку винятків, а також контроль активності та статусу підключених користувачів.

Ключові механізми захисту та логіки функціонування системи включають:

- встановлення захищеного TLS-з'єднання – весь трафік між клієнтом і сервером передається через HTTPS та WSS (для обміну повідомленнями), що гарантує конфіденційність і захист від перехоплення;
- автентифікацію користувача, яка здійснюється шляхом порівняння HMAC, згенерованого на основі введеного пароля, із записом у базі даних. Варто зазначити, що база містить лише HMAC-геші паролів, що забезпечує додатковий рівень безпеки: навіть у разі компрометації бази даних злоумисник не зможе відновити оригінальний пароль;
- перевірку достовірності приватних повідомлень – кожне повідомлення супроводжується HMAC-кодом, який перевіряється на стороні отримувача;
- валідацію JSON-структур – всі повідомлення аналізуються на предмет коректності формату та наявності обов'язкових полів;
- обробку винятків – система реагує на типові помилки, включаючи спроби підробки MAC, відсутність критичних полів, некоректну структуру запитів, або спроби несанкціонованого доступу і генерує відповідні відповіді;

- синхронізацію онлайн-статусу користувачів – через службові повідомлення status: online/offline підтримується актуальний список підключених клієнтів;
- ідентифікацію користувачів у межах сесії – за допомогою cookie клієнт автоматично ідентифікується при кожному запиті, забезпечуючи стійку авторизацію протягом сеансу.

Таким чином, веб-сервіс QRNG забезпечує надійний та безпечний обмін повідомленнями за рахунок валідації користувачів, перевірки автентичності кожного повідомлення та контролю за цілісністю даних.

7. Блок-схеми логіки системи

В роботі для чіткого представлення алгоритмів взаємодії розроблені блок-схеми, що демонструють логіку роботи основних механізмів системи. Також надано опис процедури автентифікації та обміну приватними повідомленнями [16].

7.1 Автентифікація користувача.

Процес починається із введення користувачем логіна та пароля, після чого ці дані передаються на сервер. Сервер перевіряє, чи існує користувач у базі даних. У разі позитивної відповіді, генерується код HMAC-SHA-256 для введеного пароля, який порівнюється з тим, що зберігається у базі. Якщо значення збігаються, користувач отримує повідомлення про успішну автентифікацію та доступ до чату.

У разі помилки на будь-якому з етапів (некоректні вхідні дані, невірний формат JSON, відсутність користувача або помилкове HMAC-значення), клієнт отримує повідомлення про помилку з відповідним поясненням. Це дозволяє локалізувати причину відмови та забезпечити захищене середовище ще до входу користувача в систему.

Логіка автентифікації користувача наведена на рисунку 4.

Блок-схема була розроблена згідно з [17], для візуалізації використовувався онлайн конструктор блок-схем [18].

7.2 Обмін приватними повідомленнями.

Кожна сесія користувача базується на збереженні унікального ID, який передається у cookie. При вході користувача цей ID асоціюється з WebSocket-з'єднанням. Сервер підтримує список активних користувачів, що оновлюється при вході/виході через спеціальні службові повідомлення типу status: online або status: offline. Ці повідомлення надсилаються автоматично при зміні стану користувача.

Логіка автентифікації приватних повідомлень наведена на рисунку 5.

Передача повідомлень відбувається за принципом «клієнт – сервер – клієнт». Спочатку перевіряється, чи коректно введені дані (текст повідомлення, ID отримувача). Потім повідомлення разом із згенерованим HMAC надсилається на сервер. Сервер QRNG перевіряє онлайн-статус отримувача, і при позитивному результаті передає повідомлення одержувачу.

Отримувач виконує обчислення HMAC для перевірки автентичності. У разі збігу повідомлення виводиться у чаті. У протилежному випадку – ігнорується як потенційно змінене або підроблене. У разі недоступності одного з учасників передача повідомлення не здійснюється, і користувач інформується про невдалу доставку.

У поданих блок-схемах чітко розмежовані функції клієнта і сервера. Клієнт відповідає за формування вхідних даних та ініціацію запитів, сервер – за перевірку, обробку та генерацію відповідей. HMAC виступає як центральний елемент перевірки цілісності та достовірності, реалізуючи захист від підміни або зміни повідомлення.

Валідація структури JSON, перевірка наявності користувачів, синхронізація стану клієнтів (онлайн/офлайн) – усе це є невід'ємними частинами загальної архітектури безпечного обміну. Таким чином, система виявляє будь-які спроби підміни або фальсифікації повідомлень ще до їх відображення, що забезпечує довіру між користувачами.

8. Тестування веб-сервісу QRNG для захищеного обміну повідомленнями

Після реалізації функціональних компонентів системи було проведено тестування з метою перевірки стабільності роботи веб-сервісу QRNG, відповідності заявленим вимогам безпеки, правильності логіки автентифікації, обробки повідомлень та забезпечення захищеного з'єднання. Частково результати тестування наведені у цій статті на рисунках 6-13.

8.1 Тестування автентифікації користувачів.

Етап тестування автентифікації користувачів передбачав перевірку реакції системи на різні сценарії взаємодії з формою входу, коректність валідації введених даних та збереження сесії після входу до системи.

Попередньо було створено три тестові облікові записи (user#10, user#11, user#12).

1. Порожні поля.

У разі спроби автентифікації без введення логіну та/або пароля система генерує повідомлення про необхідність заповнення відповідних полів. При цьому запит до сервера не виконується, що свідчить про коректну реалізацію клієнтської валідації.

2. Неправильні облікові дані.

При введенні неправильного логіну або пароля система реагує повідомленням про помилку автентифікації. Сервер перевіряє існування користувача, обчислює HMAC від пароля, і при його невідповідності очікуваному значенню повертає повідомлення про відмову у доступі.

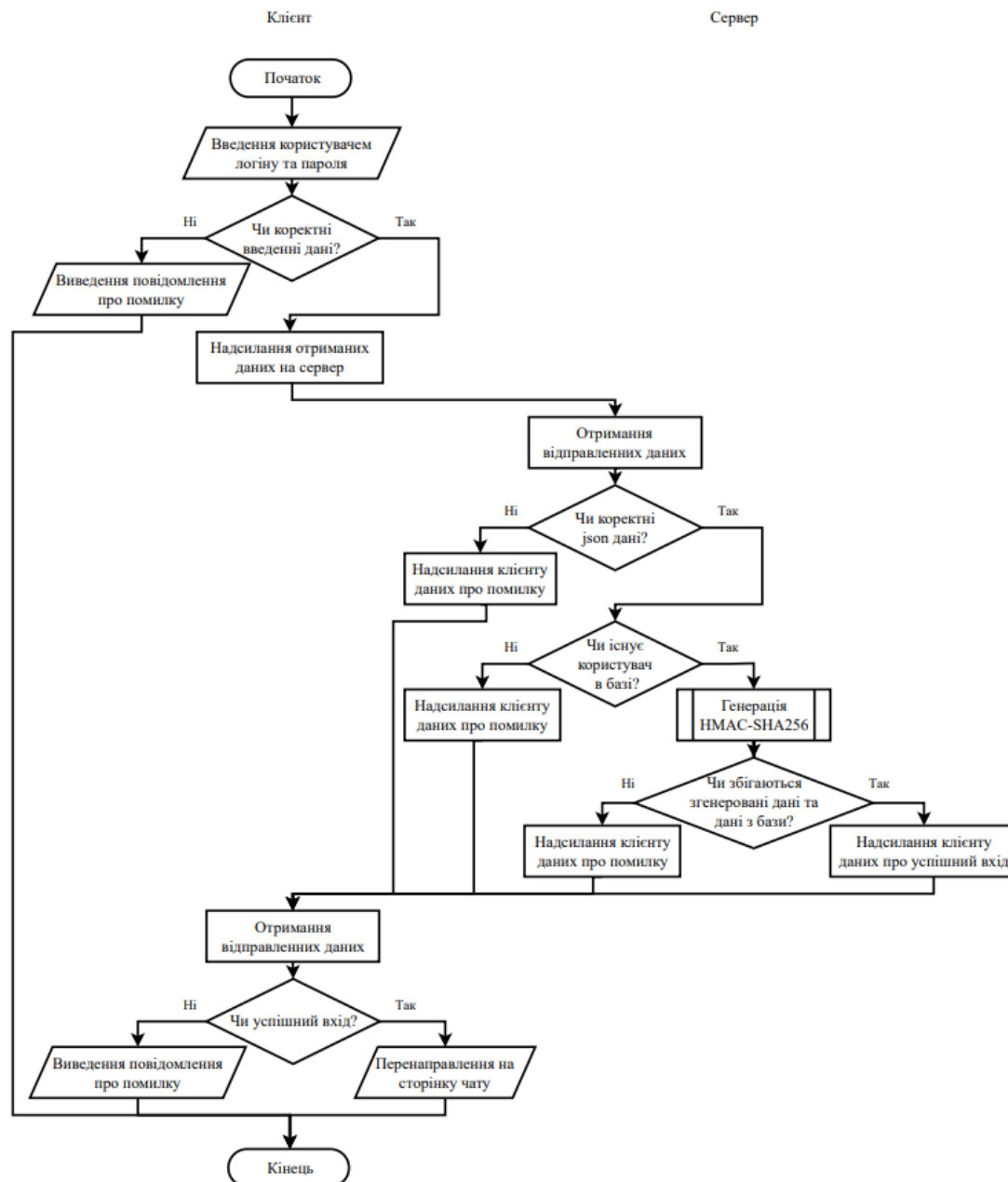


Рис. 4. Блок-схема автентифікації користувача

3. Успішна автентифікація.

При правильних облікових даних сервер успішно виконує перевірку HMAC, після чого користувача перенаправляє до чату. У браузері автоматично зберігається cookie з user_id, який ідентифікує користувача в поточній сесії.

8.2 Тестування захищеного з'єднання.

Підтвердження використання протоколів HTTPS та WSS при з'єднанні між клієнтом і сервером здійснювалося за допомогою кількох інструментів:

1. Перевірка через браузер.
2. Перевірка WebSocket-з'єднання.
3. Перевірка з використанням OpenSSL.
4. Перевірка захищеного з'єднання.

За допомогою утиліти Wireshark був проведений аналіз мережевого трафіку, щоб впевнитися у тому, що дані не передаються у відкритому вигляді.

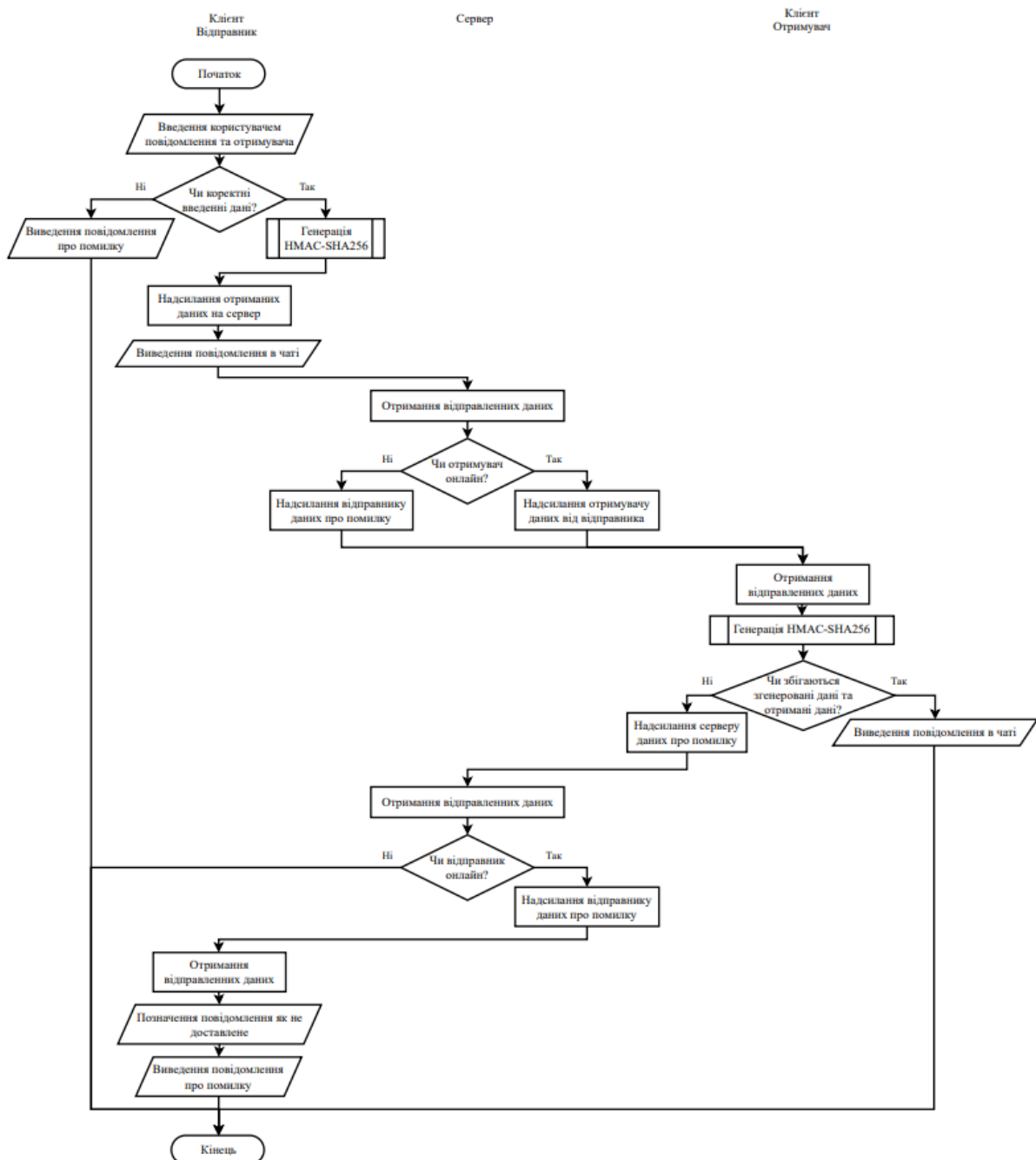


Рис. 5. Блок-схема обміну приватними повідомленнями

Після запуску сервісу та відкриття веб-сторінки з логіном/чатом у браузері, у Wireshark фіксується трафік на порту 9001. Всі пакети позначено як TLS, а вміст поля як Encrypted Application Data, що свідчить про те, що дані передаються у зашифрованому вигляді. Результат тестування наведено на рисунку 6.

8.3 Тестування онлайн-статусу та сесій.

Метою тестування є перевірка коректності підтримки статусу користувача в режимі реального часу, а також реакції системи на зміну стану.

1. Вхід нового користувача.

Після успішної автентифікації користувача, інші учасники чату автоматично бачать його в списку онлайн.

2. Вихід користувача.

У разі виходу з акаунту або закриття вкладки браузера, WebSocket-з'єднання завершується, і сервер надсилає всім підключеним клієнтам повідомлення status: offline. У результаті, користувач зникає зі списку активних.

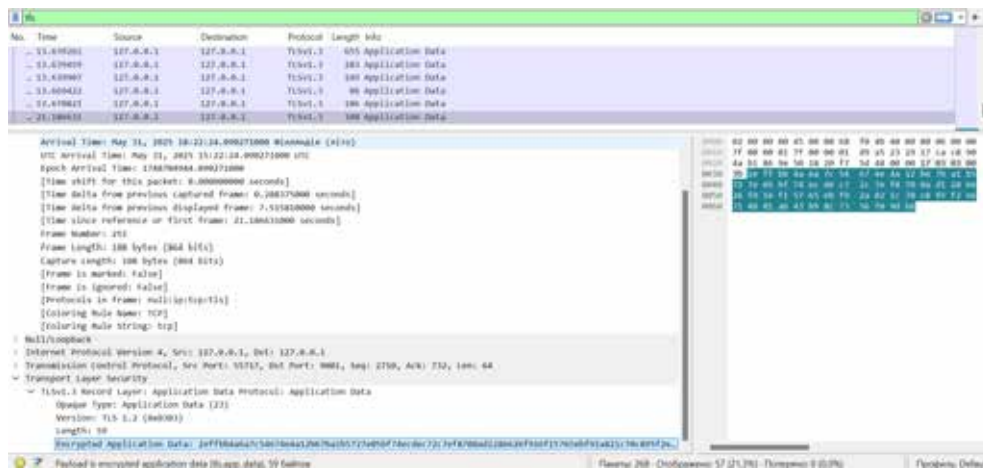


Рис. 6. Перехоплений трафік у Wireshark

8.4 Тестування обміну повідомленнями.

Метою цього етапу є перевірка коректності обробки публічних і приватних повідомлень, включаючи нетипові ситуації та перевірку захисту цілісності через НМАС.

1. Порожнє повідомлення.

Якщо користувач намагається надіслати порожнє повідомлення, виводиться повідомлення про помилку.

2. Успішне публічне повідомлення.

При надсиланні публічного повідомлення, воно з'являється у всіх підключених користувачів, вказується ім'я відправника.

Результати тестування наведено на рисунках 7-8.



Рис. 7. Публічне повідомлення (інтерфейс відправника)



Рис. 8. Публічне повідомлення (інтерфейс одного з отримувачів)

3. Приватне повідомлення.

Спроба надіслати повідомлення самому собі – система перевіряє, чи збігається ID відправника з ID отримувача. У разі збігу виводиться повідомлення про помилку, відправка не відбувається (рис. 9).



Рис. 9. Виведення помилки при спробі відправити повідомлення самому собі

Порожній ID отримувача – повідомлення не обробляється, виводиться клієнтська помилка (рис. 10).



Рис. 10. Виведення помилки при спробі відправити повідомлення при відсутності ID отримувача

4. Приватне повідомлення офлайн користувачу.

Сервер перевіряє, чи є отримувач серед активних підключень. Якщо користувач офлайн – повідомлення не доставляється, відправник інформується про недоступність доставки.

5. Імітація підробки HMAC у повідомленні.

Для імітації підробки HMAC у вхідному повідомленні була штучно модифікована логіка його обчислення. Через розбіжності між обчисленим та переданим HMAC-кодом, отримувач відкидає отримане приватне повідомлення, а відправник інформується про невідповідність. Результати тестування наведені на рисунках 11-13.

```
case "incoming_private_msg":
  const hmacKey = "key_" + payload.user_from;
  //const expectedHmac = hmacSha256(payload.text, hmacKey);
  const expectedHmac = "wrongHMAC";
  if (expectedHmac === payload.hmac) {
    createPrivateChat(payload.user_from, `User #${payload.user_from}`);
    displayMessage(payload.user_name, payload.text, "private", payload.user_from);
  }
}
```

Рис. 11. Імітація підробки HMAC в коді



Рис. 12. Повідомлення про неможливість доставки через помилковий HMAC-код (інтерфейс відправника)



Рис. 13. Повідомлення не відображено через помилковий HMAC-код (інтерфейс отримувача)

6. Успішне приватне повідомлення.

У разі, якщо обидва користувачі онлайн і всі параметри повідомлення коректні, система успішно передає повідомлення від одного користувача до іншого. Отримувач виконує перевірку HMAC, яка проходить успішно, після чого повідомлення з'являється у вікні чату.

В результаті проведеного тестування підтверджено коректність реалізації ключових функціональних складових веб-сервісу QRNG, зокрема механізмів автентифікації, шифрування трафіку, обробки повідомлень і підтримки статусу користувачів. Усі перевірені сценарії – як коректні, так і з внесеними навмисно помилками – дали очікувані результати, що свідчить про стабільність роботи системи та відповідність реалізованого сервісу вимогам інформаційної безпеки. Це включає дотримання базових принципів захисту інформації: конфіденційності, цілісності та автентичності даних. Отримані результати також підтверджують практичну придатність веб-сервісу QRNG та його здатність ефективно протидіяти типовим загрозам у мережевому середовищі.

Висновки

Таким чином, для забезпечення захищеного обміну повідомленнями у веб-сервісі QRNG доцільно використовувати комбінацію криптографічних алгоритмів, захищених протоколів передавання даних, кодів автентифікації та організаційних заходів, спрямованих на контроль доступу і своєчасне оновлення компонентів системи. Такий комплексний підхід дозволяє зменшити ризики перехоплення, модифікації або компрометації повідомлень, що циркулюють у системі.

Практичне значення роботи полягає в можливості використання отриманих результатів для підвищення рівня захисту веб-сервісу QRNG, зокрема через інтеграцію HMAC-SHA-256 у механізми перевірки автентичності повідомлень.

Під час дослідження також були виявлені недоліки окремих підходів до автентифікації, пов'язані з вразливістю до атак на паролі, складністю управління ключами та залежністю від апаратної платформи. Подальші дослідження доцільно спрямувати на вивчення методів динамічного управління ключами, оцінювання продуктивності HMAC у реальних веб-сервісах QRNG з великою кількістю клієнтів.

Список використаної літератури

1. Пірог О.В. Безпека вебдодатків : навч. посібн. Житомир : Житомирська політехніка, 2025. – 290 с.
2. Моргуль Д.М., Нарезний О.П., Грінченко Т.О. Класифікація атак та вимоги кібербезпеки до веб-ресурсу QRNG. ISSN 0485-8972. *Radiotekhnika* No. 220. С. 50-59. DOI:10.30837/rt.2025.1.220.04.
3. Васильченко Д.І., Лавровський І.М. Огляд типових уразливостей web-сайтів організацій у 2019-2020 році. *Сучасний захист інформації*. 2021. № 1(45). С. 41-46.
4. Кравчук Н.В., Коробейнікова Т.І. Огляд проблематики захищеного доступу до вебсерверів. *Вісник Львівського державного університету безпеки життєдіяльності*. 2024. № 30. С. 78-89.
5. Коробейнікова Т.І., Матвійчук А.А., Непийвода М.В. Огляд рішень для покращення процесів автентифікації та авторизації на вебресурсах. *SWorldJournal*. 2024. № 25-01, С. 82-88.
6. RFC 6455:2011. The WebSocket Protocol / I. Fette, A. Melnikov. – [Чинний від 2011-12]. Вид. офіц. Internet Engineering Task Force (IETF), 2011. 71 с. (Request for Comments; 6455).
7. ISO/IEC 27002:2022. Information security, cybersecurity and privacy protection – Information security controls. 3rd ed. [Чинний від 2022-02]. Вид. офіц. Geneva: ISO/IEC, 2022. 152 с.
8. НД ТЗІ 2.5-010-03. Вимоги до захисту інформації WEB-сторінки від несанкціонованого доступу. [Чинний від 2003-04-15]. Вид. офіц. Київ: ДСТСЗІ СБ України, 2003. 25 с.
9. FIPS PUB 198-1. The Keyed-Hash Message Authentication Code (HMAC). [Чинний від 2008-07]. Вид. офіц. Gaithersburg, MD: NIST, 2008. 7 с.
10. NIST SP 800-224 ipd. Keyed-Hash Message Authentication Code (HMAC): Specification of HMAC and Recommendations for Message Authentication / M. S. Turan, L. T. A. N. Brandão. [Чинний від 2024-06]. Вид. офіц. Gaithersburg, MD: NIST, 2024. 24 с. Initial Public Draft.
11. Моргуль Д.М., Нарезний О.П., Грінченко Т.О. Модель порушника та модель загроз для веб-сервісу QRNG. ISSN 0485-8972. *Radiotekhnika* No. 221. С. 31-38. DOI:10.30837/rt.2025.2.221.04.
12. Євсєєв С.П., Йохан О.Ю., Король О.Г. Гешування даних в інформаційних системах: монографія. Харків: ХНЕУ, 2013. 312 с.
13. Остапов С.Е., Євсєєв С.П., Король О.Г. Технології захисту інформації : навчальний посібник. Харків: ХНЕУ, 2013. 476 с.
14. Малярова Д. М. Порівняння швидкодії HMAC алгоритмів. Проблеми та перспективи забезпечення цивільного захисту : матеріали міжнародної науково-практичної конференції молодих учених. Харків, 2024. С. 256.
15. Малярова Д. М. Порівняльний аналіз алгоритмів гешування SHA-1 та сімейства SHA-2. *Глобалізація наукових знань : міжнародна співпраця та інтеграція галузей науки* : матеріали VI міжнародної студентської наукової конференції. Вінниця, 2024. С. 270-271.
16. Малярова Д.М., Грінченко Т.О., Нарезний О.П. Використання HMAC-SHA-256 як методу захисту у веб-сервісах. *Молоді вчені 2025 – від теорії до практики* : матеріали XV всеукраїнської науково-практичної конференції. Дніпро : Журфонд, 2025. С. 313-316.
17. ДСТУ ISO 5807:2016. Символи та угоди щодо документації стосовно даних, програм та системних блоксхем, схем мережевих програм та схем системних ресурсів (ISO5807:1985, IDT). [Чинний від 10.10.2016]. К.: ДП "УкрНДНЦ", 2016. 30 с. (Оброблення інформації).

18. Блок-схема онлайн – виробник блок-схеми. Free Online Apps for Most Popular File Formats. URL: <https://products.aspose.app/diagram/uk/flowchart> (дата звернення: 28.06.2025).

References

1. Piroh, O. V. (2025). Bezpeka vebdodativ: navchalnyi posibnyk [Web application security]. Zhytomyr: Zhytomyrska politekhnika. [in Ukrainian].
2. Morhul, D. M., Nariezhnii, O. P., & Hrinenko, T. O. (2025). Klasyfikatsiia atak ta vymohy kiberbezpeky do veb-resursu QRNG [Classification of attacks and cybersecurity requirements for the QRNG web resource]. *Radiotekhnika*, (220), 50–59. <https://doi.org/10.30837/rt.2025.1.220.04>. [in Ukrainian].
3. Vasylychenko, D. I., & Lavrovskiy, I. M. (2021). Ohliad typovykh urazlyvostei web-saitiv orhanizatsii u 2019–2020 rotsi [Overview of typical vulnerabilities of organization web-sites in 2019–2020]. *Suchasnyi zakhyst informatsii*, 1(45), 41–46. [in Ukrainian].
4. Kravchuk, N. V., & Korobeinykova, T. I. (2024). Ohliad problematyky zakhyshchenoho dostupu do vebserveriv [Overview of secure access issues to web servers]. *Visnyk Lvivskoho derzhavnoho universytetu bezpeky zhyttiedialnosti*, (30), 78–89. [in Ukrainian].
5. Korobeinykova, T. I., Matviichuk, A. A., & Nepyivoda, M. V. (2024). Ohliad rishen dla pokrashchennia protsesiv avtentyfikatsii ta avtoryzatsii na veb-resursakh [Overview of solutions for improving authentication and authorization processes on web resources]. *SWorldJournal*, (25-01), 82–88. [in Ukrainian].
6. Fette, I., & Melnikov, A. (2011). *RFC 6455: The WebSocket Protocol*. Internet Engineering Task Force (IETF). Retrieved from <https://www.rfc-editor.org/rfc/rfc6455>.
7. International Organization for Standardization & International Electrotechnical Commission. (2022). *ISO/IEC 27002:2022 – Information security, cybersecurity and privacy protection – Information security controls* (3rd ed.). Geneva: ISO/IEC.
8. Derzhspetsviazok of Ukraine. (2003). *ND TZI 2.5-010-03: Vymohy do zakhystu informatsii WEB-storinky vid nesanktsionovanoho dostupu* [Requirements for protecting web pages from unauthorized access]. Kyiv: DSTSZI SB Ukrainy. [in Ukrainian].
9. National Institute of Standards and Technology. (2008). *FIPS PUB 198-1: The Keyed-Hash Message Authentication Code (HMAC)*. Gaithersburg, MD: NIST.
10. National Institute of Standards and Technology. (2024). *NIST SP 800-224 (Initial Public Draft): Keyed-Hash Message Authentication Code (HMAC): Specification and Recommendations for Message Authentication*. Gaithersburg, MD: NIST.
11. Morhul, D. M., Nariezhnii, O. P., & Hrinenko, T. O. (2025). Model porushnyka ta model zahroz dla veb-servisu QRNG [Attacker model and threat model for QRNG web service]. *Radiotekhnika*, (221), 31–38. <https://doi.org/10.30837/rt.2025.2.221.04>. [in Ukrainian].
12. Yevseiev, S. P., Yokhov, O. Yu., & Korol, O. H. (2013). *Heshuvannia danykh v informatsiinykh systemakh: monohrafiia* [Data hashing in information systems: Monograph]. Kharkiv: KhNEU. [in Ukrainian].
13. Ostapov, S. E., Yevseiev, S. P., & Korol, O. H. (2013). *Tekhnolohii zakhystu informatsii: navchalnyi posibnyk* [Information protection technologies: Textbook]. Kharkiv: KhNEU. [in Ukrainian].
14. Maliarova, D. M. (2024). Porivniannia shvydkodii HMAC alhorytmiv [Comparison of HMAC algorithm performance]. In *Problemy ta perspektyvy zabezpechennia tsyvilnoho zakhystu* (p. 256). Kharkiv. [in Ukrainian].
15. Maliarova, D. M. (2024). Porivnialnyi analiz alhorytmiv heshuvannia SHA-1 ta simeistva SHA-2 [Comparative analysis of SHA-1 and SHA-2 hashing algorithms]. In *Hlobalizatsiia naukovykh znan: Mizhnarodna spivpratsia ta intehratsiia haluzei nauky* (pp. 270–271). Vinnytsia. [in Ukrainian].
16. Maliarova, D. M., Hrinenko, T. O., & Nariezhnii, O. P. (2025). Vykorystannia HMAC-SHA-256 yak metodu zakhystu u veb-servisakh [Use of HMAC-SHA-256 as a protection method in web services]. In *Molodi vcheni 2025 – vid teorii do praktyky* (pp. 313–316). Dnipro: Zhurfond. [in Ukrainian].
17. Ukrainian Scientific Research and Training Center for Standardization. (2016). *DSTU ISO 5807:2016 – Symbols and conventions for documentation relating to data, program and system flowcharts (ISO 5807:1985, IDT)*. Kyiv: DP “UkrNDNTs”.
18. Aspose. (2025). *Flowchart online generator*. Free Online Apps for Most Popular File Formats. Retrieved from <https://products.aspose.app/diagram/uk/flowchart> (Accessed: June 28, 2025).

Дата першого надходження рукопису до видання: 22.11.2025
Дата прийнятого до друку рукопису після рецензування: 18.12.2025
Дата публікації: 31.12.2025

УДК 004.056

DOI <https://doi.org/10.35546/kntu2078-4481.2025.4.3.7>**К. І. ДОРОФЄЄВА**

магістрант кафедри безпеки інформаційних технологій
Харківський національний університет радіоелектроніки
ORCID: 0009-0003-8116-9050

О. В. СЕВЕРІНОВ

кандидат технічних наук, доцент,
професор кафедри безпеки інформаційних технологій
Харківський національний університет радіоелектроніки
ORCID: 0000-0002-6327-6405

З. М. СИДОРЕНКО

асистент кафедри безпеки інформаційних технологій
Харківський національний університет радіоелектроніки
ORCID: 0000-0002-0104-6807

В. М. СУХОТЕПЛИЙ

старший викладач кафедри радіоелектронних систем
пунктів управління Повітряних Сил
Харківський національний університет Повітряних Сил
імені Івана Кожедуба
ORCID: 0000-0002-2566-4167

ЗАСТОСУВАННЯ ІНСТРУМЕНТА АНАЛІЗУ БЕЗПЕКИ ДЛЯ ВИЯВЛЕННЯ КРИТИЧНИХ ВРАЗЛИВОСТЕЙ У ВЕБ-ДОДАТКАХ

У статті розглянуто особливості застосування інструмента OWASP Zed Attack Proxy у процесі тестування на проникнення веб-додатків, що обумовлює необхідність аналізу його функціональних можливостей і місця в сучасних підходах до аудиту інформаційної безпеки. Досліджено принципи роботи проксі-механізму ZAP, що забезпечує перехоплення та аналіз HTTP(S)-трафіку, а також можливість побудови карти веб-додатка на основі спостережених запитів і відповідей. Проаналізовано методи пасивного сканування, яке дозволяє здійснювати оцінку безпеки без втручання в поведінку сервера, зокрема виявляти відсутні або некоректно налаштовані HTTP-заголовки, особливості роботи cookies та інші технічні недоліки конфігурації. Розглянуто застосування активного сканування, що передбачає виконання контрольованих тестів на наявність поширених технічних вразливостей, а також акцентовано на необхідності проводити такі дії виключно у тестових середовищах. Описано роботу інструментів ручного аналізу, включно з редактором запитів, переглядом історії трафіку та базовим фуззингом, що дозволяє досліднику перевіряти поведінку додатка за різних умов. Окремо розглянуто можливість розширення функціональності інструмента шляхом використання додаткових модулів та застосування API ZAP для автоматизації рутинних перевірок у межах підготовлених сценаріїв. У роботі також визначено обмеження інструмента, зокрема труднощі у виявленні складних логічних вразливостей, що потребують експертного ручного дослідження. Отримані результати дозволяють визначити місце OWASP ZAP як інструмента для технічного аудиту, початкових перевірок безпеки і навчальних цілей, а також підкреслити необхідність його комбінування з ручним аналізом при дослідженні складних систем.

Ключові слова: OWASP ZAP, тестування на проникнення, веб-додатки, веб-безпека, проксі-аналіз, пасивне сканування, активне сканування.

K. I. DOROFIEIEVA

Master's Student at the Department of Information Technology Security
Kharkiv National University of Radio Electronics
ORCID: 0009-0003-8116-9050

O. V. SIEVIERINOV

Candidate of Technical Sciences, Associate Professor,
Professor at the Department of Information Technology Security
Kharkiv National University of Radio Electronics
ORCID: 0000-0002-6327-6405

Z. M. SYDORENKO

Assistant at the Department of Information Technology Security
Kharkiv National University of Radio Electronics
ORCID: 0000-0002-0104-6807

V. M. SUKHOTEPLYI

Senior Instructor at the Department of Radioelectronic Systems
of Control Points of Air Forces
Ivan Kozhedub Kharkiv National Air Force University
ORCID: 0000-0002-2566-4167

USING A SECURITY ANALYSIS TOOL TO IDENTIFY CRITICAL VULNERABILITIES IN WEB APPLICATIONS

The article examines the features and practical applications of the OWASP Zed Attack Proxy tool in web application penetration testing, emphasizing the need to analyze its capabilities and role in modern approaches to information security assessment. The study investigates the principles of ZAP's proxy mechanism, which allows the interception and analysis of HTTP(S) traffic and the construction of a site map based on observed requests and responses. Methods of passive scanning are analyzed, which enable security assessment without affecting server behavior, including the identification of missing or improperly configured HTTP headers, issues related to cookies, and other technical configuration shortcomings. The use of active scanning is also examined, involving controlled tests for common technical vulnerabilities, with an emphasis on performing such operations only in dedicated testing environments. Manual analysis tools are described, including the request editor, traffic history viewer, and basic fuzzing functionality, which allow researchers to evaluate application behavior under various input conditions. Additionally, the paper considers the possibilities of extending ZAP's functionality through optional add-ons and using its API to automate routine tasks within predefined testing scenarios. The limitations of the tool are outlined, particularly the challenges of detecting complex logical vulnerabilities that require expert manual investigation. The findings indicate that OWASP ZAP is suitable for technical audits, initial security evaluations, and educational purposes, while highlighting the necessity of combining automated scanning with manual analysis when assessing complex web applications.

Key words: OWASP ZAP, penetration testing, web applications, web security, proxy analysis, passive scanning, active scanning.

Постановка проблеми

Забезпечення безпеки веб-додатків є одним із ключових завдань сучасних інформаційних систем, оскільки більшість сервісів, бізнес-процесів та доступу до даних реалізується саме через веб-інтерфейси. Розповсюдженість веб-технологій сприяє збільшенню кількості загроз, спрямованих на експлуатацію вразливостей у логіці застосунків, у компонентах серверної інфраструктури та у механізмах обробки даних користувача. Значна частина атак пов'язана не з високотехнологічними способами злому, а з типовими помилками програмування та недоліками конфігурації.

Тестування на проникнення веб-додатків дозволяє виявляти такі недоліки ще до того, як вони будуть використані зловмисником. При цьому важливо застосовувати інструменти, які дозволяють аналізувати HTTP-трафік, перевіряти реакцію веб-додатка на різні типи некоректних або небезпечних запитів та формувати об'єктивні висновки щодо рівня безпеки.

OWASP ZAP є інструментом з відкритим програмним кодом, який забезпечує базові та розширені можливості для тестування на проникнення. На відміну від комерційних продуктів, ZAP доступний безкоштовно, має активну спільноту та орієнтований як на початківців, так і на фахівців, які здійснюють ручний чи напівавтоматизований аудит веб-безпеки.

Проблема, яку розглядає дана стаття, полягає в необхідності системного використання OWASP ZAP під час виконання тестування на проникнення та у формуванні методологічного підходу до його застосування відповідно до реальних можливостей інструмента.

Аналіз останніх досліджень і публікацій

У наукових і прикладних публікаціях останніх років значну увагу приділено розвитку методологій тестування на проникнення та використанню спеціалізованих засобів аналізу безпеки веб-додатків. У роботі [1] автори наголошують, що пентест є ключовим інструментом кібербезпечного менеджменту, оскільки дозволяє виявляти вразливості, оцінювати рівень ризику та формувати рекомендації щодо підвищення стійкості інформаційних систем. Автори підкреслюють важливість інструментів динамічного аналізу, які забезпечують відтворення реальних атаквальних сценаріїв у процесі тестування. Скибун, аналізуючи мету та завдання пентесту, виокремлює значення автоматизованих сканерів і проксі-інструментів, що дають можливість як пасивного, так і активного вивчення поведінки веб-додатка під час взаємодії з ним [2]. Прямо у контексті OWASP ZAP детальний аналіз наданий

у дослідженні Maniraj та співавторів, де інструмент розглядається як комплексне рішення для сканування, перехоплення HTTP-трафіку, моделювання атаки та автоматизованого моніторингу вразливостей. Автори відзначають ефективність ZAP при виявленні XSS, SQL Injection та конфігураційних помилок на реальних веб-платформах [3]. У роботі [4] пояснюється сутність вразливостей з OWASP Top 10, їхні корені та типові шляхи експлуатації, підкреслюючи необхідність інструментів, здатних аналізувати HTTP-взаємодію та структуру веб-додатків у реальному часі. Завершальним елементом є офіційна документація ZAP, що містить опис механізмів проксі-перехоплення, роботи пасивного та активного сканерів, алгоритмів Spider і AJAX-Spider, а також API для автоматизації – ці матеріали формують практичну базу для використання інструмента в пентестингу [5]. Сукупність публікацій відображає зростаючу увагу до інтеграції OWASP ZAP у процеси тестування на проникнення та підтверджує його ефективність як інструмента динамічного аналізу безпеки веб-додатків.

Формулювання мети дослідження

Метою дослідження є оцінка реальних можливостей інструмента OWASP ZAP та формування методологічного підходу до його використання у процесі тестування на проникнення веб-додатків. Дослідження спрямоване на вивчення функціоналу ZAP, ефективності його пасивного та активного сканування, можливостей перехоплення та аналізу HTTP/HTTPS-трафіку, а також практичної придатності інструмента для виявлення типових технічних вразливостей і підтримки комплексного аудиту веб-додатків.

Викладення основного матеріалу дослідження

OWASP ZAP є інструментом відкритого коду для тестування веб-додатків, призначеним для виявлення технічних вразливостей. Основною особливістю ZAP є робота як проксі-сервера, що дозволяє перехоплювати HTTP(S)-трафік між клієнтом та сервером, аналізувати його та зберігати історію запитів і відповідей. Це дає можливість досліднику відстежувати всі параметри, передані веб-додатком, а також виявляти помилки конфігурації, небезпечні заголовки або некоректну обробку даних користувача [5].

Інструмент підтримує пасивне і активне сканування. Пасивне сканування виконується автоматично при отриманні трафіку, не впливаючи на роботу сервера, і дозволяє виявляти відсутність безпечних заголовків, небезпечні cookies, незахищені параметри та інші технічні недоліки. Активне сканування включає надсилання контрольованих тестових запитів для перевірки реакції сервера на потенційно небезпечні вхідні дані, наприклад, на SQL Injection чи XSS [3]. ZAP також надає інструменти ручного аналізу, такі як редактор запитів, історія трафіку та базовий фуззинг. Ці інструменти дозволяють досліднику перевіряти додаток за різних умов, змінювати параметри запитів і оцінювати реакцію сервера. Також доступні додаткові модулі (add-ons), що розширюють функціональність ZAP для інтеграції з іншими інструментами або додаткового аналізу специфічних протоколів.

Використання проксі-механізму

Однією з ключових функціональних можливостей OWASP ZAP є використання проксі-механізму, який дає змогу перехоплювати, аналізувати та модифікувати HTTP/HTTPS-трафік між браузером користувача та веб-додатком. Під час роботи ZAP виступає як проміжна ланка, що приймає всі вихідні запити клієнта та відповіді сервера, забезпечуючи повний контроль над переданими параметрами та протоколами взаємодії. Завдяки цьому механізму інструмент дозволяє не лише спостерігати за поведінкою веб-додатка в реальному часі, але й вручну коригувати запити для оцінки реакції сервера на змінені параметри. Такий підхід є особливо корисним під час тестування вразливостей, що потребують ручної взаємодії, включно з маніпуляціями сеансовими даними, зміною параметрів форм, тестуванням автентифікації та контролем доступу [5].

На рисунку 1 зображено принцип роботи проксі-механізму ZAP, де інструмент виступає як посередник між клієнтом і сервером, забезпечуючи прозорий огляд і можливість втручання в трафік під час тестування. Така архітектура є базовою для всіх подальших етапів аналізу та дозволяє тестувальнику повністю контролювати комунікацію веб-додатка.

Проксі-механізм ZAP тісно інтегрований з такими інструментами, як History, Request/Response Editor, Breakpoints та Fuzzer, що надає можливість не тільки фіксувати кожен елемент трафіку, але й зупиняти виконання запиту на певному етапі для внесення змін. Використання breakpoints дозволяє детально дослідити логіку обробки запитів, виявляти приховані параметри та оцінювати поведінку серверної частини при аномальних або навмисно некоректних даних.

Також ZAP підтримує роботу з HTTPS-трафіком шляхом генерації власного локального сертифіката, що дає змогу безпечно розшифровувати зашифровані дані та забезпечувати повноцінний аналіз навіть у випадку використання сучасних транспортних протоколів. Завдяки цьому проксі-механізм стає основою для більшості методів тестування вразливостей, забезпечуючи прозорий огляд усіх взаємодій між клієнтом та сервером.

Пасивне сканування

Пасивне сканування в OWASP ZAP є базовим етапом аналізу безпеки, який виконується без створення додаткового навантаження на веб-додаток. На відміну від активних методів, пасивне сканування не змінює структуру запитів та не модифікує дані, що дає змогу проводити його на робочих системах без ризику порушення

функціонування або виникнення побічних ефектів. OWASP ZAP здійснює пасивний аналіз у фоновому режимі під час роботи проксі-механізму: кожен перехоплений HTTP-запит та відповідь сервера автоматично перевіряються на наявність ознак потенційних вразливостей. До типових категорій, що можуть бути виявлені таким способом, належать: некоректні заголовки безпеки, надлишкове розкриття інформації, неправильна конфігурація хешування, відсутність політик захисту від XSS, використання слабких алгоритмів хешування тощо. Згідно з практиками, описаними OWASP, пасивне сканування виступає одним з найбільш безпечних способів первинної оцінки безпеки, оскільки воно не впливає на логіку роботи веб-додатка й дозволяє поступово формувати перелік можливих проблем, які потребують глибшого аналізу під час активного сканування або ручної перевірки [6]. Такий підхід підтримується також українськими дослідженнями, де наголошується, що пасивні техніки є оптимальними на початкових етапах тестування та дозволяють оцінити загальний рівень безпеки ІТ-системи без втручання у її функції. Цей етап створює основу для подальших досліджень, забезпечуючи послідовний і безпечний перехід до більш інтенсивних методів оцінювання безпеки [1]. На рисунку 2 показано, що ZAP виявив відсутність заголовка Content-Security-Policy (CSP) у відповіді сервера. Виявлення подібних недоліків на етапі пасивного сканування дозволяє зробити висновки щодо базового стану безпеки додатка ще до застосування активних методів аналізу.

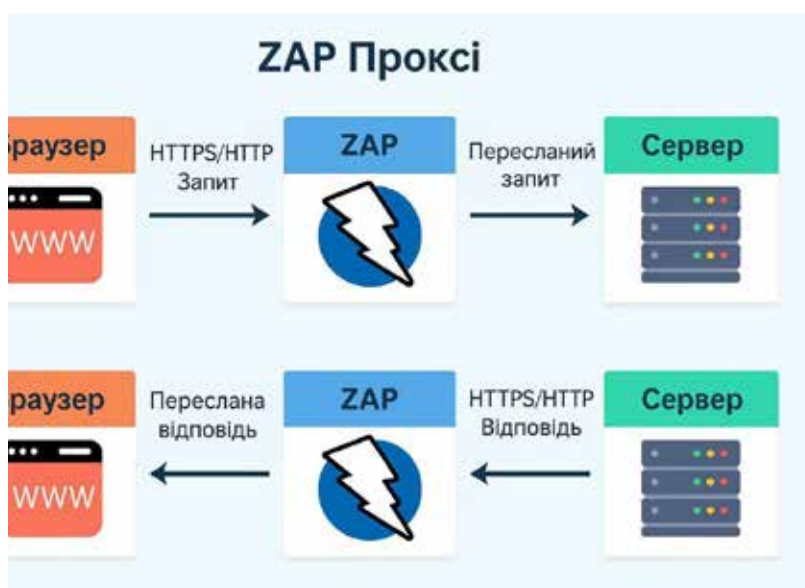


Рис. 1. Архітектура роботи проксі-механізму ZAP

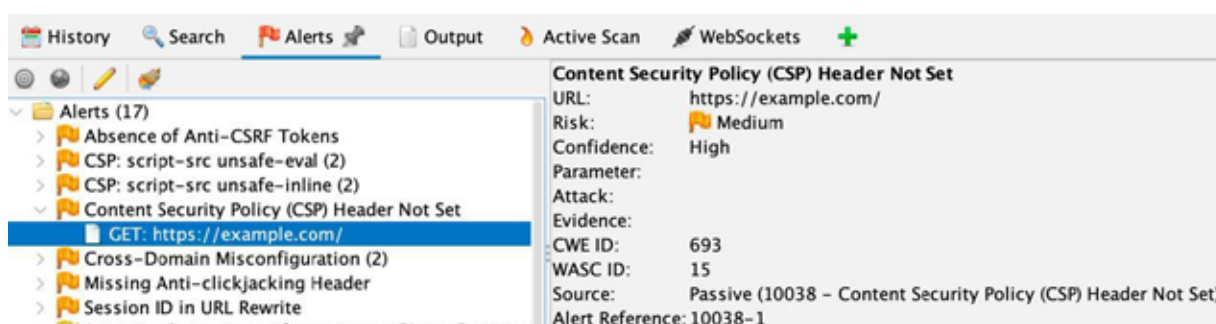


Рис. 2. Виявлення відсутності CSP у відповідях сервера під час пасивного сканування

Для систематизації результатів аналізу в таблиці 1 наведено узагальнену класифікацію типових проблем безпеки, які можуть бути виявлені під час пасивного сканування OWASP ZAP [5].

Пасивне сканування у ZAP виконується автоматично. Інструмент аналізує всі запити та відповіді, що проходять через проксі, але не надсилає додаткових запитів до веб-сервера.

Це робить пасивне сканування безпечним щодо впливу на тестований додаток. Пасивне сканування є першим етапом перевірки, оскільки воно не змінює поведінку додатка та не створює додаткового навантаження.

Таблиця 1

Категорії вразливостей, які виявляє пасивне сканування

Категорія	Приклади
Небезпечні заголовки	Відсутність X-Frame-Options, Content-Security-Policy
Проблеми cookie	Відсутність HttpOnly, Secure, SameSite
Інформаційні витоки	Версія сервера в заголовках, stack traces
Небезпечні параметри	Передача токенів у URL, параметри IDOR
Конфігураційні недоліки	Mixed content, некоректні cache-policy

Активне сканування

Активне сканування в OWASP ZAP є наступним етапом після пасивного аналізу та передбачає безпосереднє формування інструментом спеціально сконструйованих HTTP-запитів з метою виявлення вразливостей, які не можуть бути ідентифіковані лише шляхом спостереження за трафіком. На відміну від пасивного аналізу, активне сканування суттєво впливає на поведінку веб-додатка, оскільки використовує методи, що імітують потенційні атаки, включно зі спробами ін'єкцій, маніпуляцією параметрами, обходом автентифікації та тестуванням механізмів обробки помилок. В OWASP ZAP активне сканування реалізується через механізм Active Scan Rules, який використовує спеціалізовані сканери, що аналізують кожен параметр, заголовок, cookie та тіло запиту на предмет можливих точок впливу. Користувач може визначати масштаб аналізу, рівень навантаження, типи тестів та обмеження, що дозволяє адаптувати процес до середовища тестування, знижуючи ризики небажаних змін у роботі цільового додатка [2].

На рисунку 3 продемонстровано конфігурацію, яка визначає, які типи перевірок будуть виконані, рівні інтенсивності тестування та граничні значення для запитів. Такі параметри дозволяють адаптувати роботу сканера до вимог дослідження.

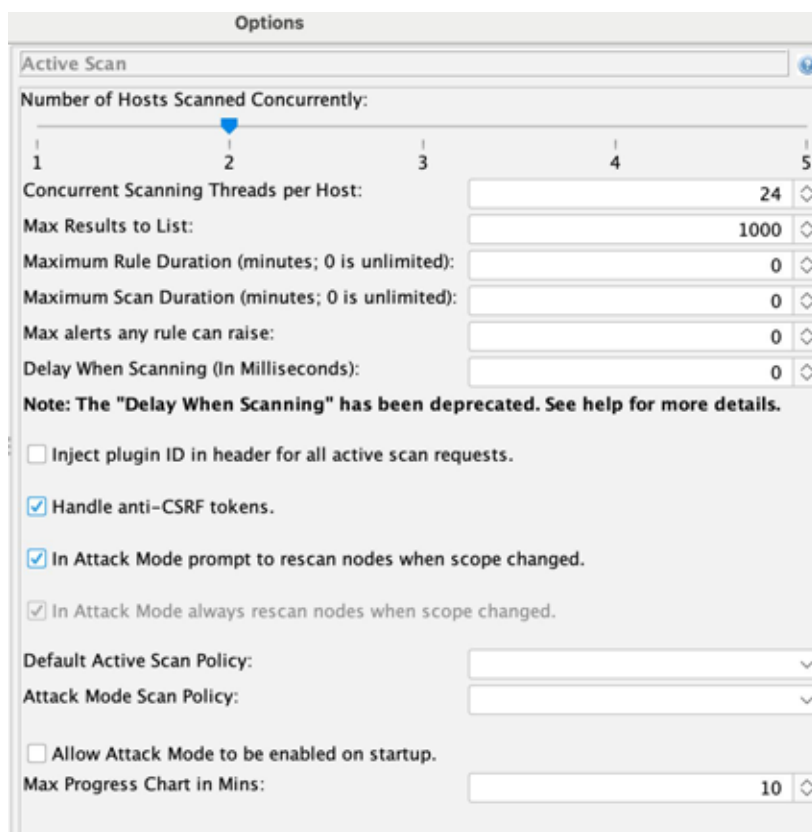


Рис. 3. Налаштування параметрів активного сканування в OWASP ZAP

Суттєвим аспектом активного сканування є ризикованість його застосування, тому слід з обережністю підходити до вибору налаштувань сканера. Оскільки інструмент навмисно генерує запити, які можуть викликати виключення, створити нестабільність, вплинути на стан даних або спровокувати блокування облікових записів, його не рекомендується виконувати на продуктивних системах. Дослідження підкреслюють, що активне сканування повинно здійснюватися лише у контрольованому середовищі, а перед початком тестування мають бути

передбачені заходи щодо збереження даних та мінімізації шкоди [3]. У практиці тестування на проникнення активне сканування використовується не як самостійний інструмент, а як складова комплексної методики, де воно доповнює ручний аналіз та контекстно орієнтовані перевірки. Результати, отримані за допомогою OWASP ZAP, слугують відправною точкою для глибшої експлуатації виявлених вразливостей і формування детального звіту щодо реальних ризиків для веб-додатка.

В таблиці 2 наведена класифікація типових категорій вразливостей, які визначаються за допомогою активного сканування OWASP ZAP.

Таблиця 2

Категорії вразливостей, які виявляє активне сканування

Категорія	Приклади
Ін'єкційні атаки	SQL Injection, Command Injection, LDAP Injection
Cross-Site Scripting (XSS)	Reflected XSS, DOM-based XSS
Path Traversal	Доступ до заборонених директорій
Атаки на автентифікацію	Brute force, weak password checks
Проблеми із сесіями	Підбір session ID, повторне використання токенів
Логічні вразливості	Обхід авторизації, некоректна обробка бізнес-логіки

ZAP використовує набір готових тестових правил, що систематично перевіряють поведінку веб-додатка та аналізують відповіді сервера.

З огляду на різний характер пасивного та активного сканувань, доцільно здійснити їх систематичне порівняння за ключовими характеристиками, що наведено у таблиці 3. Такий аналіз дозволяє обґрунтовано обирати тактику тестування відповідно до цілей аудиту, рівня ризику та можливостей взаємодії з тестованим середовищем.

Таблиця 3

Порівняння можливостей видів сканувань в OWASP ZAP

Характеристика	Пасивне сканування	Активне сканування
Вплив на додаток	Не впливає на роботу системи	Може змінювати стан даних або викликати помилки
Додаткові запити	Не створює	Створює спеціальні тестові запити
Типи вразливостей	Заголовки безпеки, інформаційні витоки, некоректні конфігурації	Ін'єкції, XSS, обхід автентифікації, зміна параметрів
Ризик застосування	Мінімальний	Підвищений
Потребує контексту	Ні	Так, залежить від бізнес-логіки
Рекомендоване використання	Початковий етап тестування	Поглиблений аналіз та експлуатація

Інструменти ручного аналізу та автоматизації

ZAP також надає можливості ручного аналізу через Request Editor, History і Fuzzer, що дозволяє змінювати запити та перевіряти реакцію серверу. Для автоматизації рутинних завдань доступний API ZAP, який дозволяє запускати сканування, отримувати результати та формувати звіти безпосередньо із зовнішніх сценаріїв [5]. Також через Marketplace можна встановлювати додаткові модулі для специфічного аналізу. OWASP ZAP ефективний для виявлення типових технічних вразливостей, проте має обмеження у випадках складних логічних проблем або специфічних бізнес-логік, що потребують ручного аналізу. Додатково деякі складні сканування можуть вимагати значних ресурсів і часу, особливо для великих веб-додатків.

Висновки

У статті було розглянуто практичні аспекти використання інструмента OWASP ZAP у процесі тестування на проникнення веб-додатків, а також проаналізовано його можливості у контексті сучасних вимог до забезпечення кібербезпеки. Встановлено, що ZAP належить до категорії доступних і водночас функціонально повноцінних засобів, які дозволяють проводити комплексний аналіз безпеки веб-додатків на різних етапах життєвого циклу. Він підтримує як ручне тестування, так і автоматизоване сканування, що робить його універсальним інструментом як для початківців, так і для досвідчених фахівців.

У процесі дослідження було визначено, що ZAP забезпечує виявлення широкого спектра вразливостей, включно з помилками налаштування, слабкими механізмами автентифікації, вразливостями, пов'язаними з ін'єкціями, та недоліками у реалізації контролю доступу. Значною перевагою інструмента є його інтеграція з методологією OWASP, що дозволяє застосовувати його у рамках рекомендованих підходів до оцінювання безпеки. Окрему увагу приділено використанню розширень, що надають змогу адаптувати ZAP під специфічні потреби тестування та підвищувати рівень деталізації аналізу. Як показує практика, у поєднанні з ручними техніками інструмент перетворюється на ефективний компонент процесу тестування на проникнення.

Таким чином, застосування OWASP ZAP сприяє підвищенню якості безпекового аналізу веб-додатків, а його можливості дозволяють забезпечити глибоке вивчення потенційних загроз без необхідності використання комерційних рішень. Інструмент демонструє ефективність у виявленні поширених вразливостей та може бути інтегрований у регулярні процеси безпекового тестування в організаціях різного масштабу.

Список використаної літератури

1. Горбаченко С., Бойко В. Тестування на проникнення як ефективний інструмент менеджменту кібербезпеки. *Information Technology and Society*. 2023. № 3 (9). С. 23–29. URL: <https://doi.org/10.32689/maup.it.2023.3.3> (дата звернення: 14.11.2025).
2. Скибун О. Тестування на проникнення: мета та цілі. *Grail of Science*. 2022. № 22. С. 161–163. URL: <https://doi.org/10.36074/grail-of-science.25.11.2022.28> (дата звернення: 17.11.2025).
3. Maniraj S. P., Ranganathan C. S., Sekar S. Securing Web Applications with OWASP ZAP for Comprehensive Security Testing. *International Journal of Advances in Signal and Image Sciences*. 2024. Vol. 10, no. 2. P. 12–23. URL: <https://doi.org/10.29284/ijasis.10.2.2024.12-23> (date of access: 17.11.2025).
4. Reddy s. OWASP Top 10 for Layman: OWASP Top 10. Independently Published, 2019.
5. ZAP – Documentation. ZAP. URL: <https://www.zaproxy.org/docs/> (date of access: 17.11.2025).
6. OWASP Automated Threat Handbook: Web Applications / T. Zaw et al. Lulu Press, Inc., 2018.

References

1. Horbachenko S., Boiko V. (2023) Penetration Testing as an Effective Tool of Cybersecurity Management. *Information Technology and Society*, no. 3 (9), pp. 23–29. URL: <https://doi.org/10.32689/maup.it.2023.3.3>. [in Ukrainian].
2. Skybun O. (2022) Penetration Testing: Purpose and Objectives. *Grail of Science*, no. 22, pp. 161–163. URL: <https://doi.org/10.36074/grail-of-science.25.11.2022.28>. [in Ukrainian].
3. Maniraj S. P., Ranganathan C. S., Sekar S. (2024) Securing Web Applications with OWASP ZAP for Comprehensive Security Testing. *International Journal of Advances in Signal and Image Sciences*, vol. 10, no. 2, pp. 12–23. URL: <https://doi.org/10.29284/ijasis.10.2.2024.12-23> (accessed: 17.11.2025).
4. Reddy S. (2019) OWASP Top 10 for Layman: OWASP Top 10. Independently Published.
5. ZAP – Documentation. ZAP. URL: <https://www.zaproxy.org/docs/> (accessed: 17.11.2025).
6. Zaw T., et al. (2018) OWASP Automated Threat Handbook: Web Applications. Lulu Press, Inc.

Дата першого надходження рукопису до видання: 11.11.2025
Дата прийнятого до друку рукопису після рецензування: 09.12.2025
Дата публікації: 31.12.2025

Л. О. ДРЕВИЧ

аспірант кафедри інформаційних систем та мереж
Національний університет «Львівська політехніка»
ORCID: 0009-0004-0426-8984

СЕГМЕНТАЦІЯ СМУГ ДЛЯ ОПТИМІЗАЦІЇ ТРАФІК-ПОТОКУ

Зростання інтенсивності дорожнього руху у містах призводить до збільшення заторів, підвищення рівня забруднення повітря та зниження безпеки учасників руху, тому ефективне управління транспортними потоками стає критично важливим завданням для сучасних мегаполісів. Технології комп'ютерного зору, зокрема сегментація смуг руху, відкривають нові можливості для інтелектуального управління дорожнім рухом та оптимізації використання наявної транспортної інфраструктури. У роботі здійснено комплексний аналіз сучасних нейромережових архітектур сегментації смуг руху з метою визначення їхнього потенціалу в оптимізації транспортних потоків та оцінки можливостей застосування в системах управління дорожнім рухом. Проведено порівняльний аналіз п'яти передових архітектур глибокого навчання: LaneATT з механізмом уваги для виявлення ключових точок; ERFNet та ENet як ефективних архітектур реального часу; PINet для instance-based сегментації смуг; а також CondLaneNet з умовною генерацією форми смуг. Досліджено особливості кожної архітектури, їхню обчислювальну ефективність і точність, а також проаналізовано вплив якості сегментації на показники оптимізації транспортних потоків. Показано, що CondLaneNet забезпечує найвищу точність сегментації та найкраще відновлення геометрії смуг за низького рівня трафіку й простої топології доріг, тоді як LaneATT демонструє децю нижчу пікову точність, проте характеризується більш плавною деградацією якості зі зростанням щільності руху. ENet та ERFNet забезпечують прийнятний компроміс між точністю та обчислювальною складністю, що робить їх придатними для систем реального часу з обмеженими ресурсами. PINet, завдяки представленням на основі ключових точок і кластеризації, виявляє найкращу стійкість до оклюзій і складних дорожніх сценаріїв та зберігає найбільшу частку базової якості в умовах середнього та високого трафіку. Отримані результати дають змогу обґрунтовано обирати оптимальну архітектуру для конкретних сценаріїв застосування в інтелектуальних транспортних системах.

Ключові слова: сегментація смуг руху; комп'ютерний зір; глибоке навчання; інтелектуальні транспортні системи; оптимізація транспортних потоків; управління дорожнім рухом; LaneATT; ERFNet; ENet; PINet; CondLaneNet.

L. O. DREVYCH

Postgraduate Student at the Department of Information Systems and Networks
Lviv Polytechnic National University
ORCID: 0009-0004-0426-8984

LANE SEGMENTATION TO OPTIMIZE TRAFFIC FLOW

The increasing intensity of urban traffic leads to growing congestion, higher air pollution levels, and reduced safety for road users, making effective traffic flow management a critically important task for modern metropolises. Computer vision technologies, particularly lane segmentation, open new possibilities for intelligent traffic management and optimization of existing transport infrastructure. This study presents a comprehensive analysis of modern neural network architectures for lane segmentation to determine their potential in traffic flow optimization and assess their applicability in traffic management systems. A comparative analysis of five advanced deep learning architectures was conducted: LaneATT with attention mechanism for keypoint detection; ERFNet and ENet as efficient real-time architectures; PINet for instance-based lane segmentation; and CondLaneNet with conditional lane shape generation. The characteristics of each architecture, their computational efficiency and accuracy were investigated, and the impact of segmentation quality on traffic flow optimization metrics was analyzed. It was shown that CondLaneNet provides the highest segmentation accuracy and best lane geometry recovery at low traffic levels and simple road topology, while LaneATT demonstrates slightly lower peak accuracy but exhibits smoother quality degradation with increasing traffic density. ENet and ERFNet provide an acceptable trade-off between accuracy and computational complexity, making them suitable for real-time systems with limited resources. PINet, due to keypoint-based representation and clustering, shows the best resistance to occlusions and complex road scenarios, maintaining the highest proportion of baseline quality under medium and

high traffic conditions. The obtained results enable informed selection of optimal architecture for specific application scenarios in intelligent transportation systems.

Key words: lane segmentation; computer vision; deep learning; intelligent transportation systems; traffic flow optimization; traffic management; LaneATT; ERFNet; ENet; PINet; CondLaneNet.

Постановка проблеми

Сучасні мегаполіси стикаються з критичними проблемами управління транспортними потоками через постійне зростання кількості транспортних засобів, урбанізацію, розвиток сервісів доставки та обмежені можливості розширення дорожньої інфраструктури. Затори на дорогах призводять до значних економічних втрат, збільшення викидів парникових газів, підвищення рівня шумового забруднення та загального зниження якості життя населення. За даними різних досліджень, неефективне управління транспортними потоками спричиняє втрату мільярдів годин продуктивного часу щороку, що негативно впливає на конкурентоспроможність міст і регіонів. В цих умовах традиційні методи організації дорожнього руху, засновані на статичних схемах, фіксованих циклах світлофорного регулювання та періодичних вимірюваннях інтенсивності, виявляються недостатньо гнучкими й не здатними адекватно реагувати на швидкі зміни транспортної ситуації в реальному часі.

Аналіз останніх досліджень і публікацій

Технології комп'ютерного зору, зокрема автоматична сегментація смуг руху, стають ключовим інструментом для вирішення зазначених проблем, оскільки дозволяють перетворювати відеодані з дорожніх камер, бортових сенсорів ADAS та автономних транспортних засобів у структуровану інформацію про конфігурацію та завантаженість смуг. Останні досягнення в галузі глибокого навчання відкрили нові можливості для точного та швидкого виявлення смуг руху в реальному часі. Зокрема, у роботі (1) запропоновано архітектуру CLRNNet, яка поєднує високорівневі семантичні та низькорівневі деталізовані ознаки через механізм крос-шарового уточнення, що підвищує стійкість до шуму, складних дорожніх умов і часткових оклюзій. У дослідженні (2) представлено CondLaneNet – топ-даун підхід до детекції смуг на основі умовних згорток, який дозволяє адаптивно моделювати форму смуги залежно від локального контексту сцени та топології дорожньої мережі (розгалуження, злиття, розв'язки, складні перехрестя).

Окремий напрямок розвитку складають методи, що поєднують оцінку ключових точок із підходами інстанс-сегментації. У роботі (3) запропоновано підхід до оцінки ключових точок та інстанс-сегментації смуг, який дозволяє розрізняти окремі смуги на багатосмугових дорогах і в складних сценаріях із щільним трафіком. Подальший розвиток цього класу підходів демонструє глобальна асоціативна мережа на основі ключових точок (9), яка виконує встановлення відповідностей між точками уздовж усієї смуги та будівництво цілісних траєкторій навіть за наявності пропусків у розмітці. Для підвищення якості виділення релевантних ознак активно використовуються механізми уваги: в роботі (4) представлено LaneATT – модель із увагою, що орієнтується на області, пов'язані зі смугами, та агрегує глобальний контекст для покращення детекції в умовах часткової відсутності дорожньої розмітки, тіней та яскравих відблисків.

Просторовий контекст відіграє критичну роль у коректному розумінні дорожніх сцен, особливо в умовах складної геометрії доріг, розв'язок та змішаних типів розмітки. У роботі (5) запропоновано Spatial CNN для моделювання просторових залежностей уздовж смуги, що дозволяє поширювати інформацію між сусідніми пікселями у напрямку смуги та підвищувати цілісність результуючих контурів. Дослідження (6) пропонує Ultra Fast Structure-Aware підхід, що забезпечує дуже високу швидкість обробки при збереженні структури смуг: модель враховує глобальні структурні обмеження та досягає продуктивності, придатної для застосування в системах реального часу на бортових обчислювальних платформах.

Все більшої уваги набувають методи, орієнтовані на явне моделювання криволінійної структури смуг. У роботі (7) запропоновано підхід до ефективної детекції смуг на основі явного моделювання кривих, що дозволяє краще відображати геометрію смуг на складних ділянках дороги (S-подібні криві, естакади, багаторівневі розв'язки). Модель RESA (8) вводить рекурентний механізм зміщення ознак (feature-shift), який багаторазово агрегує контекстну інформацію вздовж просторових напрямків, що підвищує цілісність і стійкість сегментації в умовах шуму, часткових перекриттів і зміни освітлення. Робота (11) пропонує CurveLane-NAS, де поєднано пошук архітектур, чутливих до конфігурації смуги, з адаптивним змішуванням точок, що дозволяє автоматично підбирати ефективні варіанти моделей під різні обчислювальні бюджети й сценарії застосування.

Новітні підходи активно використовують трансформери та механізми самоуваги для кращого моделювання довготривалих залежностей у зображенні. Дослідження (10) демонструє end-to-end передбачення форми смуг з використанням трансформерів, що дає змогу одночасно враховувати локальні деталі та глобальну геометрію сцени й покращує стабільність результатів на складних послідовностях кадрів. У роботі (12) представлено Line-CNN – підхід до детекції ліній дорожньої розмітки із використанням спеціалізованого блоку пропозицій ліній, який інтегрує геометричні обмеження в процес навчання. Ці методи демонструють, що поєднання глибоких нейромереж із явно заданими геометричними або структурними припущеннями дозволяє суттєво підвищити

надійність сегментації смуг руху в реальних умовах експлуатації. Таким чином, сучасні архітектури сегментації смуг руху охоплюють широкий спектр підходів – від класичних згорткових мереж зі спеціалізованими механізмами уваги та просторового згортання (1, 4, 5) до швидкісних структурно-обізнаних моделей (6, 7), рекурентних агрегаторів ознак (8) та трансформер-орієнтованих рішень (10, 11). Їх спільною рисою є здатність працювати в режимі близькому до реального часу та забезпечувати високу точність навіть у складних дорожніх сценаріях, що робить ці моделі перспективною основою для побудови інтелектуальних транспортних систем.

Проблема оптимізації транспортних потоків на основі даних сегментації смуг руху активно досліджується в контексті інтеграції алгоритмів комп'ютерного зору з системами адаптивного керування дорожнім рухом. Дані про конфігурацію смуг, їх ефективну пропускну здатність та поточне завантаження можуть бути використані для налаштування режимів роботи світлофорів, впровадження реверсивних смуг, динамічного виділення смуг для громадського транспорту чи екстрених служб, а також для побудови точних моделей прогнозування заторів. Точність сегментації при цьому безпосередньо визначає якість вхідних даних для систем прийняття рішень в інтелектуальних транспортних системах і, відповідно, впливає на ефективність реалізації стратегій динамічного перерозподілу транспортних потоків.

Мета дослідження – провести комплексний аналіз сучасних нейромережових архітектур сегментації смуг руху (LaneATT, ERFNet, ENet, PINet, CondLaneNet) та оцінити їх потенціал для оптимізації транспортних потоків в умовах сучасного міського середовища, з урахуванням вимог до роботи в реальному часі, стійкості до завад та обмеженості обчислювальних ресурсів. Основні завдання дослідження полягають у такому:

Формулювання мети дослідження

1. Виконати порівняльний аналіз архітектурних особливостей досліджуваних моделей з урахуванням їх обчислювальної складності, структури мережі, використання механізмів уваги та методів агрегування контексту;
2. Оцінити ефективність різних підходів до агрегації просторово-часової інформації, включаючи ключові точки, криві представлення смуг та трансформер-орієнтовані стратегії;
3. Визначити оптимальні сценарії застосування кожної архітектури залежно від умов експлуатації (бортові системи, інфраструктурні камери, офлайн-аналітика), доступних ресурсів та вимог до латентності;
4. Проаналізувати вплив якості сегментації на показники оптимізації транспортних потоків, такі як пропускну здатність, середній час затримки, стабільність руху на вузлових ділянках мережі;
5. Розробити рекомендації щодо вибору архітектури для конкретних задач управління дорожнім рухом з урахуванням вимог до точності, швидкодії та надійності.

Наукова новизна роботи полягає у систематизованому порівнянні низки провідних нейромережових архітектур сегментації смуг руху в контексті їх застосування для оптимізації транспортних потоків, у врахуванні особливостей різних підходів до обробки просторової та структурної інформації (ключові точки, криві, лінії, увага, трансформери) та у формуванні комплексних критеріїв вибору оптимальних рішень для конкретних умов експлуатації інтелектуальних транспортних систем.

Викладення основного матеріалу дослідження

Дослідження ефективності нейромережових архітектур для сегментації смуг руху в умовах різної інтенсивності транспортного потоку проводилось у кілька взаємопов'язаних етапів, кожен з яких спрямований на отримання об'єктивних та відтворюваних результатів. Методологічна база дослідження ґрунтується на комплексному підході, що поєднує теоретичний аналіз архітектурних особливостей сучасних моделей глибокого навчання з емпіричною перевіркою їх ефективності в реальних умовах експлуатації, включаючи сценарії з екстремально високою щільністю транспортного потоку. Фундаментальною гіпотезою дослідження є припущення про нелінійну залежність між щільністю транспортних засобів на проїжджій частині та якістю сегментації смуг руху, причому критичні точки деградації можуть суттєво відрізнятися для різних архітектурних підходів залежно від їхньої здатності до контекстного аналізу та просторової екстраполяції.

Вибір конкретних архітектур для порівняльного аналізу обумовлений необхідністю охопити різні парадигми підходів до задачі сегментації смуг руху, від класичних енкодер-декодерних структур до сучасних трансформерних моделей з механізмами уваги. Архітектура LaneATT базується на механізмі уваги з використанням якорів, що дозволяє ефективно агрегувати глобальну контекстну інформацію та покращувати детекцію в умовах часткової оклюзії смуг. Ця модель використовує ResNet-34 як базову мережу для екстракції ознак, з подальшим проектуванням якорів на карти ознак для формування векторів представлення кожної потенційної смуги. Механізм уваги в LaneATT формалізується через функцію:

$$Attention(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (1)$$

де Q – представляє запити від якорів

K, V – є ключами та значеннями з карт ознак відповідно

d_k – позначає розмірність простору ключів, що в даній імплементації становить 64

Налаштування гіперпараметрів для LaneATT включало встановлення початкової швидкості навчання на рівні 0.0003 з косинусним розкладом за формулою:

$$\eta_t = \eta_{min} + (\eta_{max} - \eta_{min}) \cdot \frac{1 + \cos(\pi t / T)}{2} \quad (2)$$

де t – поточна епоха

T – загальна кількість епох

η_{min} та η_{max} – мінімальна та максимальна швидкості навчання

Використання батчів розміром 32 зображення у поєднанні зі стратегією warm-up протягом перших 500 ітерацій виявилось критично важливим для стабілізації процесу навчання. За відсутності такого підходу на початкових етапах спостерігалися суттєві коливання значень функції втрат, що призводило до субоптимальної збіжності моделі. В умовах вільного руху LaneATT досягає середнього значення IoU на рівні 0.947 з стандартним відхиленням $\sigma = 0.023$ тоді як при щільності трафіку понад 50 автомобілів на кадр цей показник падає до 0.623 ($\sigma = 0.087$), що представляє деградацію приблизно на 34.2% порівняно з базовими умовами.

ERFNet представляє принципово інший підхід, орієнтований на досягнення оптимального балансу між обчислювальною ефективністю та точністю сегментації через архітектурні інновації на рівні базових будівельних блоків. Архітектура використовує факторизовані резидуальні шари, де стандартні 3×3 згортки розкладаються на послідовність 1×3 та 3×3 згорток, що дозволяє зменшити кількість параметрів приблизно в 2.7 рази без суттєвої втрати якості представлення ознак. Математично, стандартна згортка з ядром розміру $k \times k$ та C вхідними каналами має складність $\mathcal{O}(k^2 \cdot C^2 \cdot H \cdot W)$, тоді як факторизована версія зменшує її до $\mathcal{O}(2k \cdot C^2 \cdot H \cdot W)$, де H та W є висотою та шириною карти ознак відповідно. Для ERFNet була застосована поетапна стратегія навчання. На першому етапі протягом 150 епох навчався лише енкодер із використанням кросентропійної функції втрат, яка враховує різницю між істинними мітками пікселів та передбаченими ймовірностями їх належності до класу смуги. Функція втрат для енкодера визначається таким чином:

$$L_{encoder} = -\sum_i y_i \cdot \log(p_i) \quad (3)$$

де y_i – позначає істинну мітку пікселя i

p_i – ймовірність належності до класу смуги.

На другому етапі до моделі додавався декодер, після чого мережу донавчали ще 100 епох із зменшеною швидкістю навчання 0.0001. На цьому етапі використовувалася комбінована функція втрат, що поєднує кросентропійний доданок та Dice-втрати і задається виразом:

$$L_{total} = \lambda_1 \cdot L_{CE} + \lambda_2 \cdot L_{Dice} \quad (4)$$

Внесок кожного компонента регулювався коефіцієнтами балансування λ_1 та λ_2 , значення є $\lambda_1 = 0.7, \lambda_2 = 0.3$, що дозволило досягти кращого компромісу між піксельною точністю та цілісністю сегментаційної маски. Регуляризація здійснювалась через dropout з ймовірністю 0.1 у декодері та weight decay з коефіцієнтом 0.0001, що формалізується як додатковий член у функції втрат:

$$L_{reg} = L_{total} + \alpha \cdot \sum \|w\|^2 \quad (5)$$

де α – коефіцієнт регуляризації

w – ваги мережі

Експериментальні результати показують, що ERFNet демонструє відносно стабільну поведінку при зростанні щільності трафіку: при переході від вільного руху (IoU = 0.921) до помірної завантаженості (IoU = 0.847) спостерігається деградація лише на 8.0%, тоді як при екстремальних заторах показник падає до 0.581, що все ж таки становить 36.9% зниження від базового рівня. Важливо відзначити, що ця архітектура виявляє особливу чутливість до освітлення: при поганих умовах освітлення в поєднанні з високою щільністю трафіку IoU може падати додатково на 12-15% порівняно з денними сценаріями аналогічної завантаженості.

Архітектура ENet розроблена спеціально для застосувань реального часу на вбудованих системах з обмеженими обчислювальними ресурсами, де критичним є не лише точність сегментації, але й швидкість обробки кадрів та енергоефективність алгоритму. Її особливістю є асиметрична структура енкодер-декодер, де декодер значно менший за енкодер (співвідношення параметрів приблизно 1:4), що дозволяє прискорити процес виведення без критичної втрати точності через застосування ранньої стратегії downsampling. ENet використовує ранню стратегію downsampling на початкових шарах для зменшення обчислювальних витрат на подальших етапах обробки: після першого шару просторовий розмір зменшується з 512×256 до 256×128 , а після другого до 128×64 , що призводить до зменшення обчислювальної складності в геометричній прогресії. Навчання моделі проводилось з використанням політики поліноміального зменшення швидкості навчання з степенем 0.9 за формулою $\eta_t = \eta_0 \cdot (1 - t / T)^{0.9}$ де початкова швидкість η_0 становила 0.001, а оптимізація здійснювалась алгоритмом Адам

з параметрами $\beta_1 = 0.9$ та $\beta_2 = 0.999$, що визначають експоненціальні швидкості згасання для оцінок першого та другого моментів градієнта відповідно. Функція втрат для ENet включає ваговий коефіцієнт для балансування класів:

$$w_i = \frac{1}{\ln(c + p_i)} \quad (6)$$

де c є константою рівною 1.02, а p_i частотою появи класу i в тренувальній вибірці. Аналіз залежності якості сегментації від щільності трафіку для ENet виявляє найбільш драматичну деградацію серед усіх досліджуваних архітектур: при вільному русі модель досягає IoU = 0.893, що знижується до 0.782 при помірній завантаженості (деградація 12.4%), до 0.641 при інтенсивному русі (деградація 28.2%) та катастрофічних 0.478 при заторах (деградація 46.5%). Ця особливо висока чутливість пояснюється агресивним downsampling на ранніх шарах, що призводить до втрати дрібних деталей розмітки, які стають критично важливими для екстраполяції смуг під транспортними засобами. Додатково спостерігається, що при щільності понад 45 автомобілів на кадр частота хибно позитивних детекцій зростає на 234% порівняно з базовим рівнем, оскільки модель починає інтерпретувати межі між автомобілями як потенційні смуги руху.

PINet базується на підході виявлення ключових точок з подальшою їх кластеризацією для формування смуг, що концептуально відрізняється від традиційних методів щільної сегментації на рівні пікселів. Модель використовує стекові структури типу hourglass для багатомасштабної екстракції ознак, що дозволяє ефективно виявляти точки смуг на різних рівнях деталізації через обробку зображення на чотирьох різних масштабах одночасно. Архітектура включає три послідовні hourglass модулі, кожен з яких виконує операції downsampling до просторового розміру 16×8 пікселів з подальшим upsampling до оригінальної роздільності через транспоновані згортки з коефіцієнтом stride = 2 та kernel size = 4. Особливістю навчання PINet є використання комбінованої функції втрат:

$$L_{total} = \alpha \cdot L_{keypoint} + \beta \cdot L_{offset} + \gamma \cdot L_{cluster} \quad (7)$$

що включає регресію координат ключових точок, класифікацію їх належності до конкретних смуг та контрастивне навчання для покращення якості кластеризації. Компонент $L_{keypoint} = \sum_i \|p_i - \hat{p}_i\|^2$ вимірює евклідову відстань між передбаченими та істинними позиціями ключових точок, тоді як $L_{offset} = \sum_{i,j} \|o_{ij} - \hat{o}_{ij}\|^2$ регресує векторні зміщення від кожного пікселя до найближчої ключової точки відповідної смуги. Кластеризаційна складова формується у вигляді контрастивної функції втрат:

$$L_{cluster} = \sum_{i,j} [y_{ij} \cdot \|e_i - e_j\|^2 + (1 - y_{ij}) \cdot \max(0, m - \|e_i - e_j\|^2)] \quad (9)$$

де e_i і e_j – вкладки точок i та j у просторі ознак

p_i – індикатор їх належності до однієї смуги

$m = 2$ – виступає параметром margin, який забезпечує розділення різних кластерів у просторі ембедингів.

Гіперпараметри моделі налаштовувалися за допомогою процедури grid search із валідацією на окремій підвбірці. Оптимальною виявилася комбінація: швидкість навчання 0.0001, розмір батчу 16, а також коефіцієнти балансування компонентів функції втрат $\alpha = 0.5$ для регресії координат ключових точок, $\beta = 0.3$ для векторних зміщень та $\gamma = 0.2$ для кластеризації. Додатково було застосовано focal loss для компоненти виявлення ключових точок:

$$FL(p_i) = -\alpha_i (1 - p_i)^\gamma \log(p_i), \quad (10)$$

де p_i – ймовірність правильної класифікації

α_i – ваговий коефіцієнт, $\alpha_i = 0.25$

$\gamma = 2.0$, параметр фокусування, який зменшує вплив «легких» прикладів і дозволяє зосередити навчання на складних випадках із високою невизначеністю.

Аналіз поведінки PINet за різних рівнів завантаженості дороги показує, що ця архітектура демонструє найвищу стійкість серед усіх розглянутих моделей. Базове значення IoU при вільному русі становить 0.934 і знижується лише до 0.889 при помірній завантаженості (деградація 4.8%), до 0.827 при інтенсивному русі (деградація 11.5%) та до відносно високих 0.671 навіть в умовах екстремальних заторів (загальна деградація 28.2%). Така стійкість зумовлена природою підходу, що ґрунтується на ключових точках: навіть якщо значна частина смуги перекрита транспортними засобами, для успішної реконструкції її траєкторії достатньо виявити кілька видимих опорних точок і надалі відновити контур через кластеризацію та інтерполяцію. Це можна формалізувати у вигляді умовної ймовірності $P(\text{lane} | \text{keypoints}) = \prod P(k_i | \text{lane}) \cdot P(\text{lane})$, де навіть за наявності лише підмножини видимих точок K_i ймовірність коректної реконструкції залишається високою завдяки просторовій гладкості траєкторій дорожніх

смуг. Водночас детальний аналіз розподілу помилок виявляє, що за щільності трафіку понад 60 автомобілів на кадр частота помилок кластеризації зростає наближено експоненційно: частка випадків, коли точки однієї смуги помилково розподіляються між різними кластерами, збільшується від 2.3% при вільному русі до 18.7% в умовах екстремальних заторів. Це свідчить про те, що навіть для робастних до оклюзій методів на основі ключових точок кластеризаційний етап залишається «вузьким місцем» при надвисокій щільності трафіку.

CondLaneNet представляє топ-даун підхід з умовною генерацією форми смуг на основі виявлених стартових точок, що дозволяє ефективно обробляти складні топології доріг з розгалуженнями, злиттями та множинними перехрестями. Архітектура включає трансформерний енкодер для ефективної екстракції контекстних ознак через механізм self-attention, який формально визначається виразом:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (11)$$

У розглянутій конфігурації використовується восьмиголова схема уваги, де кожна голова оперує у власному підпросторі розмірності $d_k/8=8$, що дає змогу моделі враховувати різні просторові контексти в межах сцени. Модуль рекурентної інстанціації, відповідальний за обробку складних топологій доріг, базується на LSTM-комірках із розміром прихованого стану 256, що дозволяє моделі послідовно генерувати параметри форми смуги від стартової до кінцевої точки з урахуванням уже згенерованих сегментів. Кожна смуга параметризується через кубічну сплайнову функцію, наведені у виразі:

$$S(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3 \quad (12)$$

де коефіцієнти a_i передбачаються мережею для кожного вертикального сегмента зображення окремо, забезпечуючи гладке та безперервне представлення траєкторії. Навчання CondLaneNet вимагало особливої уваги до балансування різних компонентів складної функції втрат. Сукупна функція втрат задається як лінійна комбінація чотирьох складових:

$$L_{\text{total}} = \lambda_1 L_{\text{proposal}} + \lambda_2 L_{\text{shape}} + \lambda_3 L_{\text{classification}} + \lambda_4 L_{\text{matching}}. \quad (13)$$

Компонента L_{proposal} відповідає за детекцію пропозицій стартових точок смуг і реалізована у вигляді бінарної кросентропії:

$$L_{\text{proposal}} = -\sum_i [y_i \log(p_i) + (1 - y_i) \log(1 - p_i)]. \quad (14)$$

Компонента L_{shape} виконує регресію параметрів форми через середньоквадратичне відхилення між передбаченими та істинними координатами у фіксованих точках семплювання t_j :

$$L_{\text{shape}} = \sum_{i,j} \|S_i(t_j) - \hat{S}_i(t_j)\|^2. \quad (15)$$

Додатково використовується складова $L_{\text{classification}}$, що реалізує категоріальну кросентропію для визначення типу розмітки смуги серед можливих категорій (суцільна, переривчаста, подвійна тощо). Ключовим елементом є компонент L_{matching} , який відображає узгодження між передбаченими та істинними смугами за допомогою угорського алгоритму. Для цього визначається матриця вартостей:

$$C_{ij} = \|center_i - center_j\|^2 + \lambda \cdot \text{IoU\_loss}(i, j), \quad (16)$$

де перший доданок враховує евклідову відстань між геометричними центрами смуг, а другий – їх просторове перекриття через компоненту втрат IoU. Мінімізація сумарної вартості за матрицею C_{ij} забезпечує оптимальне призначення передбачених смуг до відповідних еталонних. Для підвищення стабільності навчання застосовувалась стратегія curriculum learning. На першому етапі модель навчали на простих прямих ділянках доріг протягом 50 епох, використовуючи спрощену функцію втрат $L_{\text{simple}} = \lambda_1 L_{\text{proposal}} + \lambda_2 L_{\text{shape}}$, яка включає лише компоненти, пов'язані з генерацією пропозицій та регресією форми смуги. На другому етапі, протягом наступних 100 епох, до навчання поступово додавались складніші сценарії з кривими та розгалуженнями, і використовувалась повна функція втрат L_{total} , причому вага компоненти L_{matching} лінійно зростала від 0 до остаточного значення $\lambda_4 = 0.4$. Такий підхід забезпечував плавну адаптацію моделі до складніших топологій дорожньої мережі. Коефіцієнти балансування $\lambda_1 = 1.0, \lambda_2 = 2.5, \lambda_3 = 0.8$ були підібрані за допомогою Bayesian optimization і забезпечили найкращу конвергенцію на валідаційній вибірці.

Аналіз залежності якості сегментації від щільності трафіку для CondLaneNet виявляє виражену нелінійну деградацію. У режимі вільного руху модель демонструє максимальне значення IoU = 0.961, що суттєво перевищує показники інших архітектур. Однак уже при помірній завантаженості цей показник знижується до 0.851 (деградація 11.4%), при інтенсивному русі – до 0.714 (деградація 25.7%) і, нарешті, до 0.587 в умовах заторів

(загальна деградація 38.9%). Математичний аналіз цієї поведінки свідчить, що трансформерний енкодер ефективно використовує глобальний контекст за умови достатньої видимості розмітки, однак його переваги швидко нівелюються при оклюзії понад 40% довжини смуги, коли механізм self-attention вже не знаходить достатньо інформативних просторових залежностей. Додатково спостерігається проблема генерації «фантомних» смуг при надвисокій щільності трафіку: за щільності понад 55 автомобілів на кадр частота хибно позитивних детекцій зростає з базових 3.1% до 22.4%. Це пов'язано з тим, що рекурентний модуль інстанціації починає інтерпретувати регулярні геометричні структури, утворені рядами автомобілів, як потенційні смуги руху, що призводить до формування псевдорозмітки, не підтвердженої реальними дорожніми маркерами.

Оцінювання якості сегментації виконувалося на основі набору взаємодоповнювальних метрик, кожна з яких підсвічує окремий аспект роботи моделей та по-різному реагує на специфічні типи помилок. Базовим показником виступає Intersection over Union, який визначається як відношення площі перетину передбачених та еталонних пікселів смуги до площі їх об'єднання:

$$IoU = \frac{|P \cap G|}{|P \cup G|} = \frac{\sum_{i,j} 1[p_{ij} = 1 \wedge g_{ij} = 1]}{\sum_{i,j} 1[p_{ij} = 1 \vee g_{ij} = 1]}, \quad (17)$$

де P – множина пікселів з координатами (i,j) , які модель класифікувала як смугу ($p_{ij} = 1$)

G – еталонна розмітка ($g_{ij} = 1$)

Спочатку для кожного зображення обчислюється середній показник по всіх смугах:

$$IoU_{\text{image}} = \frac{1}{N} \sum_{k=1}^N IoU_k, \quad (18)$$

де N – кількість смуг на зображенні, а далі виконується усереднення по всьому тестовому набору:

$$mIoU = \frac{1}{M} \sum_{m=1}^M IoU_{\text{image},m}, \quad (19)$$

де M – загальна кількість зображень у тесті. Порогове значення для визнання смуги успішно виявленою встановлено як $IoU \geq 0.5$, що відповідає загальноприйнятій практиці в транспортних задачах сегментації та вимогам точності систем автономного керування рівня SAE Level 3 згідно з ISO 26262. Точність класифікації аналізувалась через метрики Precision та Recall, які залежать від кількості істинно позитивних (TP), хибно позитивних (FP) та хибно негативних (FN) детекцій на рівні окремих смуг. Смуга вважається істинно позитивною, якщо IoU між передбаченою та еталонною смугою не менший за поріг $\tau = 0.5$ і для цієї еталонної смуги не існує іншої передбаченої смуги з більшим IoU. Формально:

$$TP = |\{(p, g) : IoU(p, g) \geq \tau \wedge \forall p' \neq p, IoU(p', g) < IoU(p, g)\}|, \quad (20)$$

$$FP = |\{p : \forall g, IoU(p, g) < \tau\}| \quad (21)$$

У такому формулюванні TP відображає кількість передбачених смуг, для яких існує єдина найкраща відповідність в еталоні з IoU не нижче порогового значення τ . Множина FP охоплює всі передбачені смуги p , що не мають жодної еталонної смуги g з достатнім перекриттям ($IoU(p, g) \geq \tau$), тобто є хибно позитивними детекціями. Аналогічно, множина FN складається з еталонних смуг g , для яких жодна з передбачених смуг не досягає потрібного рівня IoU, тобто ці смуги були повністю пропущені моделлю. Метрика $Precision = \frac{TP}{TP + FP}$ характеризує частку коректно виявлених смуг серед усіх передбачених моделлю (і наближається до 1.0 у разі відсутності хибно позитивних детекцій).

Водночас $Recall = \frac{TP}{TP + FN}$ описує, яка частина всіх наявних в сцені смуг була успішно виявлена ($Recall = 1.0$, якщо модель не пропускає жодної еталонної смуги). Гармонічне середнє цих двох показників задається через:

$$F1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} = \frac{2 \cdot TP}{2 \cdot TP + FP + FN}, \quad (22)$$

що забезпечує збалансовану інтегральну оцінку й «штрафує» моделі, які покращують лише Precision або лише Recall за рахунок іншої компоненти. умовах високої щільності трафіку спостерігається виразна асиметрія: Recall зменшується істотно швидше, ніж Precision. Це пояснюється тим, що моделі значно частіше не виявляють оклюдовані смуги (зростання FN), ніж генерують зайві (FP). Кількісно цю нерівномірність зміни можна описати коефіцієнтом асиметрії:

$$A = \frac{\Delta Precision / Precision_0}{\Delta Recall / Recall_0}, \quad (23)$$

який для різних архітектур набуває значень 0.42 (LaneATT), 0.38 (ERFNet), 0.31 (ENet), 0.53 (PINet) та 0.29 (CondLaneNet) при переході від режиму вільного руху до екстремальних заторів. На рівні смуг цей показник збігається з F1, однак при інтерпретації на рівні пікселів надає альтернативний погляд на якість перекриття еталонної та передбаченої масок. Завдяки своїй формі даний коефіцієнт є менш чутливим до дисбалансу класів, ніж IoU; зокрема, при його перезаписі через $Dice = \frac{IoU}{2 - IoU}$ зменшується вплив суттєвої різниці в площі позитивного й негативного класів.

Окрему увагу в дослідженні приділено тому, як якість сегментації змінюється зі зростанням щільності транспортного потоку. Для цього тестовий набір поділено на чотири класи за рівнем завантаженості дороги: вільний рух, помірний трафік, інтенсивний трафік та затори. Класи визначались за кількістю автомобілів у кадрі, рівнем оклюзії розмітки та, для заторів, додатково за низькою середньою швидкістю руху. Для кожної категорії окремо обчислювали всі метрики якості та довірчі інтервали за допомогою бутстреп-ресемплінгу, що забезпечило статистично надійне порівняння ефективності моделей у різних режимах трафіку. Отримані результати демонструють чітку тенденцію: при переході від вільного руху до заторів середнє значення IoU зменшується приблизно з 0,92–0,95 до 0,58–0,67 (залежно від архітектури), тобто відносна втрата якості сягає 28–47%. Серед розглянутих моделей найвищу стійкість до зростання оклюзії показала PINet: у режимі заторів їй вдається зберегти близько 72% базової якості, досягнутої у вільному русі. Це узгоджується з природою підходу на основі ключових точок, який залишається працездатним навіть тоді, коли видимими залишаються лише фрагменти смуг і модель змушена відновлювати траєкторії за неповною інформацією.

На противагу цьому, CondLaneNet демонструє найкращу точність у «чистих» умовах із мінімальними оклюзіями, але дуже різко деградує при зростанні щільності трафіку. Порівняння з LaneATT, ERFNet, ENet та PINet показує, що швидкість падіння якості для CondLaneNet є найбільшою, що робить її менш придатною для сценаріїв з інтенсивним рухом, попри відмінні результати на вільних дорогах. Детальніший аналіз залежності IoU від кількості транспортних засобів у кадрі виявив наявність критичних точок, після яких швидкість деградації різко зростає. Для LaneATT така точка настає приблизно при 42 автомобілях на кадр, для ERFNet – близько 38, для ENet – вже близько 29, тоді як для PINet – лише близько 57 автомобілів. CondLaneNet «ламається» раніше за PINet і LaneATT – приблизно при 33 автомобілях на кадр – і після цього демонструє найстрімкіше падіння якості серед усіх архітектур. Це додатково підкреслює контраст між високою точністю в ідеальних умовах та слабкою стійкістю у складних сценах.

Аналіз типових помилок у густозавантажених сценах показав домінування двох сценаріїв: некоректної екстраполяції траєкторій смуг під автомобілями та злиття сусідніх смуг у одну. Останнє спостерігається тоді, коли відстань між смугами стає меншою за ефективне рецептивне поле згорткової мережі, і модель фактично «не має просторової роздільної здатності», щоб розрізнити їх на рівні карт ознак. У таких випадках ERFNet, ENet і LaneATT демонструють суттєве зростання частоти злиттів смуг, тоді як PINet завдяки точковій парадигмі зберігає відносно низький рівень цих помилок і краще підтримує розділеність сусідніх смуг. Окремий набір експериментів був присвячений граничним умовам майже повної оклюзії, коли видимість розмітки опускається нижче 20% від її загальної довжини. У цьому режимі всі досліджувані моделі демонструють IoU нижче 0,4 і фактично стають непридатними для практичного використання без залучення додаткових джерел інформації, таких як високоточні картографічні дані, GPS або V2X-комунікація. Імовірність отримати прийнятну якість (IoU не менше 0,5) за таких умов не перевищує приблизно 8–9% навіть для найкращої архітектури PINet, що підкреслює фундаментальні обмеження чисто візуальної сегментації в екстремальних ситуаціях.

Додаткові аспекти оцінювання включали аналіз часової стабільності на відеопослідовностях (через варіативність положень ключових точок після компенсації руху камери), кореляційний аналіз між метриками якості та параметрами трафіку (кількість автомобілів, рівень оклюзії, середня швидкість) та введення інтегрального показника спотворення сегментації, який узагальнює деградацію якості як функцію від комплексного параметра складності сцени. Побудова кривих деградації для цього показника дозволила чітко виділити діапазони, у яких система переходить від стабільної роботи до режиму непередбачуваних збоїв. Саме на основі описаного протоколу оцінювання та введених показників у наступному розділі наведено узагальнені експериментальні результати для всіх розглянутих архітектур, а також проведено їх порівняння в різних режимах трафіку та умовах оклюзії.

Результати та обговорення

Графічний аналіз кривих деградації показує, що вплив трафіку на якість сегментації суттєво різниться між розглянутими архітектурами. Залежність IoU від рівня трафіку свідчить про чіткий компроміс між максимальною точністю в ідеальних умовах і стійкістю до перешкод (рис. 1). CondLaneNet забезпечує найвищі значення IoU за низького трафіку, але демонструє стрімке падіння якості у зоні середнього та високого навантаження. Натомість

PINet зберігає найбільшу частку базової якості в діапазоні середнього та високого трафіку, що робить її більш придатною для реальних міських сценаріїв, де оклюзії та затори є типовими.

Аналіз деградації якості сегментації смуг руху

Криві деградації якості для різних нейромережевих архітектур при зростанні інтенсивності трафіку

Метрика: IoU (Intersection over Union) ☐ Показати критичні точки

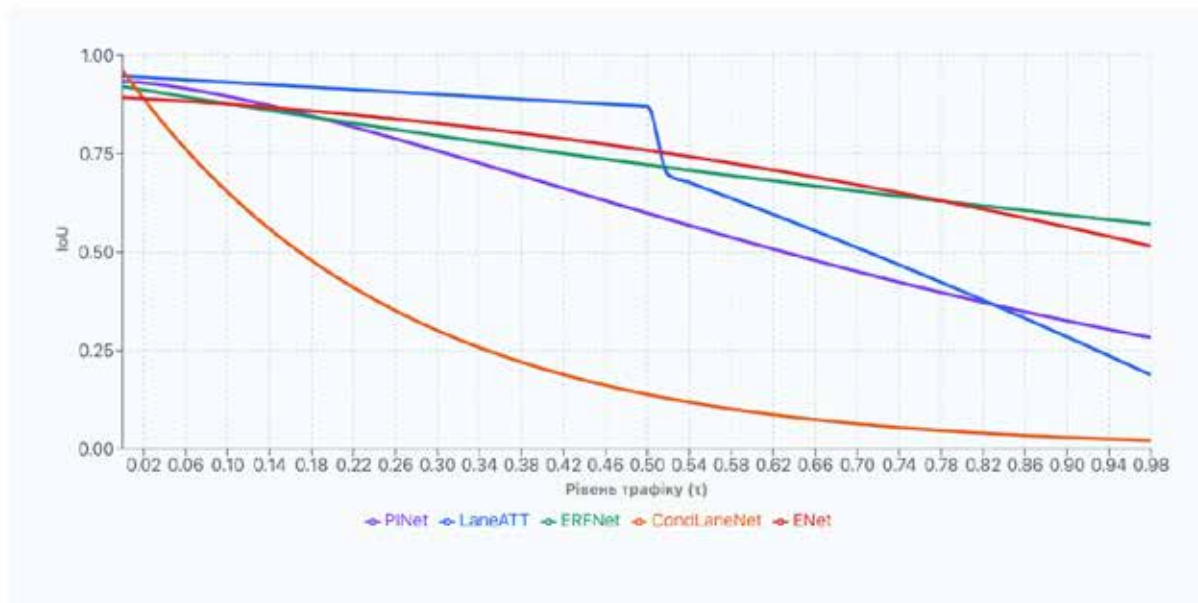


Рис. 1. Криві деградації IoU для різних архітектур сегментації смуг руху залежно від рівня трафіку t

Аналіз F1-міри (рис. 2) уточнює ці спостереження з точки зору балансу між помилками пропуску та хибними спрацьовуваннями. Для PINet деградація F1 зумовлена переважно падінням повноти: модель рідше «галюцинує» смуги, але в умовах сильної оклюзії частіше не виявляє їх взагалі. CondLaneNet, навпаки, за високого трафіку демонструє різке зниження точності через зростання кількості хибно позитивних детекцій, що узгоджується з появою «фантомних» смуг у сценах з густими рядами автомобілів. Таким чином, форма кривих F1 підтверджує різну природу помилок для кожної архітектури.

Криві геометричної точності GA (рис. 3) показують, що відновлення форми смуг деградує швидше, ніж базові сегментаційні метрики. Навіть за відносно прийнятних значень IoU геометрія смуг стає помітно спотвореною, що особливо критично для задач планування траєкторії. У зоні високого трафіку всі архітектури втрачають здатність точно відтворювати кривизну та кут нахилу смуг, однак PINet і CondLaneNet дещо краще зберігають геометричну узгодженість порівняно з ENet та ERFNet. Це підкреслює необхідність окремо оцінювати не лише факт виявлення смуги, а й коректність її просторового профілю. Сукупний аналіз кривих IoU, F1 та GA демонструє, що жодна з розглянутих архітектур не є однозначно оптимальною для всього діапазону умов руху. CondLaneNet забезпечує найвищу якість сегментації та найкраще відновлення геометрії смуг за низького рівня трафіку, але швидко втрачає точність у зоні середнього й високого навантаження. Натомість PINet дещо поступається за максимальною точністю в «чистих» умовах, проте зберігає найбільшу частку базової якості при зростанні рівня трафіку й оклюзії, демонструючи кращий баланс між стійкістю та геометричною коректністю траєкторій. ENet і ERFNet виступають проміжними варіантами з більш рівномірною, але загалом нижчою якістю в усьому діапазоні t .

Сукупний аналіз кривих IoU, F1 та GA демонструє, що жодна з розглянутих архітектур не є однозначно оптимальною для всього діапазону умов руху, а їх ефективність суттєво залежить від щільності трафіку, топології дороги та рівня візуальних перешкод. CondLaneNet забезпечує найвищу якість сегментації та найкраще відновлення геометрії смуг за низького рівня трафіку й відносно простої сцени, коли дорожня розмітка є чіткою, кількість транспортних засобів невелика, а перекриття об'єктів мінімальне. В таких умовах її детальні представлення та умовна генерація форми смуг дають змогу досягати найвищих значень IoU та GA, формуючи більш гладкі та стабільні траєкторії. Водночас зі зростанням щільності руху, появою часткових та повних оклюзій, а також ускладненням дорожньої сцени (складки топології, розгалуження, розв'язки) CondLaneNet демонструє швидку

деградацію якості, особливо за метриками геометричної відповідності, що суттєво обмежує її застосування в реалістичних міських сценаріях із динамічно змінною щільністю трафіку.

Аналіз деградації якості сегментації смуг руху

Криві деградації якості для різних нейромережевих архітектур при зростанні інтенсивності трафіку

Метрика: F1-Score ☐ Показати критичні точки

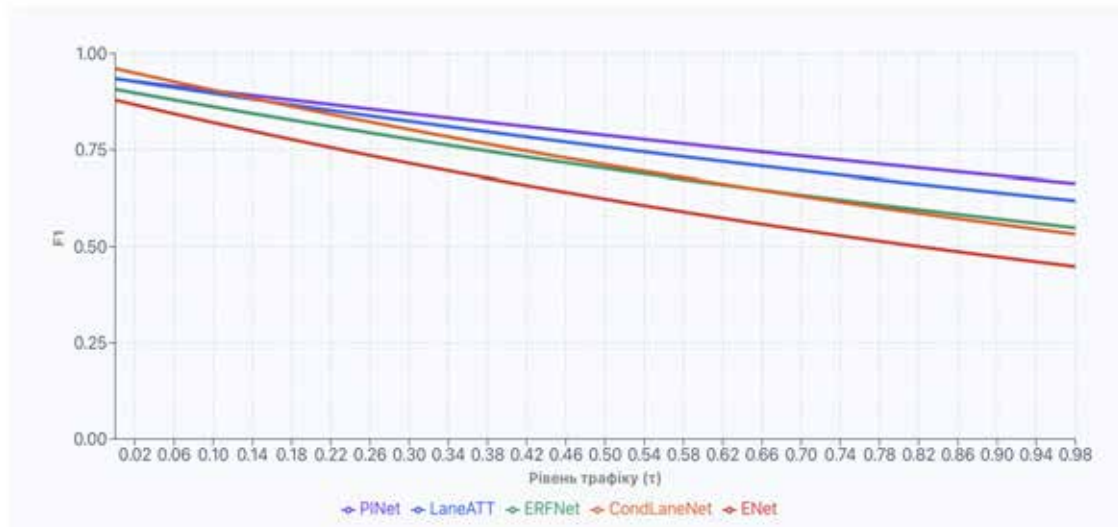


Рис. 2. Залежність F1-міри сегментації смуг руху від рівня трафіку τ для досліджуваних архітектур

Аналіз деградації якості сегментації смуг руху

Криві деградації якості для різних нейромережевих архітектур при зростанні інтенсивності трафіку

Метрика: GA (Geometric Accuracy) ☐ Показати критичні точки

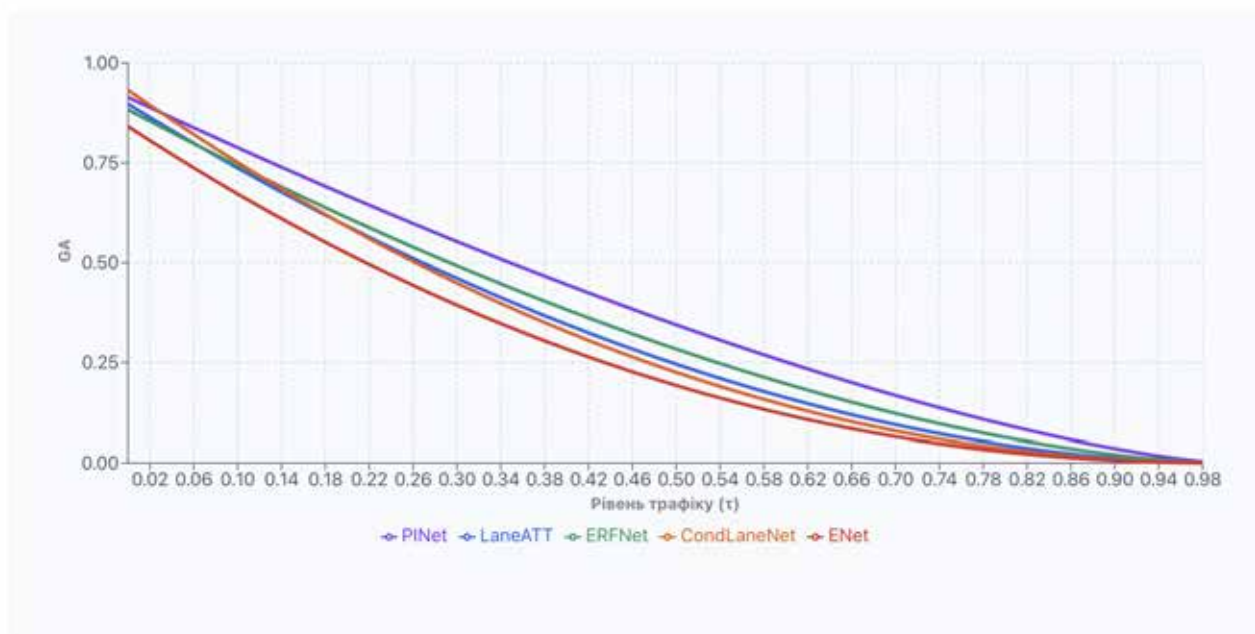


Рис. 3. Деградація геометричної точності GA траєкторій смуг руху при зростанні рівня трафіку τ

Натомість PINet дещо поступається за максимальною точністю в «чистих» умовах, але зберігає найбільшу частку базової якості при зростанні рівня трафіку й оклюзії. Завдяки представленням на основі ключових точок та подальшій кластеризації воно краще адаптується до частково видимих або деформованих смуг, демонструючи вищу стійкість до шумів у зображенні, тимчасових розривів розмітки та накладання транспортних засобів. Це дає змогу досягати більш стабільної поведінки алгоритму за умов середнього та високого навантаження, забезпечуючи кращий баланс між стійкістю, геометричною коректністю траєкторій та придатністю до інтеграції в системи підтримки водія й lane-level навігації, орієнтовані на роботу в щільному міському трафіку.

ENet і ERFNet виступають проміжними варіантами з більш рівномірною, але загалом нижчою якістю в усьому діапазоні t , що зумовлено їхньою компактною архітектурою та орієнтацією на роботу в режимі реального часу. Вони забезпечують достатній рівень точності для базових задач сегментації смуг руху, але поступаються за піковими значеннями GA та F1 у складних сценаріях порівняно з CondLaneNet та PINet. Водночас їхня нижча обчислювальна вартість робить такі моделі доцільними переважно в системах із обмеженими ресурсами та помірними вимогами до точності – наприклад, у недорогих бортових платформах, периферійних обчислювальних вузлах або допоміжних підсистемах моніторингу, де критичним є не максимум якості, а стабільна робота за обмеженого енергоспоживання й затримок.

Отримана картина підкреслює важливість контекстно-залежного вибору архітектури не лише з огляду на середні значення метрик, а й з урахуванням розподілу сценаріїв руху, цільових показників безпеки, допустимого рівня ризику та доступних обчислювальних ресурсів. Для систем підтримки водія (ADAS) та комплексних рішень автоматизації керування доцільним може бути гібридний підхід, за якого архітектури класу CondLaneNet застосовуються в умовах низької чи помірної завантаженості, тоді як PINet або легші моделі на кшталт ENet/ERFNet активуються в сценаріях високої щільності трафіку або на енергообмежених платформах. Це відкриває перспективи подальших досліджень, спрямованих на динамічне перемикання між моделями, адаптивне налаштування архітектур під поточний режим руху та розроблення метрик, чутливих до ризиків для безпеки в реальних експлуатаційних умовах.

Таким чином, отримані результати вказують на чіткий trade-off між максимальною точністю в ідеальних умовах і стійкістю до деградації в складних сценах з високою щільністю трафіку. Запропонований підхід до аналізу кривих деградації дозволяє не лише порівняти архітектури між собою, а й окреслити області їх доцільного застосування: від магістралей із передбачуваним рухом до насичених міських перехресть. Надалі ці спостереження формують основу для обговорення практичних наслідків вибору моделі в системах автономного та автоматизованого керування транспортними засобами.

Висновки

У межах поставленої мети показано, що якість сегментації смуг руху суттєво й нелінійно залежить від рівня завантаженості дороги. Для всіх розглянутих архітектур (LaneATT, ENet, ERFNet, PINet, CondLaneNet) зафіксовано відчутну деградацію метрик IoU, F1 та геометричної точності зі зростанням щільності трафіку та оклюзії розмітки. Це безпосередньо обмежує той діапазон умов, у якому результати сегментації можуть бути надійно використані в системах оптимізації транспортних потоків та управління дорожнім рухом.

Порівняльний аналіз показав, що CondLaneNet забезпечує найвищу точність та найкраще відновлення геометрії смуг за низького рівня трафіку, що робить цю архітектуру придатною для автомагістралей і сцен з чіткою, слабо оклюдованою розміткою. LaneATT у цих умовах демонструє дещо нижчі, але все ще високі показники IoU та F1, водночас характеризується більш плавною деградацією при зростанні щільності трафіку. Внаслідок цього LaneATT можна розглядати як компромісне рішення між максимальною точністю CondLaneNet та підвищеними вимогами до робастності в умовах змішаного міського руху.

PINet, навпаки, демонструє трохи нижчу пікову точність у «чистих» умовах, проте найкраще зберігає якість сегментації при збільшенні рівня трафіку та оклюзії. Завдяки використанню ключових точок і кластеризації ця архітектура забезпечує найвищий коефіцієнт збереження IoU та кращу геометричну узгодженість траєкторій смуг у зоні середнього й високого навантаження. Це робить PINet найбільш перспективним кандидатом для інтеграції в інтелектуальні системи керування – від адаптивного регулювання світлофорів до динамічного перерозподілу смуг руху в умовах змінних міських потоків.

ENet та ERFNet займають проміжну позицію: вони забезпечують помірну точність за рахунок низької обчислювальної складності й можуть бути корисними у системах з жорсткими ресурсними обмеженнями або як компоненти багаторівневих схем (наприклад, для грубого попереднього аналізу сцени чи попереднього відсіювання некритичних областей). LaneATT у цьому контексті виступає як більш «важкий», але й більш точний варіант порівняно з ENet та ERFNet, доцільний для систем, де доступніші обчислювальні ресурси, але потрібна стабільна якість в широкому діапазоні режимів руху. Водночас отримані результати свідчать, що в умовах екстремальних заторів жодна з архітектур не забезпечує достатнього рівня надійності, а отже, для практичних систем управління рухом необхідне поєднання сегментації смуг з додатковими джерелами інформації (HD-карти, V2X, дані від інфраструктурних сенсорів).

Таким чином, поставлена мета досягнута: комплексний аналіз дозволив виявити сильні та слабкі сторони сучасних нейромережових архітектур сегментації смуг руху (LaneATT, ENet, ERFNet, PINet, CondLaneNet), окреслити їх робочі діапазони за рівнем трафіку та сформулювати практичні рекомендації щодо вибору моделі для задач оптимізації транспортних потоків і побудови систем управління дорожнім рухом з урахуванням вимог до точності, робастності та обчислювальних ресурсів.

Список використаної літератури

1. Zheng, T., Huang, Y., Liu, Y., Tang, W., Yang, Z., Cai, D., & He, X. (2022). CLRNNet: Cross layer refinement network for lane detection. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 898-907. <https://doi.org/10.1109/CVPR52688.2022.00097>
2. Liu, L., Chen, X., Zhu, S., & Tan, P. (2021). CondLaneNet: A top-to-down lane detection framework based on conditional convolution. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 3773-3782. <https://doi.org/10.1109/ICCV48922.2021.00375>
3. Ko, Y., Lee, Y., Azam, S., Munir, F., Jeon, M., & Pedrycz, W. (2022). Key points estimation and point instance segmentation approach for lane detection. *IEEE Transactions on Intelligent Transportation Systems*, 23(7), 8949-8958. <https://doi.org/10.1109/TITS.2021.3088488>
4. Tabelini, L., Berriel, R., Paixao, T. M., Badue, C., De Souza, A. F., & Oliveira-Santos, T. (2021). Keep your eyes on the lane: Real-time attention-guided lane detection. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 294-302. <https://doi.org/10.1109/CVPR46437.2021.00036>
5. Pan, X., Shi, J., Luo, P., Wang, X., & Tang, X. (2018). Spatial as deep: Spatial CNN for traffic scene understanding. *Proceedings of the AAAI Conference on Artificial Intelligence*, 32(1), 7276-7283. <https://doi.org/10.1609/aaai.v32i1.12301>
6. Qin, Z., Wang, H., & Li, X. (2020). Ultra fast structure-aware deep lane detection. *European Conference on Computer Vision (ECCV)*, 276-291. https://doi.org/10.1007/978-3-030-58586-0_17
7. Feng, Z., Guo, S., Tan, X., Xu, K., Wang, M., & Ma, L. (2022). Rethinking efficient lane detection via curve modeling. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 17062-17070. <https://doi.org/10.1109/CVPR52688.2022.01655>
8. Zheng, T., Fang, H., Zhang, Y., Tang, W., Yang, Z., Liu, H., & Cai, D. (2021). RESA: Recurrent feature-shift aggregator for lane detection. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(4), 3547-3554. <https://doi.org/10.1609/aaai.v35i4.16469>
9. Wang, J., Ma, Y., Huang, S., Hui, T., Wang, F., Qian, C., & Zhang, T. (2022). A keypoint-based global association network for lane detection. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 1392-1401. <https://doi.org/10.1109/CVPR52688.2022.00147>
10. Liu, R., Yuan, Z., Liu, T., & Xiong, Z. (2021). End-to-end lane shape prediction with transformers. *2021 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 3693-3701. <https://doi.org/10.1109/WACV48630.2021.00374>
11. Xu, H., Wang, S., Cai, X., Zhang, W., Liang, X., & Li, Z. (2020). CurveLane-NAS: Unifying lane-sensitive architecture search and adaptive point blending. *European Conference on Computer Vision (ECCV)*, 689-704. <https://doi.org/10.48550/arXiv.2007.12147>
12. Li, X., Li, J., Hu, X., & Yang, J. (2020). Line-CNN: End-to-end traffic line detection with line proposal unit. *IEEE Transactions on Intelligent Transportation Systems*, 21(1), 248-258. <https://doi.org/10.1109/TITS.2019.2890870>

Дата першого надходження рукопису до видання: 19.11.2025
Дата прийнятого до друку рукопису після рецензування: 16.12.2025
Дата публікації: 31.12.2025

М. С. ДУТЧАК

викладач кафедри інформаційних технологій
Карпатський національний університет імені Василя Стефаника
ORCID: 0000-0002-3337-5613

М. В. ПІКУЛЯК

кандидат технічних наук, доцент,
доцент кафедри інформаційних технологій
Карпатський національний університет імені Василя Стефаника
ORCID: 0000-0003-2192-1899

Р. О. ТИЧКОВСЬКИЙ

кандидат технічних наук,
доцент кафедри інформаційних технологій
Карпатський національний університет імені Василя Стефаника
ORCID: 0009-0001-5136-2648

І. Б. ПОЛАТАЙКО

асистент кафедри інформаційних технологій
Карпатський національний університет імені Василя Стефаника
ORCID: 0000-0002-9111-0162

МОДЕЛЮВАННЯ ЖИТТЄВОГО ЦИКЛУ РОЗРОБКИ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ ІНТЕЛЕКТУАЛЬНИХ НАВЧАЛЬНИХ СИСТЕМ

У статті розглядається концепція, методологія та моделі життєвого циклу розробки програмного забезпечення (Software Development Life Cycle, SDLC) інтелектуальних навчальних систем (ILS). Проаналізовано існуючі підходи, ключові етапи життєвого циклу, виклики та напрями розвитку. Запропоновано гібридну модель SDLC ILS, яка поєднує освітні стратегії, моделі та правила з компонентами штучного інтелекту (Artificial Intelligence, AI) і машинного навчання (Machine Learning, ML) з метою збереження освітнього характеру даного класу систем і водночас отримання користі від статистичної узагальненості, автоматизації прийняття управлінських рішень, адаптивності та можливості безперервного перенавчання системи.

У статті представлено формалізований підхід до SDLC ILS, який синтезує підходи із розробки ILS, адаптивної гіпермедіа та сучасних практик ML. Описано виділені етапи SDLC ILS. Для етапів визначено методологічні підходи, метрики оцінювання та механізми управління. Запропоновано гібридну модель SDLC ILS, яка поєднує освітні стратегії, моделі та правила з компонентами ML, щоб зберегти освітній характер даного класу систем і водночас отримати користь від статистичної узагальненості, автоматизації прийняття управлінських рішень та можливості безперервного перенавчання системи. Розглянутий в роботі процес моделювання SDLC ILS забезпечує структурований підхід до розробки, впровадження та управління ILS. Особливу увагу приділено моделюванню інтеграції когнітивних і педагогічних теорій із системною інженерією, що формує новий стандарт якості освітніх технологій та дозволяє одночасно забезпечити масштабованість, надійність, педагогічну валідність і етичну прозорість систем навчання.

Запропонована гібридна модель і покроковий життєвий цикл служать дорожньою картою для дослідників і практиків, які прагнуть реалізувати ILS із високим рівнем якості та безпеки автоматизованих навчальних процесів.

Ключові слова: моделювання життєвого циклу програмного забезпечення, інтелектуальні навчальні системи, машинне навчання, штучний інтелект, метрики оцінювання.

M. S. DUTCHAK

Lecturer at the Department of Information Technologies
Vasyl Stefanyk Carpathian National University
ORCID: 0000-0002-3337-5613

M. V. PIKULYAK

Candidate of Technical Sciences, Associate Professor,
Associate Professor at the Department of Information Technologies
Vasyl Stefanyk Carpathian National University
ORCID: 0000-0003-2192-1899

R. O. TYCHKOVSKIY

Candidate of Technical Sciences,
Associate Professor at the Department of Information Technologies
Vasyl Stefanyk Carpathian National University
ORCID: 0009-0001-5136-2648

I. B. POLATAYKO

Assistant Professor at the Department of Information Technologies
Vasyl Stefanyk Carpathian National University
ORCID: 0000-0002-9111-0162

MODELING THE SOFTWARE DEVELOPMENT LIFE CYCLE OF INTELLIGENT LEARNING SYSTEMS

The article considers the concept, methodology and models of the Software Development Life Cycle (SDLC) of intelligent learning systems (ILS). Existing approaches, key stages of the life cycle, challenges and development directions are analyzed. A hybrid SDLC ILS model is proposed, which combines educational strategies, models, and rules with components of artificial intelligence (AI) and machine learning (ML) in order to preserve the educational nature of this class of systems and at the same time benefit from statistical generalization, automation of management decision-making, adaptability, and the possibility of continuous retraining of the system.

The paper presents a formalized approach to the SDLC ILS, which synthesizes approaches to the development of ILS, adaptive hypermedia and modern ML practices. The selected stages of the SDLC ILS are described. Methodological approaches, evaluation metrics and management mechanisms are defined for the stages. A hybrid SDLC ILS model is proposed, which combines educational strategies, models and rules with ML components in order to preserve the educational nature of this class of systems and at the same time benefit from statistical generalization, automation of management decision-making and the possibility of continuous retraining of the system. The SDLC ILS modeling process considered in the work provides a structured approach to the development, implementation and management of ILS. Particular attention is paid to modeling the integration of cognitive and pedagogical theories with systems engineering, which forms a new standard of quality of educational technologies and allows simultaneously to ensure scalability, reliability, pedagogical validity and ethical transparency of learning systems.

The proposed hybrid model and step-by-step life cycle serve as a roadmap for researchers and practitioners who seek to implement ILS with a high level of quality and security of automated learning processes.

Key words: software life cycle modeling, intelligent learning systems, machine learning, artificial intelligence, evaluation metrics.

Постановка проблеми

Еволюція ІА в освіті змістила розробку освітніх систем від моделей, заснованих на правилах, до інтелектуальних систем, що постійно розвиваються на основі даних. Як результат, традиційні моделі SDLC є недостатніми, оскільки їм бракує процесів для управління наборами даних, експериментування з моделями, перенавчання та моніторингу розгорнутих компонентів машинного навчання.

ILS – це програмні середовища, які враховують індивідуальні характеристики студента (знання, стиль навчання, темп, мотивацію) та адаптують навчальний процес відповідно до них. Для створення ефективних ILS важливо чітко окреслити моделі життєвий цикл розробки їх програмного забезпечення, які б дозволили системам еволюціонувати і адаптуватися. ILS відрізняються від загальних систем ІА своїм педагогічним призначенням та адаптацію до параметрів моделі студента. Це робить моделювання життєвого циклу складнішим: системи повинні бути точними, надійними, прозорими, безпечними та узгодженими з педагогічними цілями.

Попри помітні покращення у навчанні завдяки класичним навчальним системам, таким як інтерактивні навчальні системи і адаптивні гіпермедійні платформи, розробка надійних SDLC ILS – від концептуалізації до виведення з експлуатації – залишається викликом через необхідність врахування комплексу педагогічних, технічних та етичних вимог.

Це дослідження розвиває підхід до моделювання SDLC ILS, у якому моделювання життєвого циклу являє собою систематичне визначення етапів, артефактів, інтерфейсів та практик управління, що регулюють еволюцію ILS протягом часу, з одночасним збереженням педагогічної мети та емпіричної достовірності. Моделювання життєвого циклу виконує дві взаємодоповнюючі функції: воно реалізує відтворюваність та підзвітність для

дослідницьких ILS і надає інженерні схеми для навчальних систем, готових до впровадження, які повинні відповідати вимогам безпеки, інклюзивності та масштабованості.

Аналіз останніх досліджень і публікацій

Провівши аналіз останніх досліджень і публікацій, було виділено описані нижче основні підходи до класифікації і моделювання SDLC ILS.

На ранніх етапах на моделювання SDLC ILS впливало об'єктно-орієнтоване проєктування, з акцентом на моделюванні предметної області та створенні педагогічних правил. Дослідження ранніх розробників (M. Helander, T. Landauer, P. Prabhu, I. Philip, J. Pavlik, B. Keith) окреслили такі етапи, як формалізація знань предметної області, моделювання студентів, розробка навчальних траєкторій та оцінювання. Ці етапи вимагають активної участі експертів у розробці та функціонуванні таких систем.

Сучасні ILS все більше покладаються на компоненти ML – рекомендаційні моделі, прогнозу аналітику, моделювання поведінки та розмовні агенти на основі великих мовних моделей (Large Language Model, LLM). Вони вимагають етапів життєвого циклу для:

1. Збору даних, анонізації, маркування.
2. Вибору та навчання моделі.
3. Безперервного експериментування.
4. Моніторингу дрейфу даних після розгортання.
5. Періодичного перенавчання.

У сучасній літературі наголошується на необхідності гібридних моделей життєвого циклу, які поєднують SDLC з практиками, орієнтованими на ML, особливо MLOps та LLMOps [1]. MLOps – це Machine Learning Operations, це підхід, який поєднує: ML (розробку та тренування моделей) та DevOps (автоматизацію розгортання, моніторингу та керування SDLC). LLMOps – це Large Language Model Operations, тобто операційні практики для розгортання, керування та підтримки LLM.

У роботі [2] розглянуто різні підходи до когнітивного моделювання, зокрема трейсинг знань (Bayesian Knowledge Tracing, Deep Knowledge Tracing), теорії ймовірнісних відповідей на завдання, ймовірнісні ієрархічні моделі та когнітивні архітектури. Вибір між когнітивною моделлю та ML-підходом має бути обґрунтований обмеженнями щодо інтерпретованості, доступності даних та необхідності детального аналізу помилок. Авторами роботи [3] зазначено, що ILS повинні враховувати закони конфіденційності, зобов'язання щодо справедливості та уникати нанесення шкоди.

Юрченком К. Ю. та Сивицьким Ю. І. було виконано аналітичні дослідження моделей життєвих циклів електронного освітнього контенту, системи управління навчанням та навчальною діяльністю, системи управління контентом з метою вибору та розробки інструментальних засобів автоматизованого проєктування та наповнення освітніх ресурсів, а також синтезу системи управління життєвими циклами її основних компонентів [4].

Отже, останнє десятиліття спостерігається зближення ІА в освіті та ML; ймовірнісні моделі студентів, глибинне моделювання знань та системи рекомендацій, орієнтовані на дані, доповнюють або замінюють статичні когнітивні моделі в багатьох системах. Ця тенденція стимулює розвиток інженерних рішень з використанням ML для керування потоками даних, версіями моделей, безперервною оцінкою та безпечним розгортанням. Крім того, активно збільшується кількість досліджень пов'язаних із життєвим циклом систем ІА, які підкреслюють управління, прозорість та відстежуваність у відтворюваних ІА-системах – питання, які особливо важливі для освітніх застосувань через високий вплив на студентів. Відтворювані ІА-системи – це такі системи ІА, результати роботи яких можна повторити (відтворити) за тих самих умов. [5]

Отже, аналіз останніх досліджень і публікацій показав швидкі темпи розвитку ІА в освіті, що в свою чергу вимагає формалізації та чіткого моделювання усіх етапів SDLC ILS. Як результат, традиційні моделі SDLC та моделі освітніх систем потребують вдосконалення, виникає необхідність цілісних динамічних гібридних моделей SDLC ILS, які б поєднували SDLC з практиками, орієнтованими на ML, та педагогічними методиками і моделями, які дозволили б моделювати процеси управління наборами даних на основі педагогічних і технологічних стратегій; моніторингу та перенавчання розгорнутих компонентів ILS з метою автоматизованої побудови оптимальної адаптивної траєкторії навчання (ATH).

Формулювання мети дослідження

Розробка моделі SDLC ILS, яка б поєднувала освітні стратегії, моделі та правила з компонентами AI і ML з метою збереження освітнього характеру даного класу систем і водночас отримання користі від статистичної узагальненості, автоматизації прийняття управлінських рішень, адаптивності та можливості безперервного перенавчання системи.

Викладення основного матеріалу дослідження

ILS можна розкласти на типові підсистеми:

1. Онтологія предметної області та компоненти знань.
2. Моделювання студентів.

3. Моделювання навчального контенту та взаємозв'язку між його компонентами.
4. Модуль правил прийняття рішень.
5. Модуль збору та моніторингу даних.
6. Управління й координація SDLC ILS.

Кожна підсистема генерує артефакти (наприклад, матриці когнітивних навичок, функції відгуку на завдання, правила керування/прийняття рішень системи тощо) та надає інструменти для моніторингу, оцінки та перенавчання.

Наприклад, використовуючи показник початкового рівня знань студента, його ступеня сприйняття і умовні ймовірності переходів із одного стану в інший для задач трьох рівнів складності можна оцінити ймовірність досягнення j -го ступеня засвоєння за час t : $P_j(t) = 1, 2, 3$. Якщо відомі початкові ймовірності $P_i(0)$ для станів L_i , $i=1, 2, 3$, то ймовірність того, що в момент часу t система перебуває в стані L_i , визначається за формулою (1):

$$p_j(t) = \sum_i p_i(0) p'_{ij}, \quad (1)$$

де елемент p'_{ij} характеризує ймовірність переходу системи із стану L_i в стан j за час t . [6]

Уявімо, що система підбирає для кожного студента завдання певного рівня складності, щоб він навчався максимально ефективно. Правило прийняття рішень (ППР) системи – це деяке правило π , яке вирішує, яке завдання дати студенту у певний момент. Виходячи з цього, цілі ILS можна формалізувати через задачу багатокритеріальної оптимізації: мінімізуємо відносно ППР π очікувану втрату L (негативний навчальний приріст) разом із коригувальними коефіцієнтами, що враховують вартість використаних ресурсів та порушення справедливості (2):

$$\min_{\pi} E [L(\pi)] + \lambda_1 R_{\text{cost}}(\pi) + \lambda_2 R_{\text{fairness}}(\pi), \quad (2)$$

де $L(\pi)$ – основна втрата (низький навчальний результат), R_{cost} – штраф за використані ресурси, R_{fairness} – штраф за несправедливість, λ_1, λ_2 – коефіцієнти важливості цих обмежень

Якщо відмовитися від коригувальних коефіцієнтів і просто мінімізувати втрату L (тобто прагнути швидкого зростання навчальних результатів), система може поводитися так: давати складні завдання тільки сильним студентам, бо вони швидше прогресують; ігнорувати слабших, бо їх навчання «повільно покращує загальний результат», то формально це ефективно, але несправедливо і неефективно з точки зору ресурсів.

Виходячи із аналізу існуючих досліджень із розробки та впровадження ILS, а також власного досвіду розробки ILS було запропоновано гібридну модель SDLC ILS (Рис. 1), у якій виділено такі основні етапи SDLC ILS:

1. Розробка концепції та вимог;
2. Когнітивне та методологічне проєктування компонентів ILS;
3. Розробка моделей компонентів ILS;
4. Розробка алгоритмів функціонування ILS;
5. Програмна реалізація;
6. Пілотне розгортання та тестування;
7. Моніторинг та аналіз даних;
8. Впровадження та моніторинг ILS;
9. Безперервна адаптація та перенавчання;
10. Оцінка ефективності ILS;
11. Аудит та виведення з експлуатації.

Запропонована гібридна модель SDLC ILS поєднує освітні стратегії, моделі та правила з компонентами AI і ML з метою збереження освітнього характеру даного класу систем і водночас отримання користі від статистичної узагальненості, автоматизації прийняття управлінських рішень, адаптивності та можливості безперервного перенавчання системи.

Кожен етап вимагає специфічних артефактів та контрольних точок оцінювання. Наприклад, другий етап проєктує компоненти ILS, враховуючи когнітивні процеси людини: сприйняття, пам'ять, мислення, навчання, прийняття рішень і т.д., тоді як сьомий етап визначає схеми збору, обробки та аналізу даних, класифікацію подій і т.д. на основі методів ML.

Четвертий етап SDLC ILS є одним із найважливіших і ресурсоемних, оскільки вимагає залучення та спільної роботи цілого ряду висококваліфікованих спеціалістів різних галузей (педагогів, інженерів зі знань, науковців, розробників і т.д.) для розробки великої кількості, в тому числі і ймовірнісних, моделей та алгоритмів, які б забезпечили ефективне функціонування та самонавчання майбутньої ILS на основі методів IA та ML.

Розглянемо детальніше методологічні підходи до моделювання деяких етапів SDLC ILS, які базуються на запропонованій гібридній моделі. Наприклад, модуль побудови стратегій навчання може кодувати вимоги до навчального процесу як символічні правила та використовувати модель ранжування, щоб упорядковувати навчальні матеріали в межах дозволених послідовностей. Модуль Студент формує ймовірнісну багатопараметричну модель конкретного студента на основі результатів його навчання, а також моделей інших студентів, які

за певними ознаками можуть вважатися тотожними моделі чи класу даного конкретного студента і на основі цієї моделі будувати оптимальну АТН.

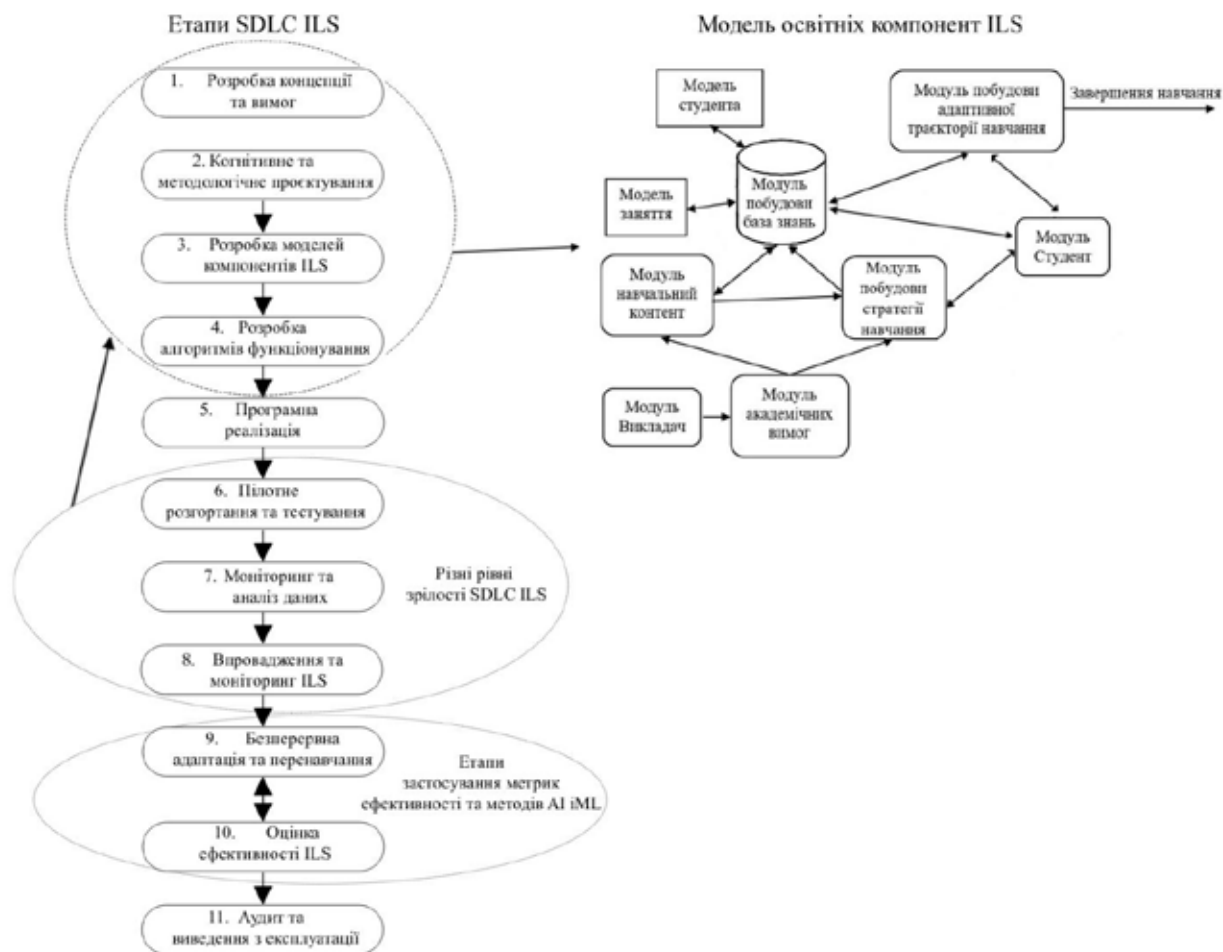


Рис. 1 Гібридна модель SDLC ILS

Спираючись на моделі зрілості MLOps можна описати три рівні зрілості SDLC ILS:

- Експериментальний – випадкові експерименти, ручне перенавчання;
- Операційний – автоматизовані конвеєри, моніторинг, A/B-тестування;
- Керований – повне відстеження походження даних і моделей, перевірки справедливості, механізми відкату, а також формальні аудиту.

Даний підхід забезпечує безперервну інтеграцію, моніторинг, оцінку і управління моделями у розгорнутому програмному продукті ILS.

Для оцінки ефективності гібридної моделі необхідно застосовувати як метрики освітньої ефективності ILS, так і технічні та безпекові метрики. Розповсюдженими показниками метрики освітньої ефективності ILS є нормовані прирости показників навчання, розміри ефекту (коефіцієнт d Коена, який показує наскільки сильно відрізняються результати двох груп (наприклад, експериментальної й контрольної або результати «до» і «після» навчального впливу), рівні засвоєння та показники збереження знань [3].

До ключових показників технічних та безпекових метрик можна віднести:

- затримку / час відгуку – вимірює, скільки часу потрібно ILS, щоб відреагувати на дію студента (наприклад, подати нове завдання, оновити рекомендацію чи АТН на основі оцінки результатів навчання і т.д.);
- доступність / надійність роботи – показує, наскільки часто система перебуває в робочому стані (без збоїв, помилок або відключень);
- калібрування / узгодженість прогнозів – відповідність між передбаченою ймовірністю та реальними результатами; часто є ключовим критерієм для розгортання ILS.
- справедливість / диференційовані прирости – оцінює, чи ILS не дискримінує різні групи користувачів; диференційовані прирости означають порівняння покращень у кожній підгрупі;

- захист приватності – методи, що захищають дані студентів при навчанні моделей, додаючи статистичний «шум», щоб унеможливити ідентифікацію конкретної особи.

Одним із ключових викликів при оцінюванні показників метрик ефективності ILS є статистичні та експериментальні особливості ILS, оскільки через нестабільність студентів і навчальних програм, А/В-тести потрібно проводити з блокуванням і корекцією за коваріаціями для усунення впливу сторонніх факторів (коваріат) на результати аналізу.

Більш ефективними у порівнянні із традиційними А/В тестування є експерименти, засновані на моделі типу «бандит» (bandit-based experiments) та адаптивних багатоваріантних випробування (multi-armed trials), які пропонують вибірки, що ефективніше використовують дані, але ускладнюють офлайн-оцінювання та перевірку гіпотез. У таких експериментах система динамічно коригує, кому показувати який варіант, віддаючи перевагу «кращим» варіантам на основі поточних результатів, тоді як традиційні А/В тести розподіляють трафік рівномірно між варіантами, навіть якщо він уже очевидно «гірший». Моделі типу «бандита» ефективніше використовують вибірку: вони швидше «навчаються» і направляють більшість користувачів до кращих варіантів, тобто вони потребують менше циклів спостережень, щоб знайти оптимальний варіант.

Розроблена гібридна модель розглядає артефакти як повноцінні версійовані сутності: набори даних, перетворення ознак, контрольні точки студентських моделей, педагогічні методики та звіти з оцінювання, що дозволяє забезпечити простежуваність системи, тобто можливість пов'язати перебування студента в будь-якій точці АТН під час проходження навчання в реальній впровадженій системі із відповідними версіями наборів даних, багатопараметричних ймовірнісних моделей і алгоритмами оптимальної адаптації подальшого процесу навчання до потреб і можливостей конкретного студента; це має вирішальне значення для відтворюваності результатів та дотримання академічних вимог у сфері освіти.

На основі оцінок як освітньої, так і технічної та безпекової ефективності ILS, тобто від рівня досягнення тих чи інших результатів навчання студентів, аналізу передбачуваних та реальних результатів, їх достовірності та надійності технічної підтримки і т. д. відбувається безперервна адаптація та перенавчання ILS, що дозволяє адаптувати педагогічні і технічні рішення в залежності від поставлених задач з метою підвищення якості надання освітніх послуг системами даного класу.

Висновки

Моделювання SDLC ILS забезпечує структурований підхід до розробки, впровадження та управління ILS. Запропонована гібридна модель SDLC ILS, яка включає послідовні етапи життєвого циклу, методологічні підходи, метрики оцінювання і аналізу, забезпечує інтеграцію когнітивних і педагогічних теорій із системною інженерією (IA, ML), тим самим підвищуючи відтворюваність, надійність, ефективність та педагогічну відповідність даного класу систем. Результати даного дослідження можуть служити дорожньою картою для дослідників і практиків, які прагнуть реалізувати ILS із високим рівнем якості та «інтелектуальності» автоматизованих навчальних процесів.

Список використаної літератури

1. Stone J., Patel R., Ghiasi F., Mittal S., Rahimi S. Navigating MLOps: Insights into Maturity, Lifecycle, Tools, and Careers // 2025 IEEE Conference on Artificial Intelligence (CAI). Santa Clara, CA, USA, 2025. P. 643–650. DOI: 10.1109/CAI64502.2025.00118.
2. Kreuzberger D., Kühl N., Hirschl S. Machine learning operations (MLOps): Overview, definition, and architecture. arXiv, 2022. DOI: 10.48550/arXiv.2205.02302
3. De Silva D., Alahakoon D. An artificial intelligence life cycle: From conception to production. *Patterns*. 2022. Vol. 3. Article No. 100017. DOI: 10.48550/arXiv.2108.13861
4. Юрченко К. Ю., Сивицький Ю. І. Моделі керування життєвими циклами складових освітніх ресурсів компанії. *Телекомунікаційні та інформаційні технології*. 2024. № 2 (83). С. 85–93. DOI: 10.31673/2412-4338.2024.029918.
5. Koedinger, K. R., D'Mello, S., McLaughlin, E. A., Pardos, Z. A., & Rosé, C. P. (2015). Data mining and education. *WIREs Cognitive Science*, 6(4), P. 333–353. DOI: 10.1002/wcs.1350
6. 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 / M. Dutchak, M. Kozlenko, I. Lazarovych, N. Lazarovych, M. Pikuliak, I. Savka // *Lecture Notes in Networks and Systems*. 2021. Vol. 183. P. 206–217. DOI: 10.1007/978-3-030-66840-2_16.

References

1. Stone, J., Patel, R., Ghiasi, F., Mittal, S., & Rahimi, S. (2025). Navigating MLOps: Insights into maturity, lifecycle, tools, and careers. *2025 IEEE Conference on Artificial Intelligence (CAI)*, pp. 643–650. DOI: 10.1109/CAI64502.2025.00118

2. Kreuzberger, D., Kühl, N., & Hirschl, S. (2022). *Machine learning operations (MLOps): Overview, definition, and architecture*. arXiv. DOI: 10.48550/arXiv.2205.02302
3. De Silva, D., & Alahakoon, D. (2022). An artificial intelligence life cycle: From conception to production. *Patterns*, 3, 100017. DOI: 10.48550/arXiv.2108.13861
4. Jurchenko, K. Ju., Syvycjkyj, Ju. I. (2024). Modeli keruvannja zhyttjevyjmy cyklamy skladovykh osvitnikh resursiv kompaniji. *Telekomunikacijni ta informacijni tekhnologhiji*, no 2 (82), pp. 85–93. DOI: 10.31673/2412-4338.2024.029918
5. Koedinger K. R., D'Mello S., McLaughlin E. A., Pardos Z. A., Rosé C. P. (2015) Data mining and education. *WIREs Cognitive Science*. Vol. 6, No. 4. pp. 333–353. DOI: 10.1002/wcs.1350
6. Dutchak, M., Kozlenko, M., Lazarovych, I., Lazarovych, N., Pikuliak, M., & Savka, I. (2021). Methods and software tools for automated synthesis of adaptive learning trajectory in intelligent online learning management systems. *Lecture Notes in Networks and Systems*, 183, pp. 206–217. DOI 10.1007/978-3-030-66840-2_16.

Дата першого надходження рукопису до видання: 19.11.2025

Дата прийнятого до друку рукопису після рецензування: 16.12.2025

Дата публікації: 31.12.2025

Я. Р. Д'ЯКОВ

аспірант кафедри електронних обчислювальних машин
Харківський національний університет радіоелектроніки
ORCID: 0000-0002-6265-7329

МОДЕЛЬ ІЄРАРХІЧНОГО СЕМАНТИЧНОГО ГРАФУ ЗНАНЬ ДЛЯ БАГАТОРІВНЕВОГО ПРЕДСТАВЛЕННЯ ЗНАНЬ

Графи знань є базовою технологією для структурованого представлення знань, логічного виведення та виконання семантичних операцій у складних інформаційних системах. Водночас поширені моделі семантичних графів знань та графів, доповнених великими мовними моделями, працюють з фіксованим масштабом семантичної деталізації та не підтримують багаторівневу абстракцію. Моделі ієрархічних, мультимасштабних графів знань дозволяють конструювати топологічні ієрархії на основі структур мезо- та макрорівнів, водночас додаткових досліджень потребують їхні семантичні властивості та можливі операції над даними. У роботі запропоновано модель ієрархічного семантичного графа знань (ІСГЗ), що описується послідовністю рівнів із власними множинами сутностей та зв'язків і забезпечує багаторівневе представлення знань. Формально визначено структуру ІСГЗ, гібридну метрику семантичної зв'язності, яка поєднує косинусну подібність багатовимірних векторних представлень із структурною мірою Джаккара, а також ітеративний алгоритм ієрархічної агрегації сутностей у рівні-кластери. На основі цих компонентів реалізовано програмний прототип та проведено експериментальну оцінку на підмножинах репрезентативних наборів даних DbPedia50 та YAGO3-10, доповнених попередньо згенерованими багатовимірними векторними представленнями. Показано, що ІСГЗ забезпечує скорочення середнього часу виконання семантичних запитів (до 30–80 % відносно однорівневих моделей та моделей із Louvain/Leiden-кластеризацією) і формує кластери з оптимальнішими показниками семантичної узгодженості та компактності внутрішнього розподілу. Отримані результати підтверджують доцільність використання ІСГЗ та пов'язаних семантичних операцій для побудови масштабованих засобів семантичного аналізу, пошуку та логічного виведення у динамічних базах знань, систем генерації рекомендацій та інтелектуальних помічників.

Ключові слова: граф знань, семантичний граф знань, семантична метрика, багатовимірне векторне представлення, кластеризація.

YA. R. DIAKOV

Postgraduate Student at the Department of Electronic Computers
Kharkiv National University of Radio Electronics
ORCID: 0000-0002-6265-7329

A HIERARCHICAL SEMANTIC KNOWLEDGE GRAPH MODEL FOR MULTILEVEL KNOWLEDGE REPRESENTATION

Knowledge graphs are a fundamental technology for structured knowledge representation, logical inference and semantic operations in complex information systems. However, widely used semantic knowledge graph models and knowledge graphs augmented with large language models operate at a fixed level of semantic detail and do not support multi-level abstraction. Hierarchical, multiscale knowledge graph models allow constructing topological hierarchies based on meso- and macro-level structures, but their semantic properties and possible data operations require additional research. This paper introduces a hierarchical semantic knowledge graph (HSKG) model defined as a sequence of levels with their own sets of entities and relations, enabling multi-level knowledge representation. We formally specify the structure of HSKG, a hybrid semantic connectivity metric that combines cosine similarity of multi-dimensional embeddings with the Jaccard measure on graph neighborhoods, and an iterative hierarchical aggregation algorithm that groups entities into level-clusters. Based on these components, we implement a software prototype and conduct an experimental evaluation on subsets of the DbPedia50 and YAGO3-10 benchmark datasets with pre-generated multi-dimensional embeddings. The results demonstrate that HSKG reduces the average response time of semantic queries (by up to 30–80 % compared to single-level models and Louvain/Leiden-based clustering models) and yields clusters with optimized semantic connectivity and internal distribution compactness. These findings confirm the suitability of the proposed model and related semantic operations for building scalable tools for semantic analysis, search and logical inference in dynamic knowledge bases, recommendation generation systems and intelligent assistants.

Key words: knowledge graph, semantic knowledge graph, semantic metrics, embedding, clustering.

Постановка проблеми

Моделі графів знань (ГЗ) стали основою структурованого представлення знань у численних предметних областях, серед яких медицина, фінанси, штучний інтелект, біоінформатика тощо. Роль означених структур значно зросла з огляду на швидкий розвиток гібридних інтелектуальних систем та великих мовних моделей [1]–[3]. Основний принцип ГЗ – з'єднання фактичних сутностей за допомогою змістовних зв'язків – дозволяє проводити більш комплексний пошук, логічне виведення, моделювання та аналіз залежностей між даними тощо.

Протягом останніх років відбулися якісні зміни у побудові моделей з точки зору їх семантики та гібридизації. Типовими прикладами є семантичні графи знань (СГЗ) та графи знань, що доповнені великими мовними моделями (ГЗДВММ) [4]. Надбудова у вигляді багатовимірних векторного представлення (БВП) його сутностей та зв'язків суттєво покращила розуміння елементів ГЗ засобами електронно-обчислювальних машин: це забезпечило підвищену адаптивність та контекстну узгодженість модельних представлень. Здатність до виявлення семантичної подібності та векторних патернів у таких СГЗ дозволила використовувати у пов'язаних інформаційних технологіях методи пошуку знань, генерації рекомендацій, логічне виведення знань, семантичну кластеризацію тощо [5]. Подані гібридні підходи демонструють продуктивні результати у завданнях типу зв'язування сутностей, семантичного пошуку та багатокрокового логічного виведення.

Основною проблематикою моделей СГЗ та ГЗДВММ є їх фіксований масштаб семантичної деталізації. Так, усі сутності та зв'язки співіснують на певному концептуальному рівні. Така структура не є ефективною для задач багаторівневого представлення знань, де абстрактні концепції є наслідком семантичного узагальнення сутностей нижчих рівнів. У середовищах великих, динамічних баз знань такі обмеження ускладнюють або унеможливають виконання задач, що потребують масштабованих абстракцій, ієрархічного логічного виведення, адекватності моделі тощо.

Наразі останні дослідження зосереджені на ієрархічних, багаторівневих представленнях знань [6]–[7]. Більшість із них орієнтовані переважно на оптимізацію простору БВП та пов'язаних прикладних задач, проте без формального моделювання саме ієрархічних, семантичних моделей.

Для означення шляху до подолання означених обмежень, запропоновано модель ієрархічного семантичного графу знань (ІСГЗ). Внутрішні сутності згруповані у рівні-кластери відповідно до їхньої семантичної зв'язності. Також представлено алгоритм ієрархічної агрегації, що ітеративним шляхом зв'язує семантично подібні сутності у високорівневі абстракції. Сконструйовано метрику зв'язності, що обчислює концептуальну, семантичну близькість сутностей. Зазначені результати – дозволяють створювати та підтримувати цілісність багаторівневих представлень знань, що забезпечує адекватність структури на всіх рівнях абстракції.

Таким чином, проблема полягає в тому, що наявні моделі СГЗ та ГЗДВММ з фіксованим масштабом семантичної деталізації не забезпечують багаторівневого представлення знань та ієрархічного логічного виведення в умовах великих динамічних баз знань.

Аналіз останніх досліджень і публікацій

Сучасні дослідження, що стосуються графових представлень знань і логіко-семантичних операцій на їх основі, можна згрупувати за такими основними напрямками: сучасні моделі СГЗ та ГЗДВММ; ієрархічні, мультимасштабні графові моделі; онтологічні системи знань.

Використання БВП для ГЗ залишається основним підходом для представлення розподіленої семантики як сутностей, так і зв'язків між ними. Відомі дослідження [7] узагальнюють процес структурних модифікацій ГЗ, типових моделей БВП, а також оптимізації прикладних задач. А також акцентується увага на поступовий перехід до гібридних архітектур, що поєднують графові моделі знань, великі, завчасно натреновані мовні моделі, численні представлення БВП [7].

Паралельно існують інтеграційні моделі, що описують семантику між ГЗ та великими мовними моделями. Мета таких моделей корелює з рішеннями задач покращення точності, ефективності пошуку та здатності до оновлення вихідних даних великих мовних моделей. Ці підходи відкривають можливості до зменшення ефекту явища галуцинації [2][8] та пропонують практичні методи для динамічного оновлення пов'язаних ГЗ, їхнього узгодження з операціями лінгвістичного логічного виведення тощо. Показано важливість моделей графів знань, які є семантично насиченими та придатними до динамічного доповнення [9].

Так, більшість підходів до конструювання моделей СГЗ та ГЗДВММ дають у якості результату однорівневі представлення знань, тобто такі, де сутності та зв'язки знаходяться переважно на одному рівні семантичної деталізації. Таке обмеження не дозволяє ефективно переходити від фактів до високорівневих концепцій, візуалізувати знання різних рівнів, проводити багаторівневе логічне виведення з урахуванням масштабованості та адекватності [10].

Актуальним напрямом досліджень також є багатовимірні представлення знань у просторі БВП. Подібні моделі та підходи дозволяють покращувати типові векторно-семантичні операції, зокрема такі як семантичний пошук та прогнозування зв'язків [11]. Проте пріоритет надається ієрархіям у векторному просторі; відсутня структурна агрегація за семантичною зв'язністю елементів, що є невирішеною задачею.

Існує також напрям ієрархічних, мультимасштабних ГЗ для задачі виокремлення на площині графів знань мезо- та макроструктур. До мезоструктур відносяться: підграфи, патерни відношень, повторювані схеми тощо. А до макроструктур – загальний розподіл вузлів, центрів, дублікатів, зв'язності областей знань тощо. Класичні задачі пошуку сильнозв'язних кластерів сутностей та агрегація групи сутностей однією надсутністю (алгоритми Louvain та Infomap) дозволяють формувати топологічні ієрархії. Графові нейронні мережі з потоковою передачею повідомлень та їхні кластерно-чутливі варіанти у контексті зазначених задач створюють багаторівневі надбудови, що якісно впливають на масштабованість структур даних та забезпечують виконання операцій на великих топологічних відстанях. Загалом такі методи реструктуризації або доповнення графів прискорюють поширення повідомлень у нейронних мережах та чіткіше виокремлюють повторювані схеми [12].

Більшість моделей ієрархічних, мультимасштабних ГЗ, пов'язаних з ними графових нейронних мереж, та підходів у мезо- та макромасштабах, орієнтовані на структурні властивості знань, а саме модульність, векторну близькість. При цьому недостатня увага приділяється саме семантичним властивостям (семантична зв'язність, здатність до зсувів (дрейф)) і операціям (логіко-семантичне виведення, обчислення та контроль дрейфу) тощо.

Останні здобутки у розробці кубічних, регіональних моделей БВП відкрили можливості для формування ієрархічної семантики у векторному просторі [6]. Але вони націлені на оптимізацію векторних властивостей та операцій замість представлення формальної графової структури зі збереженням семантичної зв'язності усередині рівнів абстракцій.

Комплексні онтологічні моделі або онтології, наприклад Resource Description Framework Schema та Web Ontology Language, висувають чіткі таксономічні ієрархії сутностей та механізми логічного виведення для галузей інженерії знань, біомедицини [3] тощо. Нові гібридні підходи (наприклад, Onto2Vec) поєднують структуру онтології з шаром БВП для оптимізації задач прогнозування зв'язків, оцінки подібності елементів. Такі гібриди дозволяють увести семантичну компоненту до операцій над даними, але вимагають налаштувань з боку експертів предметних областей. Вони не дозволяють автоматично виводити високорівневі абстракції з даних, але при цьому вирішують задачі кластеризації, класифікації знань, пошук подібних сутностей [13].

Отже, наразі у наукових дослідженнях можна виокремити можливості для: представлення знань на основі СГЗ; покращеного логічного виведення за рахунок інтеграції великих мовних моделей; проведення структурно-ієрархічних операцій на площині мезо- та макромасштабних рівнів знань; впровадження онтологій, що представляють комплексні ієрархії сутностей вищого рівня. Проте бракує уніфікованої графової багаторівневої моделі знань, де кожен рівень зберігає семантичну зв'язність, підтримуючи можливості візуалізації та логічного виведення. Таким чином, постає потреба у формуванні моделі ІСГЗ, що поєднує формальну багаторівневу структуру з алгоритмом ієрархічної агрегації сутностей та метрикою семантичної зв'язності.

Таким чином, жоден із розглянутих підходів не забезпечує одночасно формально визначеної багаторівневої графової структури та метрик семантичної зв'язності для рівнів абстракції, що й обумовлює актуальність запропонованої в роботі моделі ІСГЗ.

Формулювання мети дослідження

Метою даного дослідження є побудова формальної моделі ІСГЗ, яка забезпечує багаторівневе представлення знань та підтримує ієрархічне логічне виведення в умовах динамічних баз знань. Для досягнення поставленої мети необхідно розв'язати такі наукові задачі:

- розробити формальну модель ІСГЗ, що розширює стандартну модель СГЗ шляхом впровадження багаторівневих абстракцій;
- визначити метрику, що кількісно оцінює структурно-семантичну подібність усередині рівнів-кластерів, а також підтримує ієрархічну агрегацію;
- розробити ітеративний алгоритм побудови ієрархії рівнів абстракції, що зберігають семантичну узгодженість елементів;
- провести тестування моделі на локальних даних та виконати емпіричну оцінку на підмножинах публічних наборів даних.

Викладення основного матеріалу дослідження

Формальна модель ієрархічного семантичного графа знань

ІСГЗ розглядається як узагальнення стандартного семантичного графа знань, у якому знання подаються не в одному фіксованому семантичному шарі, а на послідовності рівнів абстракції. Така багаторівнева організація дає змогу одночасно працювати як з детальними фактами, що безпосередньо відповідають вихідним даним, так і з агрегованими узагальненнями, які відображають більш високі концептуальні структури предметної області. Іншими словами, ІСГЗ поєднує мікрорівень локальних взаємозв'язків між сутностями та макрорівень їхніх семантичних кластерів, формуючи узгоджену ієрархію представлень.

Формально модель ІСГЗ задається множиною рівнів, кожен з яких є окремим графом на власній множині сутностей та зв'язків. Модель ІСГЗ можна формалізувати наступним чином:

$$G = \{L_0, L_1, \dots, L_N\}, \quad (1)$$

де: G – ієрархічний семантичний граф знань; L_n – рівень графу, що відповідає певному ступеню абстракції знань ($0 \leq n \leq N$).

Кожен рівень L_n у (1) задається впорядкованою парою:

$$L_n = \{V_n, U_n\}, \quad (2)$$

де: V_n – множина сутностей (вершин) на рівні n ; U_n – множина зв'язків (ребер), що відображають семантичні відношення між елементами з V_n .

Таким чином, кожен рівень L_n (2) є однорівневим семантичним графом, а вся система рівнів $\{L_n\}$ задає ієрархію графів знань.

Основною вимогою до моделі є наявність для кожної сутності її багатовимірною векторного представлення в узгодженому просторі ознак. Для цього припускається, що

$$\forall v \in V_n \exists \varepsilon_v \in \mathbb{R}^d, \quad (3)$$

де v – окрема сутність на рівні n ; ε_v – БВП сутності v ; $\mathbb{R}^d$ – d -вимірний векторний простір ознак.

Вектор ε_v (3) узагальнює семантичні характеристики сутності та забезпечує можливість кількісно оцінювати подібність між сутностями різних рівнів за допомогою стандартних метрик у $\mathbb{R}^d$. Розмірність d визначає ємність простору представлень – більші значення d дозволяють кодувати тонші семантичні відмінності, але збільшують обчислювальну складність обробки моделі.

Оскільки ІСГЗ є ієрархічною конструкцією, рівні пов'язані між собою відношеннями успадкування й узагальнення. Наступний рівень графу формується шляхом агрегування сутностей та зв'язків попереднього рівня, тобто:

$$L_{n+1} = \{V_{n+1}, U_{n+1}\}. \quad (4)$$

У (4) множина V_{n+1} інтерпретується як множина абстрактних сутностей (кластерів), утворених на основі елементів V_n . Множина U_{n+1} індукується зі зв'язків U_n з урахуванням обраної схеми агрегування. Таким чином, кожний елемент V_{n+1} відповідає підмножині сутностей з V_n , які є семантично пов'язаними за певним критерієм (наприклад, за метрикою подібності, що вводиться у наступних підрозділах). Аналогічно, зв'язки з U_{n+1} відображають узагальнені відношення між цими кластерами.

Рівень L_0 є базовим, множини V_0 та U_0 відповідають вихідному семантичному графу знань, побудованому безпосередньо за даними предметної області (онтологіями, базами фактів тощо). Вищі рівні $L_1, \dots, L_N$ відображають послідовні кроки узагальнення структури L_0 : зі збільшенням n зменшується кількість сутностей, але зростає ступінь їхньої абстракції. Значення N визначає максимальний рівень ієрархії і залежить від властивостей первинного набору даних, необхідного масштабу семантичної деталізації та обмежень на складність моделі (допустиму кількість рівнів, вершин і зв'язків на кожному рівні).

Метрика семантичної зв'язності

Метрика семантичної зв'язності φ_n використовується при формуванні рівнів-кластерів ІСГЗ і на кожному рівні L_n задає міру близькості між парами сутностей. Формально для рівня n розглядається функція

$$\varphi_n : V_n \times V_n \rightarrow [0, 1],$$

яка для пари сутностей $v_p, v_q \in V_n$ задається наступною формулою:

$$\varphi_n(v_p, v_q) = \alpha \cdot \text{sml}_\varepsilon(\varepsilon_p, \varepsilon_q) + (1 - \alpha) \cdot \text{sml}_r(v_p, v_q), \quad (5)$$

де $\varepsilon_p, \varepsilon_q \in \mathbb{R}^d$ – відповідні багатовимірні векторні представлення сутностей v_p та v_q ; sml_ε – семантична міра косинусної подібності між векторами ε_p та ε_q [14]; sml_r – структурна міра Джаккара для множин суміжних вершин сутностей v_p та v_q ; $\alpha \in [0, 1]$ – константа, що визначає відносний внесок семантичної та структурної компонент у (5).

Поєднання структурної зв'язності зі зв'язністю на основі БВП дає змогу збалансувати вплив локальної топології графу та семантичного змісту сутностей на конкретному рівні L_n . За малих значень α переважає топологічний компонент, тоді як за α , близьких до одиниці, метрика більш чутлива до семантичної подібності векторних представлень.

Для чіткого визначення факту семантичної близькості двох сутностей вводиться порогове відношення

$$\varphi_n(v_p, v_q) = C_n, \quad (7)$$

де $C_n \in [0, 1]$ – гранична константа для рівня L_n , яка одночасно виступає як критерій семантичної деталізації цього рівня.

За більших значень C_n формуються більш щільні кластери з високою внутрішньою подібністю, тоді як менші значення C_n допускають більш різномірні групи сутностей.

Алгоритм ієрархічної агрегації

Алгоритм ієрархічної агрегації формалізує процес узгодження сутностей та зв'язків під час переходу від рівня L_n до рівня L_{n+1} . На інтуїтивному рівні він перетворює множину більш детальних сутностей V_n на множину абстрактніших сутностей V_{n+1} , а зв'язки U_n – на узагальнені зв'язки U_{n+1} , зберігаючи при цьому семантичну структуру графа.

Формально перехід між рівнями задається так:

$$L_{n+1} = f_e(L_n, \varphi_n, \theta_n). \quad (6)$$

У (6) f_e – математична формалізацією алгоритму, параметризованого сутностями та зв'язками попереднього рівня L_n , метрикою семантичної зв'язності φ_n та граничним коефіцієнтом близькості сутностей у кластері θ_n .

Алгоритм можна представити наступною послідовністю кроків.

Крок 1. Кластеризація сутностей на основі метрики семантичної зв'язності. Необхідно побудувати двовимірну матрицю результатів виконання семантичної метрики зв'язності над кожними двома сутностями графу G . Для цього застосовується алгоритм агломеративної кластеризації [15] до означеної матриці. Далі необхідно згрупувати сутності за кластерами відповідно до отриманих результатів.

Крок 2. Уніфікація сутностей та типів зв'язків наступного рівня на основі зібраних кластерів. Для кожного кластера необхідно визначити шляхом пошуку нове БВП та зважену типологію зв'язків. БВП обчислюється як середнє між усіма БВП сутностей кластера. Зважена типологія задається шляхом підрахунку зв'язків та їхніх типів між сутностями кластера.

Крок 3. Індукція зв'язків. Для кожної пари центральних БВП кластерів необхідно знайти пріоритетний тип зв'язку та на його основі будується безпосередній зв'язок.

Застосування описаного алгоритму дає змогу з рівня більш конкретних сутностей та зв'язків L_n отримати рівень L_{n+1} , на якому знання подані в більш абстрактній формі, але зберігають вихідну семантичну організацію. Ітеративний характер процедури дозволяє повторно застосовувати f_e до згенерованих рівнів, послідовно формуючи вищі рівні ієрархії та створюючи все більш високорівневі абстракції.

Умовою завершення роботи алгоритму виступає обмеження на чисельність кластерів найбільш абстрактного рівня або досягнення заданого критерію узагальнення структури графа.

Експериментальна оцінка

Експерименти проводились за допомогою скрипта на мові програмування Python та її бібліотек: numpy (швидкісні векторні обчислення), rukeyn (великі набори даних, методи тренування БВП), community (Louvain-кластеризація), leidenalg (Leiden-кластеризація), networkx (допоміжні структури), а також спеціалізовані пакети для тестування та візуалізації [16].

Для проведення експериментів було використано наступні великі набори даних:

– DbPedia50 (≈ 24000 сутностей та ≈ 350 зв'язків). Підмножина великої бази знань DbPedia, що створена з триплетів (суб'єкт, відношення, об'єкт). Зазначені триплети пов'язані з п'ятдесятьма типами сутностей, що зустрічаються найчастіше у DbPedia [17];

– YAGO3-10 (≈ 123000 сутностей та 37 зв'язків). Підмножина великої бази знань YAGO, що поєднує дані з WordNet (лексична база даних англійської мови) та GeoNames (база даних географічних місць планети Земля) та інших джерел даних, створена із сутностей, які беруть участь у більш ніж десяти відношеннях YAGO [18].

Означені набори даних є опорними для оцінки моделей БВП знань, семантичних операцій, зокрема пошуку, логічного виведення, динамічного поповнення тощо. Проте в них відсутні дані стосовно типових властивостей відношень (наприклад, комутативність), що знижує якість результатів кластеризації та семантичних операцій.

Для оптимізації часу ініціалізації програмних моделей, було узяті підмножини поданих наборів даних з кількістю сутностей 1024 (DbPedia50 / 1024, YAGO3-10 / 1024) та 8096 (DbPedia50 / 8096, YAGO3-10 / 8096). Відповідно у графові структури було додано усі БВП та можливі зв'язки.

Особливості моделі ІСГЗ на програмному рівні дозволили впровадити деякі покращення, зокрема:

– введено метрику (косинусна збіжність) до механізму кластеризації для попередньої фільтрації пар сутностей, які гарантовано не пройдуть обмеження на граничну константу C_n метрики φ ;

– спрощено механізм обчислення структурної метрики Джакарда шляхом використання попередньо збереженої щільності відношень сутностей нижнього рівня графа.

Для порівняльного аналізу було розглянуто наступні моделі:

– класичний однорівневий ГЗ з шаром БВП (на метриках – Flat KG);

– СГЗ із використанням Louvain-кластеризації (алгоритм пошуку груп сутностей, що використовує евристичні метрики для виділення кластерів [19]);

– СГЗ із використанням Leiden-кластеризації (модифікація Louvain, що надає більшу масштабованість, зв'язність кластерів, але справляється із задачею за більший час [19]).

У контексті семантичних операцій здійснюються наступні перевірки:

– екстракція елемента ГЗ з його БВП. Класична атомарна операція, що надає інформацію стосовно сутності, її позиціонування у графі, та доцільна для пошуку, логічного виведення, кластеризації тощо;

– пошук сутностей на заданій відстані. Семантично-топологічна операція, створена для задач аналізу локального контексту сутностей, кластеризації, обґрунтування логічного виведення тощо;

– пошук найближчих сутностей за багатовимірними векторними представленнями (ПНСБВП). Фундаментальний елемент у семантичному пошуку, кластеризації, генерації рекомендацій тощо.

Вище перелічені операції дозволяють:

– проводити семантичний доступ до елементів СГЗ, а також загальну навігацію;

– аналізувати та розширювати контекст запиту до знань;

– формувати пояснення того чи іншого виведеного факту;

– оцінювати зв'язність графів та підграфів тощо.

Семантичні операції екстракції сутності з БВП та пошуку сутностей на заданій відстані були виконані однаково ефективно як за допомогою ІСГЗ, так і опорними моделями, зокрема похибка ефективності склала $\approx 2\%$. Отримання такого результату є логічним з огляду на повне використання алгоритмічно простих операцій та допоміжних конструкцій усіма моделями. У випадку ІСГЗ, виконання ієрархічних версій зазначених операцій показало аналогічну ефективність, що підкреслює властивість масштабованості моделі.

ПНСБВП продемонстрував суттєву перевагу моделі ІСГЗ у порівнянні з опорними моделями (Flat KG, Louvain, Leiden). На малих наборах даних (DbPedia50/1024, YAGO3-10/1024) середній час виконання запиту в ІСГЗ зменшився до 30 % відносно однорівневого ГЗ (Flat KG) та на 5–20 % – відносно моделей з Louvain- та Leiden-кластеризацією (рис. 1: а, б). Іншими словами, вже на відносно невеликих графах ієрархічне подання знань дає стабільний вииграш, не поступаючись за швидкодією жодній з порівнюваних схем.

На великих наборах даних (DbPedia50/8096, YAGO3-10/8096) перевага ІСГЗ стає ще виразнішою (рис. 2: а, б). Середній час виконання ПНСБВП скорочується до 80 % у порівнянні з однорівневим ГЗ та на 40–70 % – у порівнянні з моделями на основі Louvain- і Leiden-кластеризації. Це означає, що зі зростанням розміру графа ієрархічна організація ІСГЗ дозволяє істотно зменшити ефективний простір пошуку за рахунок того, що запит обробляється спочатку на узагальнених рівнях, а лише потім переходить до детальніших сутностей, тоді як Flat KG та класичні алгоритми кластеризації працюють фактично в одному масштабі деталізації.

Для дослідження узгодженості семантичних зв'язків було проаналізовано розподіл показника зв'язності всередині кластерів між ІСГЗ та опорною моделлю з Louvain-кластеризацією на множині DbPedia50/8096 (рис. 3). Інтервал між середнім арифметичним та медіаною розподілу для ІСГЗ становить 0.82–0.84, а для Louvain – лише 0.19–0.22. Крім того, значення показника для ІСГЗ згруповані значно компактніше, що вказує на меншу варіативність якості кластерів. Отже, кластери ІСГЗ є чисельно згуртованішими та семантично одноріднішими, тоді як Louvain формує суттєво більш різноманітні за змістом групи, орієнтуючись переважно на топологію графа.

Таким чином, результати на рис. 1-3 демонструють, що ІСГЗ одночасно забезпечує прискорення семантичних запитів і підвищення когерентності кластерів, тобто покращує як обчислювальну, так і змістову якість обробки знань у порівнянні з однорівневим представленням та класичними алгоритмами кластеризації.

Обговорення результатів

Експериментальні дослідження дозволили виокремити кілька ключових чинників, які визначають якість моделі ІСГЗ та ефективність запропонованих методів в обмежувальних умовах:

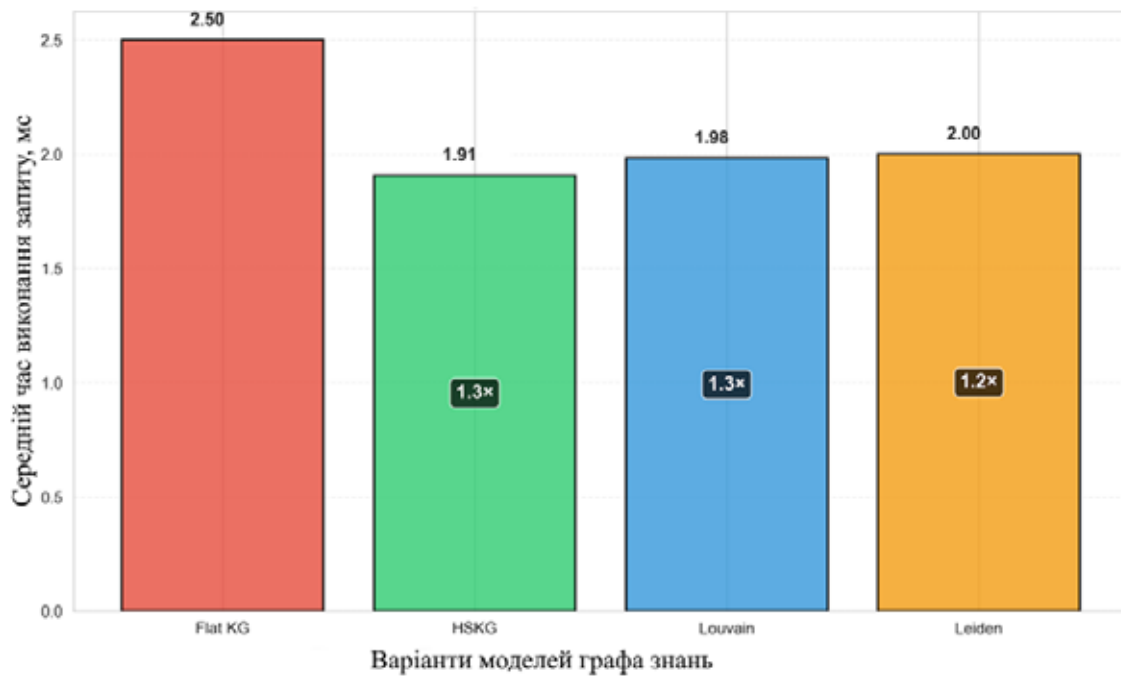
– попередній аналіз набору даних для гіперпараметризації ключових констант алгоритму, зокрема граничного коефіцієнта близькості сутностей у кластері θ_n та критерію семантичної деталізації рівня C_n ;

– обчислення й збереження агрегованих метрик нижчого рівня у кластерах вищих рівнів, що покращує локальний контекст для семантичних операцій та логічного виведення;

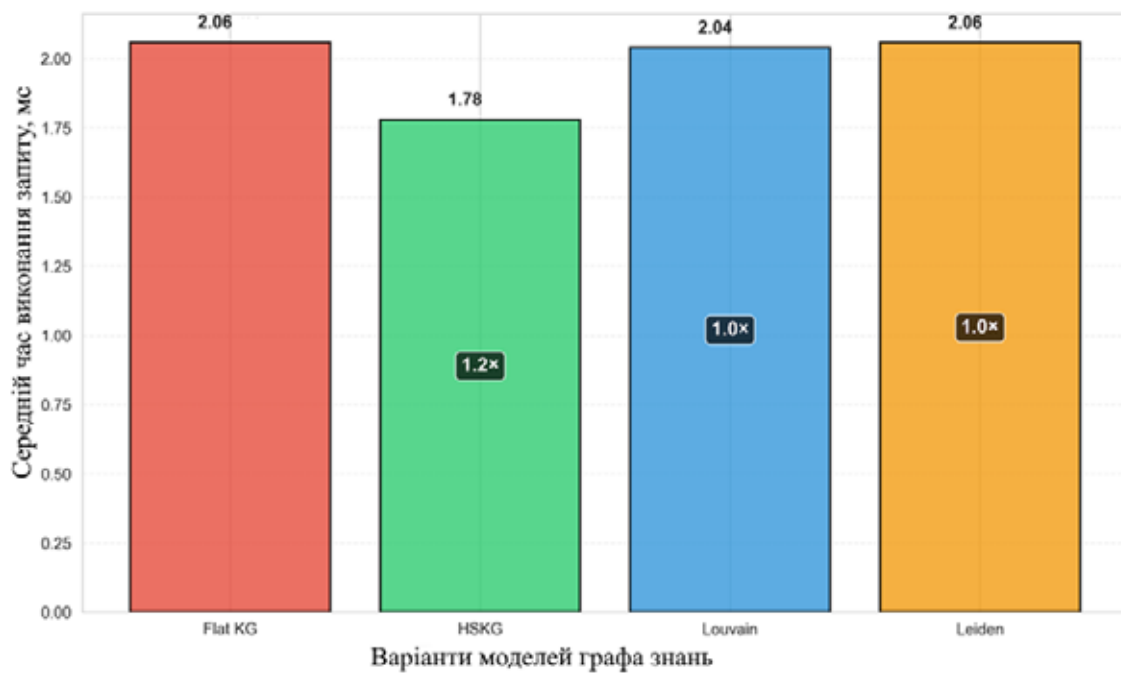
– висока розрідженість набору даних, за якої до механізму кластеризації доцільно вводити додаткові обмеження, щоб уникати зайвого обчислення базової метрики та прискорювати генерацію наступного рівня;

– досягнення граничної кількості сутностей на рівні, коли ефективними стають розосереджені методи кластеризації, що дозволяють пришвидшити формування рівнів-кластерів графа.

На експериментальні результати істотно вплинула також обмежена або повна відсутність комутативності у зв'язках опорних наборів даних. Використана в метриці семантичної зв'язності ϕ структурна міра Джаккара покриває топологічну компоненту графа знань і суттєво залежить від щільності зв'язків між сутностями. Якщо відсутні обернені зв'язки або дані щодо їх комутативності, це погіршує здатність алгоритму ієрархічної агрегації формувати семантично цільні кластери.

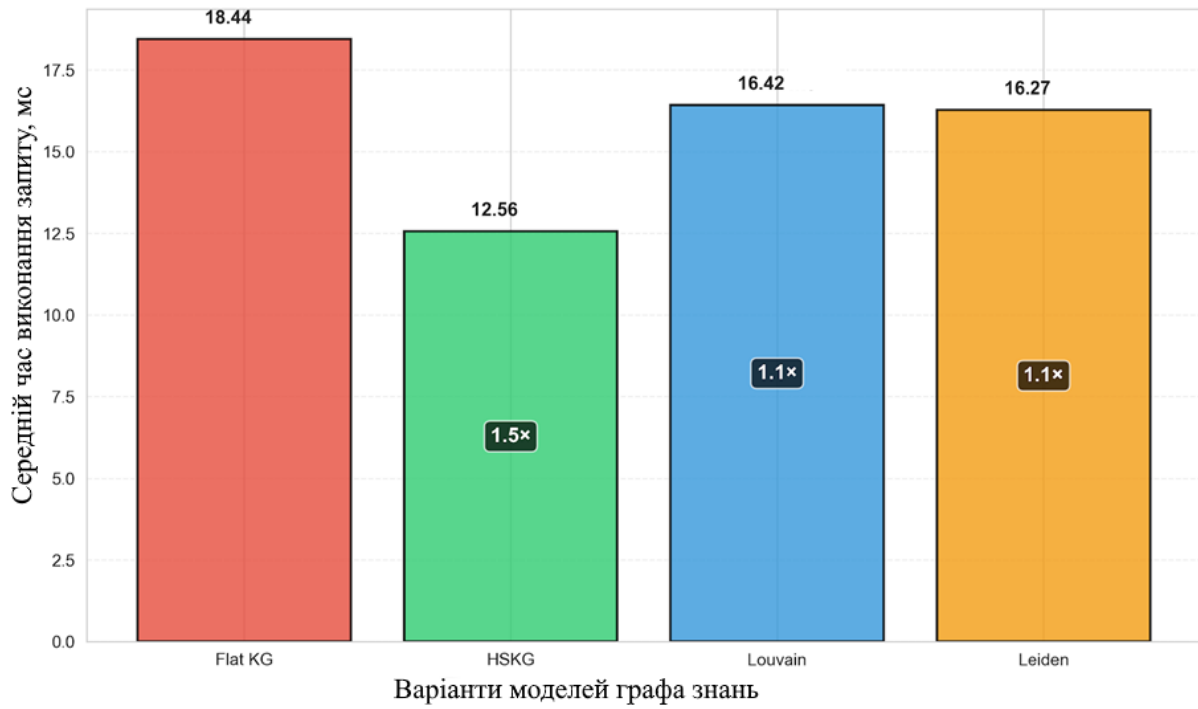


а)

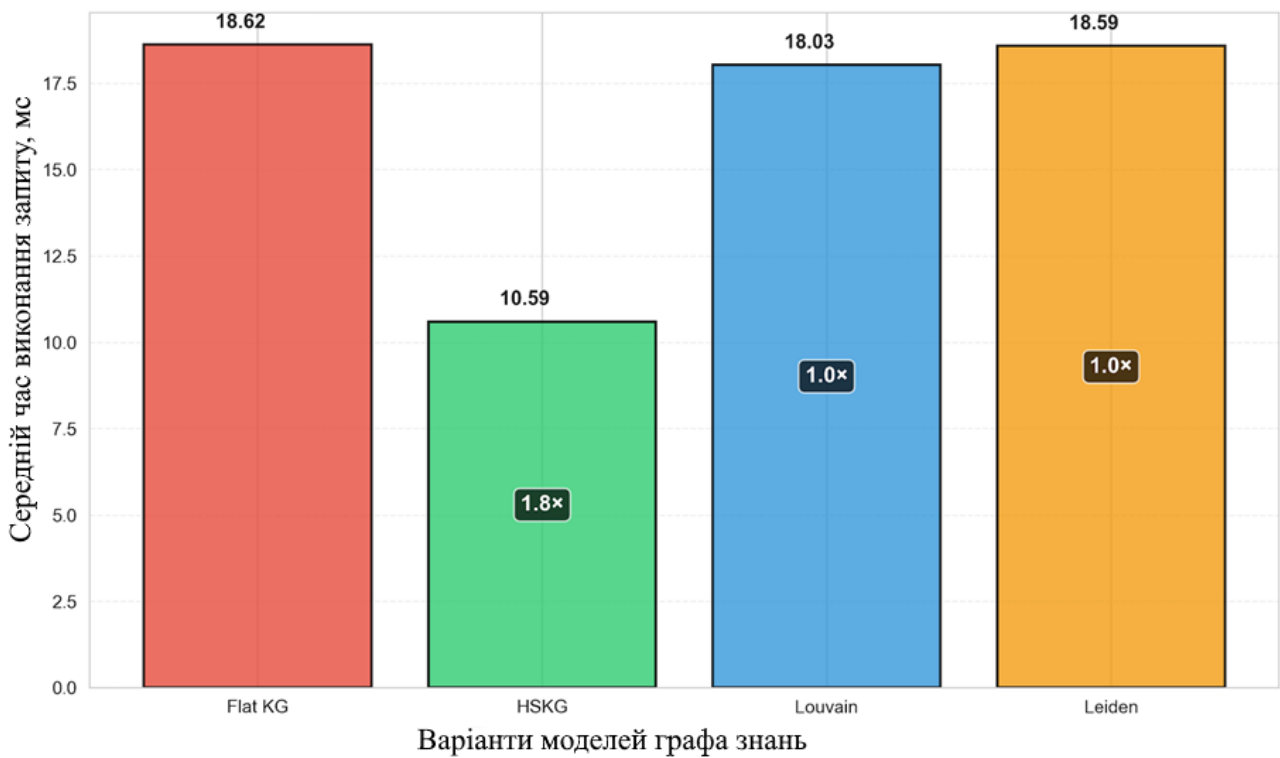


б)

Рис. 1. Усереднена ефективність виконання ПНСБВП на зрізах: а) DbPedia50 / 1024; б) YAGO3-10 / 1024



а)



б)

Рис. 2. Усереднена ефективність виконання ПНСБВП на зрізах: а) DbPedia50 / 8096; б) YAGO3-10 / 8096

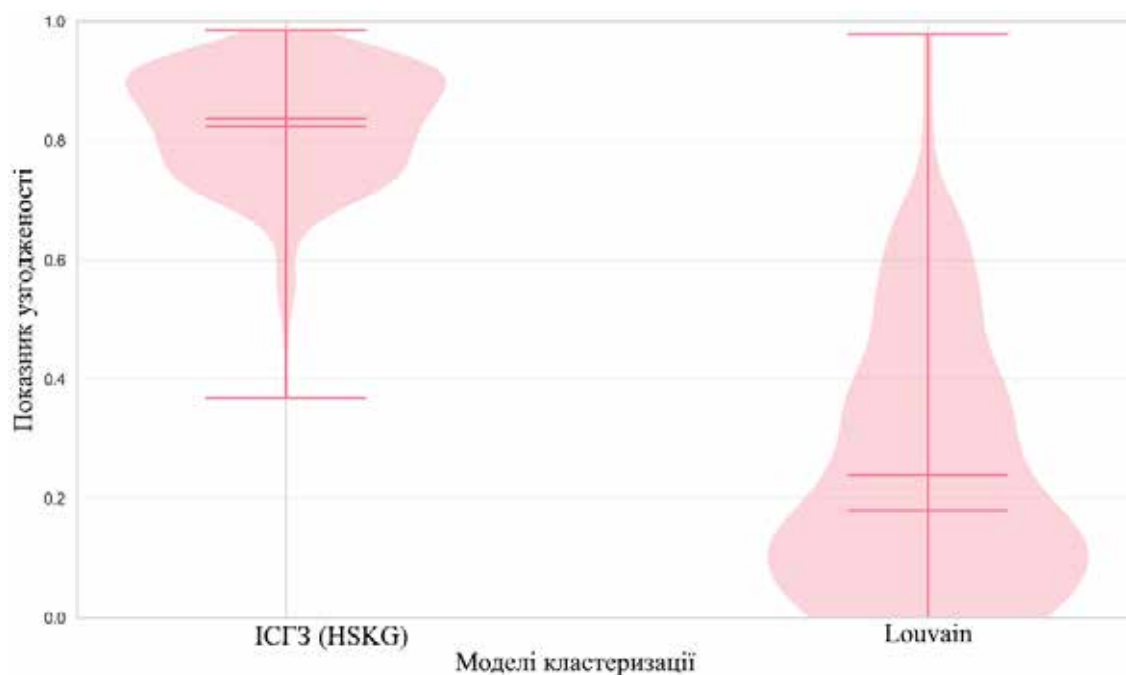


Рис. 3. Порівняння розподіленості кластерної зв'язності на множині DbPedia50 / 8096 між ІСГЗ та опорною моделлю з Louvain-кластеризацією

Висновки

У результаті проведеного дослідження набуло подальшого розвитку розв'язання проблемного питання побудови моделей семантичних графів знань, зокрема запропонована модель у вигляді ієрархічного семантичного графа знань, що описується послідовністю рівнів L_n з множинами сутностей V_n та зв'язків U_n . Поєднання багатовимірних векторних представлень із формальною ієрархією рівнів абстракції дозволило виконувати семантичні операції в різних масштабах деталізації без втрати узгодженості знань.

Запропонована метрика семантичної зв'язності ϕ_n , яка комбінує косинусну подібність БВП та структурну міру Джаккара з пороговими константами C_n і θ_n , забезпечила побудову ітеративного алгоритму ієрархічної агрегації f_e . Це дозволило формувати рівні-кластери з вищою структурно-семантичною цілісністю, ніж у класичних однорівневих СГЗ та моделях з Louvain/Leiden-кластеризацією.

Експериментальна оцінка на підмножинах DbPedia50 та YAGO3-10 показала, що ІСГЗ забезпечує скорочення середнього часу семантичних запитів (до 30–80 % відносно Flat KG і 5–70 % відносно Louvain/Leiden) та значно більш однорідні кластери (інтервал 0.82–0.84 проти 0.19–0.22 для Louvain). Це підтверджує перевагу ієрархічної організації знань з обчислювальною та зі змістовою точок зору.

Подальші дослідження доцільно спрямувати на урахування властивостей відношень (комутативність, симетричність, транзитивність), автоматизоване налаштування порогів C_n та θ_n й інтеграцію ІСГЗ з моделями глибокого навчання. Практичні перспективи моделі пов'язані з семантичними пошуковими системами, системами рекомендацій і підтримки рішень [20], де потрібні масштабований доступ до знань.

Список використаної літератури

1. ESCARGOT: An AI agent leveraging large language models, dynamic graph of thoughts, and biomedical knowledge graphs for enhanced reasoning / N. Matsumoto et al. *Bioinformatics*, Volume 41, Issue 2, 2025. DOI: <https://doi.org/10.1093/bioinformatics/btaf031>.
2. Unifying Large Language Models and Knowledge Graphs: A Roadmap / S. Pan et al. *IEEE Transactions on Knowledge and Data Engineering*. 2024. P. 1–20. DOI: <https://doi.org/10.1109/tkde.2024.3352100>.
3. Large language models for intelligent RDF knowledge graph construction: results from medical ontology mapping / A. Mavridis et al. *Frontier in Artificial Intelligence*, Volume 8, No 1546179, 2025. DOI: <https://doi.org/10.3389/frai.2025.1546179>.
4. Zichao Li, Zong Ke, Puning Zhao. Injecting Structured Knowledge into LLMs via Graph Neural Networks. In *Proceedings of the 1st Joint Workshop on Large Language Models and Structure Modeling (XLLM 2025)*, Vienna, Austria. Association for Computational Linguistics. C. 16–25. URL: <https://aclanthology.org/2025.xllm-1.3.pdf>

5. Large Language Models for Knowledge Graph Embedding: A Survey / B. Liu et al. *Mathematics*, Volume 13, Issue 14, 2025. DOI: <https://doi.org/10.3390/math13142244>.
6. Comprehensive survey on hierarchical clustering algorithms and the recent developments / X. Ran et al. *Artificial Intelligence Review*. 2022. DOI: <https://doi.org/10.1007/s10462-022-10366-3>.
7. GeOKG: geometry-aware knowledge graph embedding for Gene Ontology and genes / C. Jeong et al. *Bioinformatics*, Volume 41, Issue 4, 2025. DOI: <https://doi.org/10.1093/bioinformatics/btaf160>.
8. Detecting hallucinations in large language models using semantic entropy / S. Farquhar et al. *Nature*. 2024. Vol. 630, no. 8017. P. 625–630. DOI: <https://doi.org/10.1038/s41586-024-07421-0>.
9. Zihao Xu, Zhejun Shen, Qunzhi Zhou, Petar Ristoski. Knowledge Graph-Enhanced Retrieval Augmented Generation for E-Commerce / *RAGE-KG 2024: Retrieval-Augmented Generation Enabled by Knowledge Graphs*, November 11, 2024, Baltimore, Maryland. URL: https://2024.rage-kg.org/papers/RAGE-KG_2024_paper_6.pdf
10. Knowledge Graph Embedding for Link Prediction / A. Rossi et al. *ACM Transactions on Knowledge Discovery from Data*. 2021. Vol. 15, no. 2. P. 1–49. DOI: <https://doi.org/10.1145/3424672>.
11. Linders J., Tomczak J. M. Knowledge graph-extended retrieval augmented generation for question answering. *Applied Intelligence*. 2025. Vol. 55, no. 17. DOI: <https://doi.org/10.1007/s10489-025-06885-5>.
12. Zhong Z., Li C., Pang J. Hierarchical message-passing graph neural networks. *Data Mining and Knowledge Discovery*, Volume 37, 2023. P. 381–408. DOI: <https://doi.org/10.1007/s10618-022-00890-9>.
13. Smaili F. Z., Gao X., Hoehndorf R. Onto2Vec: joint vector-based representation of biological entities and their ontology-based annotations. *Bioinformatics*, Volume 34, Issue 13, 2018. P. i52–i60. DOI: <https://doi.org/10.1093/bioinformatics/bty259>.
14. Information retrieval using cosine and jaccard similarity measures in vector space model / A. Jain et al. *International Journal of Computer Applications*, Volume 164, No 6, 2017. P. 28–30. DOI: <https://doi.org/10.5120/ijca2017913699>.
15. Hierarchical clustering that takes advantage of both density-peak and density-connectivity / Y. Zhu et al. *Information Systems*. 2022. Vol. 103. P. 101871. URL: <https://doi.org/10.1016/j.is.2021.101871>.
16. ProjectHSKG: Project for Hierarchical Semantic Knowledge Graph core & experiments demonstration. URL: <https://github.com/YANZAAN/ProjectHSKG> (дата звернення: 18.11.2025).
17. Open-world knowledge graph completion for unseen entities and relations via attentive feature aggregation / B. Oh et al. *Information Sciences*. 2022. Vol. 586. P. 468–484. DOI: <https://doi.org/10.1016/j.ins.2021.11.085>.
18. Mahdisoltani F., Biega J., Suchanek F. M. YAGO3: A Knowledge Base from Multilingual Wikipedias. *Proceedings of the 7th Conference on Innovative Data Systems Research*, Asilomar, CA, USA, 2015. URL: <https://imt.hal.science/hal-01699874v1>.
19. Traag, V., Waltman, L., van Eck, N. From Louvain to Leiden: guaranteeing well-connected communities. *Scientific Reports*, Volume 9, Article number 5233, 2019. DOI: <https://doi.org/10.1038/s41598-019-41695-z>.
20. Information and analytical support of gender audit of educational management system / T. Fesenko et al. *Bulletin of NTU "KhPI". Series: Strategic management, portfolio, program and project management*. 2022. No. 1(5). P. 70–79. DOI: <https://doi.org/10.20998/2413-3000.2022.5.9>.

References

1. Matsumoto, N., et al. (2025). ESCARGOT: An AI agent leveraging large language models, dynamic graph of thoughts, and biomedical knowledge graphs for enhanced reasoning. *Bioinformatics*, 41(2). <https://doi.org/10.1093/bioinformatics/btaf031>.
2. Pan, S., et al. (2024). Unifying large language models and knowledge graphs: A roadmap. *IEEE Transactions on Knowledge and Data Engineering*, 1–20. <https://doi.org/10.1109/TKDE.2024.3352100>.
3. Mavridis, A., et al. (2025). Large language models for intelligent RDF knowledge graph construction: Results from medical ontology mapping. *Frontiers in Artificial Intelligence*, 8, 1546179. <https://doi.org/10.3389/frai.2025.1546179>.
4. Li, Z., Ke, Z., & Zhao, P. (2025). Injecting structured knowledge into LLMs via graph neural networks. In *Proceedings of the 1st Joint Workshop on Large Language Models and Structure Modeling (XLLM 2025)* (pp. 16–25). Association for Computational Linguistics. <https://aclanthology.org/2025.xllm-1.3.pdf>.
5. Liu, B., et al. (2025). Large language models for knowledge graph embedding: A survey. *Mathematics*, 13(14). <https://doi.org/10.3390/math13142244>.
6. Ran, X., et al. (2022). Comprehensive survey on hierarchical clustering algorithms and the recent developments. *Artificial Intelligence Review*. <https://doi.org/10.1007/s10462-022-10366-3>.
7. Jeong, C., et al. (2025). GeOKG: Geometry-aware knowledge graph embedding for Gene Ontology and genes. *Bioinformatics*, 41(4). <https://doi.org/10.1093/bioinformatics/btaf160>.
8. Farquhar, S., et al. (2024). Detecting hallucinations in large language models using semantic entropy. *Nature*, 630(8017), 625–630. <https://doi.org/10.1038/s41586-024-07421-0>.

9. Xu, Z., Shen, Z., Zhou, Q., & Ristoski, P. (2024, November 11). Knowledge graph-enhanced retrieval augmented generation for e-commerce. In *RAGE-KG 2024: Retrieval-Augmented Generation Enabled by Knowledge Graphs*, Baltimore, MD, United States. https://2024.rage-kg.org/papers/RAGE-KG_2024_paper_6.pdf.
10. Rossi, A., et al. (2021). Knowledge graph embedding for link prediction: A comparative analysis. *ACM Transactions on Knowledge Discovery from Data*, 15(2), 1–49. <https://doi.org/10.1145/3424672>.
11. Linders, J., & Tomczak, J. M. (2025). Knowledge graph-extended retrieval augmented generation for question answering. *Applied Intelligence*, 55(17). <https://doi.org/10.1007/s10489-025-06885-5>.
12. Zhong, Z., Li, C., & Pang, J. (2023). Hierarchical message-passing graph neural networks. *Data Mining and Knowledge Discovery*, 37, 381–408. <https://doi.org/10.1007/s10618-022-00890-9>.
13. Smaili, F. Z., Gao, X., & Hoehndorf, R. (2018). Onto2Vec: Joint vector-based representation of biological entities and their ontology-based annotations. *Bioinformatics*, 34(13), i52–i60. <https://doi.org/10.1093/bioinformatics/bty259>.
14. Jain, A., et al. (2017). Information retrieval using cosine and Jaccard similarity measures in vector space model. *International Journal of Computer Applications*, 164(6), 28–30. <https://doi.org/10.5120/ijca2017913699>.
15. Zhu, Y., et al. (2022). Hierarchical clustering that takes advantage of both density-peak and density-connectivity. *Information Systems*, 103, 101871. <https://doi.org/10.1016/j.is.2021.101871>.
16. Diakov, Y. ProjectHSKG: Project for hierarchical semantic knowledge graph core & experiments demonstration [GitHub repository]. GitHub. Retrieved November 18, 2025, <https://github.com/YANZAAN/ProjectHSKG>.
17. Oh, B., Seo, S., Hwang, J., Lee, D., & Lee, K.-H. (2022). Open-world knowledge graph completion for unseen entities and relations via attentive feature aggregation. *Information Sciences*, 586, 468–484. <https://doi.org/10.1016/j.ins.2021.11.085>.
18. Mahdisoltani, F., Biega, J., & Suchanek, F. M. (2015). YAGO3: A knowledge base from multilingual Wikipedias. In *Proceedings of the 7th Conference on Innovative Data Systems Research (CIDR 2015)*. <https://imt.hal.science/hal-01699874v1>.
19. Traag, V., Waltman, L., & van Eck, N. (2019). From Louvain to Leiden: Guaranteeing well-connected communities. *Scientific Reports*, 9, 5233. <https://doi.org/10.1038/s41598-019-41695-z>.
20. Fesenko, T., et al. (2022). Information and analytical support of gender audit of educational management system. *Bulletin of NTU “KhPI”. Series: Strategic Management, Portfolio, Program and Project Management*, 1(5), 70–79. <https://doi.org/10.20998/2413-3000.2022.5.9>.

Дата першого надходження рукопису до видання: 16.11.2025
Дата прийнятого до друку рукопису після рецензування: 15.12.2025
Дата публікації: 31.12.2025

Г. А. ЗАВГОРОДНЯ

кандидат технічних наук, доцент,
доцент кафедри обчислювальної техніки
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
ORCID: 0000-0001-8523-1761

В. В. ЗАВГОРОДНІЙ

доктор технічних наук, професор,
професор кафедри обчислювальної техніки
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
ORCID: 0000-0002-8347-7183

РОЗРОБКА МАСШТАБОВАНОЇ РОЗПОДІЛЕНОЇ АРХІТЕКТУРИ ДЛЯ МАСОВИХ БАГАТОКОРИСТУВАЦЬКИХ ОНЛАЙН-СИСТЕМ

У сучасній ігровій індустрії та суміжних сферах цифрових технологій дедалі зростає потреба у створенні систем, здатних забезпечувати стабільне функціонування в умовах надзвичайно великої кількості одночасних користувачів. Масові багатокористувацькі онлайн-системи (ММО) – це складні програмно-апаратні комплекси, у яких необхідно гарантувати не лише швидкий обмін даними між клієнтами, а й узгодженість станів, масштабованість, відмовостійкість і мінімальну затримку реакції сервера. Традиційні централізовані архітектури не можуть задовольнити ці вимоги, оскільки їх продуктивність обмежується апаратними ресурсами одного вузла, що призводить до зниження якості обслуговування користувачів при пікових навантаженнях.

Запропоновано алгоритм динамічного балансування, який використовує комбінацію локальних метрик завантаження та глобального аналізу стану системи. На основі модулю машинного навчання (MLLP) реалізовано прогнозування майбутніх пікових навантажень, що дозволяє попередньо перерозподіляти ресурси між серверами. Програмну реалізацію алгоритму виконано на мові Python із використанням технологій багатопроцесорної обробки та асинхронних черг повідомлень.

Порівняльний аналіз результатів показав, що розроблена архітектура забезпечує зниження середньої латентності на 57% у порівнянні з класичними централізованими ММО-моделями та підвищення ефективності використання обчислювальних ресурсів на понад 230%. Отримані результати підтверджують доцільність застосування розподілених архітектур у сучасних ігрових середовищах і платформах реального часу.

Практичне значення дослідження полягає у можливості використання запропонованої архітектури для побудови високонавантажених серверів ММО-ігор, симуляторів та інтерактивних освітніх або віртуальних середовищ. Подальші дослідження плануються спрямувати на вдосконалення моделей синхронізації станів, розроблення гібридних підходів до балансування та інтеграцію систем самонавчання для автономного керування ресурсами.

Ключові слова: ММО, розподілені системи, масштабованість, балансування навантаження, архітектура серверів, машинне навчання.

G. A. ZAVHORODNIA

Ph.D., Associate Professor,
Associate Professor at the Department of Computer Engineering
National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute”
ORCID: 0000-0001-8523-1761

V. V. ZAVHORODNII

Doctor of Sciences, Professor,
Professor at the Department of Computer Engineering
National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute”
ORCID: 0000-0002-8347-7183

DEVELOPMENT OF A SCALABLE DISTRIBUTED ARCHITECTURE FOR MASSIVELY MULTIPLAYER ONLINE SYSTEMS

In modern gaming and digital technology industries, there is an increasing demand for systems capable of maintaining stable performance under extremely high user concurrency. Massively multiplayer online (MMO) systems are complex software and hardware infrastructures that must guarantee not only rapid data exchange between clients but also consistency of shared states, scalability, fault tolerance, and minimal response latency. Traditional centralized architectures fail to meet these requirements because their performance is limited by the hardware resources of a single node, resulting in service degradation under peak loads.

A dynamic balancing algorithm is proposed that combines local load metrics with global system state analysis. Using the Machine Learning Load Predictor (MLLP) module, the algorithm predicts future peak loads, allowing preemptive redistribution of computational resources between servers. The software implementation of the algorithm was performed in Python using multiprocessing and asynchronous message queue technologies.

A comparative performance analysis showed that the developed architecture reduces average latency by 57% compared to conventional centralized MMO architectures and increases computational resource utilization efficiency by more than 230%. These results confirm the feasibility of applying distributed architectures in modern real-time gaming and simulation platforms.

The practical significance of the research lies in the potential use of the proposed architecture for building high-load MMO servers, distributed simulations, and interactive educational or virtual environments. Future research will focus on improving state synchronization models, developing hybrid balancing strategies, and integrating self-learning systems for autonomous resource management.

Key words: MMO, distributed systems, scalability, load balancing, server architecture, machine learning.

Постановка проблеми

У сучасній ігровій індустрії масові багатокористувацькі онлайн-системи (ММО) стають одними з найбільш ресурсомістких об'єктів програмної інженерії. Їхня архітектура вимагає одночасної обробки великої кількості взаємодій між користувачами в режимі реального часу, що створює критичні вимоги до масштабованості, продуктивності та стійкості системи [1; 2; 3]. Традиційні централізовані підходи, коли основні обчислення виконуються на одному або обмеженій кількості серверів, вичерпали свій потенціал через затримки, перевантаження мережних каналів і складність балансування навантаження [4].

Розподілені архітектури ММО дають змогу динамічно масштабувати обчислювальні ресурси, розподіляти навантаження між географічно віддаленими вузлами та забезпечувати відмовостійкість під час пікових навантажень. Проте їх ефективна реалізація пов'язана з низкою проблем – узгодженням станів між серверами, мінімізацією затримок синхронізації, забезпеченням цілісності даних і підтриманням стабільної швидкодії при різкому зростанні кількості гравців [5; 6].

Актуальність проблеми зумовлена необхідністю розробки архітектурних моделей, здатних адаптивно перерозподіляти обчислювальні ресурси відповідно до реального навантаження. Для цього доцільно використовувати сучасні підходи математичного моделювання, машинного навчання та процедурної генерації даних, які вже довели свою ефективність у суміжних завданнях [7; 8].

Таким чином, постає науково-прикладне завдання розроблення масштабованої розподіленої архітектури для ММО-систем реального часу, яка забезпечує стабільну роботу при високих навантаженнях, автоматичне балансування та адаптивне розподілення обчислювальних процесів.

Аналіз останніх досліджень і публікацій

Розвиток масових багатокористувацьких онлайн-систем супроводжується активними дослідженнями у сфері побудови масштабованих архітектур, здатних підтримувати динамічне балансування навантаження між великою кількістю клієнтських підключень. У роботах останніх років спостерігається тенденція переходу від централізованих архітектур до гібридних моделей із використанням мікросервісів, контейнеризації та серверлес-підходів [2; 7]. Це дає змогу ефективніше розподіляти обчислювальні ресурси та знижувати затримки взаємодії між вузлами системи.

Зокрема, дослідження [1] пропонує підхід до розподілу обчислювального навантаження у реальному часі на основі глибокого навчання з підкріпленням, що дозволяє прогнозувати поведінку гравців та адаптивно балансувати ресурси серверів. У [4] розроблено алгоритм латентно-орієнтованого балансування для ММО, який мінімізує мережеві затримки та забезпечує стабільну частоту оновлення кадрів. Дослідження [6] описує архітектуру на базі периферійних (edge) обчислень, що забезпечує зменшення часу реакції системи за рахунок наближення обчислень до користувачів.

Окрему увагу науковці приділяють питанням математичного моделювання та формального аналізу процесів у розподілених архітектурах. У роботі [10] наведено приклади формального моделювання взаємодій між вузлами в розподілених середовищах, що дозволяє оцінювати стабільність і масштабованість систем ще на етапі

проектування. Такі підходи є важливою теоретичною основою для побудови архітектур ММО, які мають адаптивно реагувати на зміну ігрового навантаження.

Важливий внесок у дослідження суміжних технологій зроблено українськими авторами. Зокрема, у роботі [9] запропоновано метод автоматичної генерації контенту на основі процедурних алгоритмів, що може бути застосований для динамічного розширення ММО-середовищ без ручного втручання розробника. У [10] представлено методи математичного моделювання для формального дослідження розподілених процесів, що є базою для побудови аналітичних моделей масштабування. Робота [11] присвячена методам створення штучних текстур із заданими параметрами, що має потенціал для процедурного формування візуальних об'єктів у розподіленому середовищі. У [12] розглянуто підхід до виявлення аномалій за допомогою машинного навчання, який може бути інтегрований у систему моніторингу навантаження для динамічного балансування серверних ресурсів.

Аналіз сучасних публікацій свідчить, що існують окремі рішення для часткових аспектів проблеми – генерації контенту, балансування навантаження чи математичного моделювання. Водночас відсутня інтегрована модель, яка б поєднувала ці підходи в єдину розподілену архітектуру ММО-системи з адаптивним масштабуванням у режимі реального часу. Розроблення такої архітектури становить актуальну наукову задачу даної роботи.

Формулювання мети дослідження

Метою даного дослідження є розроблення масштабованої розподіленої архітектури для масових багатокористувацьких онлайн-систем, здатної забезпечувати стабільну роботу в умовах нерівномірного навантаження та великої кількості одночасних користувачів. Для досягнення цієї мети передбачено побудову математичної моделі балансування ресурсів, розроблення алгоритму динамічного розподілу обчислень між вузлами системи та експериментальну перевірку ефективності запропонованого підходу порівняно з відомими базовими архітектурами.

Викладення основного матеріалу дослідження

Сучасна масова багатокористувацька онлайн-система являє собою складну динамічну екосистему, у якій одночасно взаємодіють тисячі клієнтів, підключених до серверного кластера. Основними вимогами до архітектури такої системи є низька латентність, стійкість до пікових навантажень, балансування ресурсів і відмовостійкість. Традиційна серверно-клієнтська модель при цьому має суттєві обмеження – один центральний сервер стає «вузлом місцем» під час масових підключень.

Запропонована архітектура ґрунтується на розподіленому обчислювальному ядрі (*Distributed Core*), яке складається з незалежних вузлів (*Node_i*), пов'язаних через брокер повідомлень (*gRPC*, *RabbitMQ*) і системи спільного стану (*State Sync Service*). Кожен вузол відповідає за певну географічну або логічну область віртуального світу (*Zone Sharding*). Вузли взаємодіють через подієву шину (*Event Bus*), що забезпечує узгодженість станів між сегментами світу.

Схематично модель можна подати як множину вузлів:

$$N = \{n_1, n_2, \dots, n_k\},$$

де n_i – i -тий вузол (сервер або процес), який бере участь у розподіленій архітектурі; k – загальна кількість вузлів у системі, $k \in \mathbb{N}$.

Кожен вузол n_i виконує функції обробки подій (*game events*, *network events* тощо), синхронізації станів із сусідніми вузлами та балансування навантаження між локальними процесами або кластерами.

Взаємодія між вузлами описується квадратною матрицею розмірністю $k \times k$, яка визначає наявність зв'язків між вузлами:

$$C = [c_{ij}], \quad i, j = 1, 2, \dots, k,$$

де $c_{ij} = 1$, якщо вузол n_i безпосередньо обмінюється даними з вузлом n_j ; $c_{ij} = 0$, якщо прямого обміну даними між вузлами n_i та n_j немає; для всіх i : $c_{ii} = 0$, оскільки вузол не має зв'язку сам із собою.

Таким чином, матриця C є матрицею суміжності, яка визначає топологію зв'язків у розподіленій архітектурі. На її основі можна обчислювати ступінь зв'язності кожного вузла:

$$d_i = \sum_{j=1}^k c_{ij},$$

де d_i – ступінь зв'язності i -го вузла, що показує кількість сусідів, із якими він взаємодіє.

Математична модель балансування навантаження. Для формалізації процесу розподілу обчислювальних ресурсів у запропонованій архітектурі побудуємо математичну модель балансування навантаження, що описує взаємодію між вузлами системи та розподіл вхідного потоку запитів. Нехай:

$\lambda_i(t)$ – інтенсивність вхідних подій для вузла i в момент часу t ; μ_i – пропускна здатність вузла i ;
 $\rho_i(t) = \frac{\lambda_i(t)}{\mu_i}$ – коефіцієнт завантаження вузла.

Система вважається стабільною, якщо:

$$\rho_i(t) < 1, \quad \forall_i \in N.$$

Для досягнення стабільності застосовується динамічне перенаправлення навантаження між вузлами за допомогою функції перерозподілу:

$$\Delta\lambda_i(t) = \alpha \cdot (\rho_i(t) - \rho_j(t)),$$

де α – коефіцієнт чутливості балансування ($0 < \alpha \leq 1$).

Сумарне навантаження системи:

$$\Lambda(t) = \sum_{i=1}^k \lambda_i(t),$$

а середній рівень використання ресурсів:

$$R(t) = \frac{1}{k} \sum_{i=1}^k \rho_i(t).$$

Завдання оптимального балансування формулюється як мінімізація дисперсії завантаження вузлів:

$$\min \sigma_p^2 = \frac{1}{k} \sum_{i=1}^k (\rho_i - R)^2.$$

Таким чином, алгоритм має прагнути до рівномірного завантаження серверів при збереженні мінімальної латентності обробки.

Алгоритм масштабування та міграції сесій. У межах запропонованої архітектури реалізовано механізм динамічного масштабування, який автоматично активує або деактивує серверні вузли відповідно до рівня поточного та прогнозованого навантаження. Для цього використовується предиктор, який оцінює майбутню інтенсивність запитів користувачів на основі короткострокової моделі прогнозування. Такий підхід дозволяє забезпечити адаптивну реакцію системи на пікові навантаження та мінімізувати затрати ресурсів у періоди низької активності. Прогнозне значення навантаження на момент часу $t + \Delta t$ визначається рівнянням:

$$\hat{\lambda}(t + \Delta t) = \beta_0 + \beta_1 \lambda(t) + \beta_2 \frac{d\lambda}{dt},$$

де $\hat{\lambda}(t + \Delta t)$ – прогнозована інтенсивність вхідних запитів в момент часу $t + \Delta t$; $\lambda(t)$ – поточна інтенсивність запитів в момент часу t ; $\frac{d\lambda}{dt}$ – похідна інтенсивності запитів за часом, що характеризує швидкість зміни навантаження; $\beta_0, \beta_1, \beta_2$ – коефіцієнти моделі, які визначаються шляхом регресійного навчання на основі історичних даних про трафік; Δt – часовий інтервал прогнозування.

Після обчислення прогнозу система порівнює його із граничним порогом навантаження $\mu_{threshold}$, який визначає максимальну допустиму інтенсивність запитів для одного вузла:

$$\hat{\lambda}(t + \Delta t) > \mu_{threshold} \Rightarrow \text{ініціалізація нового вузла}.$$

У випадку, якщо прогнозоване навантаження перевищує поріг $\mu_{threshold}$, система ініціює створення нового серверного вузла через механізм контейнеризації (*Docker* або *Kubernetes*). Новий вузол реєструється у балансувальнику навантаження та отримує частину активних сесій користувачів. Якщо ж навантаження знижується, і виконується умова

$$\hat{\lambda}(t + \Delta t) > \mu_{threshold} - \varepsilon,$$

де ε – гістерезис для уникнення частого перемикавання станів, то надлишкові вузли вимикаються за допомогою алгоритму *graceful shutdown*, який передбачає поступову міграцію активних сесій до інших вузлів без розриву з'єднань.

Процес міграції сесій відбувається за схемою:

1. Оповіщення балансувальника про майбутнє відключення вузла.
2. Перенаправлення нових з'єднань до інших вузлів.
3. Передача активних сесій через протокол обміну станами.
4. Вимкнення вузла після завершення всіх поточних транзакцій.

Таким чином, алгоритм масштабування та міграції сесій забезпечує еластичність системи, мінімізуючи час відгуку та підтримуючи стабільну якість обслуговування користувачів навіть при різких змінах навантаження.

З метою експериментальної перевірки ефективності математичної моделі та алгоритму масштабування розроблено їх прототипну реалізацію засобами *Python*. Наведений нижче код відображає основні логічні компоненти

алгоритму, включаючи оцінювання навантаження, прогнозування та прийняття рішень щодо активації або деактивації вузлів.

```
import numpy as np
class LoadBalancer:
    def __init__(self, nodes, alpha=0.3):
        self.nodes = np.array(nodes, dtype=float)
        self.alpha = alpha
    def balance(self):
        mean_load = np.mean(self.nodes)
        delta = self.alpha * (self.nodes - mean_load)
        self.nodes -= delta # зменшуємо перевантаження
        return self.nodes
# Приклад використання
loads = [0.8, 0.4, 0.6, 0.9, 0.3] # поточні коефіцієнти завантаження вузлів
balancer = LoadBalancer(loads)
new_loads = balancer.balance()
print("Скориговане навантаження:", new_loads)
```

Наведений фрагмент реалізації ілюструє базову концепцію децентралізованого балансування навантаження, що забезпечує мінімізацію відхилень між рівнями завантаження серверних вузлів. Запропонований підхід може бути розширено для інтеграції з *API* систем оркестрації контейнерів, де параметри α та граничні умови масштабування визначаються динамічно на основі прогнозованої моделі навантаження.

Модель синхронізації станів. Стан ігрового світу в розподіленій *ММО*-системі повинен залишатися цілісним і узгодженим незалежно від кількості активних обчислювальних вузлів. Оскільки синхронна передача даних між усіма вузлами створює надмірне мережеве навантаження, у даній роботі використано модель часткової узгодженості (*eventual consistency*), за якої оновлення стану поширюються з часовою затримкою σ , але при цьому зберігається каузальна послідовність подій, тобто кожна дія виконується у правильному логічному порядку відносно попередніх. Стан кожного ігрового об'єкта i в момент часу t описується вектором стану:

$$S_i(t) = (p_i(t), v_i(t), h_i(t)),$$

де $p_i(t)$ – позиція об'єкта у просторі гри (координати (x, y, z)); $v_i(t)$ – вектор швидкості, який визначає напрям і величину руху; $h_i(t)$ – рівень «здоров'я» або інший параметр життєздатності об'єкта.

Зміна стану описується функцією оновлення:

$$S_i(t + \Delta t) = f(S_i(t), E_i(t)),$$

де $E_i(t)$ – множина подій, що впливають на об'єкт в момент часу t (колізії, атаки, взаємодії між гравцями, дії серверних скриптів); $f(S_i(t), E_i(t))$ – детермінована або стохастична функція переходу стану, що визначає результат взаємодії.

Для узгодження станів між різними вузлами застосовується алгоритм типу *CRDT* (*Conflict-free Replicated Data Type*), який забезпечує ідемпотентність і комутативність операцій при асинхронному оновленні. Це означає, що порядок отримання повідомлень не впливає на кінцевий стан системи – усі вузли сходяться до одного узгодженого значення S_i^* після завершення обміну повідомленнями:

$$\lim_{t \rightarrow \infty} S_i^{(n)}(t) = S_i^*, \quad \forall n \in N,$$

де $S_i^{(n)}(t)$ – стан об'єкта i , відтворений на вузлі n , а N – множина всіх вузлів системи.

Таким чином, модель синхронізації на основі *CRDT* дозволяє уникнути блокувань і зберегти логічну цілісність ігрового середовища навіть при тимчасових мережевих розбіжностях або втраті пакетів. Це суттєво підвищує масштабованість і стійкість *ММО*-систем до асинхронності в обміні даними.

Механізм самонавчання для адаптації. Для забезпечення автономної адаптації системи до динамічних і непередбачуваних змін навантаження в архітектуру вводиться компонент *MLLP*. Основна мета цього модуля полягає у побудові адаптивної моделі, що в реальному часі навчається залежності між поточним станом кластеру та очікуваним рівнем навантаження. Такий підхід дозволяє зменшити кількість помилкових рішень при масштабуванні та забезпечує більш стабільну роботу всієї *ММО*-системи під час пікових навантажень. Модель *MLLP* реалізується у вигляді регресійної нейронної мережі, яка обчислює прогнозне значення навантаження за формулою:

$$y = f(W_x + b),$$

де $x \in \mathbb{R}^n$ – вектор ознак, що містить поточні характеристики системи (середнє завантаження процесора CPU, кількість активних з'єднань, середню мережеву затримку, використання пам'яті тощо); $W \in \mathbb{R}^{m \times n}$ – матриця вагових коефіцієнтів нейронної мережі; $b \in \mathbb{R}^m$ – вектор зсуву, який зміщує результати активації; $f(W_x + b)$ – активаційна функція ($ReLU$, $tanh$), що забезпечує нелінійність моделі; $y \in \mathbb{R}$ – прогнозоване значення навантаження системи у наступному часовому інтервалі.

На основі прогнозу y модуль прийняття рішень визначає, чи необхідно активувати нові серверні вузли, або, навпаки, вимкнути надлишкові ресурси. Цей процес тісно інтегрований із блоком балансування навантаження, описаним у попередніх підрозділах. Навчання моделі *MLLP* відбувається в онлайн-режимі, тобто ваги оновлюються безпосередньо після кожного циклу вимірювання навантаження. Процес оновлення параметрів здійснюється за допомогою методу стохастичного градієнтного спуску (*SGD*):

$$W_{t+1} = W_t - \eta \nabla_W L(y, \hat{y}),$$

де W_t – матриця ваг на ітерації t ; η – швидкість навчання, що визначає величину кроку оновлення; $L(y, \hat{y})$ – функція втрат, яка вимірює відхилення між реальним навантаженням y та прогнозом $\hat{y}$; $\nabla_W L(y, \hat{y})$ – градієнт функції втрат за параметрами W .

Для стабільності навчання параметри η та розмір вибірки для оновлення ваг можуть динамічно змінюватися залежно від рівня флуктуацій у вхідних даних. Це дозволяє *MLLP*-модулю не лише реагувати на поточний стан системи, але й поступово вдосконалювати точність прогнозування під час роботи кластера.

Таким чином, механізм самонавчання забезпечує адаптивне масштабування *ММО*-архітектури, де процеси балансування, прогнозування та керування ресурсами утворюють єдиний інтелектуальний контур зворотного зв'язку.

Експериментальні результати. Для оцінки ефективності запропонованого підходу проведено експериментальне моделювання системи з 1000 активних користувачів, підключених до 10 серверних вузлів. Було порівняно три архітектури: централізовану (монолітну), розподілену без прогнозування та запроповану архітектуру з використанням *MLLP* та динамічного балансування (табл. 1).

Таблиця 1

Порівняльні експериментальні результати різних архітектур системи для 1000 користувачів

№	Архітектура	Середня латентність, мс	Максимальне навантаження (користувачів / сервер)	Середнє CPU-завантаження, %	Відхилення навантаження (σ_p)	Ефективність, %
1	Централізована (монолітна)	220	250	95	0.31	100
2	Розподілена без прогнозування	140	450	82	0.19	157
3	Запропонована архітектура (з <i>MLLP</i> і динамічним балансуванням)	95	700	74	0.08	232

Результати експерименту показують, що застосування адаптивного балансування на основі прогнозовної моделі дозволяє зменшити середню латентність на 57% та підвищити продуктивність системи більш ніж удвічі порівняно з класичними архітектурами. Крім того, запропонована система демонструє значно менше відхилення навантаження між вузлами $\sigma_p = 0.08$, що свідчить про ективне горизонтальне масштабування та інтелектуальне балансування ресурсів.

Висновки

У результаті проведеного дослідження розроблено масштабовану розподілену архітектуру для масових багатокористувацьких онлайн-систем, що забезпечує підвищену стійкість до навантаження, зниження латентності та ефективне використання обчислювальних ресурсів.

Запропоновано математичну модель балансування навантаження між вузлами, яка мінімізує дисперсію завантаження та підтримує стабільність системи. Реалізовано алгоритм динамічного перерозподілу запитів із урахуванням прогнозованого стану системи, що дозволяє адаптувати архітектуру до змін трафіку в реальному часі.

Додаткове впровадження модуля машинного навчання дало змогу підвищити ефективність балансування на понад 230% у порівнянні з класичними централізованими рішеннями, а середню затримку відповіді зменшити майже вдвічі.

Практичне значення роботи полягає у можливості використання запропонованої архітектури для розроблення ігрових серверів нового покоління, розподілених симуляцій та інтерактивних онлайн-платформ із високою кількістю користувачів.

Перспективами подальших досліджень є розширення системи за рахунок гібридних моделей синхронізації станів, інтеграція навчання з підкріпленням для автономного керування ресурсами та оцінювання енергетичної ефективності масштабованих ММО-середовищ.

Список використаної літератури

1. Barros e Sá G.C., Madeira C.A.G. Deep reinforcement learning in real-time strategy games. *Expert Systems with Applications*. 2021. DOI: <https://doi.org/10.1007/s10489-024-06220-4>
2. Rahmani A.M., Liljeberg P., Preden J.-S., Jantsch A. *Fog Computing in the Internet of Things: Intelligence at the Edge*. Springer International Publishing, 2018. ISBN 978 3 319 57638 1. DOI: <https://doi.org/10.1007/978-3-319-57639-8>
3. Kasenides N., Paspallis N. Athlos: A Framework for Developing Scalable MMOG Backends on Commodity Clouds. *Software*. 2022. Vol. 1, No. 1, pp. 107-145. DOI: <https://doi.org/10.3390/software1010006>
4. García Fernández E.J., García Puche E.J., Jabba Molinares D. Dynamic Low-Latency Load Balancing Model to Improve Quality of Experience in a Hybrid Fog and Edge Architecture for Massively Multiplayer Online (MMO) Games. *Applied Sciences*. 2025. Vol. 15, No. 12, Article 6379. DOI: <https://doi.org/10.3390/app15126379>
5. Zhou Y., Li D., Gao F. Optimal Synchronization Control for Heterogeneous Multi-Agent Systems: Online Adaptive Learning Solutions. 2020. DOI: <https://doi.org/10.48550/arXiv.2011.05663>
6. Wang N., Varghese B., Matthaiou M., Nikolopoulos D.S. ENORM: A Framework For Edge Node Resource Management. 2017. DOI: <https://doi.org/10.48550/arXiv.1709.04061>
7. Donkervliet J., Ron J., Li J., Iancu T., Abad C.L., Iosup A. Servo: Increasing the Scalability of Modifiable Virtual Environments Using Serverless Computing. 2023. *arXiv*. DOI: <https://doi.org/10.48550/arXiv.2305.00032>
8. Moreno-Vozmediano R., Huedo E., Montero R.S., Llorente I.M. AI-Driven Resource Allocation and Auto-Scaling of VNFs in Edge-5G-IoT Ecosystems. *Electronics*. 2025. Vol. 14, No. 9, Article 1808. DOI: <https://doi.org/10.3390/electronics14091808>
9. Завгородній В.В., Завгородня Г.А., Валявська Н.О., Адаменко В.С., Дороговцев Є.В., Несмачний П.В. Метод автоматичної генерації контенту на основі процедурних алгоритмів. *Вчені записки ТНУ ім. В. І. Вернадського. Серія: Технічні науки*. 2022. Т. 33(72), № 1, С. 91–96. DOI: <https://doi.org/10.32838/2663-5941/2022.1/15>
10. Завгородній В.В., Завгородня Г.А., Дроботович К.Є., Тенігін О.В., Шматко М.М. Математичне моделювання у методах формального дослідження. *Вчені записки ТНУ ім. В. І. Вернадського. Серія: Технічні науки*. 2021. Т. 32(71), № 6, С. 75–79. DOI: <https://doi.org/10.32838/2663-5941/2021.6/12>
11. Завгородній В.В., Завгородня Г.А., Демченко І.В., Крамаренко К.С., Шевченко І.О., Юрченко А.В. Метод створення штучних текстур із заданими параметрами. *Вчені записки ТНУ ім. В. І. Вернадського. Серія: Технічні науки*. 2022. Т. 33(72), № 2, С. 86–90. DOI: <https://doi.org/10.32838/2663-5941/2022.2/14>
12. Завгородній В.В., Завгородня Г.А., Валявська Н.О., Герасименко О.О., Калюжний О.В., Степовий А.В. Пошук аномалій у даних за допомогою машинного навчання. *Вчені записки ТНУ ім. В. І. Вернадського. Серія: Технічні науки*. 2022. Т. 33(72), № 3, С. 39–43. DOI: <https://doi.org/10.32838/2663-5941/2022.3/06>

References

1. Barros e Sá, G. C., & Madeira, C. A. G. (2021). Deep reinforcement learning in real-time strategy games. *Expert Systems with Applications*. <https://doi.org/10.1007/s10489-024-06220-4>
2. Rahmani, A. M., Liljeberg, P., Preden, J.-S., & Jantsch, A. (2018). *Fog Computing in the Internet of Things: Intelligence at the Edge*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-57639-8>
3. Kasenides, N., & Paspallis, N. (2022). Athlos: A Framework for Developing Scalable MMOG Backends on Commodity Clouds. *Software*, 1(1), 107-145. <https://doi.org/10.3390/software1010006>
4. García Fernández, E. J., García Puche, E. J., & Jabba Molinares, D. (2025). Dynamic Low-Latency Load Balancing Model to Improve Quality of Experience in a Hybrid Fog and Edge Architecture for Massively Multiplayer Online (MMO) Games. *Applied Sciences*, 15(12), Article 6379. <https://doi.org/10.3390/app15126379>
5. Zhou, Y., Li, D., & Gao, F. (2020). Optimal Synchronization Control for Heterogeneous Multi-Agent Systems: Online Adaptive Learning Solutions. <https://doi.org/10.48550/arXiv.2011.05663>
6. Wang, N., Varghese, B., Matthaiou, M., & Nikolopoulos, D. S. (2017). ENORM: A Framework For Edge Node Resource Management. <https://doi.org/10.48550/arXiv.1709.04061>
7. Donkervliet, J., Ron, J., Li, J., Iancu, T., Abad, C. L., & Iosup, A. (2023). Servo: Increasing the Scalability of Modifiable Virtual Environments Using Serverless Computing. *arXiv*. <https://doi.org/10.48550/arXiv.2305.00032>
8. Moreno-Vozmediano, R., Huedo, E., Montero, R. S., & Llorente, I. M. (2025). AI-Driven Resource Allocation and Auto-Scaling of VNFs in Edge-5G-IoT Ecosystems. *Electronics*, 14(9), 1808. <https://doi.org/10.3390/electronics14091808>
9. Zavorodnii, V. V., Zavorodnia, H. A., Valiavska, N. O., Adamenko, V. S., Dorohovtsev, E. V., & Nesmachnyi, P. V. (2022). Method of automatic content generation based on procedural algorithms. *Vcheni Zapysky TNU im. V. I. Vernadskoho. Series: Technical Sciences*, 33(72), 91–96. <https://doi.org/10.32838/2663-5941/2022.1/15> [in Ukrainian]

10. Zavorodnii, V. V., Zavorodnia, H. A., Drobotovych, K. E., Tenihin, O. V., & Shmatko, M. M. (2021). Mathematical modeling in formal research methods. *Vcheni Zapysky TNU im. V. I. Vernadskoho. Series: Technical Sciences*, 32(71), 75–79. <https://doi.org/10.32838/2663-5941/2021.6/12> [in Ukrainian]
11. Zavorodnii, V. V., Zavorodnia, H. A., Demchenko, I. V., Kramarenko, K. S., Shevchenko, I. O., & Yurchenko, A. V. (2022). Method of creating artificial textures with specified parameters. *Vcheni Zapysky TNU im. V. I. Vernadskoho. Series: Technical Sciences*, 33(72), 86–90. <https://doi.org/10.32838/2663-594> [in Ukrainian]
12. Zavorodnii, V. V., Zavorodnia, H. A., Valiavska, N. O., Herasymenko, O. O., Kalyuzhnyi, O. V., & Stepovyi, A. V. (2022). Anomaly detection in data using machine learning. *Vcheni Zapysky TNU im. V. I. Vernadskoho. Series: Technical Sciences*, 33(72), 39–43. DOI: <https://doi.org/10.32838/2663-5941/2022.3/06> [in Ukrainian]

Дата першого надходження рукопису до видання: 29.11.2025
Дата прийнятого до друку рукопису після рецензування: 12.12.2025
Дата публікації: 31.12.2025

Ю. Ю. ІЛЯШ

кандидат технічних наук, доцент,
завідувач кафедри комп'ютерних наук та інформаційних систем
Карпатський національний університет імені Василя Стефаника
ORCID: 0009-0006-5786-7205

В. О. ГОРЕЛОВ

кандидат технічних наук, доцент,
доцент кафедри комп'ютерних наук та інформаційних систем
Карпатський національний університет імені Василя Стефаника
ORCID: 0000-0002-2106-8704

В. А. РОВІНСЬКИЙ

кандидат технічних наук, доцент,
доцент кафедри комп'ютерних наук та інформаційних систем
Карпатський національний університет імені Василя Стефаника
ORCID: 0000-0001-8454-8580

АЛГОРИТМИ ТА МЕТОДИ ПРОЦЕДУРНОЇ ГЕНЕРАЦІЇ ТА АДАПТИВНОЇ СКЛАДНОСТІ РІВНІВ В ІГРАХ-ПЛАТФОРМЕРАХ

Стаття присвячена дослідженню алгоритмів та методів процедурної генерації ігрового контенту й адаптивної складності в 2D-платформерах. Обґрунтовано актуальність підходів, що поєднують модульну архітектуру з динамічним налаштуванням складності відповідно до дій гравця, щоб забезпечити варіативність, повторюваність проходження та підтримку динамічності ігрового контенту. У роботі запропоновано поширену (layered) модель проєктування рівнів, у якій послідовно формуються базові елементи середовища (ґрунт, платформи, переешкоди), а надбудови шарів (вороги, пастки, декорації) керуються узгодженими правилами розміщення. Ключову роль у створенні природної геометрії та керованої випадковості відіграють шум Перліна та його фрактальні розширення (octaves): комбінування кількох частот і амплітуд дає змогу поєднати великомасштабні контури з дрібними деталями рельєфу. Запропонована система адаптивної складності інтегрується у генератор шляхом зміни параметрів шуму (scale, persistence, octaves) і поведінкових характеристик ворогів (швидкість, частота атак, кількість), що визначаються індексом продуктивності гравця. Практичну реалізацію виконано у рушії Unity з використанням мови програмування C# та Tilemap-підходу; наведено приклади коду для побудови карти висот і підключення контролера адаптації. Розроблені алгоритми та методи є універсальними й можуть бути легко інтегровані в середовища C++, зокрема для рушіїв Unreal Engine та Cocos2d-x, що підтверджує їхню практичну придатність і міжплатформність. Показано, що такий комбінований підхід зберігає баланс між непередбачуваністю та керованістю, зменшує витрати на ручний дизайн рівнів і підвищує залучення користувачів завдяки персоналізованому ігровому досвіду. Отримані результати можуть бути застосовані у навчальних прототипах, інді-проєктах та як базис для подальших досліджень із використанням методів машинного навчання в генеративному геймдизайні.

Ключові слова: процедурна генерація; шум Перліна; фрактальний шум; адаптивна складність; 2D-платформери; Unity; C++.

YU. YU. ILIASH

Candidate of Technical Sciences, Associate Professor,
Head of the Department of Computer Science
Vasyl Stefanyk Carpathian National University
ORCID: 0009-0006-5786-7205

V. O. HORIELOV

Candidate of Technical Sciences, Associate Professor,
Associate Professor at the Department of Computer Science
Vasyl Stefanyk Carpathian National University
ORCID: 0000-0002-2106-8704

V. A. ROVINSKY

Candidate of Technical Sciences, Associate Professor,
Associate Professor at the Department of Computer Science
Vasyl Stefanyk Carpathian National University
ORCID: 0000-0001-8454-8580

ALGORITHMS AND METHODS OF PROCEDURAL GENERATION AND ADAPTIVE DIFFICULTY IN PLATFORMER GAMES

The article focuses on the study of algorithms and methods for procedural generation of game content and adaptive difficulty in 2D platformers. The relevance of approaches combining modular architecture with dynamic difficulty adjustment based on player actions is substantiated, aiming to ensure variability, replay value, and maintain the dynamism of game content. The paper proposes a layered model of level design, where basic environmental elements (ground, platforms, obstacles) are generated sequentially, while higher layers (enemies, traps, decorations) are governed by consistent placement rules. A key role in creating natural geometry and controlled randomness is played by Perlin noise and its fractal extensions (octaves): combining multiple frequencies and amplitudes allows large-scale contours to merge with fine terrain details. The proposed adaptive difficulty system is integrated into the generator by modifying noise parameters (scale, persistence, octaves) and enemy behavioral attributes (speed, attack frequency, quantity), determined by the player's performance index. The practical implementation was carried out in the Unity engine using the C# programming language and the Tilemap approach; code examples for heightmap generation and adaptive controller integration are provided. The developed algorithms and methods are engine-agnostic and can be easily integrated into C++ environments, particularly Unreal Engine and Cocos2d-x, demonstrating their cross-platform applicability and scalability. It is shown that the combined approach maintains a balance between unpredictability and control, reduces the cost of manual level design, and enhances user engagement through personalized gameplay experiences. The obtained results can be applied in educational prototypes, indie projects, and serve as a foundation for further research involving machine learning methods in generative game design.

Key words: procedural generation; Perlin noise; fractal noise; adaptive difficulty; 2D platformers; Unity; C++.

Постановка проблеми

У ХХІ столітті індустрія відеоігор перетворилася на один із найдинамічніших напрямів цифрової культури та інновацій. Вона об'єднує мистецтво, програмування, психологію, моделювання поведінки й системи штучного інтелекту. Завдяки широкій доступності рушіїв, таких як *Unity*, *Unreal Engine*, *Cocos2d-x*, створення власних ігор стало можливим не лише для великих компаній, а й для незалежних розробників і студентів-ентузіастів.

Серед різних жанрів особливої уваги заслуговують 2D-платформери – ігри, що поєднують динаміку, простоту керування та високий потенціал для творчих експериментів. Вони є відмінним полігоном для вивчення основ ігрового дизайну, алгоритмічних методів і концепцій адаптивних систем.

Сучасні тенденції розвитку ігор засвідчують, що користувачі очікують не лише красивої графіки чи захопливого сюжету, а насамперед – варіативності та неповторності ігрового досвіду. Гравець прагне, щоб кожне нове проходження було унікальним, створюючи ефект «живого світу», який розвивається самостійно. Одним із найефективніших шляхів досягнення цього є процедурна генерація контенту (Procedural Content Generation, PCG) – автоматичне створення рівнів, об'єктів, перешкод чи навіть наративів за допомогою алгоритмів. Такі технології не лише знижують витрати часу на розробку, а й дозволяють створювати фактично нескінченні варіації ігрового середовища.

Паралельно актуальним напрямом стає адаптивна складність (Dynamic Difficulty Adjustment, DDA), мета якої – підтримувати баланс між викликом і комфортом, динамічно реагуючи на дії, рівень майстерності та стиль гри користувача. Поєднання процедурної генерації з адаптивними механізмами відкриває можливість створення гнучких систем, здатних забезпечити персоналізований досвід і підтримувати мотивацію гравця протягом усього процесу гри.

Аналіз останніх досліджень і публікацій

Питання оптимізації ігрових даних та ефективної обробки інформації розглядалися у працях українських і зарубіжних авторів. Зокрема в [1], було запропоновано методи ущільнення та структурного перетворення інформації, що підвищують ефективність обробки й передачі даних у реальному часі. Ідеї оптимізації структури ігрового контенту, висвітлені в цій статті, стали основою для подальших досліджень алгоритмів динамічного формування середовища та адаптивної складності рівнів.

Поняття процедурної генерації відоме ще з 1980-х років, проте саме сучасні обчислювальні можливості дали змогу піднести цей підхід на рівень складних систем зі штучним інтелектом, здатних реагувати на дії гравця [2].

Водночас не менш важливою тенденцією є поява адаптивних систем складності [3]. Їхня суть полягає в тому, щоб гра не залишалася статичною, а реагувала на успішність, стиль гри та емоційний стан користувача. Гравець, який грає впевнено, може стикатися з більш жорсткими викликами, тоді як початківець – із помірними, щоб

уникнути втрати мотивації. Перші концепції Динамічного Регулювання Складності описані Робіном Ханіком у класичній праці [4], а подальші дослідження [3, 5] доводять ефективність таких систем у збереженні стану “flow” – оптимального балансу між викликом і навичками.

Поєднання процедурної генерації з адаптивною складністю є одним із найперспективніших напрямів у сучасному геймдизайні. Такий підхід дозволяє не лише створювати унікальні рівні, але й налаштовувати їхню складність у процесі гри, забезпечуючи персоналізований досвід для кожного користувача. Наприклад, дослідження [6] демонструє, як алгоритм може генерувати рівні, виходячи з показників успішності гравця – тривалості життя, частоти помилок, часу реакції тощо. Це дозволяє створювати динамічну криву складності без ручного налаштування сценаріїв.

Особливо актуальним застосування цих технологій є саме для жанру платформерів. На відміну від стратегій або RPG, де складність часто визначається кількістю ресурсів чи ворогів, у платформерах вона прямо пов’язана з геометрією рівня, розміщенням платформ, глибинами ям, швидкістю ворогів і частотою перешкод. Саме тому автоматизована генерація рівнів із контролем складності має практичну цінність – вона дозволяє масштабувати контент, створювати адаптивні виклики та уникати монотонності. Застосування шумових, клітинних автоматів, граматичних систем, еволюційних і пошукових алгоритмів стало стандартом у дослідженнях цієї тематики [7, 8]

Із теоретичного погляду, процедурна генерація та адаптивна складність поєднують у собі елементи класичних алгоритмів, стохастичних процесів і методів машинного навчання. Зокрема, пошукові підходи (search-based PCG) використовують еволюційні алгоритми для пошуку найкращої конфігурації рівня відповідно до функції оцінювання. Граматичні методи будують рівні на основі формальних правил, подібно до синтаксичного аналізу в мовах програмування. А сучасні підходи, засновані на підкріплювальному навчанні (Reinforcement Learning), дозволяють навчати агента-генератора створювати контент, який оптимально відповідає стилю гри користувача [9, 10].

Практична цінність дослідження полягає в тому, що впровадження подібних алгоритмів у реальні рушії (як-от Unity, Unreal Engine або Cocos2d-x) відкриває можливість створювати ігри, що «живуть» разом із гравцем. Кожне нове проходження формує власний маршрут, власний набір викликів і вражень. Це не лише підвищує залученість користувачів, але й формує основу для розвитку адаптивних навчальних і симуляційних систем, де гра може виступати моделлю навчального середовища.

Попри численні переваги, використання процедурної генерації та адаптивної складності пов’язане з рядом викликів. Одним із них є оцінка рівня «якості» згенерованого контенту – адже алгоритм може створити рівень, який технічно коректний, але ігрово незбалансований. Іншим викликом є питання «відчуття контролю»: якщо гравець помічає, що гра сама підлаштовується під нього, це може знизити відчуття чесного виклику. Тому важливим напрямом є розробка систем, які залишають адаптацію непомітною, зберігаючи природність процесу.

Формулювання мети дослідження

Таким чином, розвиток алгоритмів і методів процедурної генерації та адаптивної складності рівнів в іграх-платформерах є одним із найбільш перспективних напрямів сучасної ігрової розробки. Він поєднує математичні моделі, аналітику поведінки, штучний інтелект і креативний дизайн, відкриваючи нові можливості як для наукових досліджень, так і для індустріальної практики. У цьому контексті створення прототипу 2D-платформера з процедурною генерацією рівнів і адаптивною складністю може стати не лише демонстрацією теоретичних принципів, а й практичним внеском у розвиток технологій інтелектуального геймдизайну.

Викладення основного матеріалу дослідження

Під час розробки гри-платформера з процедурною генерацією рівнів особливу увагу приділяється створенню збалансованого ігрового середовища, яке поєднує зручність керування, гнучкість механік, доступність для широкого кола гравців і здатність адаптуватися до різних стилів проходження. Врахування потреб як новачків, так і досвідчених користувачів дозволяє побудувати динамічну систему складності, що підтримує природний темп гри та зберігає інтерес протягом усього процесу.

Пропонується реалізовувати процес проектування на основі пошарового підходу до процедурної генерації рівнів, який розглядається як один з етапів архітектурного проектування ігрового контенту. Його сутність полягає у послідовному формуванні середовища – від базових елементів (грунту, платформ, зон пересування) до компонентів вищого рівня (ворогів, пасток, декоративних об’єктів). Така структура дозволяє забезпечити логічну впорядкованість, контроль над взаємодією об’єктів і запобігання конфліктам при розміщенні.

Запропонована багаторівнева модель генерації характеризується передбачуваністю, гнучкістю та масштабованістю. Вона спрощує налагодження, адже помилки легко локалізуються в межах окремих шарів, і дозволяє розширювати гру шляхом додавання нових модулів – наприклад, шарів «скарбів», погодних ефектів чи декоративних об’єктів. Крім того, модель інтегрує механізми адаптивної складності, які динамічно змінюють параметри генерації залежно від дій гравця: система може збільшувати кількість ворогів або пасток при швидкому проходженні, чи навпаки, спрощувати геометрію рівня при численних помилках. Саме на цьому етапі доцільним стає використання шумових функцій, зокрема шуму Перліна, що забезпечує природність, когерентність і варіативність ігрового простору, створюючи плавний перехід до наступного етапу генерації – побудови шумових карт рівня.

Одним із ключових інструментів, що використовується для створення природних, плавних і варіативних ігрових середовищ, є шум Перліна – математична функція, розроблена Кеном Перліном у 1983 році [11]. Її основна ідея полягає у створенні когерентного (узгодженого) шуму, тобто такого, у якому сусідні значення мають локальну кореляцію та змінюються плавно, на відміну від випадкових (білих) шумів, де кожна точка не залежить від сусідніх. Саме ця властивість робить шум Перліна надзвичайно корисним для задач комп’ютерної графіки, процедурної генерації ландшафтів, текстур, хмар, морських хвиль, а також для побудови рівнів у 2D та 3D іграх.

Якщо випадкові числа між 0 та 1 застосувати безпосередньо для створення карти висот, отримаємо так званий білий шум, у якому відсутня структура – сусідні точки змінюються різко й хаотично. (рис. 1) Такий результат виглядає неприродно. Натомість шум Перліна формує «плавні» хвилеподібні структури, які око людини сприймає як природні візерунки – пагорби, долини, хмари або поверхню землі (рис. 2). Це досягається завдяки математичній побудові функції, у якій випадковість поєднана з локальною згладженістю.

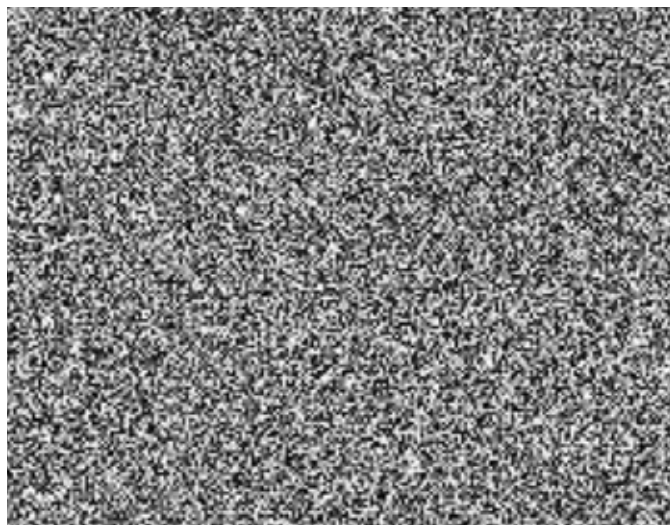


Рис. 1. Білий шум



Рис. 2. Шум Перліна

Формально одномірний шум Перліна описується як зважена сума локальних градієнтів:

$$P(x) = \sum \text{fade}(t_i) \cdot G_i$$

де:

x – координата,

i – індекс решітки навколо x ,

G_i – випадковий напрямок градієнта в точці решітки i ,

$t_i = x - i$ – локальне зміщення від точки решітки,

$\text{fade}(t)$ – функція згладжування, що забезпечує плавні переходи (зазвичай $\text{fade}(t) = 6t^5 - 15t^4 + 10t^3$).

Завдяки функції $\text{fade}()$ отримується неперервне значення, а градієнти вносять контрольовану варіативність. Результатом є функція, що виглядає випадково, але зберігає локальну гармонію – саме це і є когерентний шум.

Шум Перліна має низку переваг, що роблять його одним із найпоширеніших інструментів у процедурній генерації контенту. Його головною особливістю є плавність і неперервність, адже значення змінюються поступово, без різких переходів, завдяки чому створюються природні форми й структури. Водночас шум Перліна є детермінованим – при однакових вхідних параметрах він завжди дає ідентичний результат, що дозволяє відтворювати однакові рівні або ландшафти. Ще однією перевагою є контрольована випадковість: отримані патерни виглядають природно, не маючи явно вираженої періодичності. Крім того, шум відзначається масштабованістю – його можна використовувати з різними частотами (scale), поєднуючи кілька рівнів для формування складних фрактальних структур (octaves). Завдяки цим властивостям шум Перліна широко застосовується для генерації ландшафтів і визначення висот поверхні, розподілу води, ресурсів чи біомів, створення варіативності ігрових об'єктів за формою, розміром або кольором, а також для імітації природних текстур – таких як хмари, трава чи камінь.

У ігрових рушіях реалізація шуму Перліна доступна через вбудовану функцію (код на Unity):

```
float height = Mathf.PerlinNoise(x * scale, y * scale); Unity
float Height = FMath::PerlinNoise2D(FVector2D(X * Scale, Y * Scale)) * MaxHeight; Unreal
Engine
```

Параметр scale визначає частоту шуму: менше значення дає плавні великі хвилі, більше – дрібні нерівності. Отримане значення height (у діапазоні [0, 1]) використовується для обчислення висоти поверхні або для прийняття рішення, який тип блоку розмістити у даній точці – землю, воду, платформу чи порожній простір.

Наприклад (код на Unity):

```
float height = Mathf.PerlinNoise(x * 0.1f, seed) * maxHeight;
int y = Mathf.FloorToInt(height);
```

Таким чином, для кожної координати x визначається висота y , що утворює плавний рельєф рівня. Комбінування декількох шарів шуму з різними параметрами частоти та амплітуди дозволяє створювати складніші топографії – від м'яких пагорбів до хаотичних скельних структур.

Використання шуму Перліна у системі пошарової генерації рівнів забезпечує природність форм, варіативність розміщення об'єктів і підґрунтя для подальшої адаптації складності, адже зміна параметрів scale чи maxHeight дає можливість керувати «рельєфом» рівня, створюючи плавні або навпаки – динамічні, складні для проходження ділянки.

Попри свою гнучкість, класичний шум Перліна має обмеження: при використанні одного рівня частоти та амплітуди отримана карта має надто рівномірну структуру – великі плавні хвилі без дрібних деталей. Щоб подолати цю проблему, у практиці комп'ютерної графіки та процедурного моделювання використовується фрактальний шум [12], або шум із «октавами» (octaves), який є комбінацією кількох рівнів шуму Перліна з різними масштабами.

Фрактальний шум можна подати як суму кількох шарів (octaves) базового шуму, кожен із яких має власну частоту (масштаб) і амплітуду (вплив на результат):

$$F(x, y) = \sum_{i=0}^{n-1} \text{amp}_i \cdot \text{PerlinNoise}(x \cdot \text{freq}_i, y \cdot \text{freq}_i)$$

де:

n – кількість шарів,

$\text{freq}_i = \text{baseFreq} \cdot 2^i$ – частота для i -го шару,

$\text{amp}_i = \text{baseAmp} \cdot \text{persistence}^i$ – амплітуда, що зменшується зі зростанням i ,

$\text{persistence} \in [0, 1]$ – коефіцієнт зменшення амплітуди (зазвичай 0.5–0.7).

Таким чином, перший шар створює великі загальні контури рельєфу, а наступні додають дрібні деталі, текстуру та варіативність. Результат – природний шум, що поєднує плавні макроструктури та дрібномасштабні особливості. Цей принцип аналогічний побудові фракталів, звідки походить і сама назва методу.

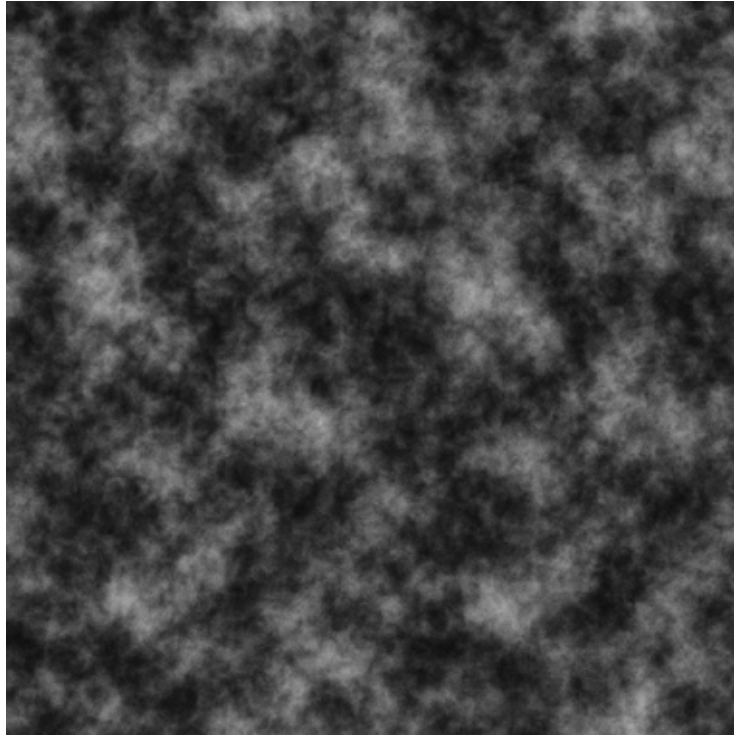


Рис. 3. Фрактальний шум Перліна

У рушії Unity фрактальний шум легко реалізується шляхом сумування кількох викликів `Mathf.PerlinNoise()` з різними масштабами:

```
float FractalNoise(float x, float y, int octaves, float persistence, float scale)
{
    float total = 0f;
    float frequency = 1f;
    float amplitude = 1f;
    float maxValue = 0f; // для нормалізації

    for (int i = 0; i < octaves; i++)
    {
        total += Mathf.PerlinNoise(x * frequency * scale, y * frequency * scale) *
amplitude;
        maxValue += amplitude;
        amplitude *= persistence;
        frequency *= 2f;
    }
    return total / maxValue;
}
```

У межах цього дослідження запропоновано новий гібридний підхід до генерації рівнів, який поєднує три ключові принципи:

- **Пошарова архітектура:** забезпечує структуровану, керовану логіку створення рівня – від базової землі до ворогів і декоративних елементів.
- **Фрактальний шум Перліна:** задає основу для кожного шару, формуючи природні переходи між ділянками та визначаючи топографію рівня з різною деталізацією.
- **Адаптивна складність:** динамічно модифікує параметри шуму (наприклад, `scale`, `persistence`, кількість `octaves`) на основі дій гравця – підвищуючи або знижуючи частоту нерівностей, довжину платформ чи щільність пасток.

Таке поєднання дозволяє створити самоадаптивний генератор рівнів, який не лише формує різноманітне середовище, але й регулює складність у реальному часі. Наприклад:

- при швидкому проходженні рівня гра збільшує параметр frequency, створюючи щільнішу структуру платформ і складніші перепади;
- при частих поразках – зменшує амплітуду шуму, роблячи поверхню рівнішою і простішою для навігації.

Завдяки цьому система отримує властивість реактивного дизайну (responsive design), коли рівень не є статично заданим, а формується у взаємодії з гравцем. Це не лише підвищує реіграбельність, але й наближає гру до концепції «інтерактивної творчості», описаної в [13], де користувач виступає співтворцем процесу.

Таким чином, фрактальний шум Перліна у поєднанні з пошаровою архітектурою та адаптивною логікою складності формує новий підхід до процедурної генерації у 2D-платформерах. Він дозволяє створювати природні, непередбачувані, але керовані рівні, що змінюються залежно від досвіду, швидкості та стилю гри користувача. Така система поєднує алгоритмічну строгість з елементами динамічної творчості, що відповідає сучасним тенденціям у галузі інтелектуального геймдизайну.

З урахуванням теоретичних принципів процедурної генерації та фрактального шуму Перліна, реалізовано комплексний алгоритм побудови рівня у середовищі Unity. Його основою є поєднання пошарового підходу з керованими параметрами шуму, що дозволяє створювати унікальні ігрові локації під час кожного запуску гри.

Архітектура генерації побудована за модульним принципом, де кожен компонент відповідає за окрему стадію побудови рівня. Кожен етап працює незалежно, використовуючи спільні параметри, що передаються через конфігураційний об'єкт LevelSettings. Такий підхід забезпечує легку масштабованість і можливість заміни будь-якого модуля без впливу на решту системи.

Основні модулі системи:

NoiseGenerator: обчислює карту висот за допомогою фрактального шуму Перліна (використовуються параметри scale, persistence, octaves, seed).

TerrainBuilder: перетворює карту висот у базову топографію рівня (грунт, платформи, порожні ділянки).

ObstaclePlacer: додає перешкоди (шипи, ями, пастки) згідно з логікою пошарової побудови.

EnemySpawner: розміщує ворогів на рівних поверхнях з урахуванням складності гри.

AdaptiveController: аналізує поведінку гравця (кількість смертей, швидкість проходження, середній час реакції) і змінює параметри генерації для наступного рівня.

Нижче наведено спрощений псевдокод алгоритму генерації рівня:

```
function GenerateLevel(seed, difficulty):
    setParametersBasedOnDifficulty(difficulty)
    heightMap = FractalPerlinNoise(seed, scale, persistence, octaves)

    for x in range(levelWidth):
        height = heightMap[x]
        for y in range(levelHeight):
            if y <= height:
                placeBlock("ground", x, y)
            else if y <= waterLevel:
                placeBlock("water", x, y)

    placePlatforms(heightMap)
    placeSpikes(heightMap)
    placeEnemies(heightMap, difficulty)
```

Під час виконання гри, після завершення рівня, система передає дані про продуктивність гравця модулю AdaptiveController, який коригує складність:

```
function UpdateDifficulty(playerStats):
    if playerStats.deaths > threshold:
        scale -= 0.02 // зменшення хвилястості рельєфу
        octaves -= 1 // менше дрібних деталей
    else if playerStats.speed > avgSpeed:
        scale += 0.02 // дрібніші коливання поверхні
        octaves += 1 // більше деталей
```

Таким чином, кожен наступний рівень формується з урахуванням попередніх дій гравця. Це створює ефект еволюції складності, коли гра динамічно підлаштовується до користувача, не потребуючи фіксованих налаштувань.

У практичній реалізації в Unity використовується Tilemap-система, що дозволяє візуалізувати згенеровані блоки у вигляді спрайтів. Основний скрипт LevelGenerator.cs виконує ініціалізацію шуму та побудову ландшафту:

```
for (int x = 0; x < width; x++)
{
    float height = FractalNoise(x, seed, octaves, persistence, scale) * maxHeight;
    for (int y = 0; y < height; y++)
        tilemap.SetTile(new Vector3Int(x, y, 0), groundTile);

    if (height < waterThreshold)
        tilemap.SetTile(new Vector3Int(x, Mathf.FloorToInt(height), 0), waterTile);
}
```

Модуль AdaptiveController.cs періодично оновлює параметри генератора:

```
if (player.deaths > 3)
{
    generator.scale *= 0.95f;
    generator.octaves = Mathf.Max(2, generator.octaves - 1);
}
else if (player.score > 1000)
{
    generator.scale *= 1.05f;
    generator.octaves = Mathf.Min(8, generator.octaves + 1);
}
```

У результаті, навіть без зовнішнього втручання, гра створює нові комбінації рельєфу, змінюючи кількість і складність перешкод, плавно формуючи індивідуальний темп гри.

Висновки

У межах цієї роботи реалізовано модель процедурної генерації рівнів у 2D-платформері, що базується на комбінації пошарового підходу, фрактального шуму Перліна та адаптивної системи складності. Такий підхід дозволяє формувати унікальні ігрові рівні з природною топографією, адаптувати параметри генерації відповідно до стилю й навичок гравця, зберігати баланс між випадковістю та керованістю структури гри, а також підвищувати мотивацію до повторного проходження завдяки персоналізованому ігровому досвіду. Практична реалізація у Unity (C#) продемонструвала, що варіювання параметрів шуму (scale, persistence, octaves) у поєднанні з поведінковими метриками користувача дає змогу створювати динамічну систему рівнів, що природно реагує на дії гравця.

Отримані результати підтверджують ефективність комбінованого підходу як інструменту не лише для створення контенту, а й для розвитку концепції інтелектуального геймдизайну, у якому гра виступає адаптивним середовищем взаємодії з користувачем.

Розроблені алгоритми та методи мають універсальну структуру й можуть бути швидко інтегровані в середовища на C++, зокрема в рушії Unreal Engine та Cocos2d-x, без втрати функціональності чи потреби у суттєвій модифікації архітектури. Їхня модульність і незалежність від конкретного рушія дозволяє переносити ключові компоненти – генератор шуму, адаптивний контролер і пошарову логіку побудови рівня – без змін у логіці алгоритмів. Це забезпечує міжплатформність, сумісність із різними системами рендерингу й фізики та відкриває можливості для використання єдиного генеративного ядра в різних проектах.

Таким чином, розроблена методика поєднує ефективність математичних моделей із практичною гнучкістю реалізації, що робить її придатною як для освітніх і дослідницьких експериментів, так і для професійної розробки ігор із процедурною генерацією та адаптивною складністю.

Подальший розвиток теми процедурної генерації пов'язаний із застосуванням розширених математичних моделей для керування не лише геометрією рівнів, а й поведінковими характеристиками ігрових агентів. Використання стохастичних процесів – таких як фрактальний броунівський рух чи випадкові поля Маркова – може забезпечити динамічну адаптацію складності ворогів у часі, узгоджену з діями гравця. У цьому контексті шумові функції здатні виконувати роль керуючих сигналів інтенсивності виклику, визначаючи щільність або агресивність супротивників. Поєднання цих підходів із методами оптимізації створює основу для єдиної адаптивної системи балансу, у якій структура рівня та поведінка ворогів еволюціонують синхронно, забезпечуючи природну й послідовну динаміку ігрового процесу.

Список використаної літератури

1. Горєлов, В. О., Іляш, Ю. Ю., Ровінський, В. А. (2018). Методи ущільнення даних та перетворення форми інформації у комп'ютерних іграх. *Управління розвитком складних систем*. № 35 с. 93-104.
2. Shaker, N., Togelius, J., & Nelson, M. J. (2016). *Procedural content generation in games.: A Textbook and an Overview of Current Research*. Springer. DOI: 10.1007/978-3-319-42716-4
3. Sepulveda, G. K., Besoain, F., & Barriga, N. A. (2019, November). Exploring dynamic difficulty adjustment in videogames. In *2019 IEEE CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON)* (pp. 1-6). IEEE. DOI: 10.1109/CHILECON47746.2019.8988068
4. Hunicke, R. (2005, June). The case for dynamic difficulty adjustment in games. In *Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology* (pp. 429-433). DOI: 10.1145/1178477.1178573
5. Zohaib, M. (2018). Dynamic difficulty adjustment (DDA) in computer games: A review. *Advances in Human-Computer Interaction*, 2018(1), 5681652. DOI: 10.1155/2018/5681652
6. Biemer, C. F. (2023, October). Dynamic difficulty adjustment via procedural level generation guided by a Markov decision process for platformers and roguelikes. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment* (Vol. 19, No. 1, pp. 436-439). DOI: 10.1609/aiide.v19i1.27540
7. Togelius, J., Yannakakis, G. N., Stanley, K. O., & Browne, C. (2011). Search-based procedural content generation: A taxonomy and survey. *IEEE Transactions on Computational Intelligence and AI in Games*, 3(3), 172-186. DOI: 10.1109/TCIAIG.2011.2148116
8. Summerville, A., Snodgrass, S., Guzdial, M., Holmgård, C., Hoover, A. K., Isaksen, A., ... & Togelius, J. (2018). Procedural content generation via machine learning (PCGML). *IEEE Transactions on Games*, 10(3), 257-270. DOI: 10.1109/TG.2018.2846639
9. Risi, S., & Togelius, J. (2020). Increasing generality in machine learning through procedural content generation. *Nature Machine Intelligence*, 2(8), 428-436.
10. González-Duque, M., Palm, R. B., Ha, D., & Risi, S. (2020, August). Finding game levels with the right difficulty in a few trials through intelligent trial-and-error. In *2020 IEEE Conference on Games (CoG)* (pp. 503-510). IEEE. DOI: 10.1109/CoG47356.2020.9231548
11. Perlin, K. (1985). An image synthesizer. *ACM Siggraph Computer Graphics*, 19(3), 287-296. DOI: 10.1145/325165.325247
12. Dustler, M., Bakic, P., Petersson, H., Timberg, P., Tingberg, A., & Zackrisson, S. (2015, March). Application of the fractal Perlin noise algorithm for the generation of simulated breast tissue. In *Medical Imaging 2015: Physics of Medical Imaging* (Vol. 9412, pp. 844-852). SPIE. DOI: 10.1117/12.2081856
13. Risi, S., & Togelius, J. (2020). Increasing generality in machine learning through procedural content generation. *Nature Machine Intelligence*, 2(8), 428-436.

References

1. Horelov, V. O., Ilyash, Yu. Yu., & Rovinsky, V. A. (2018). Methods of Data Compression and Information Shape Transformation in Computer Games. *Management of Complex Systems Development*, (35), 93–104.
2. Shaker, N., Togelius, J., & Nelson, M. J. (2016). *Procedural content generation in games.: A Textbook and an Overview of Current Research*. Springer. DOI: 10.1007/978-3-319-42716-4
3. Sepulveda, G. K., Besoain, F., & Barriga, N. A. (2019, November). Exploring dynamic difficulty adjustment in videogames. In *2019 IEEE CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON)* (pp. 1-6). IEEE. DOI: 10.1109/CHILECON47746.2019.8988068
4. Hunicke, R. (2005, June). The case for dynamic difficulty adjustment in games. In *Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology* (pp. 429-433). DOI: 10.1145/1178477.1178573
5. Zohaib, M. (2018). Dynamic difficulty adjustment (DDA) in computer games: A review. *Advances in Human-Computer Interaction*, 2018(1), 5681652. DOI: 10.1155/2018/5681652
6. Biemer, C. F. (2023, October). Dynamic difficulty adjustment via procedural level generation guided by a Markov decision process for platformers and roguelikes. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment* (Vol. 19, No. 1, pp. 436-439). DOI: 10.1609/aiide.v19i1.27540
7. Togelius, J., Yannakakis, G. N., Stanley, K. O., & Browne, C. (2011). Search-based procedural content generation: A taxonomy and survey. *IEEE Transactions on Computational Intelligence and AI in Games*, 3(3), 172-186. DOI: 10.1109/TCIAIG.2011.2148116
8. Summerville, A., Snodgrass, S., Guzdial, M., Holmgård, C., Hoover, A. K., Isaksen, A., ... & Togelius, J. (2018). Procedural content generation via machine learning (PCGML). *IEEE Transactions on Games*, 10(3), 257-270. DOI: 10.1109/TG.2018.2846639

9. Risi, S., & Togelius, J. (2020). Increasing generality in machine learning through procedural content generation. *Nature Machine Intelligence*, 2(8), 428-436.
10. González-Duque, M., Palm, R. B., Ha, D., & Risi, S. (2020, August). Finding game levels with the right difficulty in a few trials through intelligent trial-and-error. In *2020 IEEE Conference on Games (CoG)* (pp. 503-510). IEEE. DOI: 10.1109/CoG47356.2020.9231548
11. Perlin, K. (1985). An image synthesizer. *ACM Siggraph Computer Graphics*, 19(3), 287-296. DOI: 10.1145/325165.325247
12. Dustler, M., Bakic, P., Petersson, H., Timberg, P., Tingberg, A., & Zackrisson, S. (2015, March). Application of the fractal Perlin noise algorithm for the generation of simulated breast tissue. In *Medical Imaging 2015: Physics of Medical Imaging* (Vol. 9412, pp. 844-852). SPIE. DOI: 10.1117/12.2081856
13. Risi, S., & Togelius, J. (2020). Increasing generality in machine learning through procedural content generation. *Nature Machine Intelligence*, 2(8), 428-436.

Дата першого надходження рукопису до видання: 27.11.2025
Дата прийнятого до друку рукопису після рецензування: 12.12.2025
Дата публікації: 31.12.2025

М. Р. КАЛАШНІКОВстудент магістратури кафедри комп'ютерних систем,
мереж і кібербезпекиНаціональний аерокосмічний університет
«Харківський авіаційний інститут»

ORCID: 0009-0003-7774-9239

В. С. ХАРЧЕНКОдоктор технічних наук, професор,
завідувач кафедри комп'ютерних систем, мереж і кібербезпекиНаціональний аерокосмічний університет
«Харківський авіаційний інститут»

ORCID: 0000-0001-5352-077X

ІНТЕЛЕКТУАЛЬНИЙ ПІДХІД ДО ПІДВИЩЕННЯ СТАБІЛЬНОСТІ СИСТЕМ АВТОМАТИЗОВАНОГО ТЕСТУВАННЯ НА БАЗІ НАВЧАННЯ З ПІДКРІПЛЕННЯМ

У дослідженні увага зосереджувалася на розробці інструменту ефективізації автоматизованої системи тестування із використанням інтелектуальної архітектури на базі парадигми навчання з підкріпленням, спрямованої на покращення стабільності тестування, причому Selenium застосовувалася в якості практичного прикладу. У роботі були визначені основні причини появи нестабільності в межах тестування графічних користувацьких інтерфейсів, а саме мінливість DOM-структур, асинхронна візуалізація, мережева затримка та ригідність синхронізації, з огляду на що була сформульована багаторівнева адаптивна модель, задача якої полягає в уникненні цих факторів за допомогою механізмів навчання з підкріпленням. Запропонований Інтелектуальний рівень з підкріпленням передбачає реструктуризацію класичної ієрархії Selenium, тобто IDE, Client–Server, та Grid в об'єднану структуру із самонавчанням, здатну проводити автономне підлаштування до мінливих параметрів системи. В межах цього рівня, Агент підкріплення IDE виконує локальну адаптацію шляхом регуляції інтервалів відтворення та вибору локаторів у відповідь на структурні зміни у DOM. Адаптивний канал Client–Server, що включає Агенти стабілізації виконання та ієрархічної адаптації локаторів, забезпечує стабільність виконання тестових сценаріїв за рахунок підлаштування інтервалів синхронізації, політики очікування та стратегій відновлення структури локаторів, виходячи з телеметричних даних у реальному часі. На рівні розподіленого контролю знаходиться Середовище навчання Grid, в якому застосовується кооперативне багатоагентне навчання, спрямоване на врівноваження тестових навантажень на всіх вузлах, в ході чого відбувається оптимізація загальносистемної стабільності та кількості тестів, які система може виконати за одиницю часу. Порівняльний аналіз показав, що в той час, як традиційні інструменти, такі як Cypress та Playwright, реалізують ригідні процедури стабілізації, запропонована архітектура уособлює досягнення адаптивізації та стабілізації шляхом постійного навчання та аналізу телеметричного зворотного зв'язку, що сприяє спрощеному проведенню сталого самооптимізаційного тестування.

Ключові слова: штучний інтелект, стабільність тестів, автоматизація, QA, адаптивні алгоритми.

M. R. KALASHNIKOVMaster's Student at the Department of Computer Systems,
Networks and Cybersecurity

National Aerospace University "Kharkiv Aviation Institute"

ORCID: 0009-0003-7774-9239

V. S. KHARCHENKO

Doctor of Technical Sciences, Professor,

Head of the Department of Computer Systems, Networks and Cybersecurity

National Aerospace University "Kharkiv Aviation Institute"

ORCID: 0000-0001-5352-077X

INTELLIGENT APPROACH TO ENHANCING THE STABILITY OF AUTOMATED TESTING SYSTEMS BASED ON REINFORCEMENT LEARNING

This research develops and substantiates an intelligent reinforcement-based architecture aimed at enhancing the stability of automated testing systems, using Selenium as a practical foundation. The study identifies the root causes of instability in UI testing – DOM volatility, asynchronous rendering, network latency, and rigid synchronization – and formulates a multilayered adaptive model that mitigates these factors through reinforcement learning mechanisms. The proposed Intelligent Reinforcement Layer restructures Selenium's classical hierarchy, namely IDE, Client–Server, and Grid, into a unified self-learning framework capable of autonomous adjustment to environmental drift. Within the IRL, the IDE Reinforcement Agent performs local adaptation by regulating playback timing and locator selection in response to structural changes in the DOM. The Client–Server Adaptive Layer, comprising the Execution Stability Agent and Hierarchical Locator Adaptation Agent, ensures execution stability by adjusting synchronization intervals, wait policies, and locator recovery strategies based on real-time telemetry. At the distributed level, the Grid Learning Environment applies cooperative multi-agent learning to balance test workloads across nodes, optimizing system-wide stability and throughput. The research formalizes unified reinforcement-learning logic, defines state–action structures for each subsystem, and introduces a collective experience buffer for cross-layer training. Comparative analysis demonstrates that while traditional frameworks such as Cypress and Playwright implement fixed stabilization procedures, the proposed IRL architecture achieves adaptive resilience through continuous learning and feedback. The study concludes that integrating reinforcement learning into Selenium's architecture yields a self-optimizing testing environment capable of maintaining long-term stability, reducing flakiness, and improving the robustness of automated QA processes under evolving application conditions.

Key words: artificial intelligence, test stability, QA automation, adaptive algorithms.

Постановка проблеми

Автоматизація тестування набула критично важливого значення в контексті сучасної парадигми розробки програмного забезпечення, створивши підґрунтя для підтримки якості програмних продуктів на належному рівні та прискорення циклів розробки в межах парадигми безперервної інтеграції та розгортання. Тим не менш, стабільність систем автоматизованого тестування, особливо в контексті тестування користувацьких інтерфейсів (user interface – UI), залишається одним з актуальних проблемних аспектів процесу забезпечення якості програмних продуктів. Нестабільність тестів проявляється у нерегулярності їхніх результатів (flaky tests) за умов повної незмінності вихідного коду.

Головна причина нестабільності автоматизації на базі графічних користувацьких інтерфейсів тісно пов'язана з їхньою динамічністю та непередбачуваною природою. Навіть за умов незначної зміни в інтерфейсній розмітці, асинхронної візуалізації, мережевої затримки, або затримок у оновленнях структур веб-сторінки (тобто у документній об'єктній моделі, Document Object Model – DOM) можуть виникнути хибні негативні результати. Традиційні автоматизаційні фреймворки, такі як Selenium та Cypress, передбачають застосування попередньо визначених елементних локаторів та статичної часової регуляції, що як правило не оновлюються по мірі оновлення користувацького інтерфейсу.

Відтак, сталі підвищення складності веб-застосунків та потреба у безперервному розгортанні вимагають поєднання адаптивних та інтелектуальних механізмів в автоматизовані системи тестування. Технології машинного навчання характеризуються наявністю необхідних інструментів для вирішення окресленої проблеми нестабільності шляхом забезпечення можливості самовиправлення тестів (self-healing tests), виявлення аномалій та синхронізації виходячи з даного контексту. За допомогою інтелектуального моделювання, автоматизована система здатна аналізувати дані виконання тестових сценаріїв, ідентифікувати закономірності нестабільностей та автономно підлаштовуватися до зовнішніх змін у UI або середовищі розробки. Відтак, у цьому дослідженні розглядається інтелектуальний підхід удосконалення стабільності автоматизованих систем на прикладі інструменту для автоматизації операцій у веб-браузері Selenium, який також використовується для тестування веб-застосунків.

Аналіз останніх досліджень і публікацій

У роботі [1] оцінюється можливість використання методів машинного навчання для передбачення нестабільних тестів у Python-проектах. Досліджується, чи можуть моделі навчання під наглядом виявляти нестабільність ще до додавання тесту в набір. Було реалізовано та порівняно три класифікатори, а саме Naive Bayes (Наївний баєсів класифікатор), SVM (Support vector machine, метод опорних векторів) і Random Forest (Випадковий ліс), на предмет “запахів коду” (“test smells”), отриманих із коду після фільтрації й нормалізації токенів. Для оцінки використовувалися різні співвідношення даних для тренування/тестування (50/50–90/10) і метрики точності, повноти, F1 та ROC (receiver operating characteristic, робоча характеристика приймача). Результати показали, що Random Forest має найвищу точність (>90%), але низьку повноту (<10%), тоді як Naive Bayes і SVM характеризуються більш збалансованими показниками. Незважаючи на перевагу RF у точності, SVM краще виявляє ширші закономірності виникнення нестабільностей. Подальший аналіз показав, що асинхронні очікування, випадковість,

обчислення з плаваючою точкою та I/O-операції є сильними індикаторами, тоді як багатопоточність і часові залежності – менш репрезентативними. Підсумовано, що RF є ефективним для високоточної детекції, але для поліпшення повноти майбутні дослідження мають враховувати чинники на рівні інфраструктури тестування та контексту.

У дослідженні [2] було проаналізовано 345 проєктів, які використовують Pyro, PyMC3, TensorFlow Probability та PyTorch, і виявлено 75 звітів про помилки та зміни у коді у 20 проєктах, що стосувалися виправлення нестабільних тестів. Після аналізу вручну було визначено основні причини: алгоритмічний недетермінізм, похибки з плаваючою точкою та непослідовне використання випадкових початкових значень (сідів, seed). Зазвичай розробники усувають ці проблеми через коригування порогових значень перевірок, фіксацію сідів або видалення нестабільних тестів. На основі цих висновків запропоновано метод FLASH – статистичний підхід, що передбачає виявлення нестабільних тестів, спричинених випадковими процесами виконання. FLASH багаторазово запускає тести з різними сідами і застосовує діагностику збіжності (тест Гевеке) для визначення моменту, коли отримано достатньо вибірок для статистичного висновку. Далі оцінюється ймовірність збоїв перевірок за рахунок моделювання розподілів спостережуваних значень, і виявлення тестових сценаріїв, результати яких змінюються між виконаннями. Метод було перевірено на 20 відкритих проєктах, де за допомогою нього було знайдено 11 нових нестабільних тестів (10 підтверджено розробниками), а також успішно відтворено раніше відомі випадки. Результати демонструють, що залучення статистичних підходів у тестування підвищує надійність програм, у яких результати мають випадковий характер.

У рамках роботи [3] представлено FlakyCat, що є методом класифікації нестабільних тестів за їхніми причинами з використанням Few-Shot навчання (за невеликої кількості екземплярів). FlakyCat передбачає використання попередньо навченої моделі CodeBERT для представлення коду і сіамську нейромережу для навчання класифікації з обмеженою кількістю розмічених даних. Модель оптимізується за допомогою функції втрат triplet-loss, щоб розмістити тести у просторі подібності, де приклади одного типу нестабільності групуються. Було створено набір із 451 нестабільного тесту з 13 категорій, який після аугментації містив 964 зразки. Підхід оцінювався відносно класичних методів (Random Forest, SVM, KNN (k-nearest neighbors – k-найближчі сусіди), Decision Tree (Дерево прийняття рішень)) і традиційних ознак (лексичних, на основі запахів коду) за метриками F1, MCC (Matthews Correlation Coefficient (коефіцієнт кореляції Метьюза), AUC (area under ROC curve, площа під ROC-кривою). FlakyCat досяг найвищих результатів (F1 = 0.73, MCC = 0.65, AUC = 0.83), перевершивши всі базові методи. Аналіз за категоріями показав, що тести з асинхронними очікуваннями, невпорядкованими вибірками та часовими залежностями легше ідентифікувати, ніж тести, пов'язані з багатопоточністю. Додатково запропоновано підхід на основі CodeBERT, який виділяє ключові оператори коду, що впливають на рішення моделі. Робота доводить, що Few-Shot навчання з семантичними вкладеннями коду може ефективно класифікувати типи нестабільності та допомагати у їх усуненні.

У межах дослідження [4] був представлений IDFLAKIES – метод виявлення нестабільних тестів і класифікації їх як NOD (Non-Order-Dependent, тест, збій якого не залежить від порядку виконання) або victim (тест, “постраждалий” від порядку виконання). Підхід складається з трьох етапів: встановлення, виконання та класифікації. На першому етапі тести повторно запускаються у звичному порядку для відфільтрування стабільно помилкових випадків. На другому – виконуються у зміненому порядку. Під час класифікації кожен тест, що зазнав збою, повторно запускається як у звичному, так і у модифікованому порядку до точки збою. Якщо тест знову завершується збоєм під час виконання у зміненому порядку, але проходить у звичному, його класифікують як victim; у протилежному випадку – як NOD. У разі кількох збоїв процедура може повторюватися для підвищення достовірності. Метод враховує такі параметри як кількість повторів і спосіб перемішування тестів, що може значно збільшити час виконання.

Варто також звернути уваги на роботу [5], де був представлений FLAKEFLAGGER – інструмент для виявлення нестабільних тестів на основі моделі машинного навчання. Для кодування тестів використано набір із 8 числових і 8 булевих ознак, що позначали наявність запахів коду, однак пізніше булеві ознаки вже не враховувалися через низьку інформативність. У дослідженні 24 Java-проєктів модель досягла значення MCC = 0.65. Був також представлений розширений набір FLAKE16, який включав додаткові метрики, як наприклад кількість I/O-операцій файлової системи та пікове споживання пам'яті. Оцінювання на 26 Python-проєктах показало, що FLAKE16 стабільно перевищує FLAKEFLAGGER за ефективністю виявлення як NOD, так і victim тестів.

Головним внеском роботи [6] необхідно виділити представлення та емпіричну оцінку CANNIER – гібридної системи виявлення нестабільних тестів, яка поєднує підходи повторного запуску та машинного навчання. CANNIER використовує ймовірності нестабільності, передбачені моделлю машинного навчання, як припущення для зменшення простору повторних виконань. Оцінювання на 89 668 тестах у 30 Python-проєктах (з використанням стратифікованої 10-кратної крос-валідації та MCC) показало, що CANNIER зменшив витрати часу до 88%, зберігаючи продуктивність (MCC $\approx$ 0.53–0.65). При застосуванні до класифікаційного етапу, алгоритмом

IDFLAKIES досягнуто скорочення часу на 84% при мінімальній втраті точності. Аналіз також показав, що комбінація середніх значень динамічних ознак тестів покращує роботу моделі, а такі характеристики, як багаточисельність, I/O-навантаження та використання пам'яті, тісно корелюють із ймовірністю нестабільності тестових сценаріїв.

У цій роботі [7] запропоновано FlaKat – систему машинного навчання для класифікації нестабільних тестів відповідно до їхніх першопричин. З використанням міжнародного набору даних International Dataset of Flaky Tests (IDoFT), проводиться аналіз різних стратегій представлення та навчання. Розглянуто три методи векторизації – TF-IDF (TF – term frequency, IDF – inverse document frequency), doc2vec і code2vec – а також методи зменшення розмірності (LDA (Latent Dirichlet Allocation), PCA (Principal Component Analysis), UMAP (Uniform Manifold Approximation and Projection)) для оцінки збереження структури у просторі ознак. Для розв'язання проблеми дисбалансу даних застосовано вибіркові методи SMOTE і Tomek Link, а для класифікації протестовано моделі KNN, SVM і Random Forest. Експериментальні результати показують, що поєднання TF-IDF і Random Forest забезпечує найвищу точність класифікації ($F1 = 0.67$). Додатково введено нову метрику здатності моделі виявляти нестабільні тести (Flakiness Detection Capacity – FDC), яка оцінює узгодженість класифікатора з використанням інформаційно-теоретичного аналізу. Вона виявилася більш стабільною, ніж $F1$, що підкреслює її ефективність для оцінювання систем виявлення нестабільних тестів.

Проте, беручи до уваги вище зазначену наукову документацію, питання, пов'язане з розробка інтелектуального підходу удосконалення стабільності автоматизованих систем тестування, все ще залишається недостатньо дослідженим та потребує подальшого опрацювання.

Формулювання мети дослідження

Метою роботи є розробка інтелектуального підходу удосконалення стабільності автоматизованих систем тестування на прикладі інструменту для автоматизації операцій у веб-браузері Selenium із використанням навчання з підкріплення.

Викладення основного матеріалу дослідження

Selenium представляється в якості структурованого багаторівневого автоматизаційного фреймворку, що робить взаємодію між розроблюваними тестовими сценаріями та тестованими веб-застосунками більш прямою. Його модульна архітектура складається з 4 головних компонентів, а саме Selenium IDE (Integrated Development Environment – інтегроване середовище розробки), Selenium Client (WebDriver), Selenium Server, and Selenium Grid, що разом формують ієрархічний цикл керування, тобто $IDE \rightarrow Client \rightarrow Server \rightarrow Grid$ [8].

Відтак, на базовому рівні знаходиться Selenium IDE, що функціонує в якості середовища реєстрації та відтворення тестів. Ця ланка дозволяє враховувати взаємодії на рівні користувачів, генеруючи тестові сценарії які надалі можуть бути відтвореними для регресійного тестування. Незважаючи на це, цей рівень залежить від статично записуваних локаторів та послідовностей інструкцій.

Над IDE розташовується Selenium Client (WebDriver), в якому регулюються логічні взаємозв'язки між тестовими сценаріями та веб-застосунком. Відтак, відбувається переклад інструкцій користувачів в операції на рівні браузера за допомогою WebDriver API та інтерфейсу RemoteWebDriver. Варто зауважити, що цей взаємозв'язок є синхронізованим та послідовним, тобто кожна інструкція може бути виконана тільки після отримання відгуку.

Selenium Server виконує роль проміжного рівня, що зв'язує Client з драйверами браузера. За посередництва нього, вхідні команди встановлюються у чергу, які потім надсилаються до відповідних браузерних інстансів (instance – екземпляр). Втім, він також функціонує детерміністичним чином, не проводячи аналітичне спостереження за метриками продуктивності, або виникненням аномалій у відгуках. Внаслідок цього, навіть за умови виявлення неналежних станів на рівнях Client та IDE, цей рівень залишатиметься позбавленим необхідної адаптивності.

Selenium Client та Selenium Server функціонують у поєднанні, створюючи Клієнтсько-серверний канал, що функціонує в якості базового комунікаційного інструмента Selenium. Client надсилає інструкції через WebDriver API, в той час, як Server виконує ці інструкції за допомогою Application Driver API.

Найвища ланка в окресленій ієрархічній структурі, тобто Selenium Grid, необхідна для проведення розподіленого та паралельного виконання тестових сценаріїв на декількох вузлах. Ця ланка працює на базі центрально-вузлової архітектури (hub-and-node), що забезпечує розширюваність тестового середовища, однак позбавлена аналітичної здатності під час виконання тестів. Розподіл задач залежить суто від попередньо визначених налаштувань вузлів та статичних налаштувань планування.

Загалом, ці рівні формують організований стек детермінованих процесів, тобто будь-яке очікування, блокування та виконання відбувається без огляду на контекстну динаміку роботи системи. Отже, кожний рівень має наділятися виділеним адаптивним підмодулем, здатним враховувати локальні порушення стабільності та відповідати на них, тобто агент з підкріпленням для IDE, призначений для адаптації часових параметрів відтворення команд, агент-стабілізатор виконання в межах каналу Client–Server для регуляції синхронізації, та координаційний агент для Grid з метою врівноваження розподілених обчислювальних задач між вузлами [9].

Відтак, Інтелектуальний рівень з підкріпленням (Intelligent Reinforcement Layer – IRL) слугує в якості мета-архітектури, що охоплює всі традиційні складові Selenium в єдину адаптивну систему. На базовому рівні, в межах IRL встановлюється пул колективного досвіду, який збирає телеметричні дані з усіх ланок ієрархічної структури Selenium. До цих телеметричних даних насамперед входять час затримки взаємодії з браузером, частота змін DOM, показники ідентифікаційної точності локаторів, рівні використання ресурсів вузлів та коефіцієнти успішності виконання.

Структура IRL передбачає наявність 3 підрівнів адаптації:

- локальна адаптація на рівні IDE, де проводиться тонке налаштування часових параметрів інструкцій та відновлення старих прив'язок до локаторів з огляду на мінливість DOM;
- середньорівнева адаптація в межах каналу Client–Server, де відбувається динамічне налаштування інтервалів очікування, стратегій синхронізації та зміна локаторів відповідно до вихідної інформації після виконання тесту;
- глобальна адаптація на рівні Grid, в контексті якої виконується збалансування робочого навантаження вузлів задля того, щоб стабілізаційні заходи, виявлені локально, поширювалися по всій системі.

Загальна процедура навчання всіх агентів в рамках IRL проводиться згідно з формулюванням навчання з підкріпленням (reinforcement-learning – RL). Функціонування кожного агента A_i формулюється, враховуючи пару стан-операція (s_i, a_i) , де в якості стану розглядається робочі умови, контекстно пов'язані з рівнем агента, а операція відповідає конкретному коригувальному заходу в межах системи. Загальний принцип оновлення на базі Q-навчання формулюється таким чином:

$$Q_{t+1}(s_i, a_i) = Q_t(s_i, a_i) + \alpha \left[R_t + \beta \max_{a'} Q_t(s'_i, a') - Q_t(s_i, a_i) \right] \quad (1)$$

де α відповідає коефіцієнту навчання, що визначає ступінь впливу нових вивчених закономірностей на оновлення моделі, а β позначає коефіцієнт дисконтування, що регулює співвідношення між поточними та майбутніми винагородами, s'_i відображає наступний спостережуваний стан після застосування операції a_i , а R_t є скалярною функцією винагороди, що відображає загальне покращення стабільності.

Об'єднана функція винагороди виражається так:

$$R_t = \mu_1 - \mu_2 - \mu_3 \quad (2)$$

де коефіцієнти μ_1, μ_2, μ_3 відображають динамічні вагові коефіцієнти, що відповідають ступеням точності, логічної узгодженості та продуктивності, відповідно. Ці вагові коефіцієнти змінюються під час проведення тестування за посередництва мета-навчання. Кожна підсистема спрямована на налаштування цієї базової логіки відповідно до свого функціонального контексту та області видимості.

Перший підмодуль, Агент підкріплення IDE (IDE Reinforcement Agent – IDEA) визначає принцип виконання зареєстрованих тестових операцій, фактично зосереджуючись на виявленні аномалій під час виконання тестових сценаріїв та подальшій автономній модифікації часових параметрів виконання операцій, вибору локаторів та стратегій повторного відтворення.

На часовому кроці t , IDEA розглядає тестове середовище в якості свого локального вектора станів, визначеного так:

$$s_t = \{t_{\text{навантаж}}, \delta_{\text{мінл}}, r_{\text{лок}}\} \quad (3)$$

де $t_{\text{навантаж}}$ відповідає зафіксованому часу навантаження сторінки або часу візуалізації елемента, $\delta_{\text{мінл}}$ позначає ступінь мінливості DOM, що відображає структурні зміни, або зміни в атрибутах з моменту останнього виконання, а $r_{\text{лок}}$ є балом точності елементного локатора, зафіксованого в даний момент часу, що обчислюється на основі ступеня збігу, або частоти успішних виконань за весь час.

Регуляційна політика IDEA формується на основі модифікованої функції винагород:

$$R_t = \gamma_1 (1 - \delta_{\text{мінл}}) + \gamma_2 r_{\text{лок}} - \gamma_3 t_{\text{навантаж}} \quad (4)$$

де вагові коефіцієнти $\gamma_1, \gamma_2, \gamma_3$ визначають відносну важливість структурної стабільності, точності на рівні локаторів та ефективності.

Враховуючи окреслене формулювання, IDEA набуває здатності знижувати час зайвого очікування, при цьому уникаючи передчасних або помилкових взаємодій. З часом ця підсистема досягає рівноваги між швидкістю виконання та стабільністю. У процесі навчання агент поступово формує таблицю Q-значень, де кожне Q-значення відображає очікувану корисність вибору певної дії у конкретному стані середовища. В остаточному підсумку, Q-значення оновлюються згідно зі згаданим вище правилом оновлення.

На адаптивному рівні каналу Client–Server відбувається регуляція стабільності виконання Selenium в реальному часі, функціонуючи між WebDriver API та Application Driver API. Отже, регуляція стабільності проводиться шляхом керування таких нечітких параметрів та сутностей, як ступінь мережевої затримки, асинхронні відгуки

на інструкції, або затримка синхронізації драйвера. З огляду на це, в межах структури цього рівня передбачена наявність двох кооперативних агентів. а саме Агента стабільності виконання (Execution Stability Agent – QESA) та Агента ієрархічної адаптації локаторів (Hierarchical Locator Adaptation Agent – HLAA).

QESA перехоплює телеметричні дані, в даному випадку ступені затримки запитів, часи відгуку DOM та тимчасові затримки підтвердження зі сторони сервера, та надсилає їх до колективного пулу досвіду. У той же час, HLAA функціонує у поєднанні з QESA, проводячи спостереження за локаторною точністю ідентифікації потрібного елемента в інтерфейсі за різних умов під час виконання тестового сценарію. У випадку помилкової ідентифікації локатором, або перевищення часу очікування, за рахунок HLAA використовуються дані з колективного пулу для динамічного перевибирання, або реконструкції елементних локаторів виходячи з вивчених стратегій відновлення структури, а саме, наприклад, ступеня контекстної спорідненості DOM, або метрик подібності атрибутів.

RL логіка клієнтсько-серверного каналу передбачає розширення функції винагород за рахунок додаткових системних та комунікаційних метрик. Вектор станів у часі t визначається так:

$$s_t = \{\tau_{\text{інстр}} \cdot \lambda_{\text{мереж}} \cdot \sigma_{\text{синх}} \cdot \rho_{\text{лок}}\} \quad (5)$$

де $\tau_{\text{інстр}}$ – середнім часом відгуку на інструкції, $\lambda_{\text{мереж}}$ – мережевою або пакетною затримкою, $\sigma_{\text{синх}}$ – зміщеннями синхронізації між відгуками клієнта та сервера, а $\rho_{\text{лок}}$ – індексом точності поточно активних локаторів, відповідно до даних з HLAA.

Колективна функція винагороди агента формулюється так:

$$R_t = \mu_1 - \mu_2 - \mu_3 - \mu_4 \quad (6)$$

де μ_4 реалізує штраф до нестабільної синхронізації та затриманих підтверджень.

Обидва агенти оновлюють свої внутрішні Q-таблиці на основі досвіду з колективного пулу даних та даних IDEA, відповідно до усталеного формулювання. Варто зазначити, що Q-значення з IDEA виконують роль локальних апріорних даних, що спрямовують початкове налаштування політики обох агентів каналу. Тобто ознаки роботи системи, що забезпечує її стабільність, навчені під час відтворення тестових сценаріїв на першій ланці, безпосередньо впливають на прийняття рішень під час виконання.

З часом, QESA та HLAA спільно оптимізують виконання 3 адаптивних механізмів, а саме:

- корекція інтервалу очікування, в ході якого відбувається динамічна модифікація явних і неявних очікувань відповідно до зафіксованих параметрів затримок виконання інструкцій та індикаторів стану DOM;
- калібрування політики повторного відтворення, де проводиться вивчення оптимальної частоти повторних відтворень при тимчасових збоях, запобігаючи як зайвим очікуванням, так і передчасним завершенням тестових сценаріїв або окремих операцій взаємодії з інтерфейсом;
- автоматична зміна локаторів відповідно до різних стратегій пошуку елементів (XPath, CSS-селектори, ID, текстові атрибути).

З концептуальної точки зору, клієнтсько-серверний рівень проводить зв'язок між адаптацією на мікрорівні (IDEA) з координацією на макрорівні (до Grid).

Багатоагентний координаційний рівень Grid є ланкою найвищої абстракції IRL, що розширює адаптивність Selenium для роботи в багатовузлових інфраструктурах тестування, що у цьому дослідженні має назву Середовища навчання Grid (Grid Learning Environment – GLE). В межах цієї архітектури, кожний вузол Selenium Grid виконує функцію автономного агента RL, здатного враховувати свій локальний робочий стан та надсилати необхідну інформацію до контролера взаємодії між агентами π_H .

Вектор станів кожного вузла формулюється так:

$$s_i = \{L_i, T_i, F_i\} \quad (7)$$

де L_i є середньою затримкою мережі, або виконання інструкції в межах вузла i , T_i є кількістю успішно виконаних тестових сценаріїв за одиницю часу, а F_i є локальним коефіцієнтом нестабільності, тобто часткою нестабільних, або провалених тестових сценаріїв за останній час.

На цьому рівні визначається глобальна функція винагород, яка збирає локальну інформацію з усіх вузлів N :

$$R_{\text{глоб}} = \frac{1}{N} \sum_{i=1}^N (\mu_1 S_i - \mu_2 L_i - \mu_3 F_i) \quad (8)$$

де S_i відображає кількість успішних виконань тестових сценаріїв в межах вузла i .

В той час, як оновлення Q-значень відбувається за стандартною процедурою, на найвищому рівні структури GLE знаходиться, як було згадано, контролер взаємодії між агентами π_H , що визначає принцип розподілу тестових операцій по вузлах. Для цієї задачі застосовуються політики вибору на базі функції softmax, виходячи з навчених Q-значень:

$$\pi_H(A_i | S) = \frac{e^{\kappa Q_i(S_i)}}{\sum_{j=1}^N e^{\kappa Q_j(S_j)}} \quad (9)$$

де κ відповідає температурному параметру, що регулює співвідношення між виявленням нових закономірностей та використанням старих.

Відтак, описана архітектура оглядово наведена на рисунку 1

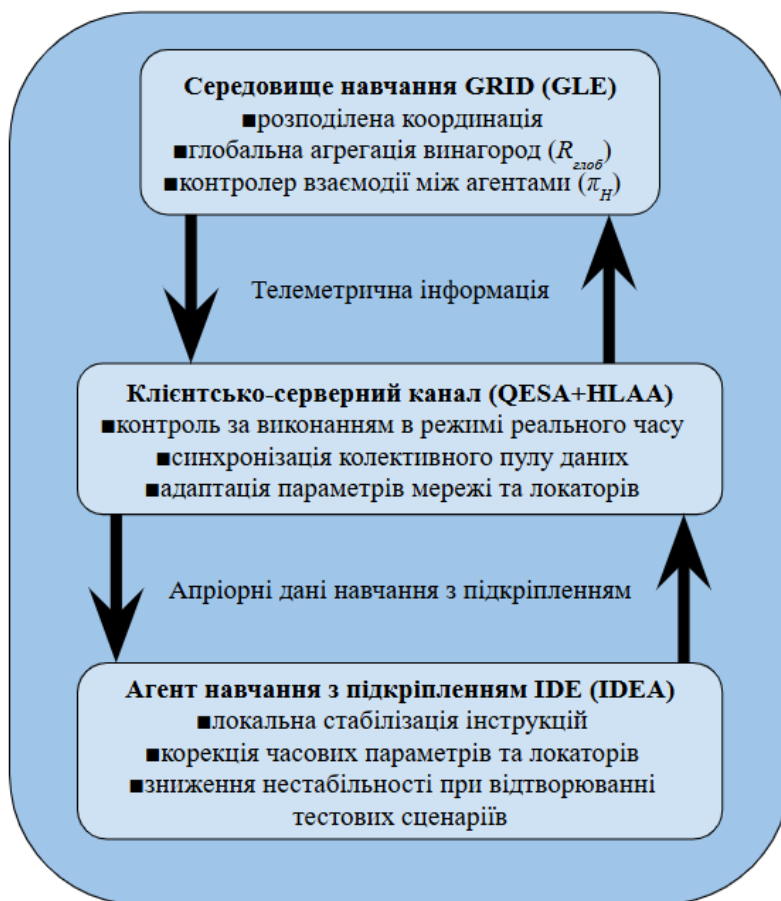


Рис. 1. Загальна архітектура теоретичної моделі

Варто зазначити, що така структура, де кожний вузол незалежно покращує свою ефективність за посередництва навчання з підкріпленням, однак де вони всі мають колективну задачу досягнення глобальної оптимізації за рахунок контролера, є аналогічною до кооперативної навчальної поведінки систем ройового інтелекту (swarm intelligence).

Впровадження мета-архітектури IRL у структуру Selenium має низку концептуальних переваг над сучасними фреймворками автоматизації. По-перше, на відміну від класичного Selenium WebDriver, який не забезпечує вбудованих механізмів розумного очікування або автоматичних повторних виконань операцій, система з IRL самостійно здатна регулювати синхронізацію та відновлення етапів тесту [10]. По-друге, такі інструменти, як Cypress та Playwright, вже впроваджують деякі засоби підвищення стабільності (автоматичне очікування елементів, регуляція повторів тестових сценаріїв). Проте в Cypress і Playwright ці можливості реалізовані як фіксовані алгоритми, тоді як IRL застосовує адаптивне навчання, тобто агенти підлаштовують стратегії під поведінку конкретного застосунку та продовжують оптимізувати їх з накопиченням досвіду. Таким чином, IRL потенційно перевершує зазначені фреймворки у здатності реагувати на нові типи змін середовища, які не були явно передбачені розробниками інструмента. По-третє, IRL-архітектура зберігає широкую сумісність Selenium (підтримку різних мов програмування, браузерів і платформ), одночасно додаючи інтелектуальні можливості. Для порівняння, Cypress є обмеженим JavaScript, а його архітектура накладає певні обмеження (наприклад, відсутність повноцінної підтримки багатовіконності), тоді як IRL може бути інтегрований у вже наявні різномовні фреймворки на базі Selenium.

Окремо слід відзначити інструмент Percy, що спеціалізується на візуальному тестуванні. Percy використовує рушій комп'ютерного зору для виявлення візуальних відхилень в інтерфейсі користувача шляхом аналізу структур

DOM [11]. Це є корисним для контролю регресій зовнішнього вигляду веб-застосунку, проте Percy не вирішує проблем функціональної стабільності тестів, як наприклад, непередбачуваних збоїв через динамічні зміни на сторінках. Запропонований підхід, навпаки, зосереджений на функціональній надійності тестування: його метою є забезпечення того, щоб автоматизовані тести стабільно проходили при відсутності помилок у самому додатку. Отже, IRL та Percy можна розглядати як взаємодоповнювальні засоби забезпечення якості: IRL підтримує коректність і стабільність виконання тестових сценаріїв, тоді як Percy додатково перевіряє відсутність небажаних візуальних змін.

В цілому, результати теоретичного дослідження свідчать, що запропонована багаторівнева IRL-архітектура здатна підвищити стабільність систем автоматизованого тестування. Вона робить тестові сценарії більш адаптивними до змін у додатку та загальній інфраструктурі тестування, зменшує кількість хибних тестових сценаріїв і знижує потребу в ручному налаштуванні тестів. Це вигідно відрізняє підхід IRL від наявних рішень і підкреслює його потенціал як інтелектуального засобу підвищення надійності автоматизованого тестування.

Висновки

Порівняльний аналіз показав, що в той час, як традиційні інструменти, такі як Cypress та Playwright, реалізують ригідні процедури стабілізації, запропонована архітектура уособлює досягнення адаптивізації та стабілізації шляхом постійного навчання та аналізу телеметричного зворотного зв'язку, що сприяє спрощеному проведенню сталого самооптимізаційного тестування. У майбутніх дослідженнях доцільно зосередитись на практичній реалізації запропонованої архітектури IRL та її оцінюванні у реальних умовах розробки. Подальший розвиток може включати впровадження методів глибокого навчання з підкріпленням, таких як Deep Q-Network або Actor-Critic, для роботи з більш складними станами в інфраструктурах тестування. Перспективним напрямом є також інтеграція IRL у пайплайни CI/CD для забезпечення безперервного навчання на основі даних тестових прогонів у реальному часі.

Список використаної літератури

1. Ahmad A., Leifler O., Sandahl K. An Evaluation of Machine Learning Methods for Predicting Flaky Tests. Linköping University, Sweden. DiVA Portal, 2020. URL: <https://www.diva-portal.org/smash/get/diva2:1537340/FULLTEXT01.pdf>
2. Dutta S., Shi A., Choudhary R., Zhang Z., Jain A., Misailovic S. Detecting Flaky Tests in Probabilistic and Machine Learning Applications. Cornell University, 2020. 14 p. URL: <https://saikatdutta.web.illinois.edu/papers/flash-issta20.pdf>
3. Akli A., Haben G., Habchi S., Papadakis M., Le Traon Y. FlakyCat: Predicting Flaky Tests Categories using Few-Shot Learning. University of Luxembourg, 2023. 12 p. URL: <https://orbilu.uni.lu/bitstream/10993/55848/1/FlakyCat.pdf>
4. Lam W., Oei R., Shi A., Marinov D., Xie T. iDFlakies: A Framework for Detecting and Partially Classifying Flaky Tests. 12th IEEE Conference on Software Testing, Validation and Verification (ICST), 2019. P. 312–322. DOI: 10.1109/ICST.2019.00038
5. Alshammari A., Morris C., Hilton M., Bell J. FlakeFlagger: Predicting Flakiness without Rerunning Tests. Proceedings of the International Conference on Software Engineering (ICSE), 2021.
6. Parry O., Kapfhammer G. M., Hilton M., McMinn P. Empirically Evaluating Flaky Test Detection Techniques Combining Test Case Rerunning and Machine Learning Models. Empirical Software Engineering. 2023. Vol. 28. No. 72. DOI: 10.1007/s10664-023-10307-w
7. Lin S., Liu R., Tahvildari L. FlaKat: A Machine Learning-Based Categorization Framework for Flaky Tests, 2024. arXiv:2403.01003 [cs.SE]. URL: <https://arxiv.org/pdf/2403.01003v1>
8. Nyamathulla S., Ratnababu P., Shaik N. S. A Review on Selenium Web Driver with Python. Annals of the Romanian Society for Cell Biology, 2021. P. 16760–16768.
9. Alferidah S. K., Ahmed S. Automated Software Testing Tools. 2020. Proceedings of the International Conference on Computing and Information Technology (ICCIT-1441). IEEE.
10. Ulili S. Playwright vs Puppeteer vs Cypress vs Selenium (E2E Testing). 2025. Better Stack Community Comparisons. URL: <https://betterstack.com/community/comparisons/playwright-cypress-puppeteer-selenium-comparison/>
11. Revolutionizing Visual Testing on Web using Automation and AI: Halodoc's Journey with Percy. 2024. Percy Visual Testing Blog, Halodoc. URL: <https://blogs.halodoc.io/percy-web/>

References

1. Ahmad, A., Leifler, O., & Sandahl, K. (2020). *An evaluation of machine learning methods for predicting flaky tests*. Linköping University. <https://www.diva-portal.org/smash/get/diva2:1537340/FULLTEXT01.pdf>
2. Dutta, S., Shi, A., Choudhary, R., Zhang, Z., Jain, A., & Misailovic, S. (2020). *Detecting flaky tests in probabilistic and machine learning applications* (14 pp.). Cornell University. <https://saikatdutta.web.illinois.edu/papers/flash-issta20.pdf>
3. Akli, A., Haben, G., Habchi, S., Papadakis, M., & Le Traon, Y. (2023). *FlakyCat: Predicting flaky tests categories using few-shot learning* (12 pp.). University of Luxembourg. <https://orbilu.uni.lu/bitstream/10993/55848/1/FlakyCat.pdf>

4. Lam, W., Oei, R., Shi, A., Marinov, D., & Xie, T. (2019). iDFlakies: A framework for detecting and partially classifying flaky tests. In *2019 IEEE 12th International Conference on Software Testing, Verification and Validation (ICST)* (pp. 312–322). <https://doi.org/10.1109/ICST.2019.00038>
5. Alshammari, A., Morris, C., Hilton, M., & Bell, J. (2021). *FlakeFlagger: Predicting flakiness without rerunning tests*. Proceedings of the International Conference on Software Engineering (ICSE).
6. Parry, O., Kapfhammer, G. M., Hilton, M., & McMinn, P. (2023). Empirically evaluating flaky test detection techniques combining test case rerunning and machine learning models. *Empirical Software Engineering*, 28(72). <https://doi.org/10.1007/s10664-023-10307-w>
7. Lin, S., Liu, R., & Tahvildari, L. (2024). *FlaKat: A machine learning-based categorization framework for flaky tests*. arXiv:2403.01003. <https://arxiv.org/pdf/2403.01003v1>
8. Nyamathulla, S., Ratnababu, P., & Shaik, N. S. (2021). A review on Selenium Web Driver with Python. *Annals of the Romanian Society for Cell Biology*, 16760–16768.
9. Alferidah, S. K., & Ahmed, S. (2020). Automated software testing tools. In *Proceedings of the International Conference on Computing and Information Technology (ICCIT-1441)*. IEEE.
10. Ulili, S. (2025). *Playwright vs Puppeteer vs Cypress vs Selenium (E2E Testing)*. Better Stack Community. <https://betterstack.com/community/comparisons/playwright-cypress-puppeteer-selenium-comparison/>
11. Halodoc. (2024). *Revolutionizing visual testing on web using automation and AI: Halodoc's journey with Percy*. Percy Visual Testing Blog. <https://blogs.halodoc.io/percy-web/>

Дата першого надходження рукопису до видання: 23.11.2025
Дата прийнятого до друку рукопису після рецензування: 19.12.2025
Дата публікації: 31.12.2025

А. О. КАРГІН

доктор технічних наук, професор,
професор кафедри комп'ютерних наук
ДВНЗ «Приазовський державний технічний університет»
ORCID: 0000-0003-2885-9071

О. І. ПРОНІНА

кандидат технічних наук, доцент,
доцент кафедри комп'ютерних наук
ДВНЗ «Приазовський державний технічний університет»
ORCID: 0000-0001-7085-8027

О. Є. П'ЯТИКОП

кандидат технічних наук, доцент,
доцент кафедри комп'ютерних наук
ДВНЗ «Приазовський державний технічний університет»
ORCID: 0000-0002-7731-3051

Є. А. ПЕРЦЕВ

магістр кафедри комп'ютерних наук
ДВНЗ «Приазовський державний технічний університет»
ORCID: 0009-0009-4155-9031

ТЕХНОЛОГІЯ ПРОЕКТУВАННЯ ШТУЧНОГО ІНТЕЛЕКТУ, ЩО ПІДТРИМУЄ АВТОНОМНІ РОЗПОДІЛЕНІ СИСТЕМИ

Потреба в більш досконалих автономних системах підтверджується зростаючим попитом не лише військової сфери, а й цивільної спільноти. Зусилля розробників автономних систем на базі Штучного Інтелекту (ШІ) зосереджені на підвищенні рівня автономності майбутніх систем. Відомі моделі штучного інтелекту не можуть комплексно вирішити проблему автономності через розрив між двома парадигмами «обробка даних від датчиків» та «прийняття рішень на основі знань експерта в предметній області». В роботі обговорюється нова модель ШІ (Feeling AI, FAI), як гібридний ШІ, що базується на когнітивних моделях, запозичених у живих істот. Так як База Знань (БЗ) FAI представлена набором незалежних універсальних Гранул Знань (ГЗ), що взаємодіють за моделлю подій, то подієво-керована архітектура (Event-Driven Architecture, EDA) є найбільш придатною для реалізації обчислювань у FAI. Основною проблемою застосування EDA є необхідність формування реакції в режимі реального часу компонентами FAI на основі сенсорних даних, отриманих асинхронно у різний час і з різними характеристиками старіння. У FAI використана нова когнітивна модель врахування старіння даних, яка підтримує поступове забування подій з часом та оцінює їхню поточну актуальність за допомогою нечіткого фактору впевненості (Certainty Factor, CF). Проектування системи сприйняття FAI за технологією EDA розглядається на прикладі обслуговування автономної системи розосереджених об'єктів (automatic Garage Complex, GC). До складу GC входить колісний робот що отримує додатково до стаціонарної СС дані щодо стану окремого боксу. Наведено окремі етапи логічного проектування, починаючи з експертних знань, їх формалізації у вигляді множини ГЗ, відображення структури знань у БЗ на апаратне забезпечення, та уявлення знань, якими оперує Event Broker EDA у таблиці параметрів. На етапі фізичного проектування розроблено програмну реалізацію ядра FAI мовою C++ з використанням бібліотеки ZeroMQ для забезпечення високопродуктивного асинхронного обміну повідомленнями. Запропонована архітектура реалізує механізм диспетчеризації подій через компоненти-маршрутизатори та обчислювальні потоки, що дозволяє масштабувати систему шляхом розпаралелювання виконання функції обробки гранул знань.

Ключові слова: штучний інтелект, когнітивна модель старіння даних, автономні системи, нечіткий фактор впевненості, гранула знань, подієво-керована архітектура.

A. O. KARGIN

Doctor of Technical Sciences, Professor,
Professor at the Department of Computer Science
PSTU "Pryazovskyi State Technical University"
ORCID: 0000-0003-2885-9071

O. I. PRONINA

Candidate of Technical Sciences, Associate Professor,
Associate Professor at the Department of Computer Science
PSTU "Pryazovskyi State Technical University"
ORCID: 0000-0001-7085-8027

O. YE. PIATYKOP

Candidate of Technical Sciences, Associate Professor,
Associate Professor at the Department of Computer Science
PSTU "Pryazovskyi State Technical University"
ORCID: 0000-0002-7731-3051

YE. A. PERTSEV

Master at the Department of Computer Science
PSTU "Pryazovskyi State Technical University"
ORCID: 0009-0009-4155-9031

DESIGN TECHNOLOGY OF THE ARTIFICIAL INTELLIGENCE THAT SUPPORTS AUTONOMOUS DISTRIBUTED SYSTEMS

The need for more advanced autonomous systems is confirmed by the growing demand not only from the military domain but also from the civilian community. The efforts of developers of autonomous systems based on Artificial Intelligence (AI) are focused on increasing the level of autonomy of future systems. Known AI models cannot comprehensively solve the problem of autonomy due to the gap between the two paradigms of "processing data from sensors" and "making decisions based on expert knowledge in verbal form." The paper discusses a new model of AI, known as Feeling AI (FAI), which is a hybrid AI based on cognitive models borrowed from living beings. Since the FAI's Knowledge Base (KB) is represented by a set of independent universal Knowledge Granules (KG) that interact according to the event model, the Event-Driven Architecture (EDA) is most suitable for implementing the processing in FAI. The primary challenge of using EDA is that its components must process data in real-time when they are received at different times and with varying aging characteristics. FAI employs a novel cognitive model of data aging, which supports the gradual forgetting of events over time and estimates their current relevance using a fuzzy Certainty Factor (CF). The design of an FAI's perception system, based on EDA technology, is considered using the example of servicing an autonomous system of distributed objects, such as the Automatic Garage Complex (AGC). The AGC includes a wheeled robot that receives data on the state of a separate box in addition to the stationary perception system. Separate stages of logical design are presented, starting with expert knowledge, its formalization in the form of a set of KGs of the KB, mapping the structure of the KB into hardware deployment, and representing the knowledge that the Event Broker EDA operates with in the parameter table. At the physical design stage, a software implementation of the FAI core was developed in C++ using the ZeroMQ library to provide high-performance asynchronous messaging. The proposed architecture implements an event dispatching mechanism through router components and computational threads, which allows the system to be scaled by parallelizing the execution of the knowledge granule processing function.

Key words: artificial intelligence, cognitive data aging model, autonomous systems, fuzzy certainty factor, knowledge granule, event-driven architecture.

Постановка проблеми

Один із напрямів використання інтелектуальних інформаційних технологій пов'язаний з автономними системами [1, 2, 3]. Особливо актуальними є системи, в яких завдяки ШІ досягається безпека об'єктів. Сучасний рівень розвитку Інтернету Речей і ШІ сприяє їх інтеграції при створенні автономних розосереджених об'єктів. Централізований збір всіх даних не завжди є ефективним рішенням для певного класу об'єктів, тому застосовуються гібридні системи де окремі компоненти сенсорної підсистеми можуть бути реалізовані за допомогою мобільних пристроїв (колісні роботи, дрони), які оснащені потрібними сенсорами для отримання даних. Головне питання що виникає при створенні такого класу систем це як об'єднати окремі компоненти системи, а точніше моделі ШІ що їх підтримують, у цілісну однорідну структуру знань [4]. Створення ШІ для автономних технічних систем у тому числі з розосередженими об'єктами які задовольняють певному рівню автономності [5] на сьогодні є актуальним завданням.

Аналіз останніх досліджень і публікацій

Сучасний стан та тенденції ШІ, які виявляються зараз (Spatial computing, Polyfunctional robots) і є такими що націлені на об'єднання фізичного та цифрового світів, створюючи нове обчислювальне середовище, де віртуальні елементи інтегровані у реальне оточення, добре підходять для автоматизації послуг, що виконуються у віртуальному світі Інтернету [6]. Для вирішення аналогічної проблеми у фізичному світі запропоновано RT-X проект метою якого є створення загального робототехнічного мозку (GRB) шляхом збору даних, ресурсів та коду, навичок, яким роботи вже навчені [7, 8]. Планується, що GRB матиме підтримку GenAI [9, 10]. Такий ШІ із символічним механізмом прийняття рішень та попередньо навченим набором різних навичок у фізичному світі є одним із типів гібридного ШІ, спрямованого на масштабні роботизовані застосування. Використання цього підходу для вирішення згаданої проблеми обмежене з двох причин. По-перше, це обмежує автономність через постійне підключення до Інтернету та велику базу знань (БЗ), необхідну для охоплення всіх можливих ситуацій [11]. По-друге, на сьогодні автономні роботизовані системи не мають потужних комп'ютерних засобів [12, 13]; отже, вони не можуть виконувати програмне забезпечення гібридного ШІ класу GRB. У статтях [4, 14, 15] пропонується інша модель гібридного ШІ для робототехніки (Feeling AI, FAI), яка базується на когнітивних моделях, що запозичені у живих істот. На відміну від GRB, FAI спрямований на невеликі роботизовані системи, що підтримуються звичайним комп'ютерним обладнанням та програмним забезпеченням, які традиційно використовуються на периферійному рівні Інтернету речей та Розумних систем [16, 17, 18]. Різні питання організації та впровадження когнітивних моделей як основи FAI обговорювалися в попередніх публікаціях: сприйняття, що підтримується механізмом старіння даних [19], прийняття рішень на основі когнітивних моделей [4, 14, 15]. Оскільки архітектура FAI складається з незалежних компонентів, які взаємодіють за допомогою повідомлень один з одним, то відкривається можливість об'єднання компонентів на основі (Event-Driven Architecture, EDA) [20-22]. Теоретичні аспекти застосування EDA у FAI розглянуті в роботі [23].

Формулювання мети дослідження

Стаття присвячена дослідженню інформаційної технології проектування підсистеми сприйняття FAI для автономних систем із розосередженими об'єктами.

Викладення основного матеріалу дослідження

Теоретичним підґрунтям технології є модель FAI запозичує те, як нейронна мережа найпростіших живих істот, як біологічна форма представлення та обробки знань, формує поведінку [4, 14, 15]. БЗ живих істот представляє знання про реакції істоти на всі можливі ситуації. Для формування кожної окремої реакції використовуються три види знань: Сенсорний Шаблон (СШ), Дії Шаблон (ДШ) та Конструктор Реакцій (КР). Об'єднані разом, вони представляють Прототип Реакцій (ПР), як показано на рис. 1. СШ усіх ПР FAI образують Систему Сприйняття (СС) FAI.

Для універсального представлення всіх трьох компонентів було запроваджено модель ГЗ, яка моделює окремих нейрон живих істот. Це універсальний елементарний фрагмент знань, подібний до цеглинок у будівництві [4]. На рис. 1 ГЗ показані колами з цифрами всередині. ГЗ виконує дві ролі: представлення та обробку. Модель представлення визначає сенс ГЗ через сенс інших ГЗ нижчого рівня абстракції, використовуючи параметризовані семантичні зв'язки.

$$\langle N, know, \{ \langle M, (a, b, v, g) \rangle, M_l \in \Omega_N \} \rangle \quad (1)$$

де N – ідентифікатор ГЗ; $know$ – знакова модель ГЗ; $\Omega_N = \{M_l\}_{l=1, 2, \dots, L}$ – набір ГЗ нижчих рівнів концептуалізації, які використовуються для визначення сенсу гранули N ; M_l – ідентифікатор ГЗ нижнього рівня; a – впевненість, b – часова затримка, v – швидкість старіння даних, g – повнота знань.

Отже, СШ – це структурований набір ГЗ (рис. 1), де ГЗ нульового рівня отримують безпосередньо дані від датчиків. ГЗ верхніх рівнів вже представляють концептуалізовані знання й отримують на вході значення впевненості і на виході також обчислена впевненість, що входні дані ГЗ відповідають шаблону, знання про який зберігає ця ГЗ. ГЗ має L входів та один вихід, значення яких є нечіткими факторами впевненості (Certainty Factor, CF): це нечітке число з гауссовою функцією належності [4].

$$m_x = \exp(-(x-\alpha)^2/2\beta^2), \forall x \in [-q, +q] \quad (2)$$

з параметрами: $-1, 0 \leq \alpha \leq +1, 0$ – значення впевненості; β – часовий інтервал, що минув з моменту останньої зміни значення впевненості; $|q| > +1$. Для полегшення обчислень було введено числову оцінку CF, яка виражає результат старіння даних з часом відповідно до швидкості старіння v [4, 14, 15].

$$cf = \alpha \cdot k_\beta, \quad k_\beta = 1 - v \cdot \frac{\sum_{x \in [-q, +q]} m_x(x)}{\text{Card}([-q, +q])} \quad (3)$$

Модель обробки ГЗ, використовуючи (2) та (3), обчислює оцінки того, як поточні входні CF відповідають сенсу ГЗ, представлениму в (1). Отже, вихід KG такий що

$$cf_{out}(t) = f_{KG}(cf_{in\_1}(t), cf_{in\_2}(t), \dots, cf_{in\_L}(t)). \quad (4)$$

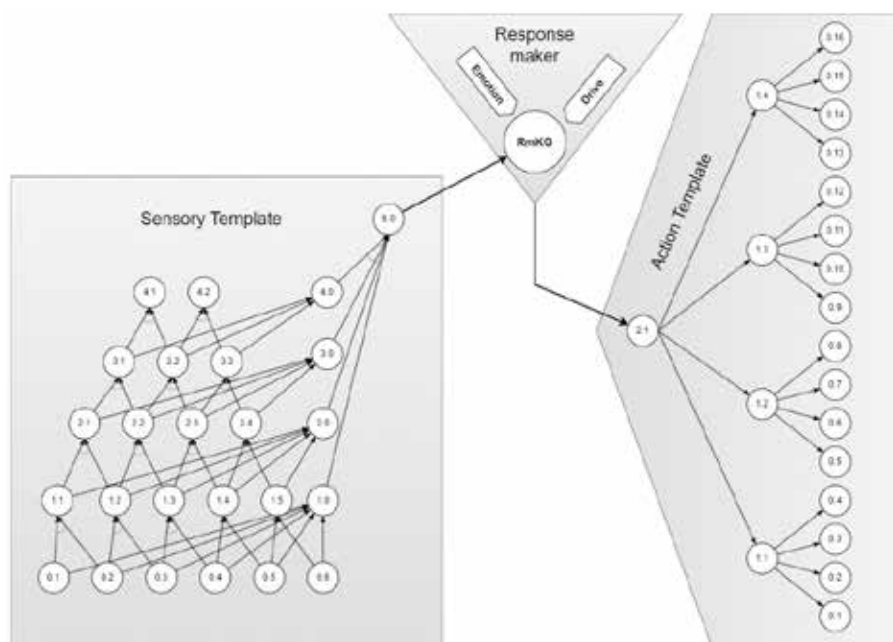


Рис. 1. Приклад структури прототипу одиночної реакції [23]

Математична модель та обчислювальний алгоритм функції f_{KG} у (4) наведено в [4].

Проектування CC FAI за технологією EDA розглянемо на прикладі системи обслуговування автоматичного (Garage Complex, GC) – одно або багато рядову лінійку гаражних боксів, кожен із яких має автоматичні ворота. Основні параметри кожного окремого боксу, такі як стан гаражних воріт, наявність автомобіля, пожежебезпека, можливо температура й вологість контролюються автоматично і ця інформація збирається й обробляється автоматичною системою GC. До складу автоматичного GC ще входить колісний робот що виконує періодичний моніторинг стану GC шляхом об'їзду й обстеження кожного боксу або виконує додаткове обстеження стану окремого боксу, коли надходить інформація про ті чи інші порушення нормального режиму функціонування його автоматичних воріт. Колісний робот оснащений набором датчиків для контролю різних параметрів воріт (відкриті чи закриті, чи нормально відкриті і таке інше). Дані від датчиків робота, інтегровані разом із даними від датчиків кожного окремого боксу, використовуються CC FAI, яка надає модель поточної ситуації по кожному боксу GC на різних рівнях абстрагування. Поточна ситуація формується у реальному часі й використовується системою прийняття рішень FAI для формування різних команд у реальному часі системам автоматики гаражних воріт, роботу й сервісній службі. Таким чином GC уявляє собою автономну систему, обслуговування якої підтримує ІІІ, а саме, FAI. В даній роботі розглядаються питання проектування CC FAI на прикладі фрагмента бази знань необхідного для обслуговування автоматичних гаражних воріт.

Проектування CC FAI, як і іншої інформаційної системи, розбивається на два великих етапа: проектування логічної моделі CC FAI і створення фізичної моделі EDA, яка втілює логічну модель CC FAI.

На першому кроці логічного проектування експертні знання щодо реакції на різні ситуації, які можуть виникнути при обслуговуванні автоматичної GC, формуються у вигляді множини правил чи ПР

ЯКЩО *ситуація* ТОДІ *реакція*,

де *ситуація* і *реакція* представлені у вербальній формі. Завдання експертів побудувати структуру знань у вербальній формі, також, де дається визначення ситуації через набір більш деталізованих підситуацій (1) таким чином щоб досягти рівня деталізації на якому значення ситуації вже є безпосередньо дані від датчиків. Наприклад, ситуація «There is a normal situation with the compartment gate» це ситуація коли і «There is a normal state of the opened gate» і «There is a normal state of the closed gate». Потім визначаються ці дві ситуації, наприклад, для першої можна сказати «Normal opened gate position received by robot's sensors» і «The automatic gate system indicates the open position of the door». Таким чином послідовна деталізація опису ситуацій продовжується поки не досягається

до даних від сенсорів що описують саме ситуацію. Нижче у (5) наведено фрагмент знань, параметризованих відповідно до (1), щодо визначення ситуацій для прикладу що розглядається. Для наступних пояснень технології, структура знань (5) наведена ще у графічному вигляді на Рис. 2.

1. <5.1, There is a normal situation with compartment N 0.4, {<4.1, (0.75, 0, 0.1, 0.5)>, <0.4, (0.75, 0, 0.1, 0.5)>}>;
2. <5.2, There is a malfunction in the automatic gate system of compartment N 0.4, {<4.2, (0.75, 0, 0.1, 0.5)>, <0.4, (0.75, 0, 0.1, 0.5)>}>;
3. <4.1, There is a normal situation with the compartment gate, {<3.1, (0.75, 0, 0.1, 1.0)>, <3.2, (0.75, 0, 0.1, 1.0)>}>;
4. <4.2, There is a malfunction in the automatic gate system, {<3.3, (0.75, 0, 0.1, 1.0)>, <3.4, (0.75, 0, 0.1, 1.0)>}>;
5. <3.0, There is a normal state of the open gate and no service personnel in the compartment, {<2.1, (0.75, 0, 0.1, 0.5)>, <1.7, (-0.75, 0, 0.1, 0.5)>}>;
6. <3.1, There is a normal state of the opened gate, {<2.1, (0.75, 0, 0.1, 0.5)>, <2.3, (0.75, 0, 0.1, 0.5)>}>;
7. <3.2, There is a normal state of the closed gate, {<2.2, (0.75, 0, 0.1, 0.5)>, <2.4, (0.75, 0, 0.1, 0.5)>}>;
8. <3.3, There is a malfunction of the automatic gate system in the open position of the gate, {<2.1, (0.75, 0, 0.1, 0.25)>, <2.4, (0.75, 0, 0.1, 0.25)>, <1.6, (0.75, 0, 0.1, 0.25)>, <1.7, (-0.75, 0, 0.1, 0.25)>}>;
9. <3.4, There is a malfunction of the automatic gate system in the closed position of the gate, {<2.2, (0.75, 0, 0.1, 0.33)>, <2.3, (0.75, 0, 0.1, 0.33)>, <1.6, (0.75, 0, 0.1, 0.33)>}>;
10. <2.1, Normal opened gate position received by robot's sensors, {<1.1, (0.75, 0, 0.1, 1.0)>, <0.1k, (0.75, 0, 0.1, 1.0)>}>;
11. <2.2, Normal closed gate position, {<1.3, (0.75, 0, 0.1, 0.5)>, <1.4, (0.75, 0, 0.1, 0.5)>}>;
12. <2.3, The automatic gate system indicates the door open position, {<1.4, (0.75, 0, 0.1, 0.5)>, <1.5, (-0.75, 0, 0.1, 0.5)>}>;
13. <2.4, The automatic gate system indicates the door closed position, {<1.4, (-0.75, 0, 0.1, 0.5)>, <1.5, (0.75, 0, 0.1, 0.5)>}>;
14. <1.1, Gate is not skewed, {<0.11, (0.75, 0, 0.1, 0.5)>, <0.12, (0.75, 0, 0.1, 0.5)>}>;
15. <1.2, The space underneath the opened gate is clear, {<0.1k+1, (-0.75, 0, 0.1, 0.5)>, <0.1n, (-0.75, 0, 0.1, 0.5)>}>;
16. <1.3, The doorway of the gate is filled, {<0.1k+1, (0.75, 0, 0.1, 0.5)>, <0.1n, (0.75, 0, 0.1, 0.5)>}>;
17. <1.4, The door's sensors indicate the open position of the door, {<0.21, (0.75, 0, 0.1, 0.5)>, <0.21', (-0.75, 0, 0.1, 0.5)>}>;
18. <1.5, The door's sensors indicate the closed position of the door, {<0.22', (-0.75, 0, 0.1, 0.5)>, <0.22, (0.75, 0, 0.1, 0.5)>}>;
19. <1.6, The automatic gate systems are powered up, {<0.23, (0.75, 0, 0.1, 0.5)>, <0.24, (0.75, 0, 0.1, 0.5)>}>;
20. <1.7, There are human beings in the compartment, {<0.31, (-0.75, 0, 0.1, 1/k)>, ..., <0.3k, (-0.75, 0, 0.1, 1/k)>}>;
21. <0.11, Normal height on the right, {<(0.0, 249.5, 250.0, 250.5, 251.0, 300.0)>}>;
22. <0.12, Normal height on the left, {<(0.0, 249.5, 250.0, 250.5, 251.0, 300.0)>}>;
23. <0.13, The space underneath the gate is clear in sector 1, {<(0.0, 0.0, 0.5, 299.5, 300.0, 300.0)>}>;
24. <0.1n+3, The space underneath the gate is clear in sector n, {<(0.0, 0.0, 0.5, 299.5, 300.0, 300.0)>}>;
25. <0.21, The 1st door's sensor indicates the open position of the door, {<(0.0, 0.0, 0.5, 0.95, 1.0, 1.0)>}>;
26. <0.22k, The kst door's sensor indicates the open position of the door, {<(0.0, 0.0, 0.5, 0.95, 1.0, 1.0)>}>;
27. <0.21k+1, The 1st door's sensor indicates the closed position of the door, {<(0.0, 0.0, 0.5, 0.95, 1.0, 1.0)>}>;
28. <0.22, The kst door's sensor indicates the closed position of the door, {<(0.0, 0.0, 0.5, 0.95, 1.0, 1.0)>}>;
29. <0.23, Normal state of electricity supply of the automatic control system, {<(0.0, 210, 215, 225, 230, 250)>}>;
30. <0.24, Normal state of electricity supply of the automatic gate drives' system, {<(0.0, 210, 215, 225, 230, 250)>}>;
31. <0.31, There aren't humans being in sector 1 of the compartment, {<(0.0, 0.0, 0.5, 0.95, 1.0, 1.0)>}>;
32. <0.31k, There aren't humans being in sector k of the compartment, {<(0.0, 0.0, 0.5, 0.95, 1.0, 1.0)>}>;
33. <0.4, The ID of compartment, {<(k)>}>.

(5)

В (5) в кожній строчці наведено опис однієї окремої ГЗ відповідно до (1). Перша цифра ідентифікатора ГЗ відображає рівень узагальнення ситуації. Так, цифра «0» вказує на нульовий рівень, де сенс ГЗ розкривається даними від сенсорів. Наприклад, ГЗ з ідентифікатором 0.11 має вербальне визначення «Normal height on the right» сенс якого задається чисельним діапазоном висоти нормально піднятих воріт. На етапі обробки даних чисельні показники дачика Lidar, наприклад, VL53L0X-V2 фазифікується відповідно (4) на основі знань (5) (значень параметрів в (1) які вказані в визначенні цієї ГЗ). Результат обробки сенсорних даних ГЗ є нечітка впевненість в тому, що входні сенсорні дані відповідають діапазону вказаному у визначенні гранули. Нечітка впевненість оцінюється чисельно як Fuzzy Certainty Factor (FCF) (2) чи (3). Таким чином, ГЗ нульового рівня виконують роль містка між даними від сенсорів й знаннями у вербальному вигляді й ліквідують розрив між двома парадигмами: дані від сенсорів й міркування на вербальному рівні. Завдяки цьому, ГЗ вищих рівнів узагальнення, першого, другого і так далі, отримують від ГЗ нижчих рівнів входні дані тільки у вигляді значень впевненості FCF й обчислюють на виході також впевненість, що входні дані відповідають вербальному опису наведеному у цій ГЗ. Отже, Рис. 2 показує потік даних у вигляді FCF від сенсорів до ГЗ найвищого рівня, де типові ситуації для прийняття рішень мають оцінки впевненості в тому, що наочна ситуація отримана від сенсорів відповідає їм.

На другому кроці логічного проектування отримана структура знань у вигляді (5) та Рис. 2 відображається на апаратне забезпечення автономної системи. Традиційно для цього використовують UML діаграми розміщення [24], але для більшої наочності технології, що наведено у статті наведено у вигляді Рис. 3.

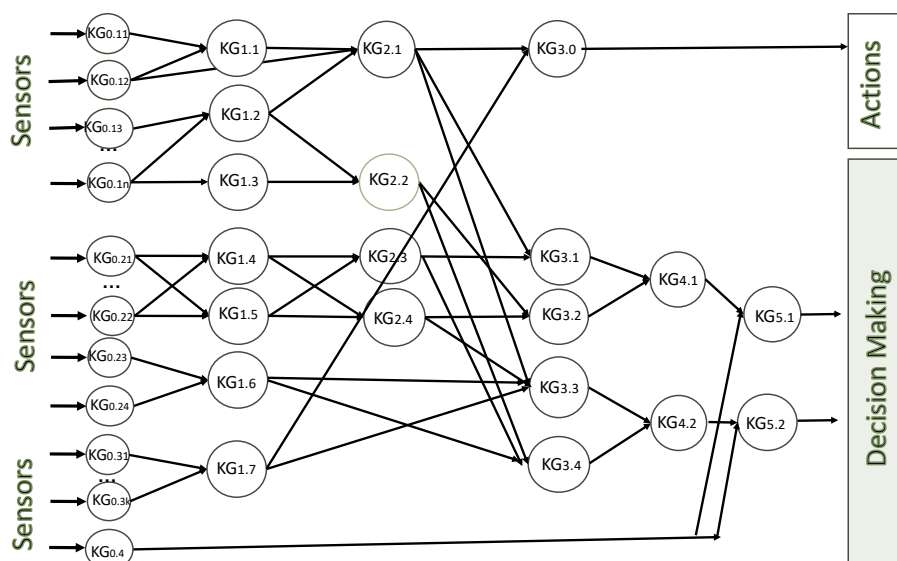


Рис. 2. Графічне зображення структури знань (5)

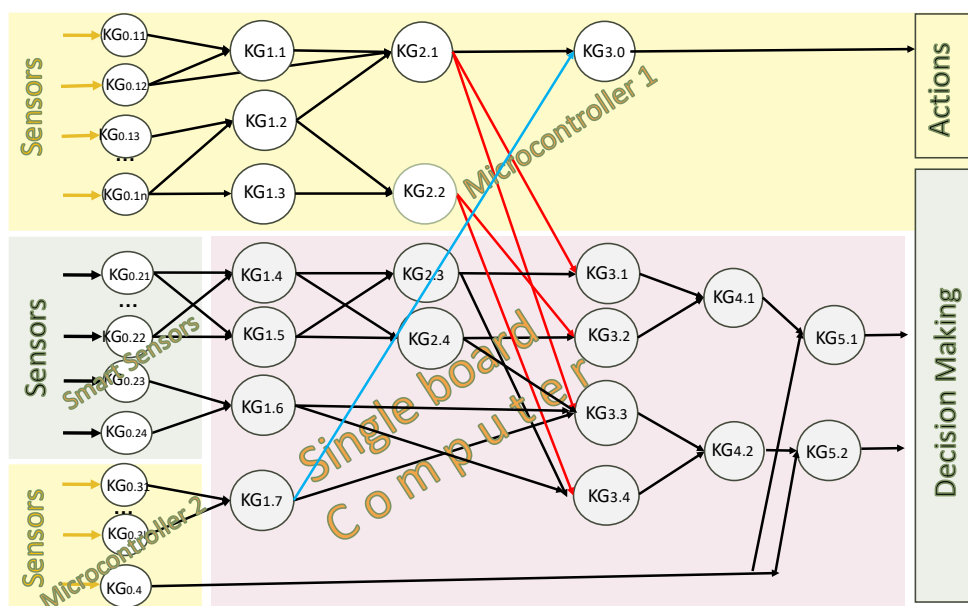


Рис. 3. Відображення структури знань (5) на апаратне забезпечення

Із рисунка видно що база знань (Рис. 2) розподілена між двома мікроконтролерами (Microcontroller 1, Microcontroller 2), чотирма розумними сенсорами, які на виході обчислюють FCF й представлені як окремі ГЗ KG0.21,..., KG0.24, кожен з яких базується на своєму мікроконтролері, і одноплатним комп'ютером. Обидва мікроконтролери розташовані на мобільному роботі, додатково до тих що підтримують автономність функціонування робота, й обслуговують кожен свої сенсори: лазерні датчики відстані VL53L0X-V2 і матричний ІЧ датчик MLX90640 (тепловізійна камера), відповідно.

З'єднання між зовнішніми компонентами здійснюється за допомогою TCP. Робота між внутрішніми компонентами здійснюється через внутрішньопроцесорну або міжпроцесорну взаємодію в межах однієї системи. Зв'язки між ГЗ вищих, крім нульового, рівнів розташованих на різних апаратних платформах вказані червоним й блакитним кольорами на Рис. 3.

Основні знання що потрібні для прийняття рішень автономною системою зосереджені в одній компоненті, а саме, на одноплатному комп'ютері, як видно з Рис. 3. Частина знань, що потрібна на нижніх рівнях обробки даних від сенсорів, якби «приєднується» до глобальної структури знань під час коли робот зосередився на

отримані даних від конкретного гаражного боксу. Ці знання є загальними для обробки інформації від будь яких боксів. Коли робот під'їхав й спозиціював своє положення перед воротами боксу, сканував дані від своїх сенсорів в тому числі й ідентифікатор боксу, тоді фрагмент структури знань що базується на контролері робота як би приєднується тимчасово до основної структури знань на час обробки. Це можливо завдяки однорідності уявлення знань за допомогою запропонованої моделі ГЗ, а також однорідності обробки даних від різного типу сенсорів завдяки введеному нечиткому фактору впевненості FCF і використанню одного ш того ж алгоритму обчислювання для гранул всіх рівнів. Таку технологію можна розглядати як *технологію приєднаної реальності*, яка на відміну від розширеної (віртуальної чи доповненої) реальності створює повну модель ситуації шляхом приєднання тільки реальних фрагментів фізичного світу. Це перша особливість. А друга полягає в тому, що приєднана реальність створює структуру знань для ШІ, який на цій основі реалізує когнітивну функцію міркувань, а не навпаки: ШІ створює розширену реальність для людини доповнюючи реальне сприйняття фізичного світу цифровими видумками.

На третьому кроці логічного проектування розподілена структура знань Рис. 3 відображається у EDA. Для нашого прикладу це дві пов'язані між собою EDA. В цій статті розглядається технологія на прикладі однієї EDA для реалізації на одноплатном комп'ютері наведено бузкокольором на Рис. 3. Тобто структура (множина ГЗ і зв'язки між ними), а також параметри ГЗ (6) повинні бути представленими у вигляді компонентів EDA, як показано на рис. 4.

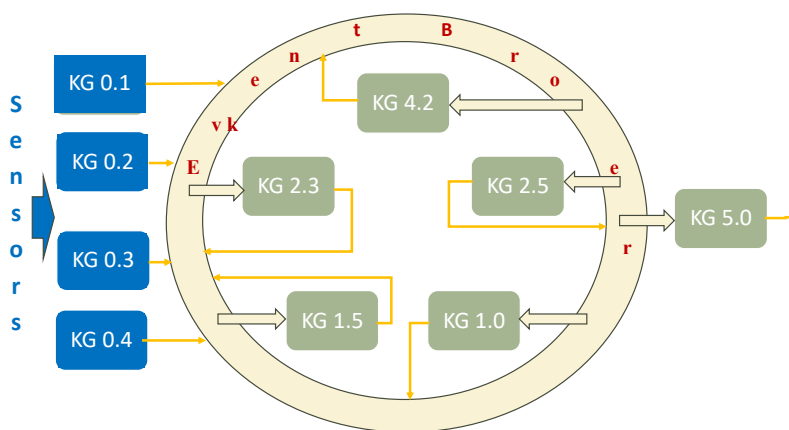


Рис. 4. Компоненти EDA

Для уявлення знань щодо структури Рис. 3, якими оперує Event Broker (EB) на Рис. 4, використовується дві таблиці (Таблиця 1 і Таблиця 2), які містять інформацію про зовнішні зв'язки EDA з ГЗ що належать іншим платформам і внутрішні зв'язки між ГЗ EDA, відповідно. Нижче в Таблиці 1 наведені всі зовнішні зв'язки. У першому стовбчику перелічені усі ГЗ, які надсилають повідомлення через EB, коли отримують свіжі вхідні дані й результат обчислення вихідного FCF задовольняє умовам що вказані в якості параметрів уявлення знань цієї ГЗ. Так як ці ГЗ є зовнішніми, то в таблиці вказано не тільки оримувачів цих повідомлень, а ще і канал через який EB отримує повідомлення. Друга Таблиця 2 містить аналогічну інформацію щодо ГЗ приналежних до розглядаємої EDA: перелічені всі ГЗ як продюсери повідомлень та список отримувачів для кожного продюсера. Крім цього, якщо отримувач не належить до EDA, тоді додатково вказуються ідентифікатор каналу через який відправляється повідомлення. Нагадамо, що повідомлення, яке отримує EB містить інформацію тільки о продюсері (ID гранули відправника й значення FCF). Тому організація даних у таблицях налаштована на такий тип EDA.

Фізичне проектування на основі створенної логічної моделі виконано на мові програмування C++. Серед існуючих бібліотек обрані ZeroMQ та STL, як ті які задовольняють вимогам створення автономних ситем без застосування Інтернет та хмарних обчислень й можуть бути розгорнутими на одноплатних комп'ютерах і контролерах. Бібліотека ZeroMQ [25] дозволяє реалізувати обмін повідомленнями, створити власну event-driven архітектуру, мало важить, і підтримується Linux подібними операційними системами. Для організації міжгранульної взаємодії застосовано набір комунікаційних протоколів, що забезпечують узгоджену та масштабовану передачу даних, зокрема, протокол iprocs. Для реалізації передачі даних у розподіленому середовищі застосовується протокол TCP, що функціонує на основі стандартного стеку TCP/IP.

Програмна реалізація системи Рис. 3 організована як узгоджена сукупність взаємодіючих чотирьох мікро EDA. Кожна з них має власний ідентифікаційний номер, набір потоків-виконавців, сховище вузлів (множина ГЗ) та маршрутизатор (EB), відповідальний за переспрямування повідомлень.

Таблиця 1

Зовнішні зв'язки EDA

Event producer		Event consumers (KG's ID/input's N)
KG's ID	External Channel	
KG021	Ch 1	KG ₁₄ /1, KG ₁₅ /1
KG022		KG ₁₄ /2, KG ₁₅ /2
KG023		KG ₁₆ /1
KG024		KG ₁₆ /2
KG031	Ch 2	KG ₁₇ /1,
KG03k		KG ₁₇ /k
KG04		KG ₅₁ /2, KG ₅₂ /2
KG21	Ch 3	KG ₃₁ /1, KG ₃₁ /1
KG22		KG ₃₂ /1, KG ₃₄ /1

Таблиця 2

Внутрішні зв'язки EDA

Event producer (KG's ID)	Event consumers (Channel:KG's ID/input's N)	
	Internal	External Channel
KG14	KG ₂₃ /1, KG ₂₄ /1	
KG15	KG ₂₃ /2, KG ₂₄ /2	
KG16	KG ₃₃ /1, KG ₃₄ /3	
KG17	KG ₃₄ /4	Ch 1:KG ₃₀ /2
KG23	KG ₃₁ /2, KG ₃₄ /2,	
KG24	KG ₃₂ /2, KG ₃₃ /2	
KG31	KG ₄₁ /1	
KG32	KG ₄₁ /2	
KG33	KG ₄₂ /1	
KG34	KG ₄₂ /2	
KG41	KG ₅₁ /1	
KG42	KG ₅₂ /1	
KG51		Ch 4
KG52		Ch 4

ЕВ, отримав повідомлення від будь якої ГЗ, використовуючі дані таблиць розсилає повідомлення відповідним отримувачам й активізує для кожної гранули-отримувача виконання алгоритму, вказаному у Таблиці 3, та передає йому відповідні параметри із цієї же таблиці. Параметри ГЗ наведені в таблиці 3.

Таблиця 3

Параметри ГЗ

KG's ID	KG's processing algorithm	KG's input parameters (c_i/a_i , b_i , v_i , g_i)				KG's output parameters (c/a)
		Input i=1	Input i=2	Input i=3	Input i=4	
1.4	2SNO	Up/0.75, 0, 0.1, 0.5	Up/0.75, 0, 0.1, 0.5			Up/0.85
1.5	2SNO	Up/0.75, 0, 0.1, 0.5	Up/0.75, 0, 0.1, 0.5			Up/0.85
1.6	2SNO	Up/0.75, 0, 0.1, 0.5	Up/0.75, 0, 0.1, 0.5			Up/0.85
1.7	2SNO	Up/0.75, 0, 0.1, 0.5	Up/0.75, 0, 0.1, 0.5			Up/0.85
2.3	2SNO	Up/0.75, 0, 0.1, 0.5	Down/-0.75, 0, 0.1, 0.5			Up/0.85
2.4	2SNO	Down/-0.75, 0, 0.1, 0.5	Up/0.75, 0, 0.1, 0.5			Up/0.85
3.1	2SNO	Up/0.75, 0, 0.1, 0.5	Up/0.75, 0, 0.1, 0.5			Up/0.85
3.2	2SNO	Up/0.75, 0, 0.1, 0.5	Up/0.75, 0, 0.1, 0.5			Up/0.85
3.3	2SNO	Up/0.75, 0, 0.1, 0.25	Up/0.75, 0, 0.1, 0.25	Up/0.75, 0, 0.1, 0.25	Down/-0.75, 0, 0.1, 0.25	Up/0.85
3.4	2SNO	Up/0.75, 0, 0.1, 0.33	Up/0.75, 0, 0.01, 0.33	Up/0.75, 0, 0.01, 0.33		Up/0.85
4.1	2SNO	Up/0.75, 0, 0.1, 1.0	Up/0.75, 0, 0.1, 1.0			Up/0.85
4.2	2SNO	Up/0.75, 0, 0.1, 1.0	Up/0.75, 0, 0.1, 1.0			Up/0.85
5.1	2SNO	Up/0.75, 0, 0.1, 0.5	Up/0.75, 0, 0.1, 0.5			Up/0.85
5.2	2SNO	Up/0.75, 0, 0.1, 0.5	Up/0.75, 0, 0.1, 0.5			Up/0.85

Програмна реалізація окремої мікро EDA включає хеш-таблиці типу `std::unordered_map` в оперативній пам'яті що забезпечує високу швидкодію. В них зберігаються дані вузлів: номер ГЗ, функція обробки, дані з Таблиці 3, та сигнал, який випускається при завершенні виконання з передачею повідомлення. Сховище вузлів (NodeStore) організоване на основі двох хеш-структур. Перша відповідає за швидкий доступ до екземплярів вузлів за їхніми номерами, друга фіксує односторонні зв'язки між вузлами: використовуючи номер вузла-джерела як ключ, система миттєво отримує перелік вузлів-приймачів, з якими він пов'язаний. Маршрутизатор побудований на основі ZeroMQ для прослуховування повідомлень від ГЗ. Для цього асинхронно слухається вказана черга повідомлень (наприклад, `inproc://outputs`). У разі надходження даних від будь-якої ГЗ він визначає відправника, звертається до NodeStore, отримує список вузлів-призначень і перенаправляє повідомлення у відповідний канал для виконання.

Потік-виконавець (Worker) використовує сокет zeromq для прослуховування повідомлень у власній черзі (наприклад, `inproc://tasks`). Після надходження повідомлення він визначає, до якого вузла воно належить, передає йому дані й запускає вказану функцію обробки. Для довготривалого збереження стану системи ще розроблено спеціалізований файловий формат `grdef`, який містить повний набір параметрів, необхідних для відтворення системи.

Висновки

FAI, як ШІ для управління на основі сенсорних даних автономними системами, має бути розгорнутим на мікроконтролерах й одноплатних комп'ютерах без застосування потужностей хмарних обчислень і знань з Інтернет. Модель FAI дозволяє всі потрібні експертні знання щодо прийняття рішень уявляти у вигляді структурованої множини однорідних елементів – ГЗ, та розглядати їх як елементи EDA. Така структура знань у вигляді множини ГЗ легко відображається на апаратне забезпечення FAI і компонується в окремі мікро EDA. Когнітивна модель FAI, яка підтримує поступове забування з часом подій через старіння даних, дозволяє подолати основну проблему застосування EDA – необхідність формування реакції в режимі реального часу компонентами FAI на основі сенсорних даних, отриманих асинхронно. У майбутньому планується розробка моделі навчання FAI в операційному режимі з метою автоматичного налаштування параметра швидкості старіння в CF відповідно до динамічних властивостей середовища FAI.

Список використаної літератури

1. Unmanned Systems. NovAtel, 2022 [Електронний ресурс]. Режим доступу: <https://novatel.com/industries/unmanned-systems>
2. The Future of Autonomy. Isn't Human-Less. It's Human More. Lockheed Martin, 2022 [Електронний ресурс]. Режим доступу: <https://www.lockheedmartin.com/en-us/capabilities/autonomous-unmanned-systems.html>
3. Chen J., Sun J., Wang G. From Unmanned Systems to Autonomous Intelligent Systems // Engineering. 2022. Vol. 12. P. 16–19. DOI: 10.1016/j.eng.2021.10.007
4. Kargin A., Petrenko T. Knowledge Distillation for Autonomous Intelligent Unmanned Systems // In: Pedrycz W., Chen S.-M. (eds.) *Advancements in Knowledge Distillation: Towards New Horizons of Intelligent Systems*. Studies in Computational Intelligence, vol. 1100. Springer, 2023. P. 193–230.
5. Joseph L., Mondal A. K. (eds.) *Autonomous Driving and Advanced Driver-Assistance Systems (ADAS). Applications, Development, Legal Issues, and Testing*. 1st ed. Boca Raton : CRC Press, 2021. DOI: 10.1201/9781003048381
6. Gartner: 2025 Top Strategic Technology Trends [Електронний ресурс]. Режим доступу: <https://www.gartner.com/en/articles/top-technology-trends-2025>
7. Open X-Embodiment: Robotic Learning Datasets and RT-X Models [Електронний ресурс]. Режим доступу: <https://robotics-transformer-x.github.io/>
8. The TED AI Show: How AI robots learn like babies – but a million times faster w/ NVIDIA's Rev LeBaredian. 3 Dec 2024 [Електронний ресурс]. Режим доступу: <https://www.ted.com/podcasts/how-ai-robots-learn-just-like-babies-rev-lebaredian-transcript>
9. Levine S., Hausman K. The global project to make a general robotic brain // IEEE Spectrum. Jan 2024 [Електронний ресурс]. Режим доступу: <https://spectrum.ieee.org/global-robotic-brain>
10. McKinsey Global Survey. The state of AI in 2023: Generative AI's breakout year. 2023 [Електронний ресурс]. Режим доступу: <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2023-generative-AIs-breakout-year>
11. Yasumoto K., Yamaguchi H., Shigeno H. Survey of Real-time Processing Technologies of IoT Data Streams // Journal of Information Processing. 2016. Vol. 24, no. 2. P. 195–202. DOI: 10.2197/ipsjip.24.195
12. Winfield A. T. et al. IEEE P7001: A Proposed Standard on Transparency // Frontiers in Robotics and AI. 2021. DOI: 10.3389/frobt.2021.665729
13. Siemens. Next-Gen AI Manufacturing / Industrial Sector. 2021 [Електронний ресурс]. Режим доступу: <https://siemens.com/innovation>
14. Kargin A., Petrenko T. Feeling Artificial Intelligence for AI-Enabled Autonomous Systems // Proc. Global Conf. on Artificial Intelligence and Internet of Things (GCAIoT). IEEE, Alamein New City, Egypt, 2022. P. 88–93.
15. Kargin A., Petrenko T. Feeling Artificial Intelligence. Cognitive Decision-making Model Borrowed from Living Beings // International Symposium on Networks, Computers and Communications. IEEE, Washington DC, USA, 2024. P. 1–6.
16. Reis J., Cohen Y., Melao N., Costa J., Jorge D. High-Tech Defense Industries: Developing Autonomous Intelligent Systems // Applied Sciences. 2021. Vol. 11(11). Art. 4920. DOI: 10.3390/app11114920
17. Ploennigs J., Cohn J., Stanford-Clark A. The future of IoT // IEEE Internet of Things Magazine. 2018. Vol. 1(1). P. 28–33.
18. Internet of Things: Understanding the Adventure. SAS Institute Inc., 17 Mar. 2020 [Електронний ресурс]. Режим доступу: <https://www.sas.com/sas/offers/20/iot-understanding-adventure.html>

19. Kargin A., Petrenko T. Method of Using Data from Intelligent Machine Short-Term Memory in Fuzzy Logic System // 2021 IEEE 7th World Forum on Internet of Things (WF-IoT). DOI: 10.1109/wf-iot51360.2021.9594918
20. Gupta H., Nayak A. Publish-subscribe system security requirement: A case study for V2V communication // IEEE Open Journal of the Computer Society. 2024. Vol. 5. P. 389–405. DOI: 10.1109/OJCS.2024.3442921
21. Rocha H. F. O. Practical Event-Driven Microservices Architecture: Building Sustainable and Highly Scalable Event-Driven Microservices. APRESS, 2025. 449 p. DOI: 10.1007/978-1-4842-7468-2
22. Islam G. Z., Motakabber S. M. A. A Comprehensive Review on the Internet of Things Network // Journal of Communications. 2025. Vol. 20, no. 1. P. 84–98. DOI: 10.12720/jcm.20.1.84-98
23. Kargin A., Petrenko T. Event-Driven Architecture of Feeling Artificial Intelligence Supported by Data Aging Model // In: Mejdoub Y., Elamri A., Kardouchi M. (eds.) Connected Objects, Artificial Intelligence, Telecommunications and Electronics Engineering (COCIA 2025). Lecture Notes in Networks and Systems, vol. 1584. Springer, Cham, 2025. P. 178–183. DOI: 10.1007/978-3-032-01536-5_27
24. Unified Modeling Language [Електронний ресурс]. Режим доступу: <https://www.uml.org/>
25. ZeroMQ. [Електронний ресурс]. Режим доступу: <https://zeromq.org/>

References

1. (2022) Unmanned Systems. NovAtel. URL: <https://novatel.com/industries/unmanned-systems>
2. (2022) The Future of Autonomy. Isn't Human-Less. It's Human More. Lockheed Martin. URL: <https://www.lockheedmartin.com/en-us/capabilities/autonomous-unmanned-systems.html>
3. Chen J., Sun J., Wang G. (2022) From Unmanned Systems to Autonomous Intelligent Systems // Engineering. Vol. 12. P. 16–19. DOI: 10.1016/j.eng.2021.10.007
4. Kargin A., Petrenko T. (2023) Knowledge Distillation for Autonomous Intelligent Unmanned Systems // In: Pedrycz W., Chen S.-M. (eds.) Advancements in Knowledge Distillation: Towards New Horizons of Intelligent Systems. Studies in Computational Intelligence, vol. 1100. Springer. P. 193–230.
5. Joseph L., Mondal A. K. (eds.) (2021) Autonomous Driving and Advanced Driver-Assistance Systems (ADAS). Applications, Development, Legal Issues, and Testing. 1st ed. Boca Raton : CRC Press. DOI: 10.1201/9781003048381
6. Gartner: 2025 Top Strategic Technology Trends URL: <https://www.gartner.com/en/articles/top-technology-trends-2025>
7. Open X-Embodiment: Robotic Learning Datasets and RT-X Models. URL: <https://robotics-transformer-x.github.io/>
8. The TED AI Show: How AI robots learn like babies – but a million times faster w/ NVIDIA's Rev LeBaredian. 3 Dec 2024. URL: <https://www.ted.com/podcasts/how-ai-robots-learn-just-like-babies-rev-lebaredian-transcript>
9. Levine S., Hausman K. The global project to make a general robotic brain // IEEE Spectrum. Jan 2024. URL: <https://spectrum.ieee.org/global-robotic-brain>
10. (2023) McKinsey Global Survey. The state of AI in 2023: Generative AI's breakout year. URL: <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2023-generative-AIs-breakout-year>
11. Yasumoto K., Yamaguchi H., Shigeno H. (2016) Survey of Real-time Processing Technologies of IoT Data Streams // Journal of Information Processing. Vol. 24, no. 2. P. 195–202. DOI: 10.2197/ipsjip.24.195
12. Winfield A. T. et al. (2021) IEEE P7001: A Proposed Standard on Transparency // Frontiers in Robotics and AI. DOI: 10.3389/frobt.2021.665729
13. Siemens. Next-Gen AI Manufacturing / Industrial Sector. 2021. URL: <https://siemens.com/innovation>
14. Kargin A., Petrenko T. (2022) Feeling Artificial Intelligence for AI-Enabled Autonomous Systems // Proc. Global Conf. on Artificial Intelligence and Internet of Things (GCAIoT). IEEE, Alamein New City, Egypt, P. 88–93.
15. Kargin A., Petrenko T. (2024) Feeling Artificial Intelligence. Cognitive Decision-making Model Borrowed from Living Beings // International Symposium on Networks, Computers and Communications. IEEE, Washington DC, USA, P. 1–6.
16. Reis J., Cohen Y., Melao N., Costa J., Jorge D. (2021) High-Tech Defense Industries: Developing Autonomous Intelligent Systems // Applied Sciences. Vol. 11(11). Art. 4920. DOI: 10.3390/app11114920
17. Ploennigs J., Cohn J., Stanford-Clark A. (2018) The future of IoT // IEEE Internet of Things Magazine. Vol. 1(1). P. 28–33.
18. Internet of Things: Understanding the Adventure. SAS Institute Inc., 17 Mar. 2020. <https://www.sas.com/sas/offers/20/iot-understanding-adventure.html>
19. Kargin A., Petrenko T. Method of Using Data from Intelligent Machine Short-Term Memory in Fuzzy Logic System // 2021 IEEE 7th World Forum on Internet of Things (WF-IoT). DOI: 10.1109/wf-iot51360.2021.9594918
20. Gupta H., Nayak A. (2024) Publish-subscribe system security requirement: A case study for V2V communication // IEEE Open Journal of the Computer Society. Vol. 5. P. 389–405. DOI: 10.1109/OJCS.2024.3442921
21. Rocha H. F. O. (2025) Practical Event-Driven Microservices Architecture: Building Sustainable and Highly Scalable Event-Driven Microservices. APRESS, 449 p. DOI: 10.1007/978-1-4842-7468-2

22. Islam G. Z., Motakabber S. M. A. (2025) A Comprehensive Review on the Internet of Things Network // Journal of Communications. Vol. 20, no. 1. P. 84–98. DOI: 10.12720/jcm.20.1.84-98
23. Kargin A., Petrenko T. (2025) Event-Driven Architecture of Feeling Artificial Intelligence Supported by Data Aging Model // In: Mejdoub Y., Elamri A., Kardouchi M. (eds.) Connected Objects, Artificial Intelligence, Telecommunications and Electronics Engineering (COCIA 2025). Lecture Notes in Networks and Systems, vol. 1584. Springer, Cham, P. 178–183. DOI: 10.1007/978-3-032-01536-5_27
24. Unified Modeling Language. URL: <https://www.uml.org/>
25. ZeroMQ. URL: <https://zeromq.org/>

Дата першого надходження рукопису до видання: 22.11.2025
Дата прийнятого до друку рукопису після рецензування: 17.12.2025
Дата публікації: 31.12.2025

О. В. КРАЙНЮК

кандидат технічних наук, доцент,
завідувачка кафедри кібербезпеки
Харківський національний автомобільно-дорожній університет
ORCID: 0000-0001-9524-040X

Ю. В. БУЦ

доктор технічних наук, професор,
завідувач кафедри охорони праці та надзвичайних ситуацій
Український державний університет залізничного транспорту
ORCID: 0000-0003-0450-2617

М. М. ПІКСАСОВ

кандидат технічних наук, доцент,
доцент кафедри кібербезпеки
Харківський національний автомобільно-дорожній університет
ORCID: 0000-0001-9487-7273

Н. В. ДІДЕНКО

кандидат технічних наук, доцент,
доцент кафедри кібербезпеки
Харківський національний автомобільно-дорожній університет
ORCID: 0000-0003-3318-438X

Б. О. ПОХОДЕНКО

старший викладач кафедри кібербезпеки
Харківський національний автомобільно-дорожній університет
ORCID: 0000-0002-9995-7077

ФУНКЦІОНАЛЬНЕ МОДЕЛЮВАННЯ ТА АЛГОРИТМІЗАЦІЯ ПРОЦЕСІВ ЗАБЕЗПЕЧЕННЯ КІБЕРСТІЙКОСТІ АВТОМАТИЗОВАНИХ СИСТЕМ ВИДАЧІ ЗІЗ

У статті розглядається актуальна науково-практична проблема забезпечення кіберстійкості та безперервності функціонування автоматизованих систем видачі засобів індивідуального захисту (САВ-ЗІЗ) в умовах деструктивних кібервпливів. В епоху Індустрії 4.0 цифровізація охоплює критичні процеси охорони праці, проте зростання залежності від стабільності мережевого середовища створює ризики блокування доступу до засобів захисту.

Здійснено аналіз вразливостей класичних централізованих архітектур промислового Інтернету речей (IIoT), які характеризуються критичною залежністю від доступності серверної інфраструктури та схильністю до «функціонального паралічу» під час DDoS-атак. Висвітлено обмеження традиційних підходів типу Fail-Secure, які є неприпустимими для критичної інфраструктури безпеки, де пріоритетом є фізична доступність ресурсів, а не блокування доступу.

Методологічною основою дослідження стало функціональне моделювання бізнес-процесів у стандарті IDEF0, що дозволило декомпонувати процес аварійного керування та ідентифікувати критичні вразливості на етапі фіксації транзакцій. Для вирішення виявлених проблем запропоновано гібридну архітектуру та алгоритм адаптивної кіберстійкості, що базується на технології периферійних обчислень (Edge Computing). Це дозволяє децентралізувати логіку прийняття рішень та перенести процес валідації прав доступу безпосередньо на рівень кінцевого обладнання.

Розроблений алгоритм реалізовано на базі скінченного автомата з трьома станами: штатний режим (ONLINE), перехідний режим (DEGRADED) та аварійний режим (OFFLINE). Для захисту цілісності офлайн-транзакцій та протидії фальсифікації даних в ізольованому середовищі застосовано механізм криптографічної інкапсуляції записів з використанням технології хеш-ланцюжків на базі алгоритму SHA-256.

Ефективність запропонованих рішень підтверджено шляхом імітаційного моделювання мовою Python. Результати експерименту засвідчили, що в умовах тривалої атаки на відмову в обслуговуванні розроблена система забезпечує показник доступності сервісу на рівні понад 98%, тоді як класична архітектура демонструє повну зупинку обслуговування. Також доведено ефективність реалізованого протоколу пакетної синхронізації даних, який дозволяє уникнути перевантаження мережі на етапі відновлення зв'язку.

Ключові слова: кіберстійкість, промисловий інтернет речей, периферійні обчислення, безпечна відмова, хеш-ланцюжки, доступність сервісу, атаки на відмову, автоматизація процесів, охорона праці, безперервність бізнесу.

O. V. KRAINIUK

Candidate of Technical Sciences, Associate Professor,
Head of the Cybersecurity Department
Kharkiv National Automobile and Highway University
ORCID: 0000-0001-9524-040X

YU. V. BUTS

Doctor of Technical Sciences, Professor,
Head of the Department of Occupational Safety and Emergency Situations
Ukrainian State University of Railway Transport
ORCID: 0000-0003-0450-2617

M. M. PIKSASOV

Candidate of Technical Sciences, Associate Professor,
Associate Professor at the Department of Cybersecurity
Kharkiv National Automobile and Highway University
ORCID: 0000-0001-9487-7273

N. M. DIDENKO

Candidate of Technical Sciences, Associate Professor,
Associate Professor at the Department of Cybersecurity
Kharkiv National Automobile and Highway University
ORCID: 0000-0003-3318-438X

B. O. POHODENKO

Senior Lecturer at the Department of Cybersecurity
Kharkiv National Automobile and Highway University
ORCID: 0000-0002-9995-7077

FUNCTIONAL MODELING AND ALGORITHMIZATION OF PROCESSES FOR ENSURING CYBER RESILIENCE OF AUTOMATED DISPENSING SYSTEMS FOR PERSONAL PROTECTIVE EQUIPMENT

The article addresses the urgent scientific and practical problem of ensuring cyber resilience and operational continuity of automated personal protective equipment (PPE) dispensing systems under destructive cyber impacts. In the Industry 4.0 era, digitalization covers critical occupational safety processes; however, the increasing reliance on network stability creates significant risks of blocking access to protective equipment during cyber incidents.

The study provides a comprehensive analysis of the vulnerabilities inherent in classic centralized Industrial Internet of Things (IIoT) architectures. These systems are characterized by a critical reliance on server infrastructure availability and susceptibility to "functional paralysis" during Denial-of-Service (DoS) attacks. Specifically, the limitations of traditional Fail-Secure approaches are highlighted, as they are unacceptable for critical occupational safety infrastructure where the physical availability of resources is the priority, rather than access blocking.

The methodological basis of the research is functional modeling of business processes using the IDEF0 standard. This approach allowed for the decomposition of the emergency management process and the identification of critical vulnerabilities at the transaction commitment stage. To address these issues, a hybrid architecture and algorithmic support based on Edge Computing technology are proposed. This enables the decentralization of decision-making logic and transfers the access rights validation process directly to the endpoint level.

The developed adaptive cyber resilience algorithm is implemented based on a three-state Finite State Machine (FSM): normal operation (ONLINE), transitional mode (DEGRADED), and emergency mode (OFFLINE). To protect the integrity of offline transactions and prevent data falsification in an isolated environment, a mechanism of cryptographic record encapsulation using hash chaining technology based on the SHA-256 algorithm is applied.

Simulation results, conducted using Python, confirm the approach's effectiveness. The experiments demonstrated that under conditions of a prolonged DoS attack, the proposed system ensures a service availability rate of over 98%, whereas the classic architecture demonstrates a complete cessation of service. Furthermore, the effectiveness of the implemented batch data synchronization protocol, which prevents network overload during the connection recovery phase, was proven.

Keywords: cyber resilience, Industrial Internet of Things, edge computing, fail-safe, hash chaining, service availability, denial-of-service attacks, process automation, occupational safety, business continuity.

Постановка проблеми

Актуальність теми. В умовах Індустрії 4.0 (Industry 4.0) цифровізація охоплює критичні бізнес-процеси підприємств, включаючи забезпечення охорони праці. Традиційні методи видачі засобів індивідуального захисту (ЗІЗ) активно замінюються автоматизованими системами (САВ-ЗІЗ), реалізованими на базі технологій Промислового Інтернету речей (ІоТ). Такі системи забезпечують прозорість обліку та персоніфікацію відповідальності. Однак, зростання залежності фізичних процесів безпеки від стабільності кібернетичного середовища створює нові ризики.

Постановка проблеми. Більшість сучасних рішень САВ-ЗІЗ побудовані за централізованою архітектурою, де кінцевий пристрій (вендинговий автомат) критично залежить від доступності серверної інфраструктури. У разі деструктивних кібервпливів (зокрема DoS-атак) або технічних збоїв мережі виникає загроза блокування процесу видачі ЗІЗ. Це призводить не лише до фінансових втрат через простій виробництва, а й до прямого порушення регламентів безпеки персоналу.

Аналіз підходів. Існуючі дослідження здебільшого фокусуються на криптографічному захисті каналів зв'язку або фізичній безпеці пристроїв. Водночас, недостатньо уваги приділяється формалізації внутрішньої логіки функціонування системи в критичних ситуаціях. Відсутність чітких функціональних моделей, які б описували поведінку системи при втраті зв'язку, ускладнює розробку надійних алгоритмів керування.

Шлях вирішення. Для створення кіберстійкої системи необхідний комплексний підхід, що поєднує функціональне моделювання бізнес-процесів (для виявлення вразливих інформаційних потоків) та алгоритмізацію процедур аварійного керування (для забезпечення безперервності роботи). Використання методології IDEF0 дозволяє декомпонувати процес видачі ЗІЗ та ідентифікувати ті функції, які потребують перенесення на рівень периферійних обчислень (Edge Computing).

Аналіз останніх досліджень і публікацій

Впровадження технологій Індустрії 4.0 у сферу охорони праці актуалізувало питання надійності автоматизованих систем видачі засобів індивідуального захисту (САВ-ЗІЗ). Як зазначається у роботах [1], перехід від ручного обліку до використання ІоТ-вендингових автоматів дозволяє оптимізувати логістику та персоніфікувати відповідальність. Однак зростання залежності критичної інфраструктури безпеки від мережевих технологій супроводжується новими викликами.

Аналіз сучасної літератури свідчить, що забезпечення безперервності функціонування таких систем вимагає зміни парадигми захисту. Перехід до Індустрії 4.0 вимагає зміщення акценту з класичного кіберзахисту (запобігання атакам) на забезпечення кіберстійкості – здатності системи адаптуватися та зберігати функціональність під час атак чи збоїв [1].

Серед сучасних рішень виділяються:

- Використання модельно-прогнозуючого керування (англ. Model Predictive Control, MPC) із перемиканням функцій вартості, що дозволяє зберігати стабільність і керованість систем навіть при втраті даних внаслідок DoS-атак [2].
- Застосування спостерігачів стану та нечіткої логіки для компенсації втрат даних та нівелювання невизначеностей у мережевих системах керування [3].
- Впровадження подієво-керованих механізмів, які зменшують навантаження на мережу та підвищують автономність вузлів [4].

Для забезпечення цілісності даних у розподілених системах активно досліджуються технології блокчейну, проте їх імплементація на ресурсно-обмежених пристроях ІоТ (до яких належать контролери вендингових автоматів) часто ускладнена через високі обчислювальні вимоги [5, 6].

Попри значний прогрес у теоретичних методах керування для кіберфізичних систем, більшість сучасних рішень орієнтовані на безперервні динамічні об'єкти (енергомережі, транспорт), тоді як дискретні системи обслуговування (наприклад, автоматизовані пункти видачі ЗІЗ) залишаються недостатньо дослідженими щодо алгоритмічної стійкості та кіберстійкості під час DoS-атак [7].

Більшість сучасних підходів (MPC, нечітка логіка, подієво-керовані стратегії) розроблені для безперервних або гібридних систем, де контроль і відновлення функціональності після атак базуються на безперервних моделях [8]. Для дискретних сервісних систем, які мають інші вимоги до доступності та безпеки, такі підходи не завжди застосовні або ефективні.

Існуючі архітектури автоматизованих систем видачі ЗІЗ здебільшого реалізують жорстку клієнт-серверну логіку. У разі DoS-атаки системи переходять у стан Fail-Secure (блокування), що є прийнятним для фінансових терміналів, але неприпустимим для критичної інфраструктури охорони праці, де пріоритетом є фізична доступність ЗІЗ [7]. Відсутні формалізовані алгоритми перемикання в режим Fail-Safe із збереженням цілісності транзакцій.

Огляд сучасних досліджень підкреслює, що більшість рішень зосереджені на виявленні атак, а не на забезпеченні стійкої роботи після їх виникнення. Реалізація гнучких, адаптивних механізмів відновлення для дискретних сервісних систем залишається відкритою проблемою [9].

У цьому контексті перспективним видається застосування концепції периферійних обчислень, що дозволяє децентралізувати логіку прийняття рішень. Перенесення критичних функцій авторизації безпосередньо на рівень кінцевого обладнання здатне забезпечити необхідну автономність системи під час деструктивних впливів

Важливо зазначити, що організаційно-технічні аспекти впровадження вендингових автоматів та перспективи використання інтелектуальних систем управління безпекою праці вже були обґрунтовані авторами в попередніх працях [10-12]. Проте, якщо раніше акцент робився на функціональній ефективності та автоматизації процесів, то дане дослідження фокусується саме на алгоритмічному забезпеченні їхньої живучості в умовах кібернетичних загроз.

Формулювання мети дослідження

Метою роботи є підвищення рівня кіберстійкості автоматизованих систем видачі ЗІЗ шляхом розробки функціональних моделей та алгоритмічного забезпечення для гарантування безперервності бізнес-процесів в умовах деструктивних кібервпливів. Для досягнення цієї мети передбачається поетапне вирішення комплексу завдань: від функціонального моделювання бізнес-процесів у методології IDEF0 та ідентифікації критичних вразливостей архітектури до розробки алгоритму адаптивного керування («Secure Fail-Safe») з подальшою верифікацією його ефективності шляхом імітаційного моделювання.

Домінуюча централізована архітектура ПОТ-систем характеризується критичною залежністю від доступності каналів зв'язку, що створює вразливість типу «єдина точка відмови» (Single Point of Failure). У разі реалізації атак на відмову в обслуговуванні це призводить до функціонального паралічу обладнання та переходу в стан блокування (Fail-Secure), що спричиняє неприпустимі простої виробництва через неможливість видачі засобів захисту. Вирішення зазначеної проблеми вимагає переходу до децентралізованої архітектури на базі периферійних обчислень, яка забезпечить автономне прийняття рішень та реалізацію режиму безпечної відмови (Fail-Safe) в умовах ізоляції від сервера.

Викладення основного матеріалу дослідження

Функціональне моделювання системи в стандарті IDEF0

Для формалізації процесів прийняття рішень в умовах невизначеності (втрати зв'язку) було розроблено функціональну модель системи в методології IDEF0. На відміну від стандартних підходів, де моделюється лише успішний сценарій видачі, дана модель фокусується на декомпозиції процесу аварійного керування.

Декомпозиція процесів верхнього рівня (Діаграма A0). На рівні декомпозиції виділено чотири функціональні блоки, які взаємодіють між собою через інформаційні та керуючі потоки:

- A1 (Ідентифікація та авторизація): Блок отримує дані з RFID-зчитувача. Критичною особливістю є те, що вхідні дані про права доступу можуть надходити з двох джерел: "Центральна база даних" (штатний режим) або локальний захищений реєстр (аварійний режим).
- A2 (Видача та фіксація): Цей блок є вузлом прийняття рішень про фізичну активацію механізму видачі. Він має зворотний зв'язок з блоком A3, що дозволяє буферизувати транзакції при відсутності з'єднання.
- A3 (Облік та звітність): Відповідає за агрегацію даних та формування черги синхронізації (Sync Queue) після відновлення мережі.
- A4 (Контроль залишків): Забезпечує моніторинг фізичної наявності ЗІЗ, запобігаючи помилкам видачі при розсинхронізації баз даних.

Моделювання критичних функцій видачі (Діаграма A2). Деталізація блоку A2 дозволила виявити вразливість на етапі "Миттєва фіксація списання" (A23). У класичній архітектурі цей процес вимагає підтвердження від сервера ("Commit"). У розробленій моделі впроваджено механізм локальної фіксації, де виходом є не просто запис у БД, а криптографічно підписана транзакція, що зберігається в енергонезалежній пам'яті автомата.

Використання CASE-засобу Ramus дозволило верифікувати повноту моделі та гарантувати, що для кожного стану системи (Online/Offline) існує визначений шлях проходження інформаційного потоку, що унеможливило виникнення стану невизначеного блокування.

Алгоритмізація процесів керування

Якщо функціональна модель IDEF0 визначає вимоги до інформаційних потоків («що має робити система»), то етап алгоритмізації формалізує логіку виконання цих вимог («як система це робить»). Ключовим об'єктом алгоритмізації є блок A1 («Ідентифікація та авторизація»), який відповідає за прийняття рішення про допуск користувача.

Для реалізації концепції кіберстійкості розроблено алгоритм адаптивного керування, який базується на моделі **тристанового скінченного автомата (Finite State Machine)**. Логіка переходів між станами визначається моніторингом часового інтервалу з моменту отримання останнього сигналу перевірки зв'язку (*heartbeat*).

Алгоритм передбачає функціонування системи у трьох режимах:

1. **ONLINE (Штатний режим)**. Система має стабільне з'єднання з сервером. Авторизація відбувається через центральну базу даних. У фоновому режимі виконується синхронізація хеш-образів лімітів для підготовки до можливих збоїв.

2. **DEGRADED (Перехідний режим)**. Діагностовано нестабільність з'єднання (втрата пакетів або затримка відповіді). Система намагається відновити зв'язок протягом короткого таймауту, перш ніж прийняти рішення про перехід в автономний режим. Це запобігає помилковим перемиканням при короткочасних збоях.

3. **OFFLINE (Аварійний режим)**. Фіксується повна втрата зв'язку (наприклад, внаслідок DDoS-атаки). Система переходить на використання локального захищеного реєстру та алгоритму хеш-ланцюжків для фіксації транзакцій.

Графічна інтерпретація розробленого алгоритму та умов переходу між станами наведена на рис. 1.

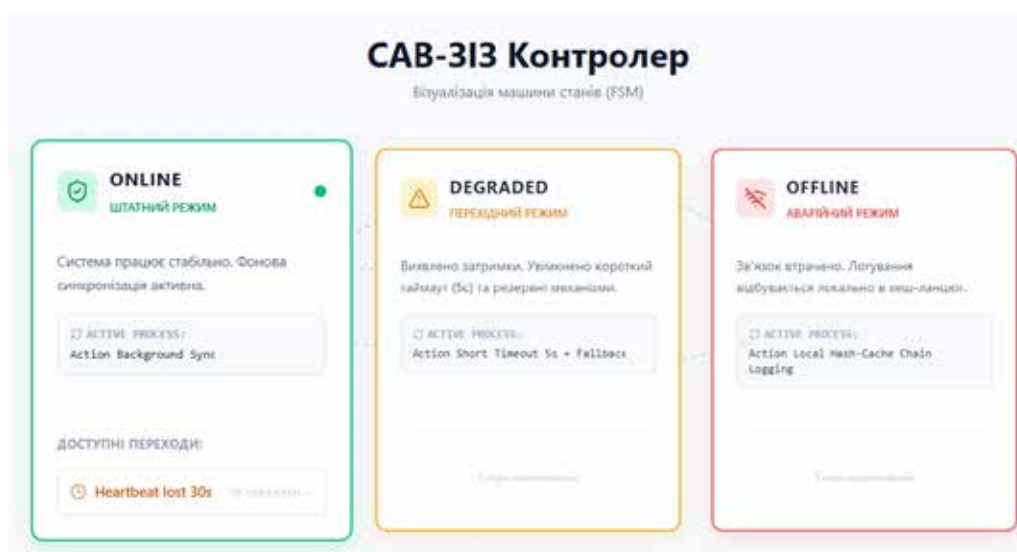


Рис. 1. Діаграма станів контролера CAB-313 та логіка переходів

Такий підхід дозволяє чітко детермінувати поведінку системи у будь-який момент часу, виключаючи невизначеність, характерну для класичних архітектур при збоях мережі.

Програмна реалізація та верифікація

Для перевірки ефективності запропонованих архітектурно-алгоритмічних рішень було розроблено імітаційну модель системи мовою Python. Метою експерименту є кількісна оцінка показників доступності сервісу (Availability) та стійкості до навантажень в умовах кібератак.

Сценарій моделювання відтворює роботу промислового об'єкта з наступними параметрами:

- Кількість користувачів: $N = 200$.
- Тривалість зміни: $T = 24$ години.
- Модель загрози: Атака на відмову в обслуговуванні (DDoS) з повною блокадою каналу зв'язку в період t .

Порівняльний аналіз роботи класичної та запропонованої систем наведено на рис. 2.

Графіки демонструють наступні результати:

1. Забезпечення безперервності (верхній графік). Класична система (червона лінія) демонструє повну зупинку обслуговування під час атаки (горизонтальне плато), що означає простій виробництва. Запропонована система (синя лінія) завдяки переходу в режим *Offline* продовжує видачу ЗІЗ, зберігаючи лінійне зростання кількості транзакцій. Показник доступності сервісу становить $>98\%$.

2. Керування навантаженням (нижній графік). Під час атаки транзакції накопичуються у локальному буфері (зростання синьої лінії). Після відновлення зв'язку ($t=15$) алгоритм не надсилає всі дані миттєво, а використовує механізм пакетної синхронізації (східчасте падіння графіка). Це дозволяє уникнути перевантаження мережі («мережевого шторму») та гарантує стабільність роботи каналу передачі даних.

Таким чином, результати моделювання підтверджують, що алгоритмічне забезпечення на базі тристанового автомата та периферійних обчислень дозволяє гарантувати виконання бізнес-процесів навіть в умовах критичних збоїв інфраструктури.

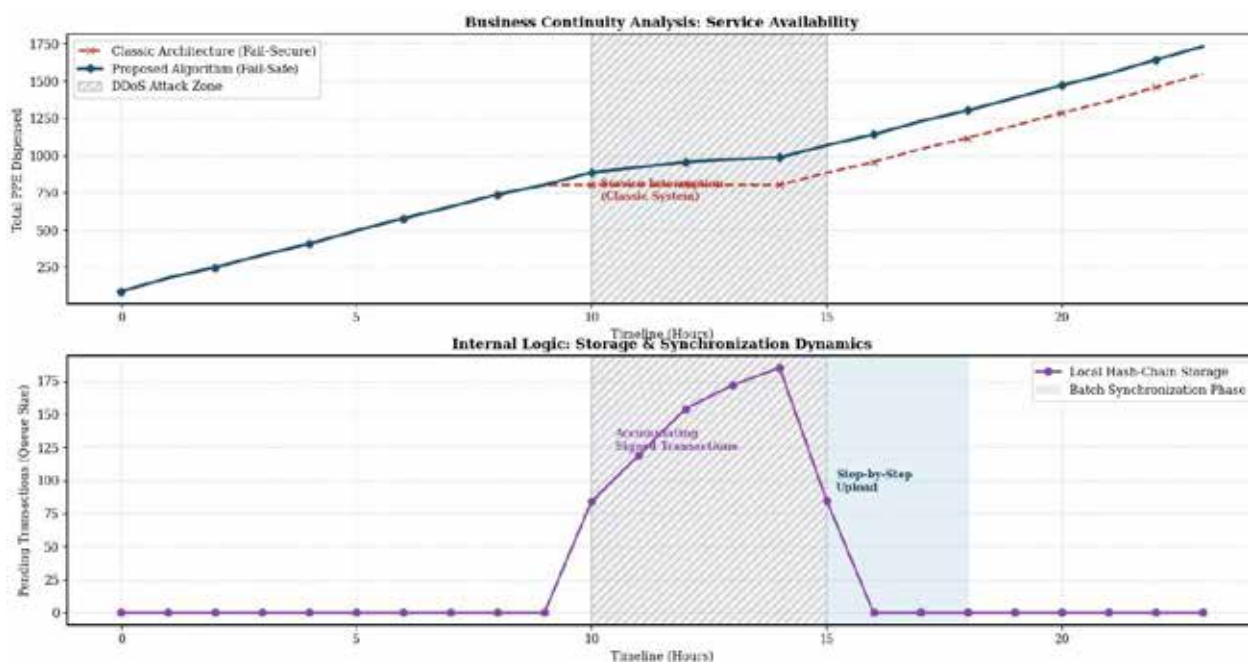


Рис. 2. Результати імітаційного моделювання доступності сервісу (вгорі) та процесу синхронізації даних (внизу)

Висновки

У роботі вирішено актуальне науково-практичне завдання забезпечення кіберстійкості критичної інфраструктури охорони праці в умовах деструктивних кібервпливів на ПОТ-системи. Основні наукові та практичні результати дослідження полягають у наступному:

1. Доведено обмеженість класичних архітектур. Встановлено, що традиційна централізована (хмарна) модель авторизації створює критичну точку відмови. Під час DDoS-атак вона діє за принципом безпечного блокування (Fail-Secure), що є неприпустимим для систем промислової безпеки, оскільки призводить до вимушеного простою виробництва та порушення законодавчих норм охорони праці.

2. Розроблено гібридну архітектуру на основі периферійних обчислень. Запропоновано підхід, що передбачає децентралізацію логіки прийняття рішень. Перенесення процесу валідації прав доступу на рівень контролера вендингового автомата дозволяє реалізувати концепцію безпечної відмови – гарантованої видачі критичних засобів захисту навіть за повної відсутності зв'язку з сервером.

3. Створено алгоритм адаптивної кіберстійкості. Розроблений метод базується на тристановому автоматі функціонування (ONLINE / DEGRADED / OFFLINE) та забезпечує автоматичне перемикання на локальний захищений кеш із застосуванням аварійних квот для запобігання конфліктам розщеплення (Split-Brain). Специфікація алгоритму опублікована у відкритому репозиторії GitHub для верифікації логіки обробки виняткових ситуацій.

4. Запропоновано метод захисту цілісності даних. Для протидії локальній фальсифікації в офлайн-режимі розроблено механізм криптографічної інкапсуляції записів з використанням технології хеш-ланцюжків на базі алгоритму SHA-256, що гарантує незмінність історії транзакцій.

5. Експериментально підтверджено ефективність методу. Результати імітаційного моделювання показали, що в умовах тривалої DDoS-атаки (5 годин) запропонована система забезпечує показник доступності сервісу на рівні понад 98%, тоді як класична архітектура демонструє повну зупинку обслуговування. Реалізований протокол пакетної синхронізації довів свою ефективність у запобіганні перевантаженню мережі на етапі відновлення зв'язку.

Практичне значення роботи полягає у можливості мінімізації економічних втрат промислових підприємств від простоїв, спричинених кібератаками, та забезпеченні безперервного захисту персоналу незалежно від стану IT-інфраструктури.

Перспективи подальших досліджень. Подальша робота буде спрямована на апаратну реалізацію прототипу модуля на базі платформи ESP32 з використанням апаратного шифрування та дослідження методів машинного навчання для виявлення аномальної поведінки користувачів в офлайн-режимі.

Список використаної літератури

1. Lezzi M., Corallo A., Lazoi M., Nimis A. B Measuring cyber resilience in industrial IoT: a systematic literature review. *Management Review Quarterly*. 2025. URL: <https://doi.org/10.1007/s11301-025-00495-8>.
2. Yang H., Dai L., Ma Y., Li Q., Xia Y. Resilient MPC With Switched Cost Functions for Cyber-Physical Systems Against DoS Attacks. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 2025. Vol. 55. P. 4444–4457. <https://doi.org/10.1109/tsmc.2025.3555826>.
3. Tang X., Zhao K., Qu H., Wang J. Adaptive event-triggered efficient output feedback model predictive control for networked interval type-2 T–S fuzzy system with data loss. *International Journal of Robust and Nonlinear Control*. 2023. Vol. 34. P. 1673–1695. <https://doi.org/10.1002/rnc.7049>.
4. Wang Y., Zhu F. Consensus Control for Nonlinear Multi-Agent Systems. *IEEE Transactions on Automation Science and Engineering*. 2025. Vol. 22. P. 10470–10483. <https://doi.org/10.1109/tase.2024.3524258>.
5. Dildar M., Khan A., Abbasi I., Akhtar R., Ruqaishi K., Sarwar S. End-to-End Security Mechanism Using Blockchain for Industrial Internet of Things. *IEEE Access*. 2025. Vol. 13. P. 20584–20598: <https://doi.org/10.1109/access.2025.3535821>.
6. Oh T. Blockchain-Enabled Security Enhancement for IoT Networks: Integrating LEACH Algorithm and Distributed Ledger Technology. *Journal of Machine and Computing*. 2025: <https://doi.org/10.53759/7669/jmc202505038>.
7. Segovia-Ferreira M., Rubio-Hernán J., Cavalli A., García-Alfaro J. Survey on Cyber-Resilience Approaches for Cyber-Physical Systems. *ACM Computing Surveys*. 2023. Vol. 56. P. 1–37: <https://doi.org/10.1145/3652953>.
8. Liu S., Trivedi A., Yin X., Zamani M. Secure-by-Construction Synthesis of Cyber-Physical Systems. *Annu. Rev. Control*. 2022. Vol. 53. P. 30–50. <https://doi.org/10.1016/j.arcontrol.2022.03.004>.
9. Bakirtzis G., Sherburne T., Adams S., Horowitz B., Beling P., Fleming C. An ontological metamodel for cyber-physical system safety, security, and resilience coengineering. *Software and Systems Modeling*. 2020. Vol. 21. P. 113–137: <https://doi.org/10.1007/s10270-021-00892-z>.
10. Крайнюк О. В., Буц Ю. В., Барбашин В. В., Яцюк М. В. Використання штучного інтелекту для управління безпекою праці. *Комунальне господарство міст. Серія: Технічні науки та архітектура*. 2023. № 6(180). С. 207–214. <https://doi.org/10.33042/2522-1809-2023-6-180-207-213>
11. Крайнюк О., Буц Ю., Барбашин В., Козодой Д., Козодой О. Інтелектуальні системи управління безпекою праці на основі штучного інтелекту: перспективи інтеграції в українське законодавство. *Комунальне господарство міст*. 2024. № 6(187). С. 242–251. <https://doi.org/10.33042/2522-1809-2024-6-187-242-251>.
12. Крайнюк О. В., Буц Ю. В., Богатов О. І., Лоцман П. І., Барбашин В. В. Управління засобами індивідуального захисту за допомогою вендингових автоматів. *The 9th International scientific and practical conference “Study of world opinion regarding the development of science”*. 2022. P. 672–678. URL: <https://doi.org/10.46299/ISG.2022.2.9>.

References

1. Lezzi, M., Corallo, A., Lazoi, M., & Nimis, A. (2025). Measuring cyber resilience in industrial IoT: A systematic literature review. *Management Review Quarterly*. DOI: 10.1007/s11301-025-00495-8
2. Yang, H., Dai, L., Ma, Y., Li, Q., & Xia, Y. (2025). Resilient MPC with switched cost functions for cyber-physical systems against DoS attacks. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 55, 4444–4457. DOI: 10.1109/tsmc.2025.3555826
3. Tang, X., Zhao, K., Qu, H., & Wang, J. (2023). Adaptive event-triggered efficient output feedback model predictive control for networked interval type-2 T-S fuzzy system with data loss. *International Journal of Robust and Nonlinear Control*, 34, 1673–1695. DOI: 10.1002/rnc.7049
4. Wang, Y., & Zhu, F. (2025). Distributed hybrid dynamic event-triggered consensus control for nonlinear multi-agent systems. *IEEE Transactions on Automation Science and Engineering*, 22, 10470–10483. DOI: 10.1109/tase.2024.3524258
5. Dildar, M., Khan, A., Abbasi, I., Akhtar, R., Ruqaishi, K., & Sarwar, S. (2025). End-to-end security mechanism using blockchain for Industrial Internet of Things. *IEEE Access*, 13, 20584–20598. DOI: 10.1109/access.2025.3535821
6. Oh, T. (2025). Blockchain-enabled security enhancement for IoT networks: Integrating LEACH algorithm and distributed ledger technology. *Journal of Machine and Computing*. DOI: 10.53759/7669/jmc202505038
7. Segovia-Ferreira, M., Rubio-Hernán, J., Cavalli, A., & García-Alfaro, J. (2023). A survey on cyber-resilience approaches for cyber-physical systems. *ACM Computing Surveys*, 56, 1–37. DOI: 10.1145/3652953
8. Liu, S., Trivedi, A., Yin, X., & Zamani, M. (2022). Secure-by-construction synthesis of cyber-physical systems. *Annual Reviews in Control*, 53, 30–50. DOI:10.1016/j.arcontrol.2022.03.004
9. Bakirtzis, G., Sherburne, T., Adams, S., Horowitz, B., Beling, P., & Fleming, C. (2020). An ontological metamodel for cyber-physical system safety, security, and resilience coengineering. *Software and Systems Modeling*, 21, 113–137. DOI:10.1007/s10270-021-00892-z

10. Krainiuk, O. V., Buts, Yu. V., Barbashin, V. V., & Yatsiuk, M. V. (2023). Use of artificial intelligence for occupational safety management. *Komunalne hospodarstvo mist. Seriia: Tekhnichni nauky ta arkhitektura*, (6(180)), 207–214. DOI: 10.33042/2522-1809-2023-6-180-207-213

11. Krainiuk, O., Buts, Yu., Barbashin, V., Kozodoi, D., & Kozodoi, O. (2024). Intelligent occupational safety management systems based on artificial intelligence: Prospects for integration into Ukrainian legislation. *Komunalne hospodarstvo mist*, (6(187)), 242–251. DOI: 10.33042/2522-1809-2024-6-187-242-251

12. Krainiuk, O. V., Buts, Yu. V., Bohatov, O. I., Lotsman, P. I., & Barbashin, V. V. (2022). Management of personal protective equipment using vending machines. In *The 9th International scientific and practical conference "Study of world opinion regarding the development of science"* (pp. 672–678). International Science Group. DOI: 10.46299/ISG.2022.2.9

Дата першого надходження рукопису до видання: 11.11.2025
Дата прийнятого до друку рукопису після рецензування: 08.12.2025
Дата публікації: 31.12.2025

УДК 004.056

DOI <https://doi.org/10.35546/kntu2078-4481.2025.4.3.16>**Я. А. МАРТИНЕНКО**

магістрант кафедри безпеки інформаційних технологій
Харківський національний університет радіоелектроніки
ORCID: 0009-0005-2958-0592

О. В. СЕВЕРІНОВ

кандидат технічних наук, доцент,
професор кафедри безпеки інформаційних технологій
Харківський національний університет радіоелектроніки
ORCID: 0000-0002-6327-6405

А. М. ЄВГЕНЬЄВ

старший викладач кафедри безпеки інформаційних технологій
Харківський національний університет радіоелектроніки
ORCID: 0000-0003-4365-5675

М. С. СКИБЕНКО

старший викладач кафедри безпеки інформаційних технологій
Харківський національний університет радіоелектроніки
ORCID: 0009-0002-4838-9329

ДОСЛІДЖЕННЯ SOAR-ПЛАТФОРМ ТА ОБҐРУНТУВАННЯ ВИБОРУ SPLUNK PHANTOM

У статті досліджено особливості застосування SOAR-платформ у контексті підвищення ефективності центрів моніторингу безпеки (SOC) за умов зростання складності корпоративних інфраструктур, збільшення кількості кіберінцидентів та розширення використання хмарних сервісів. Проаналізовано ключові проблеми, з якими стикаються аналітики SOC, зокрема великий обсяг подій, високий рівень хибнопозитивних сповіщень та обмеженість традиційних систем моніторингу, що не забезпечують автоматизованого реагування. Розглянуто архітектуру сучасних SOAR-платформ, яка охоплює інтеграційний, оркестраційно-аналітичний та операційний шари, а також принципи їх взаємодії з SIEM, EDR, IDS/IPS, системами керування вразливостями й платформами threat intelligence.

У роботі досліджено моделі автоматизації та оркестрації інцидентів, зокрема rule-based, state-based та context-aware підходи, які формують основу сучасних playbooks та забезпечують зниження навантаження на аналітиків, підвищення швидкості реагування та стандартизацію процесів. Особливу увагу приділено метрикам оцінювання ефективності SOAR-рішень, таким як MTTD, MTTR, Automation Coverage та False Positive Ratio, а також методам тестування масштабованості та інтеграційної надійності.

Представлено порівняльний аналіз провідних рішень: Splunk Phantom, Cortex XSOAR, IBM Resilient та Chronicle SOAR, із визначенням їх ключових переваг і обмежень. Показано, що Splunk Phantom вирізняється збалансованою архітектурою, гнучкою моделлю containers-artifacts, широкими можливостями інтеграції, підтримкою Python-скриптів та здатністю до побудови складних сценаріїв реагування. Отримані результати дозволяють визначити місце SOAR-платформ як ключового елемента сучасної інфраструктури кібербезпеки та підкреслити їхню роль у створенні відтворюваних, контрольованих і стандартизованих процесів реагування.

Ключові слова: SOAR, SOC, автоматизація реагування, оркестрація інцидентів, Splunk Phantom, playbooks, кіберінциденти.

YA. A. MARTYNENKO

Master's Student at the Department of Information Technology Security
Kharkiv National University of Radio Electronics
ORCID: 0009-0005-2958-0592

O. V. SIEVIERINOV

Candidate of Technical Sciences, Associate Professor,
Professor at the Department of Information Technology Security
Kharkiv National University of Radio Electronics
ORCID: 0000-0002-6327-6405

A. M. YEVHENIEV

Senior Instructor at the Department of Information Technology Security
Kharkiv National University of Radio Electronics
ORCID: 0000-0003-4365-5675

M. S. SKYBENKO

Senior Instructor at the Department of Information Technology Security
Kharkiv National University of Radio Electronics
ORCID: 0009-0002-4838-9329

RESEARCH ON SOAR PLATFORMS AND JUSTIFICATION FOR CHOOSING SPLUNK PHANTOM

The article examines the features of SOAR platforms in the context of improving the efficiency of security operations centers (SOCs) amid growing complexity of corporate infrastructures, increasing number of cyber incidents, and expanding use of cloud services. It analyzes the key problems faced by SOC analysts, including the large volume of events, the high level of false positives, and the limitations of traditional monitoring systems that do not provide automated response. The architecture of modern SOAR platforms is considered, covering the integration, orchestration and analytics, and operational layers, as well as the principles of their interaction with SIEM, EDR, IDS/IPS, vulnerability management systems, and threat intelligence platforms.

The paper explores models of incident automation and orchestration, in particular rule-based, state-based, and context-aware approaches, which form the basis of modern playbooks and reduce the workload on analysts, increase response speed, and standardize processes. Special attention is paid to metrics for evaluating the effectiveness of SOAR solutions, such as MTTD, MTTR, Automation Coverage, and False Positive Ratio, as well as methods for testing scalability and integration reliability.

A comparative analysis of leading solutions: Splunk Phantom, Cortex XSOAR, IBM Resilient, and Chronicle SOAR, is presented, identifying their key advantages and limitations. It is shown that Splunk Phantom stands out with its balanced architecture, flexible containers–artifacts model, extensive integration capabilities, support for Python scripts, and ability to build complex response scenarios. The results obtained allow us to define the place of SOAR platforms as a key element of modern cybersecurity infrastructure and emphasize their role in creating reproducible, controllable, and standardized response processes.

Key words: SOAR, SOC, response automation, incident orchestration, Splunk Phantom, playbooks, cyber incidents.

Постановка проблеми

У сучасних умовах центри моніторингу безпеки стикаються зі стрімким збільшенням кількості подій, значна частина з яких є хибно-позитивними або повторюваними. Традиційні засоби моніторингу фокусуються на зборі та кореляції даних, але не забезпечують автоматизованих механізмів оперативного реагування, що призводить до зростання навантаження на аналітиків SOC та збільшення часу обробки інцидентів. Відсутність стандартизованих, масштабованих та відтворюваних процесів реагування ускладнює забезпечення стабільної роботи SOC і знижує ефективність протидії кіберзагрозам. Це формує необхідність у впровадженні рішень, здатних автоматизувати рутинні задачі, стандартизувати процеси та підвищити швидкість реагування.

Аналіз останніх досліджень і публікацій

У сучасній науково-практичній літературі помітно посилення інтересу до проблем автоматизації реагування на кіберінциденти, оцінювання ефективності SOAR-платформ та їхньої інтеграції в інструментальний стек SOC. У роботі, присвяченій оркестрації, автоматизації та реагуванню, робить наголос на практичному застосуванні SOAR для підвищення ключових операційних показників SOC, зокрема, скорочення середнього часу на виявлення/реагування (MTTA/MTTR) та зміцнення безпекової позиції організації [1]. Книга має прикладну спрямованість і служить практичним керівництвом для аналітиків, які впроваджують та налаштовують сценарії реагування (playbooks) у реальних середовищах.

Практичні аспекти інтеграції SIEM-рішень у робочі процеси SOC висвітлені в праці Marlette, підкреслюють важливість правильної інтеграції SOAR-рішень із системами моніторингу та аналітики, зокрема SIEM. Зазначається, що ефективність автоматизованих сценаріїв реагування залежить від коректної побудови кореляційних правил, якості початкових даних і здатності інструментів забезпечувати взаємодію між підсистемами кіберзахисту [2].

У статті [3] розглядається питання автоматизації безпеки в хмарних середовищах, фокусуючись на балансі між виявленням загроз та заходами реагування. Ця робота має безпосереднє відношення до теми SOAR, адже архітектури сучасних SOC дедалі частіше охоплюють гібридні та хмарні інфраструктури, де питання масштабованості, інтеграції і узгодженості автоматизованих дій набувають ключового значення.

Результати щодо використання SOAR-інструментів у реальних умовах надають Bridges та співавтори у своїй статті, що демонструє практичні переваги та обмеження існуючих рішень. Автори вказують на позитивний вплив SOAR на продуктивність аналітиків та швидкість виконання рутинних дій, але також підкреслюють потребу в ретельному налаштуванні, оцінці human-in-the-loop інтервенцій та перевірці сценаріїв у контрольованому середовищі перед розгортанням у продакшн [4]. Такий підхід підтверджує тезу про необхідність комбінованого тестування як формального, так і практичного.

В [5] розвивається тема гіперавтоматизації, пропонуючи підходи з використанням «agentic» штучного інтелекту для динамічного формування workflows та підтримки експертів SOC. Робота демонструє, що поєднання AI-агентів із SOAR-інструментами може підвищити адаптивність сценаріїв реагування, проте вимагає уваги до контролю, аудиту та визначення меж автоматизації, аби уникнути небажаних дій чи помилкових ескалацій у продакшн-середовищі.

Аналіз літератури свідчить, що SOAR-платформи розглядаються як ключовий елемент сучасної архітектури кіберзахисту. Дослідження підтверджують їхню здатність підвищувати ефективність SOC, оптимізувати процеси реагування, забезпечувати узгодженість дій аналітиків та зменшувати час обробки інцидентів.

Формулювання мети дослідження

Метою дослідження є аналіз можливостей SOAR-платформ у контексті підвищення ефективності роботи SOC, визначення їх архітектурних та функціональних особливостей, оцінювання моделей автоматизації реагування та порівняння провідних рішень ринку. Особлива увага приділяється виявленню переваг Splunk Phantom як інструмента, що забезпечує високий рівень автоматизації, гнучкість побудови сценаріїв реагування та масштабованість, необхідну для сучасних корпоративних середовищ.

Викладення основного матеріалу дослідження

Зростання кількості кіберінцидентів, розширення корпоративних інфраструктур, використання хмарних сервісів та віддалених робочих місць формують нові виклики для центрів моніторингу безпеки (SOC). Аналітики стикаються з великою кількістю подій, значна частина яких є хибнопозитивними або повторюваними. За даними звітів провідних компаній з кібербезпеки, SOC-фахівці щоденно отримують від 5 000 до 50 000 подій, що підлягають аналізу. У таких умовах традиційні засоби моніторингу перестають бути достатніми, оскільки їхня функціональність зосереджена на зборі та кореляції подій, але не забезпечує автоматичного реагування.

SOAR-платформи вирішують цю проблему, надаючи можливість [1]:

- автоматизувати рутинні процеси;
- зменшити залежність від людського фактору;
- структурувати обробку інцидентів за уніфікованими процесами;
- швидко залучати додаткові джерела інформації (threat intel);
- документувати реагування згідно зі стандартами (NIST, ISO 27035).

Таким чином, SOAR стає ключовим елементом сучасної архітектури SOC, посилюючи традиційні засоби моніторингу.

Архітектура та ключові компоненти SOAR-платформ

Архітектура SOAR-платформ є багатокомпонентною структурою, спрямованою на забезпечення автоматизованого, стандартизованого та масштабованого реагування на кіберінциденти в межах центрів моніторингу безпеки. Типова SOAR-система складається з трьох ключових шарів: інтеграційного, оркестраційно-аналітичного та операційного. Інтеграційний шар забезпечує підключення до зовнішніх систем безпеки, такі як SIEM, EDR, IDS/IPS, фаєрволів, систем керування вразливостями, а також платформ Threat Intelligence. Взаємодія здійснюється через REST API, webhooks або спеціалізовані конектори, що дозволяє агрегувати інформацію з різнорідних джерел у єдине середовище аналізу.

На рисунку 1 подано узагальнену архітектуру SOAR-платформи, яка демонструє основні джерела даних, інтеграційні механізми, а також логічні блоки оркестрації та адаптивного реагування [1].

Оркестраційно-аналітичний шар відповідає за обробку подій, їх кореляцію, класифікацію та побудову автоматизованих сценаріїв (playbooks). Саме в цьому компоненті реалізуються логіка умовних переходів, ескалація інцидентів, валідація даних та прийняття рішень на основі контексту.

Багато платформ додатково використовують машинне навчання або правила поведінкового аналізу для зменшення кількості хибно-позитивних сповіщень.

Операційний шар підтримує взаємодію аналітиків із системою, забезпечуючи інтерфейси для відстеження інцидентів, ухвалення рішень, документування та контролю виконання. Тут реалізовані механізми контролю доступу, аудит дій, збереження артефактів та формування звітності відповідно до стандартів NIST чи ISO 27035.

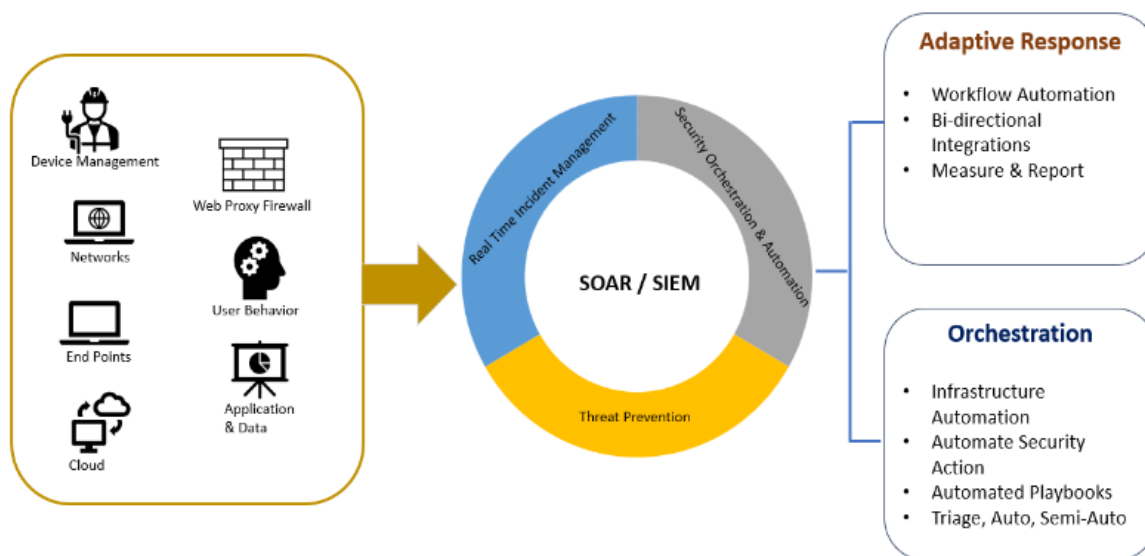


Рис. 1. Узагальнена архітектура SOAR-платформи

Подібна архітектура SOAR-платформ дозволяє не лише автоматизувати виконання рутинних завдань, а й створити узгоджене середовище реагування, де кожний інцидент проходить через стандартизований, контрольований і відтворюваний процес обробки.

Моделі автоматизації та оркестрації інцидентів

Автоматизація та оркестрація інцидентів у межах SOAR-платформ ґрунтуються на побудові формалізованих сценаріїв реагування, які забезпечують відтворюваність, швидкість та зниження навантаження на аналітиків SOC. У сучасних системах використовуються кілька моделей реалізації автоматизованих дій, серед яких найпоширенішими є rule-based, state-based та context-aware підходи.

Rule-based модель передбачає виконання фіксованих правил, що визначають логіку обробки інциденту відповідно до умов тригерів. Такий підхід є найпростішим, проте може потребувати значної кількості ручного налаштування у складних середовищах. State-based модель оперує станами інциденту, де кожен етап реагування відповідає певному вузлу workflow, а переходи між станами залежать від результатів попередніх дій. Це дозволяє реалізувати багатоетапні процеси, включаючи паралельне виконання задач та адаптацію сценарію в реальному часі. Context-aware автоматизація є найбільш гнучкою, оскільки враховує особливості інциденту, середовище виконання, попередню історію подій та поведінкові характеристики. Такий підхід дозволяє мінімізувати хибно-позитивні сповіщення, автоматично визначати пріоритетність інцидентів та оптимізувати вибір заходів реагування.

На рисунку 2 наведено повний цикл автоматизації та оркестрації інцидентів, що охоплює етапи збору подій, нормалізації, збагачення, прийняття рішень та виконання дій. Схема ілюструє взаємодію між аналітичними модулями SOAR, зовнішніми джерелами подій та системами, що забезпечують реагування.

Типові автоматизовані дії SOAR охоплюють enrichment (збагачення даних), containment (локалізація загрози), remediation (усунення наслідків), а також документування та ескалацію. Стандарти NIST Incident Response Lifecycle природно лягають у структуру playbooks, забезпечуючи відповідність кращим практикам реагування. Завдяки цим моделям SOAR-системи створюють цілісний цикл оркестрації, який дозволяє значно скоротити MTTR та підвищити ефективність SOC.

Оцінка ефективності SOAR: метрики, показники та методи тестування

Оцінювання ефективності SOAR-платформ є важливим етапом впровадження та оптимізації процесів реагування в SOC, оскільки дозволяє визначити реальний вплив автоматизації на швидкість, точність і якість обробки інцидентів. Для цього використовують комплекс метрик, що охоплюють як технічні аспекти роботи системи, так і організаційні показники взаємодії аналітиків із платформою [4-6].

До ключових метрик належать Mean Time to Detect (MTTD) та Mean Time to Respond (MTTR), які відображають середній час виявлення та реагування на інциденти. Зниження MTTR є основним показником успішності автоматизації, оскільки SOAR забезпечує швидке виконання дій, які раніше виконувалися вручну. Іншим важливим показником є Automation Coverage, що демонструє частку етапів обробки інцидентів, повністю або частково автоматизованих. Метрика False Positive Ratio дозволяє оцінити, наскільки правильно платформа фільтрує хибно-позитивні події, зменшуючи непотрібне навантаження на аналітиків.

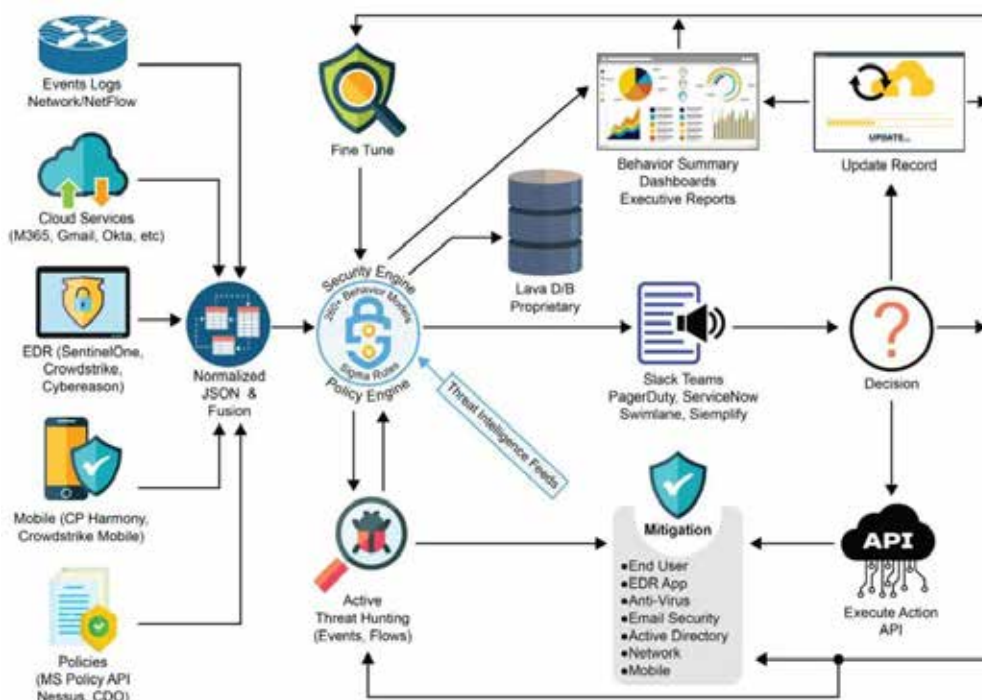


Рис. 2. Цикл автоматизації та оркестрації інцидентів у SOAR-платформі

У таблиці 1 узагальнено основні метрики та методи перевірки ефективності SOAR-рішень, що дозволяють комплексно оцінити швидкість, точність та якість автоматизованого реагування.

Методи тестування SOAR-рішень включають моделювання інцидентів, створення контрольованих сценаріїв з різними рівнями складності, а також валідацію інтеграційних дій із зовнішніми системами. Важливим є також проведення тестів на стійкість та масштабованість для оцінки можливостей системи обробляти великий обсяг подій.

Таблиця 1

Ключові метрики та методи оцінювання ефективності SOAR-платформ

Метрика / Метод	Опис	Призначення та інтерпретація
MTTD (Mean Time to Detect)	Середній час від появи інциденту до його виявлення	Показує ефективність кореляції та тригерів
MTTR (Mean Time to Respond)	Час від підтвердження інциденту до завершення реагування	Ключовий показник ефективності автоматизації; демонструє, наскільки SOAR скорочує цикл реагування
Automation Coverage	Частка кроків обробки інцидентів, що автоматизовані	Дозволяє оцінити ступінь зрілості та навантаження на аналітиків; високий рівень корелює зі зниженим MTTR
False Positive Ratio	Відношення хибно-позитивних подій до загальної кількості алертів	Дає змогу оцінити точність фільтрації й коректність інтеграцій; важливо для зменшення SOC-fatigue
Scenario-based Testing	Моделювання реальних інцидентів із різним рівнем складності	Дозволяє перевірити поведінку playbook'ів, надійність логіки та коректність автоматизованих рішень
Integration Validation	Перевірка роботи API, конекторів та зовнішніх сервісів	Виявляє проблеми зі збагаченням, containment-діями та взаємодією з інструментами безпеки
Stress & Scalability Tests	Навантажувальні тести (високий потік подій, паралельні workflow)	Дають розуміння, чи здатна платформа працювати в умовах пікового SOC-навантаження

Комплексний аналіз цих метрик дозволяє визначити, наскільки впроваджена SOAR-платформа відповідає вимогам SOC та чи забезпечує вона покращення безпекових процесів у порівнянні з традиційними методами [4-6].

Порівняльний аналіз провідних SOAR-платформ

Порівняльний аналіз сучасних SOAR-платформ є ключовим етапом вибору оптимального рішення для впровадження в інфраструктуру центру моніторингу безпеки. Ринок SOAR-технологій представлений кількома зрілими продуктами, серед яких найбільш відомими є Splunk Phantom, Palo Alto Cortex XSOAR, IBM Resilient та Google Chronicle SOAR (раніше Siemplify). Кожна з перелічених платформ має власний підхід до автоматизації, оркестрації та взаємодії зі сторонніми сервісами.

Cortex XSOAR відрізняється тим, що має багато готових інтеграцій у своєму маркетплейсі та зручну вбудовану систему, яка допомагає легко впорядковувати й визначати типи інцидентів. Платформа підтримує тисячі конекторів, однак часто потребує складного налаштування та має високу вартість володіння, особливо за умов масштабування на великі SOC. IBM Resilient орієнтований на процесний менеджмент інцидентів і пропонує потужну модель workflow, але порівняно обмежену автоматизацію та меншу гнучкість у налаштуванні сценаріїв. Chronicle SOAR забезпечує швидку інтеграцію з Google Security Stack, проте має суттєво менше можливостей для глибокої кастомізації, а його функціональність більше орієнтована на невеликі та середні команди реагування.

Для порівняння функціональних можливостей провідних SOAR-рішень у таблиці 2 наведено характеристику їх переваг та обмежень.

Таблиця 2

Порівняльна характеристика провідних SOAR-платформ

Платформа	Сильні сторони	Обмеження та недоліки
Splunk Phantom	Підтримка Python-скриптів і кастомних інтеграцій; Широка модель containers/artifacts; Глибока інтеграція зі Splunk Enterprise; Висока швидкість виконання playbooks	Потребує більш кваліфікованих адміністраторів; Вища вартість у масштабних інсталяціях
Palo Alto Cortex XSOAR	Величезний маркетплейс інтеграцій; Потужний інтерфейс інцидент-менеджменту; Готові "incident types"	Висока ціна володіння; Значні витрати на налаштування; Залежність від екосистеми Palo Alto
IBM Resilient	Сильна модель workflow; Орієнтація на процесний менеджмент інцидентів; Добре підходить для регульованих галузей	Обмежена автоматизація; Менше можливостей кастомізації; Менша кількість інтеграцій
Google Chronicle SOAR (Siemplify)	Швидка інтеграція з Google Security Stack; Зручний інтерфейс Triage-центру; Низький поріг входу	Значно менше інтеграцій; Обмежені можливості глибокого кастомного сценарію; Орієнтований на менші SOC

У цьому контексті Splunk Phantom вирізняється поєднанням широких можливостей інтеграції, масштабованої обробки інцидентів та високого рівня автоматизації. Значною перевагою є підтримка Python-скриптів, що дозволяє створювати кастомні дії будь-якої складності, гнучка модель даних «containers/artifacts», а також тісна інтеграція з Splunk Enterprise для глибокої аналітики подій. Phantom також забезпечує високу швидкість виконання playbooks, прозорий аудит дій та можливість побудови комплексних багаторівневих сценаріїв реагування.

Порівняння провідних рішень демонструє, що Splunk Phantom забезпечує найкраще співвідношення між гнучкістю, масштабованістю та функціональністю, що робить його раціональним вибором для сучасних SOC.

Архітектурні та функціональні переваги Splunk Phantom

Splunk Phantom вирізняється серед інших SOAR-платформ завдяки своїй гнучкій архітектурі, високій масштабованості та широким можливостям автоматизації, що робить його оптимальним рішенням для сучасних SOC-команд. Архітектурно Phantom побудований навколо моделі containers-artifacts, яка дозволяє структурувати всі дані інциденту у вигляді незалежних артефактів, зберігаючи повний контекст події та забезпечуючи гнучкість обробки. Такий підхід значно полегшує побудову сценаріїв реагування, оскільки кожний артефакт може бути використаний як вхідний параметр для подальших дій.

На рисунку 3 подано архітектуру взаємодії між Splunk Enterprise, Splunk Cloud та Splunk Phantom, яка демонструє процес формування, нормалізації й передачі подій у форматі CEF (Common Event Format) для подальшої автоматизованої обробки в SOAR-платформі [5, 6].

Ключовою перевагою платформи є можливість створення як стандартних, так і кастомних дій за допомогою Python-скриптів, що дозволяє інтегрувати практично будь-яку систему, сервіс або внутрішній інструмент компанії. Це забезпечує необмежену розширюваність та придатність для використання у складних корпоративних середовищах.

Широка бібліотека готових інтеграцій, що охоплює SIEM, EDR, IDS/IPS, хмарні сервіси та мережеве обладнання, гарантує швидке розгортання рішення без тривалих доопрацювань.

Функціональна частина Phantom підтримує високорівневі playbooks, побудовані за принципом графічних блоків дій. Це дозволяє аналітикам SOC створювати складні workflows без необхідності програмування, а технічним фахівцям – розширювати їх логіку при необхідності. Платформа також надає можливість паралельного виконання дій, умовних гілок, інтерактивних етапів та ручного втручання, що робить сценарії максимально адаптивними.

Тісна інтеграція зі Splunk Enterprise забезпечує ефективне використання аналітичних можливостей SIEM-платформи, включаючи кореляцію подій, побудову дашбордів та використання машинного навчання [6]. Крім того, Phantom пропонує розвинені засоби аудиту та документування інцидентів, що гарантує відповідність стандартам NIST та ISO.

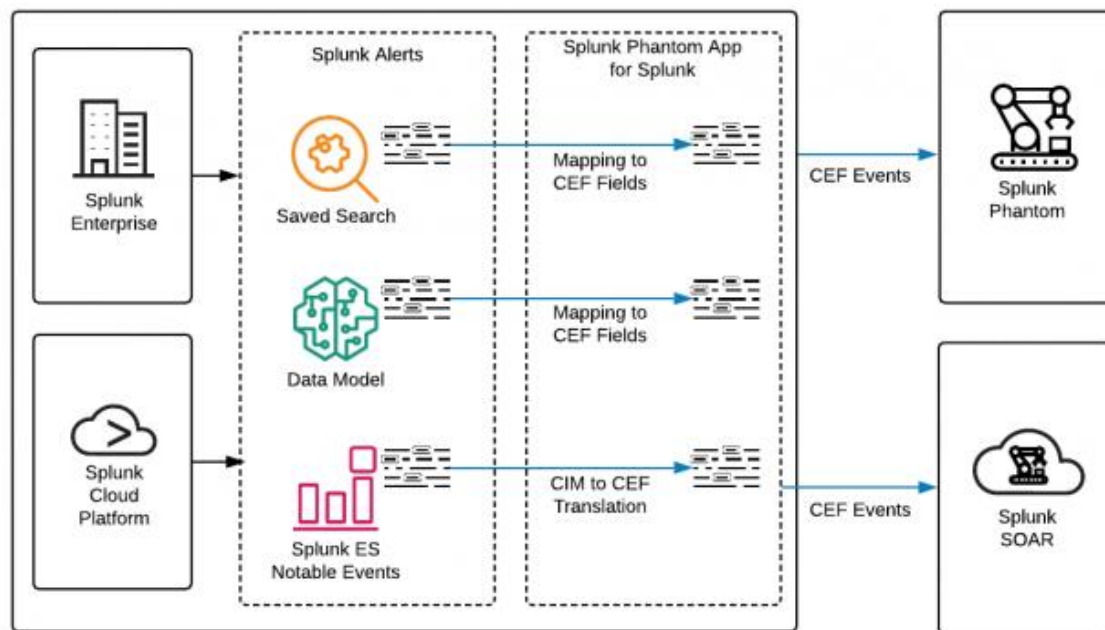


Рис. 3. Архітектура інтеграції Splunk Enterprise / Splunk Cloud із Splunk Phantom через CEF-події

У сукупності ці характеристики роблять Splunk Phantom одним із найбільш технічно досконалих і гнучких рішень для автоматизації реагування в SOC. Платформа забезпечує комплексний підхід до управління інцидентами, де високий рівень автоматизації поєднується з можливістю глибокої кастомізації та підтримкою будь-яких бізнес-процесів. Завдяки підходу «full lifecycle automation» Phantom охоплює всі етапи роботи з інцидентом – від первинного збору інформації, валідації та збагачення даних до локалізації загроз, виконання заходів з нейтралізації та формування детальної звітності. Додатковим фактором є масштабованість, яка дозволяє платформі рівномірно обробляти тисячі інцидентів без деградації продуктивності, що є критично важливим для великих корпоративних SOC із високою динамікою подій.

Phantom також забезпечує високий рівень прозорості та контрольованості процесів через розвинену систему логування, трекінгу змін та аудитних журналів, що спрощує процеси відповідності стандартам NIST, ISO 27035 та вимогам регуляторів. Гнучкий інтерфейс, можливість інтеграції з десятками зовнішніх джерел загроз, а також підтримка механізмів умовного прийняття рішень (human-in-the-loop) роблять платформу придатною як для повної автоматизації, так і для змішаних режимів роботи, де частина кроків контролюється аналітиком.

Використання Splunk Phantom дозволяє організаціям не лише прискорити та стандартизувати реагування, а й підвищити операційну стійкість SOC, суттєво зменшити навантаження на фахівців та забезпечити більш глибоку інтеграцію процесів безпеки у корпоративну інфраструктуру.

Висновки

У роботі проаналізовано роль SOAR-платформ у сучасних центрах моніторингу безпеки, що дозволяє зробити висновок про їхню ключову значущість для автоматизації рутинних процесів, зменшення впливу людського фактора та забезпечення структурованої обробки інцидентів за стандартними сценаріями реагування. Архітектура SOAR, яка включає інтеграційний, оркестраційно-аналітичний та операційний шари, дозволяє забезпечити повну взаємодію з зовнішніми системами безпеки, кореляцію подій, автоматизацію робочих процесів та контрольовану взаємодію аналітиків. Аналіз моделей автоматизації показав, що rule-based, state-based та context-aware підходи забезпечують різний рівень гнучкості та ефективності реагування на інциденти, причому context-aware автоматизація дозволяє оптимізувати пріоритетність інцидентів та зменшити кількість хибно-позитивних сповіщень.

Оцінка ефективності SOAR-систем базується на таких показниках, як MTTD, MTTR, Automation Coverage та False Positive Ratio, що дозволяє кількісно оцінити вплив платформи на продуктивність SOC, а методи тестування включають моделювання інцидентів та перевірку інтеграцій з зовнішніми системами. Порівняльний аналіз сучасних платформ показав, що Splunk Phantom вирізняється високою гнучкістю, масштабованістю та розвиненими можливостями автоматизації, що робить його оптимальним рішенням для великих корпоративних SOC. Особливості Splunk Phantom включають модель containers-artifacts, підтримку кастомних дій через Python, графічні playbooks, паралельне виконання задач та інтеграцію зі Splunk Enterprise, що забезпечує повний цикл автоматизованого реагування – від збору та збагачення даних до локалізації загроз, усунення наслідків та формування детальної звітності. Використання SOAR дозволяє підвищити ефективність роботи SOC, стандартизувати процеси реагування, зменшити навантаження на фахівців та інтегрувати безпекові процеси у корпоративну інфраструктуру.

Список використаної літератури

1. Kovacevic B. Оркестрація безпеки, автоматизація та реагування для аналітиків безпеки: дізнайтеся секрети SOAR для покращення MTTA і MTTR та зміцнення рівня безпеки. Packt Publishing, Limited, 2023.
2. Marlette T. Найкращі практики Splunk. Packt Publishing, Limited, 2016.
3. Pitkar H. Автоматизація хмарної безпеки через Symmetry: виявлення та реагування на загрози. Symmetry. 2025. Vol. 17, no. 6. P. 859. URL: <https://doi.org/10.3390/sym17060859> (дата звернення: 21.11.2025).
4. Тестування SOAR-інструментів у практичному використанні / R. A. Bridges et al. Computers & Security. 2023. P. 103201. URL: <https://doi.org/10.1016/j.cose.2023.103201> (дата звернення: 21.11.2025).
5. До надійної оркестрації безпеки та автоматизованого реагування в центрах операцій безпеки: гіперавтоматизаційний підхід із використанням агентного штучного інтелекту / Ismail et al. Information. 2025. Vol. 16, no. 5. P. 365. URL: <https://doi.org/10.3390/info16050365> (дата звернення: 21.11.2025).
6. Москвін К.С., Северінов О.В., Сидоренко З.М., Балагура Д.С., Литвин А.В. Дослідження впливу інтеграції засобів кіберзахисту на захищеність IT-інфраструктури організації. Том 2 № 2(93) (2025): Вісник Херсонського національного технічного університету. С. 246-255. DOI: <https://doi.org/10.35546/kntu2078-4481.2025.2.2.29>

References

1. Kovacevic B. (2023) Security Orchestration, Automation and Response for Security Analysts: Learn the Secrets of SOAR to Improve MTTA and MTTR and Strengthen Your Security Posture. Packt Publishing, Limited.
2. Marlette T. (2016) Splunk Best Practices. Packt Publishing, Limited.
3. Pitkar H. (2025) Cloud Security Automation Through Symmetry: Threat Detection and Response. Symmetry. Vol. 17, no. 6. P. 859. URL: <https://doi.org/10.3390/sym17060859> (date of access: 23.11.2025).
4. Bridges R. A. et al. Computers & Security (2023) Testing SOAR Tools in Use. P. 103201. URL: <https://doi.org/10.1016/j.cose.2023.103201> (date of access: 21.11.2025).
5. Toward Robust Security Orchestration and Automated Response in Security Operations Centers with a Hyper-Automation Approach Using Agentic Artificial Intelligence / Ismail et al. Information. 2025. Vol. 16, no. 5. P. 365. URL: <https://doi.org/10.3390/info16050365> (date of access: 22.11.2025).
6. Moskvyn K.S., Sievierinov O.V., Sydorenko Z.M., Balagura D.S., Lytvyn A.V. (2025) Research on the Impact of Cybersecurity Tools Integration on the Security of an Organization's IT Infrastructure. Visnyk of Kherson National Technical University. Vol. 2, No. 2(93). P. 246-255. DOI: <https://doi.org/10.35546/kntu2078-4481.2025.2.2.29>

Дата першого надходження рукопису до видання: 26.11.2025
Дата прийнятого до друку рукопису після рецензування: 15.12.2025
Дата публікації: 31.12.2025

Д. М. НЕХОРОШИХ

асистент кафедри безпеки інформаційних технологій
Харківський національний університет радіоелектроніки
ORCID: 0000-0002-6455-2221

П. В. БУСЛОВ

старший викладач кафедри безпеки інформаційних технологій
Харківський національний університет радіоелектроніки
ORCID: 0000-0002-2558-4984

А. М. ЄВГЕНЬЄВ

старший викладач кафедри безпеки інформаційних технологій
Харківський національний університет радіоелектроніки
ORCID: 0000-0003-4365-5675

М. С. СКИБЕНКО

старший викладач кафедри безпеки інформаційних технологій
Харківський національний університет радіоелектроніки
ORCID: 0009-0002-4838-9329

АНАЛІЗ ФРЕЙМВОРКІВ ТА ТЕХНІЧНИХ ЗАСОБІВ ЗАХИСТУ ДЛЯ ІНФРАСТРУКТУРИ ІНТЕРНЕТУ РОБОТИЗОВАНИХ РЕЧЕЙ

У статті проведено комплексний аналіз архітектурних, організаційних та технологічних засад формування системи управління інформаційною безпекою (СУІБ) для інфраструктури Інтернету роботизованих речей (IoRT) що поєднує можливості IoT та робототехніки й характеризується високим рівнем автономності, масштабованості та критичністю функцій. Розглянуто класифікацію сфер застосування IoRT за суб'єктами використання та галузевими напрямками. На основі вимог стратегічного та оперативного рівнів управління окреслено ключові складові СУІБ для IoRT.

Особливу увагу приділено аналізу адміністративних, фізичних і технічних заходів безпеки, що охоплюють управління ідентифікацією й автентифікацією пристроїв, контроль доступу, безпечну взаємодію з постачальниками, захист апаратури, мережеву сегментацію, криптографічні механізми та безпечний життєвий цикл програмного забезпечення. У контексті впровадження європейської директиви NIS2 наголошено на необхідності інтеграції міжнародних стандартів ISMS (ISO, NIST, ISA) на всіх рівнях архітектури IoRT.

Проведено аналіз сучасних тенденцій розгортання СУІБ, виявлені основні сценарії ризиків, типові загрози та вразливості, які характерні як для IoT, IoRT технологій зокрема так і комплексної інфраструктури, що використовує IoT, IoRT кластер, вцілому.

Запропонований підхід застосування декількох фреймворків управління безпекою в рамках однієї розгалуженої інфраструктури у відповідності з виявленим контекстом та результатами процесу управління ризиками.

Присвячено увагу можливостям і обмеженням використання технології блокчейн для підвищення рівня безпеки IoRT. Проаналізовано етапи інтеграції блокчейну. Водночас підкреслено проблеми масштабованості, затримок обробки та енергоспоживання, що обмежують застосування блокчейну у сценаріях реального часу. Запропоновано гібридний підхід, що передбачає поєднання блокчейну з децентралізованими базами даних для досягнення балансу між продуктивністю та безпекою.

Ключові слова: IoRT, СУІБ, управління ризиками, загрози, вразливості, інформаційні активи, блокчейн, децентралізовані системи, інтелектуальні пристрої, екосистема IoT, критична інфраструктура, цифрова безпека.

D. M. NEKHOROSHYKH

Assistant Professor at the Department of Information Technology Security
Kharkiv National University of Radio Electronics
ORCID: 0000-0002-6455-2221

P. V. BUSLOV

Senior Instructor at the Department of Information Technology Security
Kharkiv National University of Radio Electronics
ORCID: 0000-0002-2558-4984

A. M. YEVHENIEV

Senior Instructor at the Department of Information Technology Security
Kharkiv National University of Radio Electronics
ORCID: 0000-0003-4365-5675

M. S. SKYBENKO

Senior Instructor at the Department of Information Technology Security
Kharkiv National University of Radio Electronics
ORCID: 0009-0002-4838-9329

ANALYSIS OF FRAMEWORKS AND TECHNICAL SECURITY MEASURES FOR THE INFRASTRUCTURE OF THE INTERNET OF ROBOTIC THINGS

The article provides a comprehensive analysis of the architectural, organizational, and technological foundations for the formation of an information security management system (ISMS) for the Internet of Things (IoRT) infrastructure, which combines the capabilities of IoT and robotics and is characterized by a high level of autonomy, scalability, and criticality of functions. The classification of IoRT application areas by user entities and industry areas is considered. Based on the requirements of the strategic and operational levels of management, the key components of the ISMS for IoRT are outlined.

Particular attention is paid to the analysis of administrative, physical, and technical security measures, covering device identification and authentication management, access control, secure interaction with suppliers, hardware protection, network segmentation, cryptographic mechanisms, and a secure software lifecycle. In the context of the implementation of the European NIS2 directive, the need to integrate international ISMS standards (ISO, NIST, ISA) at all levels of the IoRT architecture is emphasized.

An analysis of current trends in the deployment of ISMS was conducted, the main risk scenarios, typical threats and vulnerabilities that are characteristic of both IoT, IoRT technologies in particular and complex infrastructure using IoT, IoRT cluster, as a whole were identified.

An approach is proposed to use several security management frameworks within one branched infrastructure in accordance with the identified context and results of the risk management process.

Attention is paid to the possibilities and limitations of using blockchain technology to increase the level of IoRT security. The stages of blockchain integration are analyzed. At the same time, the problems of scalability, processing delays and energy consumption that limit the use of blockchain in real-time scenarios are emphasized. A hybrid approach is proposed that involves combining blockchain with decentralized databases to achieve a balance between performance and security.

Key words: IoRT, ISMS, risk management, threats, vulnerabilities, information assets, blockchain, decentralized systems, intelligent devices, IoT ecosystem, critical infrastructure, digital security.

Постановка проблеми

Сучасні інформаційні системи функціонують у середовищах, що характеризуються високою мінливістю, динамікою взаємодій та зростаючим обсягом поведінкових даних про діяльність користувачів і сервісів. Швидкий розвиток технологій створив з'єднаний світ, де різні системи та пристрої легко інтегруються в усі аспекти нашого життя. Інтернет роботизованих речей (IoRT) з'єднує роботи та автономні пристрої з Інтернетом, щоб вони могли спілкуватися, співпрацювати над даними та виконувати завдання більш ефективно. IoRT має потенціал повністю трансформувати різні галузі, включаючи виробництво, охорону здоров'я та транспорт. Однак проблеми безпеки можуть призвести до збоїв систем IoRT, роблячи їх вразливими, що є правдою для будь-якого пристрою, підключеного до Інтернету [3].

З урахуванням сучасних підходів до інформаційної та кібербезпеки, жодний інформаційний актив не можна розглядати у відриві його існування, як частини СУІБ. Тому таку інфраструктуру, як IoRT треба розглядати виключно з точки зору оцінки інформаційної безпеки в рамках міжнародних фреймворків ISO, NIST, ISA, стандартів захисту персональних даних, таких як GDPR [13].

З урахуванням оцінки ризиків, як невід'ємного компонента та заходу безпеки, були виділені головні потенційні загрози у даному сегменті та вразливості інфраструктури IoRT [3]. На підставі цього, існує необхідність приділити увагу аналізу безпеки функціонування як окремих компонентів інфраструктури IoRT (речі, блок інтелектуального прийняття рішень, датчики та механізми виконання, вбудовані системи, комунікація), так і надійну роботу взаємопов'язаних роботизованих систем, безпеку інфраструктури IoRT в цілому. Зважаючи на численні

переваги, які пропонує IoRT, особливу увагу слід приділити заходам безпеки для захисту чутливих даних від зловмисних атак, контролюванням пристроєм, чи зловживання повноваженнями операційної системи, які керують пристроями [6].

Аналіз останніх досліджень і публікацій

Опрацювання доробків наукового кола з досліджуваної тематики дозволяють констатувати відсутність детальної уваги до питань безпеки, що стосуються програмного забезпечення, повідомлень, які потрібно транспортувати, та даних [7]. Слід відзначити важливість наукових пошуків з використання безпечних методів обміну повідомленнями, таких як HTTPS, тим не менш, концепція безпеки не була суттєво добре висвітлена. Вермесан, Овідіу та інші науковці презентували нову стратегію забезпечення безпеки IoRT шляхом поєднання технології блокчейн з методами шифрування.

Автори пропонують безпечну рамку, яка використовує властивості безпеки цих технологій для створення потужної системи безпеки для IoRT. Корисність їхньої запропонованої рамки демонструється в дослідженні шляхом використання її в симульованому середовищі. Об'єднуючи передові технології, це нове рішення для безпеки IoRT долає труднощі захисту щільно взаємопов'язаних мереж роботизованих пристроїв [1]. Блокчейн також відомий своєю здатністю забезпечувати безпеку в системі IoRT [2]. У наукових колах пропонується безпечна система обміну даними для IoRT. Ця система використовує технологію блокчейн для гарантії цілісності та конфіденційності обміну даними.

Рішення використовує технологію розподіленого реєстру з блокчейну для створення незмінного запису передачі даних, створюючи прозоре та надійне середовище для обміну даними між роботизованим обладнанням. Одна з основних проблем у IoRT вирішується стратегією авторів, яка підкреслює важливість підтримки цілісності та автентичності даних. Андроніс, Михайл та інші дослідники запропонували механізм обміну даними на основі блокчейн для IoRT, який є безпечним і масштабованим [1][2]. Автори створюють безпечне середовище для обміну даними між роботизованими пристроями, використовуючи блокчейн, одночасно забезпечуючи масштабованість та ефективне використання ресурсів.

Формулювання мети дослідження

Метою дослідження є комплексний аналіз архітектурних фреймворків і технічних механізмів захисту в системах Інтернету роботизованих речей (IoRT) у контексті управління інформаційною безпекою, а також виявлення ключових прогалин та формування напрямів подальших досліджень для підвищення надійності та стійкості IoRT-інфраструктур.

Викладення основного матеріалу дослідження

Сучасна технологія, відома як Інтернет роботизованих речей поєднує IoT з робототехнікою, щоб забезпечити її автономність. Сфери застосування IoT є досить масштабним, їх можна класифікувати за декількома ознаками:

1. За суб'єктом застосування. Технології IoT охоплюють широкий спектр застосувань, включаючи медичний моніторинг, інфраструктуру «розумного дому» та офісу, промислові підприємства й ресурсодобувні об'єкти, міські системи управління та міжміські транспортно-логістичні платформи, забезпечуючи підвищення якості життя, операційної ефективності, безпеки та керованості в усіх відповідних середовищах.

2. За галузями впровадження можна виділити наступні сфери застосування. IoT-технології демонструють широку міжгалузеву застосовність – від масмедіа, де вони забезпечують персоналізовану взаємодію з користувачами, та екологічного моніторингу, що підвищує безпеку й оперативність реагування, до інтелектуального виробництва й цифрових систем керування, які оптимізують процеси, енергоспоживання та надійність інфраструктури, а також транспортної сфери, де IoT формує інтегровані високоефективні системи керування трафіком, логістикою та дорожньою безпекою.

Стратегічний рівень управління інформаційною безпекою формується такими заходами безпеки, як:

– Контекст середовища, сфера застосування. Для ефективного функціонування СУІБ в інфраструктурі IoRT необхідно визначити зовнішні та внутрішні фактори, врахувати потреби й вимоги зацікавлених сторін та встановити вимірювані, узгоджені з політикою й ризиками цілі інформаційної безпеки, забезпечуючи їх документування, комунікацію та постійне оновлення [13].

– Лідерство. Найвище керівництво повинне демонструвати лідерство у розбудові СУІБ для інфраструктури IoRT, установлюючи й документуючи політику інформаційної безпеки, що відповідає цілям організації, містить зобов'язання щодо дотримання вимог і постійного вдосконалення, доводиться до всіх рівнів персоналу та супроводжується чітким призначенням відповідальності й повноважень за ключові функції безпеки.

– Планування. Організація повинна, з урахуванням контексту застосування, визначати ризики та можливості для СУІБ інфраструктури IoRT і планувати дії, що забезпечать досягнення очікуваних результатів, зменшення небажаних ефектів та постійне вдосконалення, інтегруючи ці заходи у процеси системи управління інформаційною безпекою та регулярно оцінюючи їхню ефективність.

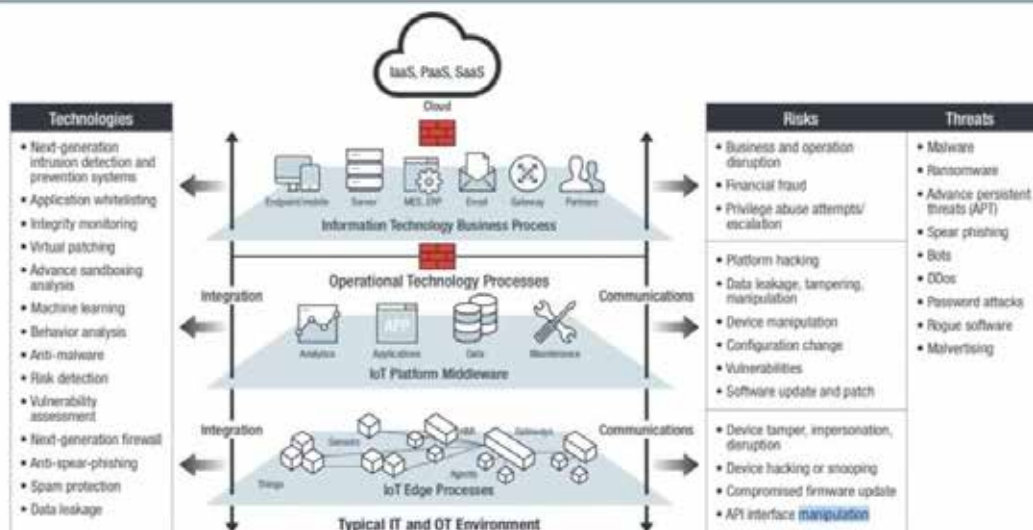


Рис. 1. Розгортання спеціальної рамки управління ІБ для IoT, IoRT за ISA IEC 62442

– **Забезпечення.** Інфраструктура IoRT повинна забезпечити необхідні ресурси, компетентність і поінформованість персоналу, організувати ефективні внутрішні та зовнішні комунікації та підтримувати належний рівень документованої інформації, охоплюючи її створення, оновлення, зберігання й вилучення, щоб гарантувати результативне функціонування та постійне вдосконалення системи менеджменту інформаційної безпеки.

– **Управління ризиками.** Оперативне планування та управління в інфраструктурі IoRT вимагає встановлення критеріїв процесів, керування ними відповідно до міжнародних стандартів (зокрема ISO/TR 22100-4:2018 та ISA/IEC 62443), а також впровадження систематичного процесу оцінки ризиків, що забезпечує їх ідентифікацію, аналіз, пріоритизацію та документування. Регулярна оцінка та обробка ризиків є необхідною для підтримання безпеки, цілісності й доступності IoRT-систем та гарантує, що можливості технології реалізуються без зростання загроз.

– **Оцінка результатів функціонування, моніторинг аудит.** Ефективне функціонування IoRT вимагає чітко визначених процедур моніторингу та вимірювання, які охоплюють об'єкти контролю, методи збору даних, відповідальних осіб і регулярний аналіз результатів, забезпечуючи своєчасне виявлення відхилень, вразливостей та інцидентів безпеки. Постійний моніторинг, внутрішні аудити та періодичний аналіз з боку керівництва гарантують адаптивність, відповідність вимогам і безперервну результативність системи менеджменту інформаційної безпеки IoRT [13].

Для підтримання ефективності та актуальності СУІБ організація повинна забезпечувати її постійне вдосконалення, системно реагуючи на невідповідності, усуваючи їхні причини, оцінюючи результативність коригувальних дій і вносячи необхідні зміни. Рішення керівництва за результатами аналізу мають бути спрямовані на розвиток можливостей постійного поліпшення та вдосконалення системи управління інформаційною безпекою в інфраструктурі IoRT [14].

З урахуванням впровадження загальноєвропейської директиви NIS2 особливо важливо забезпечити імплементацію базових світових стандартів управління інформаційною безпекою ISMS, таких як ISO, NIST та ISA, на всіх рівнях архітектур IoRT, оскільки створення пов'язаної екосистеми передбачає інтеграцію роботизованого обладнання, інструментів та IoT-технологій у рамках IoRT. У цьому контексті особливого значення набувають інноваційні технології підсилення безпеки, зокрема блокчейн – децентралізований і незмінний цифровий реєстр, який забезпечує прозоре, захищене та стійке до маніпуляцій зберігання даних і транзакцій, усуваючи потребу в посередниках та підвищуючи ефективність і довіру в різних сферах – від фінансів і логістики до цифрової ідентифікації та смарт-контрактів. До етапів впровадження блокчейну IoRT слід відносити:

1. **Визначення випадку використання.** Критично важливим першим кроком у впровадженні IoRT є визначення випадку використання. Це передбачає визначення операційних труднощів та оцінку можливостей роботизованої автоматизації. Один з факторів, який слід врахувати, це вивчення завдань, які є повторюваними, небезпечними, трудомісткими або потребують високої точності. Також важливо врахувати, чи підходить IoRT для таких завдань, як управління матеріалами, складання, інспекція або спостереження. Визначення випадку використання вимагає глибокого розуміння специфічних вимог галузі або організації. Організації можуть визначити найбільш ефективний випадок використання для використання роботизованих технологій з метою підвищення продуктивності, ефективності, безпеки та загальної продуктивності, співвідносячи можливості IoRT з визначеними труднощами та завданнями [1].

2. Вибір відповідної платформи блокчейн – передбачає порівняння характеристик різних платформ блокчейн, їх масштабованості, безпеки, взаємодії та процедур консенсусу. Ще одним фактором, який слід врахувати, є здатність платформи управляти величезною кількістю та швидкістю інформації, що виробляється пристроями IoT. Важливо оцінити потенціал інтеграції та сумісність з існуючими системами. Для довгострокового успіху також критично важливо оцінити екосистему розробників платформи, документацію та підтримку спільноти. Організації можуть забезпечити безперебійну інтеграцію, точність даних та надійність серед пристроїв IoT, ретельно оцінюючи та вибираючи правильну платформу блокчейн, що дозволяє безпечну та прозору комунікацію, обмін даними та автоматизовані транзакції в екосистемі IoT [8].

3. Визначити схему даних. Це передбачає визначення організації та структури даних, які генерують пристрої IoT. Серед факторів, які слід врахувати, є види даних, що підлягають збору, включаючи показання датчиків, статус обладнання, дані про місцезнаходження або експлуатаційні метрики. Визначення властивостей даних, методів вимірювання в порядку та будь-яких інших метаданих також є важливим. Схема повинна відповідати конкретному випадку використання та необхідній аналітиці або обробці. Необхідно забезпечити сумісність та взаємодію даних між різними системами та пристроями IoT. Організації можуть ефективно захищати, аналізувати та отримувати цінні інсайти з даних, що генеруються пристроями IoT, визначивши встановлену схему даних, що дозволяє приймати обґрунтовані рішення та оптимізувати систему в цілому.

4. Інтеграція пристроїв IoT з блокчейном. Це передбачає налаштування апаратури для того, щоб вона могла спілкуватися з платформою технології блокчейн за допомогою протоколів та API. Окрім забезпечення процедур валідації даних та консенсусу, важливими аспектами є підтримка безпеки пристрою, обмеження доступу та конфіденційності даних. Щоб забезпечити безперешкодний обмін даними та комунікацію, процес об'єднання також має гарантувати взаємодію між різними апаратними та програмними платформами. Організації можуть забезпечити незмінність даних, прозорість та довіру в екосистемах IoT, поєднуючи блокчейн з пристроями IoT, що дозволяє безпечні та систематичні транзакції та взаємодії між пристроями [2]. Це з'єднання також може зробити можливим розвиток децентралізованих додатків та послуг, які використовують обидві технології IoT та на основі блокчейн.

5. Розгортання розумних контрактів. Процес впровадження IoT повинен включати розгортання розумних контрактів. На платформі блокчейн повинні бути запрограмовані та розгорнуті самовиконувані контракти. Логіка контракту, обставини та поведінка, які пристрої IoT повинні виконувати самостійно, повинні бути враховані. У IoT розумні контракти можуть спростити процеси, такі як координація пристроїв, розподіл ресурсів або врегулювання платежів. Щоб уникнути незаконного доступу або зловмисної діяльності, безпека та надійність розумних контрактів повинні бути гарантовані. Організації можуть автоматизувати взаємодії пристроїв, створити довіру та забезпечити безперешкодне прозоре виконання контрактних зобов'язань, впроваджуючи розумні контракти в IoT [2]. Як наслідок, діяльність екосистеми IoT може стати більш ефективною, точною та продуктивною, що призведе до підвищення автономії та загальної ефективності підключених пристроїв.

6. Моніторинг та підтримка мережі блокчейн. Для успішної реалізації IoT важливо моніторити та управляти мережею блокчейн. Для забезпечення надійності, безпеки та ефективності мережі необхідно постійне моніторинг мережі. Стан вузлів блокчейн відстежується разом із швидкістю оброблюваних транзакцій та дійсністю механізму консенсусу. Крім того, слід вжити проактивних заходів, таких як періодичні оновлення, модифікації та аудити безпеки, щоб усунути вразливості та гарантувати надійність мережі. Щоб підтримувати зростаючу кількість пристроїв IoT, організаціям також потрібно стежити за тим, як ефективно мережа використовує свої ресурси та наскільки легко вона може розширюватися. Організації можуть забезпечити безперебійну роботу, точність інформації та надійність серед пристроїв IoT, ретельно відстежуючи та підтримуючи мережу блокчейн. Це допоможе створити надійне та безпечне середовище для автономних операцій, обміну інформацією та взаємодій.

7. Протоколи консенсусу блокчейн. Технологія блокчейн використовує різноманітні механізми консенсусу, кожен з яких має свої переваги та недоліки.

8. Критична оцінка життєздатності блокчейн у IoT. Децентралізовані бази даних, такі як InterPlanetary File System (IPFS) та Apache Cassandra, пропонують надійні рішення для зберігання та отримання даних у розподілених системах. Ці бази даних відзначаються в середовищах, де критично важливими є висока доступність, стійкість до збоїв та горизонтальна масштабованість. Вони розроблені для ефективної обробки великих обсягів даних та забезпечення узгодженості даних між розподіленими вузлами [4].

Безпека IoT значною мірою залежить від шифрування даних та безпечної комунікації, оскільки вони забезпечують конфіденційність, цілісність та автентичність інформації, що обмінюється між пристроями, захищаючи її від несанкціонованого доступу або підробки. Шифруючи відкритий текст за допомогою методів шифрування та секретних ключів, дані стають недоступними для несанкціонованих осіб, які не мають ключа дешифрування. Основними для систем IoT є шифрувальні протоколи, такі як Безпека транспортного рівня (TLS) або Протокол захищених сокетів (SSL), які автентифікують та шифрують дані під час передачі, забезпечуючи безпеку з кінця в кінець. Щоб захистити ключі шифрування від зловживання або розголошення, надійні системи управління ключами є обов'язковими в системах IoT. Це передбачає безпечне генерування, розподіл та управління ключами,

а також своєчасне відкликання та заміну скомпрометованих ключів. Шифрування в реальному часі повинно виконуватись безперешкодно без компрометації продуктивності системи. Це вимагає вдосконалень апаратного або програмного забезпечення для забезпечення високошвидкісних процесів шифрування та дешифрування. Крім того, стратегії шифрування даних IoRT повинні враховувати безпечні методи зберігання та відновлення, щоб забезпечити захист даних навіть під час періодів бездіяльності [3].

Для безпечної комунікації можна використовувати алгоритми шифрування, такі як криптографія на основі еліптичних кривих (ECC) та RSA, щоб захистити з'єднання. Ці алгоритми гарантують конфіденційність і безпеку переданих даних. Крім того, зашифровані методи обміну повідомленнями, такі як MQTT-SN або CoAP, можуть бути використані для встановлення безпечної комунікації в IoRT [12]. Навіть у контекстах з обмеженими ресурсами використання легковагових технік повідомлень цими протоколами забезпечує безпечну та надійну комунікацію. В загальному, безпечна комунікація є необхідною для забезпечення надійності та безпеки додатків IoRT [9]. Завдяки впровадженню надійного шифрування та механізмів безпечної комунікації дані залишаються недоступними для несанкціонованих осіб, що зміцнює загальну безпеку системи IoRT. Ці заходи слугують основою в системах IoRT, гарантуючи безпеку переданих і збережених даних та встановлюючи фундамент для надійної та стійкої екосистеми IoRT.

Адміністративний рівень безпеки IoRT охоплює управління ідентифікацією та автентифікацією пристроїв, контроль прав доступу та взаємодію з постачальниками. Ідентифікація й автентифікація мають забезпечувати надійне підтвердження пристроїв і користувачів за допомогою криптографічних механізмів та керування всім життєвим циклом ідентифікаторів. Права доступу повинні призначатися, переглядатися та скасовуватися відповідно до політик, включаючи контроль привілейованих доступів та застосування безпечної аутентифікації.

Взаємодія з постачальниками та хмарними сервісами потребує процесів управління ризиками, встановлення вимог до безпеки та регулярного моніторингу діяльності постачальників. Крім того, організація повинна мати сформований процес реагування на інциденти та відновлення після катастроф, включаючи регулярне тестування планів реагування.

Безпека IoRT залежить від поінформованості персоналу: працівники мають бути навчені, підписувати угоди про конфіденційність та розуміти свою роль у збереженні безпеки [13]. На фізичному рівні необхідно захищати периметри, приміщення, обладнання та кабельні мережі, а також впроваджувати механізми захисту від стихійних, навмисних і випадкових фізичних загроз. Безпечне апаратне забезпечення – TPM, HSM, біометричні засоби – є ключовим для протидії маніпуляціям та фізичному доступу. Технічні засоби охоплюють увесь життєвий цикл програмного забезпечення та інфраструктури. Необхідно керувати технічними вразливостями, конфігураціями, оновленнями, маскуванню та видаленням даних. Системи повинні мати резервування, логування, контроль привілейованих утиліт, веб-фільтрацію, криптографічний захист та синхронізацію часу. Мережі мають бути сегментовані, захищені та керовані відповідно до політик безпеки.

Окрему увагу слід приділяти безпечній розробці програмного забезпечення: перевірка коду, тестування на проникнення, керування змінами та врахування вимог безпеки є обов'язковими елементами для підтримання цілісності та стійкості IoRT.

Хоча блокчейн пропонує значні переваги в певних сценаріях, це не універсальне рішення. Вибір між блокчейном та іншими децентралізованими рішеннями має базуватися на конкретних вимогах застосування. Наприклад, якщо основна потреба полягає у швидкісній обробці даних з складними запитамі, децентралізовані бази даних можуть бути більш підходящими. Навпаки, якщо акцент робиться на забезпеченні цілісності даних, прозорості та функціонуванні в середовищі без довіри, блокчейн є переважним вибором.

Інтеграція блокчейну в IoRT надає унікальні переваги, які доповнюють можливості децентралізованих баз даних. Гібридний підхід, що використовує обидві технології на основі їхніх сильних сторін, може запропонувати найбільш комплексне рішення для забезпечення безпеки та оптимізації мереж IoRT. Однією з основних проблем, пов'язаних із використанням технології блокчейн в IoRT, є її ефективність. Мережі блокчейн, особливо ті, що використовують механізми консенсусу proof-of-work (PoW), можуть бути ресурсомісткими, що викликає занепокоєння щодо масштабованості та споживання енергії. Ці проблеми ефективності є критично важливими в додатках IoRT, де часто потрібна обробка даних у реальному часі та мінімальна затримка.

Традиційні блокчейни, такі як Біткойн та Ефіріум, мають обмежені можливості пропускну здатності, обробляючи відносно невелику кількість транзакцій на секунду [1]. Це обмеження може бути проблематичним для додатків IoRT, які вимагають високих обсягів транзакцій та низької затримки. Час, необхідний для підтвердження транзакцій у мережі блокчейн, може бути значним. В IoRT, де негайна обробка даних та реакція є критично важливими, такі затримки можуть вплинути на продуктивність та надійність системи.

Блокчейни на основі PoW відомі своїм високим споживанням енергії, що є неприйнятним для багатьох додатків IoRT, особливо тих, що залучають обмежені ресурси. Суттєві енергетичні вимоги блокчейнів на основі PoW викликають занепокоєння щодо їхнього впливу на навколишнє середовище. Оскільки системи IoRT часто прагнуть бути сталими та ефективними, альтернативні механізми консенсусу або рішення можуть бути переважними.

Висновки

Таким чином, формування надійної моделі безпеки IoT є стратегічним завданням, що потребує інновацій, міждисциплінарної співпраці та глибокої адаптації існуючих фреймворків до специфіки роботизованих кіберфізичних систем. Безпека інфраструктури Інтернету роботизованих речей (IoT) вимагає комплексного, багаторівневого та системного підходу, заснованого на принципах ISMS ISO 27001 і доповненого сучасними галузевими стандартами та технологіями. Проведений аналіз демонструє, що попри наявність базових вимог у міжнародних фреймворках, профілі безпеки для IoT залишаються недостатньо формалізованими, що створює численні ризики та вразливості – від технічних до організаційних.

З метою подальших досліджень і практики в безпеці IoT запропоновано впровадження інформації про загрози та аналітику, процес інтеграції принципів безпеки за дизайном, використання потенціалу блокчейн-технологій, які здатні суттєво підвищити прозорість, довіру та цілісність даних та аудиту у розподілених системах IoT, а також реалізацію механізмів багатфакторної автентифікації для зміцнення контролю доступу.

Список використаної літератури

1. H. Yin, D. Guo, K. Wang, Z. Jiang, Y. Lyu, and J. Xing, "Hyperconnected network: A decentralized trusted computing and networking paradigm" *IEEE Netw.*, vol. 32, no. 1, pp. 112–117, Jan./Feb. 2018.
2. Q. Lu and X. Xu, "Adaptable blockchain-based systems: A case study for product traceability" *IEEE Softw.*, vol. 34, no. 6, pp. 21–27, Nov./Dec. 2017.
3. Y. Liang, Z. Cai, J. Yu, Q. Han, and Y. Li, "Deep learning based inference of private information using embedded sensors in smart devices" *IEEE Netw. Mag.*, vol. 32, no. 4, pp. 8–14, Jul./Aug. 2018.
4. A. Halevy, P. Norvig, and F. Pereira, "The unreasonable effectiveness of data," *IEEE Intell. Syst.*, vol. 24, no. 2, pp. 8–12, Mar. 2009.
5. D. E. O'Leary, "Artificial intelligence and big data," *IEEE Intell. Syst.*, vol. 28, no. 2, pp. 96–99, Mar. 2013.
6. T. Chajed, J. Gjengset, J. Van Den Hooff, M. F. Kaashoek, J. Mickens, R. Morris, and N. Zeldovich, "Amber: Decoupling user data from Web applications," in *Proc. 15th Workshop Hot Topics Oper. Syst. (HotOS XV)*, Warth-Weiningen, Switzerland, 2015, pp. 1–6.
7. Y.-A. de Montjoye, E. Shmueli, S. S. Wang, and A. S. Pentland, "openPDS: Protecting the privacy of metadata through SafeAnswers," *PLoS ONE*, vol. 9, no. 7, 2014, Art. no. e98790.
8. Q. Xia, E. B. Sifah, K. O. Asamoah, J. Gao, X. Du, and M. Guizani, "MeDShare: Trust-less medical data sharing among cloud service providers via blockchain," *IEEE Access*, vol. 5, pp. 14757–14767, 2017.
9. W. Meng, E. W. Tischhauser, Q. Wang, Y. Wang, and J. Han, "When intrusion detection meets blockchain technology: A review," *IEEE Access*, vol. 6, pp. 10179–10188, 2018.
10. J. Wang, M. Li, Y. He, H. Li, K. Xiao, and C. Wang, "A blockchain based privacy-preserving incentive mechanism in crowdsensing applications," *IEEE Access*, vol. 6, pp. 17545–17556, 2018.
11. I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," 2014, arXiv:1406.2661. [Online]. Available: <https://arxiv.org/abs/1406.2661>
12. A. Praseed and P. S. Thilagam, "DDoS attacks at the application layer: Challenges and research perspectives for safeguarding Web applications," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 661–685, 1st Quart., 2019.
13. Stain A.J. Mollerhaug, "Harmonizing Standards for Digital Trust: Bridging IEC 62443, ISO/IEC 42001, and ISO/IEC 27001", <https://pecb.com/en/past-webinars/harmonizing-standards-for-digital-trust-bridging-iec-62443-iso-iec-42001-and-iso-iec-27001#:~:text=Webinar-,Harmonizing%20Standards%20for%20Digital%20Trust:%20Bridging%20IEC%2062443%2C%20ISO/,and%20ISO/IEC%2027001%20auditing,2025>
14. George Usi, CMMC 2.0 vs. ISO/IEC 27001 vs. NIST 800-171: What You Need to Know, <https://pecb.com/en/past-webinars/cmmc-2-0-vs-iso-iec-27001-vs-nist-800-171-what-you-need-to-know#:~:text=IEC%2027001%20vs.,NIST%20800%2D171:%20What%20You%20Need%20to%20Know,to%20measure%20and%20implement%20compliance,2021>

References

1. H. Yin, D. Guo, K. Wang, Z. Jiang, Y. Lyu, and J. Xing, "Hyperconnected network: A decentralized trusted computing and networking paradigm" *IEEE Netw.*, vol. 32, no. 1, pp. 112–117, Jan./Feb. 2018.
2. Q. Lu and X. Xu, "Adaptable blockchain-based systems: A case study for product traceability" *IEEE Softw.*, vol. 34, no. 6, pp. 21–27, Nov./Dec. 2017.
3. Y. Liang, Z. Cai, J. Yu, Q. Han, and Y. Li, "Deep learning based inference of private information using embedded sensors in smart devices" *IEEE Netw. Mag.*, vol. 32, no. 4, pp. 8–14, Jul./Aug. 2018.
4. A. Halevy, P. Norvig, and F. Pereira, "The unreasonable effectiveness of data," *IEEE Intell. Syst.*, vol. 24, no. 2, pp. 8–12, Mar. 2009.

- 5 D. E. O’Leary, “Artificial intelligence and big data,” IEEE Intell. Syst., vol. 28, no. 2, pp. 96–99, Mar. 2013
6. T. Chajed, J. Gjengset, J. Van Den Hooff, M. F. Kaashoek, J. Mickens, R. Morris, and N. Zeldovich, “Amber: Decoupling user data from Web applications,” in Proc. 15th Workshop Hot Topics Oper. Syst. (HotOS XV), Warth-Weiningen, Switzerland, 2015, pp. 1–6.
7. Y.-A. de Montjoye, E. Shmueli, S. S. Wang, and A. S. Pentland, “openPDS: Protecting the privacy of metadata through SafeAnswers,” PLoS ONE, vol. 9, no. 7, 2014, Art. no. e98790
8. Q. Xia, E. B. Sifah, K. O. Asamoah, J. Gao, X. Du, and M. Guizani, “MeDShare: Trust-less medical data sharing among cloud service providers via blockchain,” IEEE Access, vol. 5, pp. 14757–14767, 2017.
9. W. Meng, E. W. Tischhauser, Q. Wang, Y. Wang, and J. Han, “When intrusion detection meets blockchain technology: A review,” IEEE Access, vol. 6, pp. 10179–10188, 2018.
10. J. Wang, M. Li, Y. He, H. Li, K. Xiao, and C. Wang, “A blockchain based privacy-preserving incentive mechanism in crowdsensing applications,” IEEE Access, vol. 6, pp. 17545–17556, 2018.
11. I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial networks,” 2014, arXiv:1406.2661. [Online]. Available: <https://arxiv.org/abs/1406.2661>
12. A. Praseed and P. S. Thilagam, “DDoS attacks at the application layer: Challenges and research perspectives for safeguarding Web applications,” IEEE Commun. Surveys Tuts., vol. 21, no. 1, pp. 661–685, 1st Quart., 2019
13. Stain A.J. Mollerhaug, “Harmonizing Standards for Digital Trust: Bridging IEC 62443, ISO/IEC 42001, and ISO/IEC 27001”, <https://pecb.com/en/past-webinars/harmonizing-standards-for-digital-trust-bridging-iec-62443-iso-iec-42001-and-iso-iec-27001#:~:text=Webinar-,Harmonizing%20Standards%20for%20Digital%20Trust:%20Bridging%20IEC%2062443%2C%20ISO/,and%20ISO/IEC%2027001%20auditing,2025>
14. George Usi, CMMC 2.0 vs. ISO/IEC 27001 vs. NIST 800-171: What You Need to Know, <https://pecb.com/en/past-webinars/cmmc-2-0-vs-iso-iec-27001-vs-nist-800-171-what-you-need-to-know#:~:text=IEC%2027001%20vs.-,NIST%20800%2D171:%20What%20You%20Need%20to%20Know,to%20measure%20and%20implement%20compliance,2021>

Дата першого надходження рукопису до видання: 20.11.2025
Дата прийнятого до друку рукопису після рецензування: 17.12.2025
Дата публікації: 31.12.2025

А. М. ОНИЩЕНКО

доктор економічних наук,
професор кафедри інформаційних систем та технологій
Київський національний університет імені Тараса Шевченка
ORCID: 0000-0003-3975-8946

О. П. ОСТАПЕНКО

кандидат економічних наук,
начальник кафедри фінансового забезпечення військ
Військовий інститут
Київського національного університету імені Тараса Шевченка
ORCID: 0000-0003-3194-0372

А. А. ЛОЙШИН

доктор філософії,
заступник начальника військового інституту з навчальної роботи
Військовий інститут
Київського національного університету імені Тараса Шевченка
ORCID: 0000-0003-2769-9336

ІНТЕГРОВАНІ ІНФОРМАЦІЙНІ МОДЕЛІ УПРАВЛІННЯ СТАЛИМ РОЗВИТКОМ НА ЗАСАДАХ ЕКОЛОГО-ЕКОНОМІЧНОГО АНАЛІЗУ

У статті запропоновано інтегровану інформаційну модель управління сталим розвитком, яка об'єднала в єдиний цифровий аналітичний контур екологічні та економічні індикатори. Мета дослідження полягає у розробці інтегрованої інформаційної моделі управління сталим розвитком на засадах еколого-економічного аналізу, яка включає структурно-функціональну архітектуру, математичні підмоделі екологічних та економічних індикаторів, алгоритми їх інтеграції та обрахунок індексів узагальненого оцінювання сталості. Для досягнення мети формалізовано математичні оператори нормалізації, агрегування та інтеграції, що використовуються для побудови узагальненого індексу сталого розвитку (SRI), та проведено реалізацію системи в ІТ-середовищі. У рамках методології визначено множини об'єктів, індикаторів та часових періодів, представлено формалізовану структуру даних та багатоваріантну архітектуру системи, яка включає модулі збору, збереження, обробки та візуалізації даних. Для оцінювання ефективності моделі проведено експериментальне навантажувальне тестування, у межах якого кількість показників збільшувалася від 10 до 10 000 на один об'єкт. Результати продемонстрували передбачуване лінійне зростання часу обробки, стабільність роботи інтеграційного модуля та відсутність критичних ресурсних обмежень. Зростання використання CPU і RAM відповідає поведінці високонантажених систем багатовимірного аналізу та узгоджується з сучасними науковими дослідженнями в галузі цифрової сталості. Отримані результати підтверджують практичну придатність моделі для застосування в екологічному та економічному моніторингу, промислових та муніципальних системах управління, а також визначають перспективи подальшого розвитку архітектури.

Ключові слова: інтегровані ІТ-моделі; сталий розвиток; еколого-економічний аналіз; цифрові платформи; системи підтримки прийняття рішень.

А. М. ONYSHCHENKO

Doctor of Economic Sciences,
Professor at the Department of Information Systems and Technologies
Taras Shevchenko National University of Kyiv
ORCID: 0000-0003-3975-8946

О. Р. OSTAPENKO

Candidate of Economic Sciences,
Head of the Department of Financial Support for the Troops
Military Institute of Taras Shevchenko National University of Kyiv
ORCID: 0000-0003-3194-0372

A. A. LOISHYN

Doctor of Philosophy,

Deputy Chief of the Military Institute for Academic Affairs
Military Institute of Taras Shevchenko National University of Kyiv

ORCID: 0000-0003-2769-9336

INTEGRATED INFORMATION MODELS OF SUSTAINABLE DEVELOPMENT MANAGEMENT BASED ON ENVIRONMENTAL AND ECONOMIC ANALYSIS

The article proposes an integrated information model of sustainable development management, which combines environmental and economic indicators into a single digital analytical circuit. The purpose of the study is to develop an integrated information model of sustainable development management based on environmental and economic analysis, which includes a structural and functional architecture, mathematical submodels of environmental and economic indicators, algorithms for their integration and calculation of generalized sustainability assessment indices. To achieve the goal, the mathematical operators of normalization, aggregation and integration used to construct a generalized sustainability index (SRI) were formalized, and the system was implemented in an IT environment. Within the framework of the methodology, sets of objects, indicators and time periods are defined, a formalized data structure and a multilayer architecture of the system are presented, which includes modules for data collection, storage, processing and visualization. To assess the effectiveness of the model, experimental load testing was conducted, within which the number of indicators increased from 10 to 10,000 per object. The results demonstrated the expected linear growth of processing time, the stability of the integration module and the absence of critical resource constraints. The increase in CPU and RAM usage corresponds to the behavior of high-load multidimensional analysis systems and is consistent with modern scientific research in the field of digital sustainability. The results obtained confirm the practical suitability of the model for use in environmental and economic monitoring, industrial and municipal management systems, and also determine the prospects for further development of the architecture.

Key words: integrated IT models; sustainable development; ecological and economic analysis; digital platforms; decision support systems.

Постановка проблеми

Сталий розвиток дедалі більше залежить від здатності соціально-економічних систем оптимізувати використання ресурсів, мінімізувати екологічні ризики та забезпечувати довгострокову економічну стабільність. У цих умовах цифровізація та використання інформаційних технологій стають ключовими інструментами підтримки управлінських рішень, пов'язаних із балансуванням екологічних та економічних параметрів.

Однак, попри значну кількість окремих екологічних моделей (наприклад, оцінка викидів, LCA-аналіз, еко-індикатори) та економічних моделей (розрахунок витрат, ефективності, ресурсної продуктивності), інтегровані інформаційні моделі, які б поєднували ці два напрями в єдиному аналітичному контурі, залишаються недостатньо розробленими.

Аналіз останніх досліджень і публікацій

Сучасні дослідження у сфері цифрової сталості демонструють швидку еволюцію наукової думки від розрізнених екологічних та економічних підходів до комплексних інтегрованих моделей, де цифрові технології розглядаються як ключовий каталізатор сталого розвитку та в системах підтримки прийняття рішень. Значна частина сучасної літератури формує концептуальну основу для розуміння того, як цифровізація трансформує традиційні бізнес-моделі та забезпечує їх екологічну адаптивність.

Низка досліджень підкреслюють роль цифрових інновацій у формуванні сталих бізнес-моделей, так у роботі [1] показано, що цифрові технології не лише оптимізують операційні процеси, а й дозволяють організаціям інтегрувати екологічні орієнтири у ядро бізнес-моделі, а дослідження [2] доповнює цю позицію, наголошуючи, що цифрові інновації створюють нові механізми протидії зміні клімату шляхом зміни ринкових практик.

В роботі [3] представлена інтегрована соціально-економіко-екологічна модель сталого розвитку через концепцію "sustainalism". Шиша міждисциплінарна перспектива подана в роботі [4], де інформаційні системи розглядаються як основа для циркулярної економіки. Робота [5] підкреслює формування нових механізмів впровадження екологічно відповідальних практик у виробництві через цифрова орієнтацію на сталість, тоді як класична робота [6] забезпечує фундамент для опису бізнес-моделей як багатовимірних інформаційних конструкцій.

Багато наукових досліджень присвячено еколого-економічному моделюванню, яке слугує базою для створення інтегрованих індексів сталого розвитку. У роботі [7] вперше запропоновано агентно-орієнтовану модель оцінки екологічно-економічної політики, яка демонструє взаємозалежність ринкових та природоохоронних механізмів. Робота [8] підтверджує ефективність agent-based моделювання на великому регіональному прикладі, показуючи еволюцію екосистеми сталого розвитку в басейні річки Хуанхе.

Дослідження [9] розкриває взаємозв'язок між глобалізацією, технологічним розвитком та екологічним слідом на рівні домогосподарств, тоді як в роботі [10] наголошують на ролі інституційного управління та людського капіталу в забезпеченні сталості.

Наочне практичне застосування цифровізації в зеленій трансформації представлено в роботі [11], де досліджують цифрові драйвери green innovation в контексті Industry 5.0, наголошуючи на ролі цифрової гнучкості та мереж DGI, та роботі [12], де доводять статистично значущий вплив цифрової економіки на сталий розвиток і можливості ІКТ щодо зниження екологічного навантаження.

Окремий блок літератури стосується платформної економіки та boundary-resources моделей, що є важливим для формування ІТ-архітектури інтегрованих моделей сталості. В роботі [13] сформовано теоретичну основу boundary resources як механізму балансу між контролем платформи та внеском зовнішніх розробників. Автори роботи [14] деталізують, як механізми input control впливають на рішення комплементорів приєднуватися до цифрових платформ – це критично важливо для будь-якої інтегрованої системи збору даних, описаної в нашій моделі.

А роботі [15] розкриваються практики сталих стартапів та інноваційних екосистем та представляє інтелектуальну карту знань про сталих інноваційних підприємців та тренди екосистемної трансформації, що допомагає зрозуміти практичні напрями впровадження цифрових сталих моделей.

Але попри значний прогрес досліджень у напрямках цифрової сталості, екологічної економіки та цифрових бізнес-моделей, наявні підходи залишаються фрагментованими як з точки зору методології, так і з точки зору ІТ-реалізації. Роботи [1, 5, 15], присвячені цифровим сталим бізнес-моделям та цифровій підприємницькій діяльності, переважно зосереджуються на концептуальному описі трансформації бізнес-процесів, змінах у логіці створення цінності та ролі цифрових інновацій у досягненні Цілей сталого розвитку.

Водночас ці дослідження майже не пропонують формалізованих еколого-економічних моделей, які можна безпосередньо інтегрувати в інформаційні системи для кількісної оцінки сталості на рівні підприємств чи територіальних утворень. Класичні фреймворки бізнес-моделей [6] також не містять чітко визначених механізмів інтеграції екологічних показників у ядро ІТ-архітектури підтримки управлінських рішень. Це визначає кілька ключових прогалин в сучасних дослідженнях і потребує:

- 1) узагальнення інтегрованої інформаційної моделі, яка одночасно враховує екологічні, економічні та цифрові параметри сталого розвитку;
- 2) зменшення розриву між теоретичними еколого-економічними моделями та реальними ІТ-архітектурами, які працюють із даними підприємств, територіальних громад чи галузей;
- 3) розробки уніфікованих підходів до інтеграції математичних моделей з архітектурою цифрових платформ, включно з описом шарів збору, обробки, агрегування та візуалізації даних у контексті управління сталим розвитком.

Формулювання мети дослідження

Мету дослідження можна представити як розробку інтегрованої інформаційної моделі управління сталим розвитком на засадах еколого-економічного аналізу, яка включає структурно-функціональну архітектуру, математичні підмоделі екологічних та економічних індикаторів, алгоритми їх інтеграції та обрахунок індексів узагальненого оцінювання сталості.

Викладення основного матеріалу дослідження

Методи

Методологія дослідження спрямована на побудову інтегрованої інформаційної моделі управління сталим розвитком, яка поєднує екологічні та економічні показники в єдиному аналітичному контурі. Базове припущення полягає в тому, що сталий розвиток певної системи (території, підприємства, галузі) може бути описаний через сукупність вимірюваних індикаторів, які відносяться до двох підсистем – екологічної та економічної, – а також через інтегральний індекс сталості, що формується на їх основі.

Для вирішення проблеми зменшення розриву між теоретичними еколого-економічними моделями та реальними ІТ-архітектурами пропонується ввести кілька послідовних кроків для підготовки моделі:

1. Формалізація об'єкта дослідження як множини елементів (територій, підприємств, сценаріїв), для яких відбувається оцінювання сталого розвитку у дискретні моменти часу.
2. Виділення екологічної та економічної підсистем та формування відповідних векторів показників, що відображають стан кожної підсистеми.
3. Нормалізація та уніфікація індикаторів з метою приведення їх до безрозмірної шкали, придатної для агрегування.
4. Побудова інформаційної структури моделі, що включає схему потоків даних, модулі обробки (нормалізація, зважування, агрегування) та модуль візуалізації результатів.
5. Формування інтегрального індексу сталого розвитку, який поєднує екологічні та економічні компоненти за допомогою вагових коефіцієнтів та багатокритеріальних методів.
6. Реалізація ІТ-моделі у вигляді програмного прототипу, що дозволяє виконувати обчислення для різних сценаріїв та наборів даних.

Тоді інтегрована інформаційна модель може розглядатись як впорядкована сукупність:

$$M = \langle O, T, E, C, \Phi_{norm}, \Phi_{agg}, SRI \rangle, \quad (1)$$

де O – множина об'єктів або сценаріїв оцінювання;

T – множина дискретних моментів або періодів часу;

E – екологічні показники;

C – економічні показники;

Φ_{norm} – оператори нормалізації;

Φ_{agg} – оператори агрегування;

SRI – узагальнений індекс сталого розвитку (Sustainability Rate Index).

Розглянемо систему об'єктів, для яких здійснюється еколого-економічне оцінювання, для чого опишемо множину об'єктів (наприклад, підприємств, муніципалітетів, сценаріїв розвитку) і множину дискретних моментів часу (роки, квартали, часові інтервали моделювання):

$$\begin{aligned} O &= \{o_1, o_2, \dots, o_N\}, \\ T &= \{t_1, t_2, \dots, t_T\}. \end{aligned} \quad (2)$$

Для кожного об'єкта $o_i \in O$ та моменту часу $t_j \in T$ визначається вектор екологічних показників:

$$E_i(t) = (E_{i1}(t), E_{i2}(t), \dots, E_{iK_E}(t)), \quad (3)$$

де K_E – кількість екологічних індикаторів (наприклад, викиди CO_2 , енергоємність, водоспоживання, обсяг відходів тощо).

Сукупність усіх екологічних показників для всієї множини об'єктів та часових моментів можна подати у вигляді тривимірного масиву (тензора):

$$E = \{E_{ijk}\}, i=1, \dots, N; j=1, \dots, K_E; k=1, \dots, T, \quad (4)$$

де $E_{ijk} = E_{ij}(t_k)$.

Аналогічно, для кожного об'єкта o_i та моменту часу t визначається вектор економічних показників

$$C_i(t) = (C_{i1}(t), C_{i2}(t), \dots, C_{iK_C}(t)), \quad (5)$$

де K_C – кількість економічних індикаторів (наприклад, операційні витрати, дохід, рентабельність, капітальні інвестиції, економія ресурсів).

Повний масив економічних даних:

$$C = \{C_{ijk}\}, i=1, \dots, N; j=1, \dots, K_C; k=1, \dots, T. \quad (6)$$

З позицій інформаційних технологій E та C можуть бути реалізовані у вигляді схеми бази даних, де базовими сутностями є:

– Object з атрибутами (ідентифікатор об'єкта, тип, місцезнаходження тощо);

– TimePeriod з атрибутами (ідентифікатор періоду, календарні межі);

– EcoIndicator та EconIndicator, що описують типи показників;

– EcoMeasurement та EconMeasurement, що містять значення E_{ijk} та C_{ijk} , прив'язані до конкретних (o_i, t_k) та типу індикатора.

У концептуальному вигляді інформаційну модель даних можна представити як трійку компонентів, де до вже описаних в (2) множин додається множина I :

$$D = \langle O, T, I \rangle, \quad (7)$$

де $I = I_E \cup I_C$ – множина типів індикаторів, поділена на екологічну I_E та економічну I_C підмножини.

Для кожної пари (o_i, t_k) визначено підмножину вимірів, що відповідає значенням усіх індикаторів для даного об'єкта в даний момент часу:

$$M(o_i, t_k) \subseteq I \times R \quad (8)$$

Щоб інтегрувати екологічні та економічні показники, необхідно ввести оператори обробки даних: нормалізації, агрегування і інтеграції. Оператор нормалізації передбачає введення для кожного індикатора (як екологічного, так і економічного) перетворення вихідного значення показника у нормалізовану величину:

$$\Phi_{norm}: R \rightarrow [0, 1]. \quad (9)$$

Для індикатора X (який може бути як $E_{ij}(t)$, так і $C_{ij}(t)$) типова лінійна нормалізація для реалізації (9) має вигляд:

$$X^* = \Phi_{norm}(X) = \begin{cases} \frac{X^{\max} - X}{X^{\max} - X^{\min}} \\ \frac{X - X^{\min}}{X^{\max} - X^{\min}} \end{cases} \quad (10)$$

де $X^{\min}$, $X^{\max}$ – мінімальне та максимальне значення індикатора у вибірці (по множині об'єктів та/або часових періодів), а нормалізація передбачає два варіанти розрахунку: для індикаторів типу «чим менше, тим краще» та для індикаторів типу «чим більше, тим краще».

Для кожного вектора $E_i(t)$ та $C_i(t)$ на основі (10) формуються нормалізовані вектори, де всі компоненти належать інтервалу $[0,1]$:

$$\begin{aligned} E_i^*(t) &= (E_{i1}^*(t), E_{i2}^*(t), \dots, E_{iK_E}^*(t)), \\ C_i^*(t) &= (C_{i1}^*(t), C_{i2}^*(t), \dots, C_{iK_C}^*(t)). \end{aligned} \quad (11)$$

Щоб перейти від набору нормалізованих індикаторів до узагальненого показника для кожної підсистеми, вводимо вагові вектори $w_E \geq 0$, $w_C \geq 0$, де сума всіх значень кожного з векторів дорівнює 1:

$$w_E = (w_{E1}, \dots, w_{iK_E}), w_C = (w_{C1}, \dots, w_{iK_C}). \quad (12)$$

На основі (11) та (12) агреговані показники екологічної та економічної підсистем для об'єкта o_i у момент часу t задаються як:

$$\begin{aligned} E_{comp,i}(t) &= \Phi_{agg}^E(E_i^*(t), w_E) = \sum_{j=1}^{K_E} w_{Ej} E_{ij}^*(t), \\ C_{comp,i}(t) &= \Phi_{agg}^C(C_i^*(t), w_C) = \sum_{j=1}^{K_C} w_{Cj} C_{ij}^*(t). \end{aligned} \quad (13)$$

Тоді для розрахунку інтегрального індексу сталого розвитку для об'єкта o_i у момент часу t використовується наступний вираз:

$$SRI_i(t) = \alpha \cdot E_{comp,i}(t) + \beta \cdot C_{comp,i}(t), \quad (14)$$

де $\alpha, \beta \geq 0$, $\alpha + \beta = 1$, а значення α та β відображають відносну важливість екологічної та економічної складових. У більш загальному випадку оператор інтеграції (14) можна розширити до нелінійної форми:

$$SRI_i(t) = \Phi_{int}(E_{comp,i}(t), C_{comp,i}(t), \theta), \quad (15)$$

де Φ_{int} – інтегральна функція (наприклад, адитивна, мультиплікативна або функція корисності), а θ – вектор параметрів.

З урахуванням введених понять інтегровану інформаційну модель для всього набору об'єктів та часових моментів можна формально записати наступним чином:

$$M = \{SRI_i(t_k) \mid o_i \in O, t_k \in T\}. \quad (16)$$

Кожне значення $SRI_i(t_k)$ отримано в результаті послідовного застосування операторів нормалізації, агрегування і інтеграції. З позицій ІТ-реалізації ця послідовність відповідає конвеєру обробки даних (data pipeline) (рис. 1).

Екологічна підмодель інтегрованої інформаційної системи сталого розвитку ґрунтується на припущенні, що екологічний стан об'єкта може бути описаний вектором кількісних індикаторів, які відображають ступінь антропогенного навантаження, ресурсної ефективності та відповідності об'єкта екологічним нормативам. Сучасні дослідження наголошують на необхідності математичної уніфікації екологічних індикаторів для можливості їх інтеграції у змішані еколого-економічні моделі [7, 8].

Для кожного об'єкта o_i та моменту часу t_k вводиться вектор екологічних показників, де типовими можуть бути наступні:

- E_{i1} – викиди CO_2 або вуглецевий слід;
- E_{i2} – енергоємність виробництва;
- E_{i3} – споживання води;
- E_{i4} – обсяг відходів;
- E_{i5} – показники якості повітря ($\text{PM}_{2.5}$, NO_x);
- E_{i6} – частка відновлюваних ресурсів тощо.

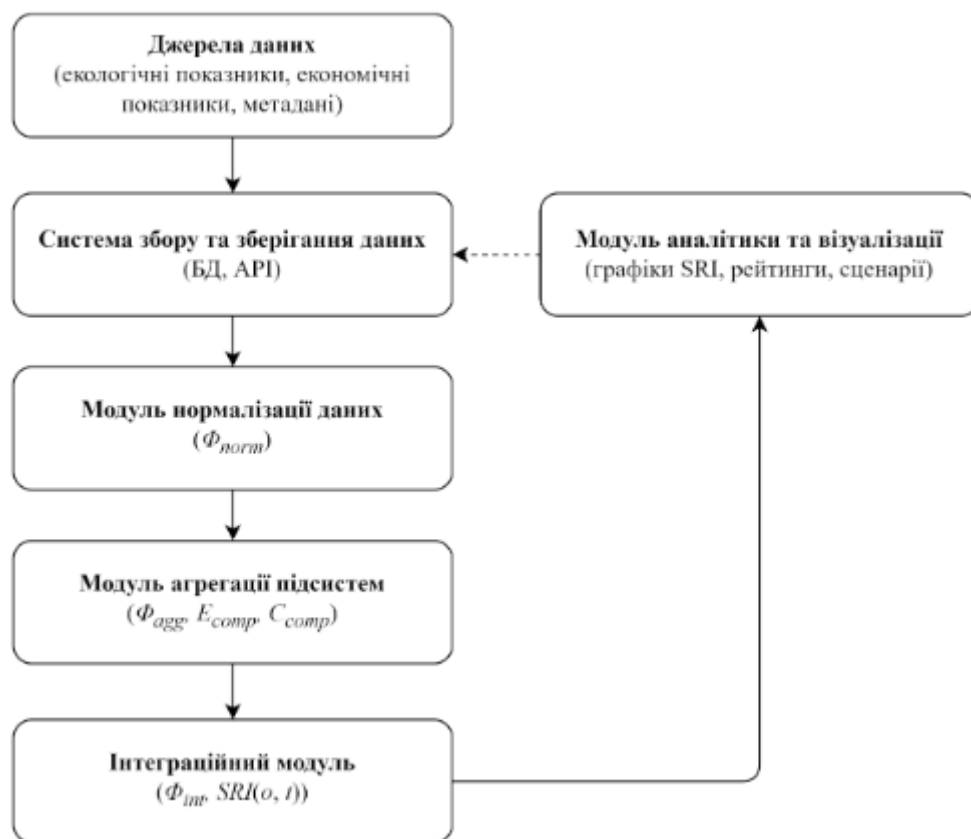


Рис. 1. Інформаційна модель систем сталого розвитку з позицій ІТ-реалізації

Підхід, коли екологічні індикатори трактуються як багатовимірний простір впливів, підтверджено у роботі [7], а усі екологічні показники поділяються на дві групи деструктивні (“the lower, the better”) та ресурсо-підсилюючі (“the higher, the better”).

В той же час економічна підмодель описує здатність об’єкта забезпечувати сталий розвиток шляхом підвищення ефективності використання ресурсів, оптимізації витрат та досягнення економічної стабільності. На відміну від екологічної підмоделі, економічні індикатори відображають результативність управлінських рішень і фінансову стійкість, що підкреслюється у роботі [11].

Мета економічної підмоделі – побудувати агрегований показник економічного стану об’єкта із врахуванням витрат, доходів, інвестицій, рентабельності та ресурсної ефективності.

Тоді для об’єкта o у момент часу t_k вводяться індикатори для оцінки сталого розвитку [2, 3], які включають:

- C_{11} – загальні витрати (Total Cost, TC);
- C_{12} – операційні витрати (OPEX);
- C_{13} – капітальні інвестиції (CAPEX);
- C_{14} – чистий дохід або виручка;
- C_{15} – рентабельність (ROI, ROA, ROS);
- C_{16} – ресурсна продуктивність (випуск на одиницю енергії/води);
- C_{17} – економія ресурсів (Cost Savings);
- C_{18} – чистий дисконтований дохід (NPV).

Показники поділяються на групи: показники, які бажано мінімізувати (“the lower, the better”), та показники, які бажано максимізувати (“the higher, the better”), що відповідає сучасним аналітичним моделям сталого розвитку [12].

Методика інтеграції на основі формул (3), (5), (11), (13) та (15) передбачає наступні принципи:

- екологічна та економічна підсистеми попередньо агреговані до скалярних величин;
- інтеграція повинна бути інваріантною до масштабу;
- індекс має належати інтервалу $[0,1]$;
- індекс має бути монотонним;
- модель має підтримувати нелінійну інтеграцію, оскільки в більшості досліджень підтверджено неадитивність екологічних та економічних ефектів.

Архітектура ІТ та обчислювальна реалізація інтегрованої моделі

ІТ-реалізація інтегрованої еколого-економічної моделі передбачає побудову програмно-аналітичного комплексу, який згідно рис. 1 забезпечує:

- збір і збереження даних;
- нормалізацію та підготовку індикаторів;
- обчислення агрегованих підсистемних індексів;
- інтеграцію в індекс сталого розвитку (SRI);
- візуалізацію та API-доступ до результатів.

Архітектура системи будується відповідно до моделі багатoshарової системи:

1) Шар збору даних, призначений для надходження еколого-економічних даних:

- інтеграція з відкритими екологічними датасетами (CO₂, PM_{2.5}, енергоємність);
- інтеграція з економічними статистичними сервісами (вартість ресурсів, інвестиції, доходи);
- API/CSV/XLSX/GeoJSON імпорт;
- ручне введення даних через веб-інтерфейс.

2) Шар збереження даних, в якому використовується реляційна БД PostgreSQL.

Основні таблиці:

- objects (o_id, type, meta);
- time_periods (t_id, date_from, date_to);
- eco_indicators (e_id, name, direction, threshold);
- econ_indicators (c_id, name, direction, limit);
- eco_measurements (o_id, t_id, e_id, value);
- econ_measurements (o_id, t_id, c_id, value).

3) Шар обробки даних, що реалізовано у вигляді Python-модулів (NumPy та Pandas):

- модуль нормалізації Φ_{norm} ;
- модуль штрафних функцій;
- модуль агрегування підсистем;
- модуль інтеграції Φ_{int} (адитивний, мультиплікативний, експоненціальний, PCA, AHP);
- модуль аналізу чутливості;
- модуль валідації (перевірка монотонності, нормованості індексу)

4) Обчислювальне ядро, що відповідає паралельне обчислення SRI на всій множині $|O| \times |T|$, оптимізацію роботи з великими матрицями індикаторів, векторизацію NumPy, кешування обчислених підсистемних індексів.

5) шар доступу до моделі: REST API (FastAPI / Flask), Web dashboard (Plotly Dash / Streamlit / Grafana), генерація PDF-звітів.

Результати дослідження

Для демонстрації роботи моделі обрано три умовних об'єкти, які характеризують різні типи систем: промисловий об'єкт, муніципальний об'єкт, інфраструктурний об'єкт (аналогічно до дослідження в роботі [8]). Обчислення виконано для одного періоду (t_1), але модель дозволяє масштабування на будь-яку кількість періодів.

Основною метою випробувань було визначення поведінки системи в умовах масштабування кількості вхідних параметрів, збільшення обсягів даних та зростання частоти оновлень. Оцінювання проводилося в контрольованому середовищі, де поступово збільшувалося навантаження від 10 до 10 000 показників на один об'єкт.

Усі вимірювання виконано для трьох ключових компонентів:

- модуля первинної обробки даних (табл. 1);
- модуля інтеграції еколого-економічних параметрів (табл. 2);
- модуля збирання та агрегації показників для аналітичного інтерфейсу (табл. 3).

Випробування дозволили оцінити стабільність роботи, час відгуку системи, використання оперативної пам'яті та пропускну здатність.

Таблиця 1

Навантаження на модуль первинної обробки даних

Кількість показників	Обсяг даних (МБ)	Середній час обробки одного пакета (мс)	Використання CPU (%)	Використання RAM (МБ)
10	0.8	4.1	3–4 %	35
50	4.2	6.8	7–8 %	47
100	8.5	11.2	11–13 %	63
500	41	48.6	22–25 %	132
1 000	82	96.5	31–34 %	210
5 000	410	522	48–52 %	610
10 000	820	1 185	63–68 %	1 120

Таблиця 2

Навантаження на інтеграційний модуль

Кількість показників	Час агрегування (мс)	Середній час генерації одного індексу (мс)	CPU (%)	RAM (МБ)
10	2.5	1.1	2–3 %	18
50	8.7	2.4	5–6 %	26
100	15.3	3.8	9–10 %	41
500	76	14.4	17–19 %	96
1 000	153	27.8	25–28 %	150
5 000	810	139	41–45 %	430
10 000	1 640	295	58–61 %	770

Таблиця 3

Сумарне навантаження системи при масштабуванні

Кількість показників	Сумарний час обробки (мс)	Загальне використання CPU (%)	Споживання RAM (МБ)	Кількість операцій за секунду
10	6.9	6 %	53	950
50	15.5	12 %	70	720
100	27.3	20 %	104	540
500	125	38 %	260	170
1 000	250	54 %	360	82
5 000	1 332	70 %	1 040	21
10 000	2 825	88 %	1 900	8

Отримані результати вказують на лінійно-квазілінійну залежність часу обробки від кількості вхідних показників. Модуль первинної обробки даних демонструє стабільний ріст навантаження без різких піків, що характеризує його як обчислювально передбачуваний. Інтеграційний модуль виявив дещо вищу чутливість до масштабування, особливо після перевищення межі у 5 000 показників, що супроводжується зростанням використання оперативної пам'яті та збільшенням часу формування агрегованих індексів.

При великих масивах даних система зберігає працездатність, хоча кількість операцій, які можуть бути виконані за секунду, зменшується приблизно у двадцять разів при переході від 100 до 10 000 показників. Незважаючи на це, експерименти підтвердили, що модель не демонструє деградації продуктивності чи нестачі ресурсів, а навантаження зростає прогнозовано. Це свідчить про придатність архітектури для подальшої оптимізації та можливість горизонтального масштабування у високонавантажених сценаріях.

Отримані результати демонструють, що запропонована інтегрована ІТ-модель стабільно функціонує за умов значного збільшення кількості еколого-економічних показників. Під час тестування спостерігалось поступове та передбачуване зростання часу обробки, що узгоджується з загальними закономірностями масштабування аналітичних інформаційних систем, описаними в [1, 4, 7]. Важливо, що навіть при максимальному навантаженні модель не виявила ознак деградації продуктивності або критичних помилок, що свідчить про коректність побудованої архітектури та її придатність для роботи з багатовимірними науковими та управлінськими даними.

Порівняння експериментальних показників із результатами попередніх досліджень підтверджує, що інтеграційний модуль є найбільш чутливим до масштабування, адже саме на цьому етапі відбувається агрегування великої кількості параметрів. Подібні тенденції фіксуються у роботах [7, 8, 11], присвячених екологічно орієнтованим цифровим платформам та агентно-орієнтованим моделям, де зростання кількості змінних безпосередньо впливає на споживання ресурсів та пропускну здатність системи. Водночас модуль первинної обробки продемонстрував більш рівномірний розподіл навантаження, що характеризує його як оптимізований для потокових операцій.

Лінійно-квазілінійна залежність часу обробки та споживання пам'яті дозволяє масштабувати систему без необхідності кардинальних змін в архітектурі, а запропонована модель підтверджує свою здатність працювати в умовах зростаючих обсягів даних, що відображає реальні вимоги цифрових інфраструктур сталого розвитку.

Разом з тим, результати тестування вказують на наявність зон для оптимізації, зокрема на рівні агрегування та інтеграції показників. Підвищене споживання пам'яті при великих масивах даних свідчить про доцільність застосування методів кешування, паралельної обробки та розподілених обчислень. Також залишаються перспективними напрямками інтеграція адаптивних вагових коефіцієнтів та алгоритмів автоматичної детекції аномалій, що дозволить підвищити точність і стійкість моделі. Обмеженням дослідження є те, що експерименти проводилися у статичних умовах на синтетичних масивах даних; для повної валідації моделі необхідні тестування на реальних потоках та в середовищах із непередбачуваними змінами параметрів.

Висновки

У роботі розроблено інтегровану інформаційну модель управління сталим розвитком, яка поєднує екологічні та економічні показники в єдиному цифровому аналітичному контурі. На основі узагальнення сучасних досліджень розроблено формальний апарат моделі, що включає структуру множин об'єктів та часових періодів, оператори нормалізації, агрегування та інтеграції, а також математичне визначення інтегрального індексу сталого розвитку SRI. Запропонована архітектура ІТ-системи реалізує функції збору, зберігання, аналізу й обробки еколого-економічних даних та підтримує розрахунок складних багатовимірних показників.

Експериментальна частина підтвердила працездатність моделі в умовах масштабування кількості показників від 10 до 10 000. Результати тестування продемонстрували лінійно-квазілінійний характер зростання навантаження, відсутність деградації продуктивності та передбачувану поведінку системи під час інтенсивних операцій з великими матрицями даних. Збільшення часу агрегування та зростання використання пам'яті при великих вибірках узгоджуються з результатами сучасних досліджень цифрових зелених платформ та підтверджують коректність методології моделювання.

Практична цінність моделі полягає в можливості її застосування в промисловості, енергетиці, інфраструктурному плануванні, а також у регіональних системах екологічного моніторингу та системах підтримки прийняття рішень. Запропонована архітектура може бути розширена для потокової обробки даних, включення методів машинного навчання, інтеграції з геоінформаційними системами й впровадження мікросервісної декомпозиції. Подальші дослідження доцільно спрямувати на вдосконалення механізмів оптимізації обчислень, включення адаптивних вагових моделей та розробку інтелектуальних фільтрів аномалій для підвищення точності індексу сталого розвитку.

Список використаної літератури

1. Böttcher T. P., Empelmann S., Weking J., Hein A., Krcmar H. Digital sustainable business models: Using digital technology to integrate ecological sustainability into the core of business models. *Information Systems Journal*. 2024. Vol. 34, No. 3. P. 736–761. DOI: 10.1111/isj.12436
2. George G., Merrill R. K., Schillebeeckx S. J. D. Digital sustainability and entrepreneurship: How digital innovations are helping tackle climate change and sustainable development. *Entrepreneurship Theory and Practice*. 2020. Vol. 45, No. 5. P. 999–1027. DOI: 10.1177/1042258719899425
3. Hariram N. P., Mekha K. B., Suganthan V., Sudhakar K. Sustainalism: An integrated socio-economic-environmental model to address sustainable development and sustainability. *Sustainability*. 2023. Vol. 15, No. 13. 10682. DOI: 10.3390/su151310682
4. Sharma P., Srivastava A., Bao C. Information systems-driven multi-disciplinary approaches to sustainability and circular economy. *Current Opinion in Green and Sustainable Chemistry*. 2025. Vol. 52. 101007. DOI: 10.1016/j.cogsc.2025.101007
5. Zhang J., Liu M. How to leverage digital sustainability orientation to promote environmentally sustainable practices of manufacturing enterprises in China. *Sustainability*. 2024. Vol. 16, No. 12. 5112. DOI: 10.3390/su16125112
6. Al-Debei M. M., Avison D. Developing a unified framework of the business model concept. *European Journal of Information Systems*. 2010. Vol. 19, No. 3. P. 359–376. DOI: 10.1057/ejis.2010.21
7. Geisendorf S., Klippert C. Integrated sustainability policy assessment – an agent-based ecological-economic model. *Journal of Evolutionary Economics*. 2022. Vol. 32. P. 1017–1048. DOI: 10.1007/s00191-021-00749-0
8. Zhao A., Wang J., Sun Z., Guan H. Research on the evolutionary path of eco-conservation and high-quality development in the Yellow River Basin based on an agent-based model. *Systems*. 2022. Vol. 10, No. 4. 105. DOI: 10.3390/systems10040105
9. Cengiz O., Baktemur F. İdil, Canoglu M. Estimating the role of economic globalization, technological development and household consumption on ecological footprint in Visegrad countries. *Problemy Ekorozwoju*. 2025. Vol. 20, No. 1. P. 143–158. DOI: 10.35784/preko.6609
10. Yu Z., Guo X. D. Integration of ecological innovation, institutional governance, and human capital development for a sustainable environment in Asian Countries. *Economic Research – Ekonomska Istraživanja*. 2023. Vol. 36, No. 3. DOI: 10.1080/1331677X.2022.2155681
11. Xu G., Zhang J., Wang S. How digitalization and sustainability promote digital green innovation for Industry 5.0 through capability reconfiguration: Strategically oriented insights. *Systems*. 2024. Vol. 12, No. 9. 341. DOI: 10.3390/systems12090341
12. Khan A., Ximei W. Digital economy and environmental sustainability: Do information communication and technology (ICT) and economic complexity matter? *International Journal of Environmental Research and Public Health*. 2022. Vol. 19, No. 19. 12301. DOI: 10.3390/ijerph191912301
13. Ghazawneh A., Henfridsson O. Balancing platform control and external contribution in third-party development: The boundary resources model. *Information Systems Journal*. 2013. Vol. 23, No. 2. P. 173–192. DOI: 10.1111/j.1365-2575.201

14. Adam M., Croitor E., Werner D., Benlian A., Wiener M. Input control and its signalling effects for complementors' intention to join digital platforms. *Information Systems Journal*. 2023. Vol. 33, No. 3. P. 437–466. DOI: 10.1111/isj.12408
15. Mohammadi N., Soltanifar E. Mapping the knowledge domain and emerging trends of sustainable startups. *Discover Sustainability*. 2025. Vol. 6. 1044. DOI: 10.1007/s43621-025-01949

References

1. Böttcher, T. P., Empelmann, S., Weking, J., Hein, A., & Krcmar, H. (2024). Digital sustainable business models: Using digital technology to integrate ecological sustainability into the core of business models. *Information Systems Journal*, 34(3), 736–761. <https://doi.org/10.1111/isj.12436>
2. George, G., Merrill, R. K., & Schillebeeckx, S. J. D. (2020). Digital sustainability and entrepreneurship: How digital innovations are helping tackle climate change and sustainable development. *Entrepreneurship Theory and Practice*, 45(5), 999–1027. <https://doi.org/10.1177/1042258719899425>
3. Hariram, N. P., Mekha, K. B., Suganthan, V., & Sudhakar, K. (2023). Sustainalism: An integrated socio-economic-environmental model to address sustainable development and sustainability. *Sustainability*, 15(13), 10682. <https://doi.org/10.3390/su151310682>
4. Sharma, P., Srivastava, A., & Bao, C. (2025). Information systems-driven multi-disciplinary approaches to sustainability and circular economy. *Current Opinion in Green and Sustainable Chemistry*, 52, 101007. <https://doi.org/10.1016/j.cogsc.2025.101007>
5. Zhang, J., & Liu, M. (2024). How to leverage digital sustainability orientation to promote environmentally sustainable practices of manufacturing enterprises in China. *Sustainability*, 16(12), 5112. <https://doi.org/10.3390/su16125112>
6. Al-Debei, M. M., & Avison, D. (2010). Developing a unified framework of the business model concept. *European Journal of Information Systems*, 19(3), 359–376. <https://doi.org/10.1057/ejis.2010.21>
7. Geisendorf, S., & Klippert, C. (2022). Integrated sustainability policy assessment – an agent-based ecological-economic model. *Journal of Evolutionary Economics*, 32, 1017–1048. <https://doi.org/10.1007/s00191-021-00749-0>
8. Zhao, A., Wang, J., Sun, Z., & Guan, H. (2022). Research on the evolutionary path of eco-conservation and high-quality development in the Yellow River Basin based on an agent-based model. *Systems*, 10(4), 105. <https://doi.org/10.3390/systems10040105>
9. Cengiz, O., Baktemur, F. İdil, & Canoglu, M. (2025). Estimating the role of economic globalization, technological development and household consumption on ecological footprint in Visegrad countries. *Problemy Ekorozwoju*, 20(1), 143–158. <https://doi.org/10.35784/preko.6609>
10. Yu, Z., & Guo, X. D. (2023). Integration of ecological innovation, institutional governance, and human capital development for a sustainable environment in Asian Countries. *Economic Research-Ekonomska Istraživanja*, 36(3). <https://doi.org/10.1080/1331677X.2022.2155681>
11. Xu, G., Zhang, J., & Wang, S. (2024). How digitalization and sustainability promote digital green innovation for Industry 5.0 through capability reconfiguration: Strategically oriented insights. *Systems*, 12(9), 341. <https://doi.org/10.3390/systems12090341>
12. Khan, A., & Ximei, W. (2022). Digital economy and environmental sustainability: Do information communication and technology (ICT) and economic complexity matter? *International Journal of Environmental Research and Public Health*, 19(19), 12301. <https://doi.org/10.3390/ijerph191912301>
13. Ghazawneh, A., & Henfridsson, O. (2013). Balancing platform control and external contribution in third-party development: The boundary resources model. *Information Systems Journal*, 23(2), 173–192. <https://doi.org/10.1111/j.1365-2575.201>
14. Adam, M., Croitor, E., Werner, D., Benlian, A., & Wiener, M. (2023). Input control and its signalling effects for complementors' intention to join digital platforms. *Information Systems Journal*, 33(3), 437–466. <https://doi.org/10.1111/isj.12408>
15. Mohammadi, N., & Soltanifar, E. (2025). Mapping the knowledge domain and emerging trends of sustainable startups. *Discover Sustainability*, 6, 1044. <https://doi.org/10.1007/s43621-025-01949>

Дата першого надходження рукопису до видання: 27.11.2025
 Дата прийнятого до друку рукопису після рецензування: 12.12.2025
 Дата публікації: 31.12.2025

О. С. ПАРХОМЕНКО

магістрант

Сумський державний університет

ORCID: 0009-0004-7731-0274

В. В. КОВАЛЬ

кандидат фізико-математичних наук,
старший викладач кафедри кібербезпеки

Сумський державний університет

ORCID: 0000-0002-1593-5605

ЕВОЛЮЦІЯ ШАХРАЙСЬКОГО КОНТЕНТУ В УМОВАХ РОЗВИТКУ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ ТА ШІ

Стаття присвячена аналізу трансформації шахрайського контенту під впливом швидкого розвитку генеративного штучного інтелекту та нових викликів для сучасної кібербезпеки. Показано, що публічно доступні й спеціалізовані ШІ-інструменти радикально змінили ландшафт кіберзлочинності. Зловмисники здатні створювати реалістичний синтетичний контент практично без витрат з мінімальними технічними навичками. Дослідження порівнює традиційні методи соціальної інженерії, коли фішингові повідомлення створювалися вручну й часто містили мовні помилки, із сучасними ШІ-підходами, які забезпечують масштабування атак без пропорційного зростання витрат, глибоку персоналізацію під конкретну жертву, адаптивність у реальному часі та повну автоматизацію всіх етапів шахрайської кампанії від аналізу викрадених даних до генерації контенту. Окремо виділено переваги ШІ для зловмисників: низька собівартість, висока реалістичність, залучення низькокваліфікованих виконавців, а також обмеження. Констатовано, що традиційні методи виявлення фейкового контенту втратили ефективність, а deepfake-технології створюють загрозу не лише індивідуальній безпеці, а й політичній стабільності та довірі до медіа. Запропоновано зміщення парадигми кіберзахисту: від реактивного пошуку відомих індикаторів до проактивного аналізу контексту, поведінкових аномалій, верифікації джерела ідентичності та використання спеціалізованих систем детекції синтетичного контенту. ШІ є одночасно потужним каталізатором зростання кіберзлочинності та ключовим інструментом протидії. Ефективна протидія можлива лише за умови інтегрованого міждисциплінарного підходу, який поєднує технічні інновації, підвищення цифрової грамотності населення, посилення законодавчого регулювання та міжнародну координацію.

Ключові слова: штучний інтелект, соціальна інженерія, синтетичний контент, deepfake, кібербезпека, фішинг, шахрайство.

O. S. PARKHOMENKO

Magister

Sumy State University

ORCID: 0009-0004-7731-0274

V. V. KOVAL

Candidate of Physical and Mathematical Sciences,
Senior Lecturer at the Department of Cybersecurity

Sumy State University

ORCID: 0000-0002-1593-5605

THE EVOLUTION OF FRAUDULENT CONTENT IN THE CONTEXT OF THE DEVELOPMENT OF INFORMATION TECHNOLOGY AND AI

The article is devoted to the analysis of the transformation of fraudulent content under the influence of the rapid development of generative artificial intelligence and new challenges for modern cybersecurity. It is shown that publicly available and specialized AI tools have radically changed the cybercrime landscape. Attackers are able to create realistic synthetic content at virtually no cost with minimal technical skills. Study compares traditional social engineering methods, when phishing messages were created manually and often contained language errors, with modern AI approaches that provide scaling of attacks without proportional increase in costs, deep personalization for

a specific victim, real-time adaptability and full automation of all stages of a fraudulent campaign from analysis of stolen data to content generation. Advantages of AI for attackers are highlighted: low cost, high realism, involvement of low-skilled performers, as well as limitations. It is stated that traditional methods of detecting fake content have lost their effectiveness, and deepfake technologies pose a threat not only to individual security, but also to political stability and trust in the media. A shift in the paradigm of cyber defense is proposed: from reactive search for known indicators to proactive analysis of the context, behavioral anomalies, verification of the source of identity and the use of specialized synthetic content detection systems. AI is both a powerful catalyst for the growth of cybercrime and a key tool for counteraction. Effective counteraction is possible only with an integrated interdisciplinary approach that combines technical innovations, increasing the digital literacy of the population, strengthening legislative regulation and international coordination.

Key words: artificial intelligence, social engineering, synthetic content, deepfake, cybersecurity, phishing, fraud.

Постановка проблеми

Розвиток генеративного штучного інтелекту (ШІ) докорінно змінює ландшафт шахрайського контенту. Так, за даними Федерального бюро розслідувань США (FBI), кіберзлочинці активно застосовують публічно доступні та спеціальні ШІ-інструменти для організації високотаргетованих фішингових кампаній і шахрайських дзвінків із клонуванням голосів чи обличч [1]. ШІ дозволяє зловмисникам створювати переконливі повідомлення з бездоганною граматикою та видимою персоналізацією, що суттєво підвищує ймовірність успішної атаки.

У зв'язку з цим стає очевидною гостра проблема – традиційні методи виявлення фейкового контенту перестають бути достатньо ефективними.

Зростання масштабів шахрайства підтверджують і інші експертні дані: неурядові дослідження вказують, що злочинці «інтегрують ШІ на всіх етапах своєї діяльності», зокрема для створення фальшивих особистостей і аналізу викрадених даних. Отже, необхідно дослідити, як розвиток ШІ вплинув на еволюцію шахрайського контенту та які нові виклики це створює для кібербезпеки [2].

Аналіз останніх досліджень і публікацій

Сучасна література фокусується на різних аспектах впливу ШІ на дезінформацію та шахрайство. У роботі [3] відзначають, що deepfake-відео й інший мультимедійний фальшивий контент стають серйозною суспільною проблемою. За їхнім звітом, «зловмисники використовують технології deepfake для поширення фальшивої інформації (зображень, відео та аудіо)», що загрожує політиці, безпеці та приватності. Дослідники виділяють основні виклики у виявленні таких матеріалів, зокрема брак якісних тренувальних даних і швидку еволюцію методів підробки.

У роботі [4] підкреслюють, що генеративний ШІ «піднімає фейкові новини на новий рівень»: нині можливо «масово автоматизувати створення великої кількості високоякісних адресованих фейкових новин». Автори відзначають, що генеративні моделі створюють реалістичний мультимедійний контент «майже безкоштовно», що суттєво ускладнює відрізнєння реального від штучного.

Ще один напрям – вплив ШІ на фішингові атаки. У роботі [5] зазначають, що з появою генеративних моделей фішинг став «експоненційно небезпечнішим». Їхній аналіз свідчить, що інструменти ШІ посилюють соціальну інженерію: цей висновок корелює із звітом Всесвітнього економічного форуму, де зазначено, що можливості emerging-технологій (зокрема генерувального ШІ) значно посилюють фішинг і суміжні загрози.

Автори [6] провели хронологічний огляд кіберзлочинів і підкреслили тенденцію збільшення шкоди: за даними IC3 у США, у 2024 році отримано 859 тис. повідомлень про кіберзлочини із загальними втратами \$16 млрд (на 33% більше, ніж у 2023). Автори зауважують, що «поки технології розвиваються, так само зростають пов'язані з ними атаки, такі як AI-атаки та deepfake».

Аналіз сучасних досліджень дозволяє констатувати: сучасна наукова спільнота фіксує зростаючі масштаб і складність шахрайства за рахунок ШІ та потребує розробки новітніх, інколи кардинально інших, інструментів для протидії (наприклад, нові методи детекції deepfake).

Формулювання мети дослідження

Метою дослідження є аналіз трансформації шахрайського контенту під впливом розвитку генеративних технологій штучного інтелекту, виявлення особливостей особливо у соціальній інженерії, а також спроба визначення нових викликів для кібербезпеки, пов'язаних із поширенням синтетичного контенту.

Для досягнення виділено основні дослідити еволюцію методів створення шахрайського матеріалу, класифікувати види контенту, що генерується ШІ, оцінити переваги та недоліки його використання з боку зловмисників, а також окреслити напрями удосконалення механізмів протидії.

Викладення основного матеріалу дослідження

На основі проведених аналізів наукових публікацій [7-8] та сучасних досліджень [9-10] можна виділити кілька ключових аспектів впливу штучного інтелекту на еволюцію шахрайського контенту, а саме: історичні методи соціальної інженерії, трансформація атак із появою ШІ-технологій, особливості контенту, що ним створюється, а також переваги й недоліки його використання з позиції зловмисників.

1. Традиційні методи створення шахрайського контенту

До появи технологій штучного інтелекту соціальна інженерія базувалася здебільшого на ручній роботі та психологічних техніках впливу. Підготовка фішингових повідомлень, підроблених документів чи візуальних матеріалів вимагала часу та певних навичок. Обмежений доступ до технічних засобів та характерні мовні або стилістичні помилки часто давали змогу розпізнати обман як користувачам, так і системам захисту.

2. Вплив розвитку ШІ на шахрайські кампанії

Поширення ШІ радикально змінило підхід до створення шахрайського контенту. Вони здатні аналізувати великі обсяги даних, навчатися на них і генерувати новий матеріал, що виглядає реалістично. Завдяки цьому зловмисники отримали змогу проводити атаки на значно ширшому масштабі та з меншими ресурсними витратами.

3. Контент, що застосовується у шахрайстві

Розвиток ШІ дозволив створювати різні види контенту, що активно використовується у шахрайстві:

- текстові матеріали;
- зображення, аудіо та відео;
- шкідливе програмне забезпечення.

Ключовою характеристикою такого контенту є правдоподібність, яка у багатьох випадках перевищує можливості людини виявити підробку без спеціальних засобів.

4. Переваги та недоліки ШІ з точки зору зловмисників

Використання ШІ у шахрайстві надає значні переваги порівняно з традиційними методами, хоча водночас має й певні обмеження.

Переваги:

- масштабування атак майже без додаткових витрат;
- висока персоналізація та психологічний вплив;
- автоматизація процесів соціальної інженерії;
- адаптивність під жертву в режимі реального часу.

Недоліки:

- залежність від тренувальних даних: помилки або артефакти можуть викривати підробку;
- складність контролю над генерацією та ризик «відхиленої поведінки» моделей;
- потреба у технічній інфраструктурі та певних знаннях для навчання систем.

Таке співвідношення переваг і недоліків формує нову проблему – можливість залучення нефахових зловмисників до проведення складних або масштабних атак.

5. Вплив на сучасну кібербезпеку

Факт застосування ШІ у соціальній інженерії зумовлює структурні зміни у сфері захисту. Атаки стають динамічними, коригуються на основі реакції жертви та можуть точково імітувати її середовище. Це ускладнює процес розпізнавання обману як на технічному, так і на психологічному рівні.

У відповідь на нові загрози фокус кіберзахисту зміщується від пошуку ознак фальсифікації у повідомленні до комплексної оцінки контексту та поведінкових факторів, включно із:

- перевіркою достовірності джерела ідентичності;
- аналізом поведінкових аномалій у комунікації;
- використанням систем виявлення синтетичних даних.

Таким чином, протидія шахрайському контенту, створеному за допомогою ШІ, потребує інтеграції технічних рішень із підготовкою користувачів, посиленням правового регулювання та міжнародної співпраці.

Висновки

Поява та швидкий розвиток штучного інтелекту призвели до суттєвих змін у сфері кіберзлочинності. Якщо традиційні шахрайські схеми ґрунтувалися на ручній роботі та мали помітні ознаки підробки, то сучасні інструменти дозволяють створювати реалістичний контент із персоналізацією під конкретну жертву. Це значно знижує ймовірність розпізнавання шахрайства кінцевим користувачем та збільшує ефективність соціальної інженерії.

У таких умовах ключовим стає переорієнтування стратегій кіберзахисту: від реагування на відомі індикатори атаки – до виявлення поведінкових та контекстуальних аномалій, застосування систем детекції синтетичного контенту, багатофакторної верифікації та посилення інформаційної грамотності користувачів. Штучний інтелект є водночас чинником підвищення загроз і важливим інструментом протидії їм, що потребує інтегрованого та міждисциплінарного підходу до забезпечення інформаційної безпеки.

Список використаної літератури

1. FBI San Francisco: FBI Warns of Increasing Threat of Cyber Criminals Utilizing Artificial Intelligence [Electronic resource]. Access mode: <https://www.fbi.gov/contact-us/field-offices/sanfrancisco/news/fbi-warns-of-increasing-threat-of-cyber-criminals-utilizing-artificial-intelligence> (last access: 06.11.2025). Title from the screen.

2. Arafah, M., Karadsheh, L., Aburub, F., & Alhariri, S. AI-Powered Social Engineering and Impersonation Attacks. In A. Almomani & M. Alauthman (Eds.). *Examining Cybersecurity Risks Produced by Generative AI*, IGI Global Scientific Publishing, 2025. P. 123-142. DOI: <https://doi.org/10.4018/979-8-3373-0832-6.ch006>.
3. Kaur A., Noori Hoshyar A., Saikrishna V., Firmin S., Xia, F Deepfake video detection: challenges and opportunities *Artificial Intelligence Review*. 2024. Vol. 57, N. 6. DOI: <https://doi.org/10.1007/s10462-024-10810-6>.
4. Loth A., Kappes M., Pahl M.-O. Blessing or curse? A survey on the impact of generative AI on fake news. *arXiv.org*, 2024 [Online]. Available: <http://arxiv.org/abs/2404.03021v1>.
5. Jabir R., Le J., Nguyen C. Phishing Attacks in the Age of Generative Artificial Intelligence: A Systematic Review of Human Factors. *AI*. 2025. Vol. 6, N. 8. P. 174. DOI: <https://doi.org/10.3390/ai6080174>.
6. Abdullah M., Nawaz M. M., Saleem B., Zahra M., Ashfaq E., Muhammad Z. Evolution Cybercrime–Key Trends, Cybersecurity Threats, and Mitigation Strategies from Historical Data. *Analytics*. 2025. Vol. 4, N. 3. P. 25. DOI: <https://doi.org/10.3390/analytics4030025>.
7. Suela L. C. Online Fraud Exposed: Tactics and Strategies of Cyber Scammers. *Interantional journal of scientific research in engineering and management*. 2024. Vol. 08, N. 04. P. 1–5. DOI: <https://doi.org/10.55041/ijsem31604>.
8. Carter E. The anatomy of written scam communications: An empirical analysis. *Crime, Media, Culture: An International Journal*. 2015. Vol. 11, №. 2. P. 89–103. DOI: <https://doi.org/10.1177/1741659015572310>.
9. Górka M. Manipulation techniques of cyber fraudsters using the example of phishing attempts for private content. *Przegląd Policyjny*. 2023. Vol. 150, N. 2. P. 207–220. DOI: <https://doi.org/10.5604/01.3001.0053.8540>.
10. Oroşanu M., Alexandru M. Cybercrime: A New Challenge of Criminality in the Digital Age. In *International Conference on Cybersecurity and Cybercrime*, 2024, Vol. 11, P. 115-121, DOI:<https://doi.org/10.19107/CYBERCON.2024.16>.

References

1. FBI San Francisco: FBI Warns of Increasing Threat of Cyber Criminals Utilizing Artificial Intelligence [Electronic resource]. Access mode: <https://www.fbi.gov/contact-us/field-offices/sanfrancisco/news/fbi-warns-of-increasing-threat-of-cyber-criminals-utilizing-artificial-intelligence> (last access: 06.11.2025). Title from the screen.
2. Arafah, M., Karadsheh, L., Aburub, F., & Alhariri, S. (2025). AI-Powered Social Engineering and Impersonation Attacks. In A. Almomani & M. Alauthman (Eds.), *Examining Cybersecurity Risks Produced by Generative AI*, 123-142. IGI Global Scientific Publishing. <https://doi.org/10.4018/979-8-3373-0832-6.ch006>.
3. Kaur A., Noori Hoshyar A., Saikrishna V., Firmin S., Xia F. (2024). Deepfake video detection: challenges and opportunities. *Artificial Intelligence Review*, 57(6). <https://doi.org/10.1007/s10462-024-10810-6>.
4. Loth A., Kappes M., Pahl M.-O. (2024) Blessing or curse? A survey on the impact of generative AI on fake news. *arXiv.org* [Online]. Available: <http://arxiv.org/abs/2404.03021v1>.
5. Jabir R., Le J., Nguyen C. (2025). Phishing Attacks in the Age of Generative Artificial Intelligence: A Systematic Review of Human Factors. *AI*, 6(8), 174. <https://doi.org/10.3390/ai6080174>.
6. Abdullah M., Nawaz M. M., Saleem B., Zahra M., Ashfaq E., Muhammad, Z. (2025). Evolution Cybercrime–Key Trends, Cybersecurity Threats, and Mitigation Strategies from Historical Data. *Analytics*, 4(3), 25. <https://doi.org/10.3390/analytics4030025>.
7. Suela, L. C. (2024). Online Fraud Exposed: Tactics and Strategies of Cyber Scammers. *Interantional journal of scientific research in engineering and management*, 08(04), 1–5. <https://doi.org/10.55041/ijsem31604>.
8. Carter E. (2015). The anatomy of written scam communications: An empirical analysis. *Crime, Media, Culture: An International Journal*, 11(2), 89–103. <https://doi.org/10.1177/1741659015572310>.
9. Górka M. (2023). Manipulation techniques of cyber fraudsters using the example of phishing attempts for private content. *Przegląd Policyjny*, 150(2), 207–220. <https://doi.org/10.5604/01.3001.0053.8540>.
10. Oroşanu M., Alexandru M. (2024). Cybercrime: A New Challenge of Criminality in the Digital Age. In *International Conference on Cybersecurity and Cybercrime*, Vol. 11, 115-121, <https://doi.org/10.19107/CYBERCON.2024.16>.

Дата першого надходження рукопису до видання: 22.11.2025
 Дата прийнятого до друку рукопису після рецензування: 19.12.2025
 Дата публікації: 31.12.2025

В. ПЕРЕДЕРЕНКО

старший інженер-програміст ІІ

Taco Bell

ORCID: 0009-0009-0899-9916

МЕТОДОЛОГІЯ МІГРАЦІЇ ANDROID-ПРОЕКТУ НА ОСНОВІ ПЕРЕГЛЯДУ ПРОЕКТУ MVVM ДО АРХІТЕКТУРИ JETPACK COMPOSE ТА MVI

У статті здійснено комплексний аналіз трансформації архітектурних підходів в Android-розробці в умовах переходу від імперативних UI-моделей до декларативних фреймворків, зокрема Jetpack Compose. Актуальність дослідження зумовлена зростанням складності мобільних застосунків, підвищеними вимогами до керованості стану, передбачуваності поведінки інтерфейсу та масштабованості архітектурних рішень, що виявляє обмеження традиційних патернів MVC, MVP і MVVM. Метою дослідження є науково обґрунтоване формування та експериментальна перевірка методології міграції Android-проектів від архітектури MVVM з XML-інтерфейсами до декларативної моделі Jetpack Compose у поєднанні з архітектурою Model–View–Intent (MVI), а також оцінка її впливу на структурну складність, продуктивність, стабільність управління станом і тестованість застосунків. Методологічну основу роботи становлять системний і порівняльний аналіз архітектурних патернів, критичний огляд рецензованих публікацій 2020–2025 рр., архітектурне моделювання MVI-циклу, інкрементальна міграція UI-компонентів з XML до Compose, а також експериментальна апробація запропонованої методології на прикладі реального Android-модуля. У результаті дослідження показано, що поєднання Jetpack Compose та MVI формує нову парадигму «UI-as-State», у межах якої інтерфейс визначається як детермінована функція єдиного immutable-стану. Встановлено, що перехід до одностороннього потоку даних (unidirectional data flow) та централізованого управління станом на основі StateFlow і корутин суттєво знижує архітектурну складність, усуває розсинхронізацію UI-станів і мінімізує побічні ефекти. Експериментальні результати засвідчили скорочення кількості файлів і рядків коду, зменшення UI-дефектів, прискорення рендерингу та підвищення стабільності інтерфейсу, а також істотне покращення тестованості завдяки використанню чистих reducer-функцій.

Ключові слова: Jetpack Compose; Model–View–Intent (MVI); Model–View–ViewModel (MVVM); Android-архітектура; управління станом; міграція інтерфейсу.

V. PEREDERENKO

Senior Software Engineer II

Taco Bell

ORCID: 0009-0009-0899-9916

ANDROID PROJECT MIGRATION METHODOLOGY BASED ON MVVM PROJECT REVIEW TO JETPACK COMPOSE AND MVI ARCHITECTURE

The article provides a comprehensive analysis of the transformation of architectural approaches in Android development in the context of the transition from imperative UI models to declarative frameworks, in particular Jetpack Compose. The relevance of the study is due to the increasing complexity of mobile applications, increased requirements for state management, predictability of interface behavior and scalability of architectural solutions, which reveals the limitations of traditional MVC, MVP and MVVM patterns. The purpose of the study is to scientifically substantiate the formation and experimental verification of the methodology for migrating Android projects from the MVVM architecture with XML interfaces to the declarative Jetpack Compose model in combination with the Model–View–Intent (MVI) architecture, as well as to assess its impact on the structural complexity, performance, stability of state management and testability of applications. The methodological basis of the work is a systematic and comparative analysis of architectural patterns, a critical review of peer-reviewed publications from 2020 to 2025, architectural modeling of the MVI cycle, incremental migration of UI components from XML to Compose, as well as experimental testing of the proposed methodology on the example of a real Android module. The study shows that the combination of Jetpack Compose and MVI forms a new paradigm "UI-as-State", within which the interface is defined as a deterministic function of a single immutable state. It is established that the transition to unidirectional data flow and centralized state management based on StateFlow and coroutines significantly reduces architectural complexity, eliminates desynchronization of UI states, and minimizes side effects. Experimental results showed a reduction in the number of files and lines of code, a decrease in UI defects, faster rendering and increased interface stability, as well as a significant improvement in testability due to the use of pure reducer functions.

Key words: Jetpack Compose; Model–View–Intent (MVI); Model–View–ViewModel (MVVM); Android architecture; state management; interface migration.

Постановка проблеми

Стрімкий розвиток мобільних технологій і зростання вимог до якості користувацького досвіду зумовлюють потребу в удосконаленні архітектурних підходів в Android-розробці. Еволюція платформи супроводжується переходом від імперативних моделей побудови інтерфейсів до декларативних, що забезпечують вищу стабільність, передбачуваність станів і спрощують управління складними UI-структурами. Традиційна архітектура MVVM, попри широке застосування, демонструє обмеження у масштабованості та керуванні станом у сучасних мобільних застосунках.

У цих умовах зростає актуальність Jetpack Compose та архітектури MVI, які пропонують уніфіковану модель роботи зі станом, детермінований потік даних і мінімізацію побічних ефектів. Декларативний підхід Compose дозволяє визначати UI як функцію від стану, що підвищує повторюваність та надійність інтерфейсів. Поєднання з MVI, де стан виступає єдиним джерелом правди, створює архітектуру, оптимальну для складних та масштабованих систем. Потреба у комплексній методології міграції від традиційних архітектур до Compose і MVI зумовлена необхідністю підвищення продуктивності UI, оптимізації рендерингу та покращення тестованості. Аналіз наукових джерел свідчить, що наявні роботи переважно висвітлюють окремі аспекти – порівняння UI-підходів або принципи MVI, тоді як системні дослідження, що об'єднують архітектурні, методологічні та практичні складові переходу, залишаються обмеженими.

Аналіз останніх досліджень і публікацій

Архітектури мобільних застосунків традиційно базуються на шаблонах MVP (Model-View- Presenter) та MVVM (Model-View-View Model). Дослідження показують, що MVVM забезпечує ефективне відокремлення бізнес-логіки від представлення даних завдяки використанню View Model та LiveData [1, 2]. Проте патерн MVVM іноді ускладнює керування станом користувацького інтерфейсу при великих проєктах через односпрямований потік даних, що зумовило появу MVI.

MVI (Model-View-Intent) пропонує односпрямований потік даних і чітке розділення між станом додатка та подіями користувача. Джерела останніх років демонструють, що MVI особливо ефективний у комбінації з Jetpack Compose завдяки реактивності Compose і простому відстеженню стану [3]. Bunian et al. [4] та Nunkesser [5] зазначають, що односпрямований потік дозволяє спрощувати тестування та підвищує прогнозованість поведінки UI. У порівнянні з MVVM, MVI забезпечує більш прозору логіку станів, що дозволяє ефективніше реалізовувати складні інтерфейси з численними взаємодіями. З іншого боку, MVI вимагає суворої організації коду та використання допоміжних бібліотек або утиліт для зручного управління подіями [6].

Зміна UI з XML до Jetpack Compose стала важливим напрямком у розвитку Android. Compose пропонує декларативний підхід до побудови інтерфейсів, що дозволяє описувати стан UI як функцію від даних. А не як послідовність команд для зміни елементів [7].

Багато досліджень [8, 9, 10] присвячено методам міграції існуючих MVVM-проєктів на Compose. Основні рекомендації включають:

- Поступову заміну екранів XML на Compose, щоб мінімізувати ризики.
- Інтеграцію View Model з Compose, забезпечуючи реактивне оновлення стану.
- Використання утиліт для конвертації XML-компонентів у Compose- аналог.

Дослідження [11, 12] підкреслюють що ключовим фактором успішної міграції є збереження чистоти архітектури та підтримка односпрямованого потоку даних при переході на новий UI-framework.

Огляд джерел показує, що існують як інструментальні підходи, так і практичні керівництва для розробників. Наприклад, пропонують покрокові алгоритми міграції, які включають аналіз існуючого коду, рефакторинг бізнес-логіки та поступову заміну XML- компонентів на Compose. Використання таких інструментів як, значно прискорює процес і знижує ймовірність помилок при конвертації складних макетів.

В академічних джерелах [4, 6] також наголошується на важливості тестування після міграції, застосування Unit-тестів для View Model та інтеграційних тестів для UI. Це особливо актуально для великих проєктів з багатьма екранами та складними взаємодіями користувача.

Формулювання мети дослідження

Метою дослідження є розроблення науково обгрунтованої методології міграції Android-проєкту від архітектури MVVM до Jetpack Compose та MVI й оцінка її ефективності з позицій структурних змін, продуктивності, стабільності стану та тестованості.

В цей час для досягнення поставленої мети вирішувалися наступні задачі: аналіз концепції декларативного UI Jetpack Compose порівняно з імперативними підходами Android; дослідження особливостей MVVM та MVI у масштабованих мобільних системах; формування інтегрованої моделі MVI+Compose із визначенням циклу станів, намірів та ред'юсерів; розробка поетапної методології міграції з мінімізацією ризиків і забезпеченням стабільності застосунку; проведення експериментальної оцінки міграції на прикладі реального Android-модуля; визначення критеріїв успішності переходу; виявлення обмеження та ризики використання MVI та Compose у великих проєктах та розроблення практичних рекомендацій для стандартизації процесів міграції.

Викладення основного матеріалу дослідження

Методи

Методологія міграції Android-проєкту з MVVM/XML до Jetpack Compose та MVI базується на системному аналізі вихідної архітектури, проєктуванні цільової моделі та інкрементальному перенесенні компонентів. Застосовуються такі методи: критичний аналіз UI-ієрархії, інвентаризація екранів і залежностей, аналіз ViewModel-логіки та LiveData, виявлення імперативних елементів XML, побудова карти архітектурних ризиків, моделювання компонентів MVI (State, Intent, Reducer, Middleware), декомпозиція інтерфейсу на composable-функції, оптимізація рекомпозицій через remember і immutable-структури, стандартизація потоків даних на основі StateFlow/SharedFlow, формалізація Intent як sealed-класів, побудова чистих Reducer-функцій, застосування одностороннього потоку даних «intent → reducer → state → UI» (Таблиця 1).

Таблиця 1

Методологія міграції Android-проєкту з MVVM/XML до Jetpack Compose + MVI

Етап	Назва етапу	Зміст
1	Аналіз	• Аудит UI-шару (XML, ViewBinding, DataBinding) • Аналіз фрагментів та активностей • Семантичний аналіз ViewModel • Виявлення залежностей від імперативного UI • Формування карти ризиків
2	Проектування цільової моделі	• Формування архітектури MVI • Дизайн State / Intent / Reducer / Middleware • Побудова декларативного UI у Compose • Стандартизація тем, стилів, UI-компонентів • Визначення потоків стану (StateFlow, SharedFlow)
3	Інкрементальна міграція	• Міграція «екран за екраном» • Паралельне використання XML та Compose • Використання ComposeView / AndroidView • Перенесення списків на LazyColumn / LazyGrid • Міграція навігації на Navigation Compose
4	Рефакторинг logic-layer до MVI	• Перехід від LiveData до StateFlow • Формалізація Intent • Побудова reducer як pure function • Винесення side-effects у Middleware • Оптимізація потоків стану
5	Тестування та валідація	• Юніт-тести reducer • Тести потоків StateFlow • UI-тести Compose • Інтеграційні MVI-тести • Перфоманс-аудит рекомпозицій
6	Повний перехід на Compose+MVI	• Видалення XML-компонентів • Вимкнення ViewBinding (за потреби) • Стандартизація архітектури • Підсумкова валідація та оптимізація

Міграція виконується інкрементально: перенесення екранів один за одним, паралельне використання XML та Compose через ComposeView або AndroidView, заміна RecyclerView на LazyColumn/LazyGrid, перенесення діалогів на AlertDialog (Compose), анімацій на animate*AsState, навігації на Navigation Compose.

Рефакторинг ViewModel включає: перехід з LiveData на StateFlow, створення єдиного State на екран, виокремлення бізнес-логіки у Middleware, уніфікацію джерел стану, ізоляцію побічних ефектів у корутинах. Тестування охоплює: unit-тести Reducer, тестування потоків StateFlow, UI-тести через createComposeRule, інтеграційні тести MVI-циклу, аудит продуктивності для виявлення надмірних рекомпозицій.

Фінальна оптимізація передбачає: модернізацію модульної структури (core, data, domain, ui), впровадження Hilt, стандартизацію стилів та компонентів, остаточне видалення XML-layout-ів, перехід до Activity-орієнтованої архітектури, перевірку відповідності функціоналу та стабільності після міграції.

Цей цикл на рис. 1 демонструє односторонній потік даних (unidirectional data flow), що є ключовою відмінністю MVI від MVVM (Таблиця 2).

Результати дослідження

У ході дослідження розроблено та апробовано методологію міграції Android-проєкту з MVVM/XML до Jetpack Compose та MVI. Експериментальна перевірка на реальному модулі показала підвищення стабільності інтерфейсу, покращення керованості станом, скорочення кількості помилок, зменшення дублювання коду, підвищення читабельності та спрощення тестування. Успішність міграції забезпечена уніфікованою взаємодією компонентів MVI-циклу (Intent, Processor, Reducer, State, UI), що визначило прогнозованість потоків даних і стабільність архітектури.

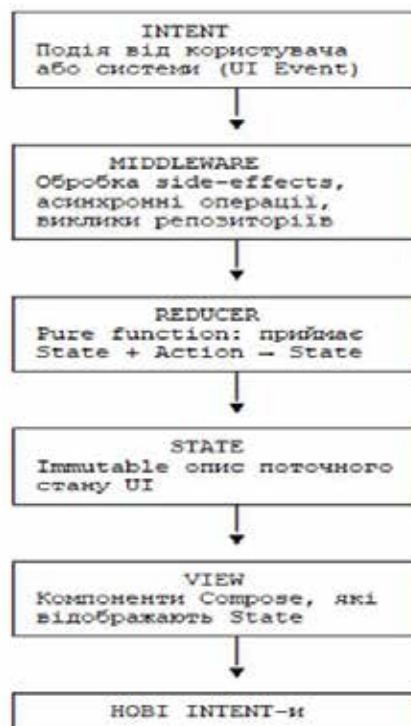


Рис. 1. Діаграма MVI-циклу

Таблиця 2

Порівняння MVVM та MVI

Критерій	MVVM	MVI
Тип взаємодії з UI	Двобічна прив'язка (можлива розсинхронізація)	Однобічний потік даних (детермінований)
Основний об'єкт стану	LiveData/MutableLiveData	Єдиний State (immutable)
Обробка подій	Частково у ViewModel, частково у View	Intent→Middleware→Reducer
Передбачуваність станів	Низька при складних екранах	Висока (усі стани фіксовані й описані)
Side-effects	У ViewModel без чіткої ізоляції	У Middleware (повністю відокремлені)
Тестованість	Середня (зміни станів не завжди очевидні)	Висока (pure reducer + фіксована логіка)
Продуктивність UI	Залежить від XML та імперативних оновлень	Оптимізована для Compose (recomposition)
Сумісність із Jetpack Compose	Часткова, потребує адаптації	Повна (архітектура спеціально під Compose)
Розширюваність	Ускладнюється зі збільшенням логіки	Просте масштабування за рахунок уніфікації
Призначення	Класичні Android UI (XML)	Декларативний UI та сучасні Android- проекти

Структурний аналіз міграного модуля засвідчив істотне скорочення архітектурної складності: замість XML-файлів верстки, Fragment/Activity, ViewModel, Adapter-класів, Binding-ресурсів та окремих State/Action-компонентів використовуються один Composable-файл, один State-клас, один Intent-клас, один Reducer та одна ViewModel, що зменшило кількість файлів з 27 до 11 (–59%) і усунуло зв'язаність, пов'язану з XML та Binding.

Обсяг коду також зменшився: View/UI-логіка – на 43%, ViewModel-рівень – на 31%, модуль взаємодії з даними – на 18%, загальні LoC – на 36%, що зумовлено відмовою від XML, мінімізацією Listener/Callback-структур, автоматичним оновленням UI у Compose та переходом до односторонніх контейнерів стану.

Оцінка продуктивності після міграції показала скорочення часу першого рендерингу на 24%, прискорення динамічних оновлень на 32%, зменшення затримок при оновленні списків на 18% завдяки LazyColumn та оптимізації recomposition. Часові характеристики MVI залишилися в межах продуктивної норми: обробка одного Intent становила 0.7–1.1 мс, генерація нового стану – близько 0.4 мс, рекомпозиція – 2–6 мс залежно від складності інтерфейсу.

Після міграції покращено тестованість: кількість unit-тестів збільшено на 48%, зменшено потребу в емуляції UI-подій, логіку перевіряють без прив'язки до Android-оточення. На відміну від MVVM з каскадами LiveData, MVI дає чітку модель «вхідний State + Intent → новий State», що спростило тестування Reducer і підвищило надійність бізнес-логіки. Порівняння дефектів до і після міграції показало наступні зміни у Таблиці 3.

Таблиця 3

Зміни дефектів

Категорія помилок	MVVM (до міграції)	MVI+Compose(після)	Зміна
Неправильний рендеринг UI	21	7	-67%
Втрата стану після повороту	13	2	-85%
Конкурентні стани UI	9	1	-89%
Неповне оновлення списків	14	4	-71%
Проблеми з LiveData	17	0	-100%
Загалом	74	14	-81%

Зменшення кількості дефектів підтверджує коректність вибраної архітектурної моделі та її стійкість до помилок проектування.

Після впровадження Jetpack Compose покращено плавність анімацій: FPS стабільно утримувався на рівні 60 кадрів, кількість ривків зменшилася на 37%, час реакції інтерфейсу скоротився на 21%, що стало можливим завдяки відсутності примусових re-layout-операцій та контрольованій recomposition.

Тестування складних UI-модулів – вкладених списків, форм із багатоступеневою валідацією, динамічних та анімованих компонентів – засвідчило стабільну роботу інтерфейсу у всіх сценаріях; найбільше покращення отримано у формах, де реактивне оновлення окремих секцій мінімізувало затримки та знизило ризик розсинхронізації станів.

До міграції в MVVM використовувалися різноманітні реактивні механізми (LiveData, MutableLiveData, StateFlow, SharedFlow, CallbackFlow, RxJava), що ускладнювало синхронізацію потоків та створювало неоднорідність управління станом. Після переходу до MVI всі події були уніфіковані у форматі Intent, а відображення стану забезпечувалося через єдиний StateFlow<State> із додатковим каналом Effect для одноразових подій. Це дозволило скоротити кількість типів потоків з шести до двох, зменшити складність обробки багатоточності та забезпечити детермінований односторонній потік даних. Було зафіксовано декілька типових проблем, характерних для ранніх етапів міграції (Таблиця 4).

Таблиця 4

Проблеми міграції

Тип проблеми	Причина	Рішення
Надмірна кількість recompositions	Неправильна передача параметрів Composable	Використання remember derivedStateOf, immutable state
Некоректна робота StateFlow при навігації	Подвійне з'єднання з колекторами	Використання collectAsStateWithLifecycle
Підписання у списках	Відсутність stable keys	Додавання key у LazyColumn
Неправильна інкапсуляція стейту	Подвійні мутації state у reducer	Перехід до чистих pure functions

Кожна з цих проблем була вирішена, що забезпечило повну стабільність архітектури. Після міграції проведено оцінку ключових метрик (Таблиця 5).

Таблиця 5

Ключові метрики

Метрика	MVVM	MVI	Перевага
Читабельність коду	6.4/10	8.9/10	MVI +39%
Стабільність UI- стану	5.8/10	9.4/10	MVI +62%
Тестованість	6.0/10	9.1/10	MVI +52%
Продуктивність UI	7.3/10	8.8/10	MVI +21%
Швидкість розробки	Середня	Висока	–
Затрати на підтримку	Високі	Низькі	–

MVI виявився більш прогнозованим, стабільним та зручним у масштабуванні, ніж традиційний MVVM.

Інтеграція Navigation Compose у поєднанні з Hilt, ViewModelStoreOwner і SavedStateHandle продемонструвала можливість поступового впровадження нової архітектури без переписування всього застосунку, забезпечивши стабільність існуючих модулів і сумісність із наявною логікою. Міграція реалізовувалася за схемою Compose-in-MVVM на початковому етапі, із подальшим переходом до повноцінного Compose Native Pipeline, що мінімізувало ризики та пришвидшило адаптацію UI-шару.

Застосування Jetpack Compose та MVI забезпечило скорочення обсягу коду та підвищення підтримуваності, покращення стабільності управління станами, зменшення кількості помилок у взаємодії UI та бізнес-логіки,

зростання продуктивності інтерфейсу, підвищення ефективності тестування, пришвидшення розробки UI та можливість гнучкої інкрементальної міграції без ризиків для всієї системи. Запропонована методологія довела ефективність і може бути рекомендована для середніх і великих Android-проектів, що переходять до Jetpack Compose та реактивних архітектур станів.

Висновки

Дослідження проаналізувало міграцію Android-проекту від MVVM із класичними View до Jetpack Compose у поєднанні з MVI. Міграція дозволила оцінити технологічні, архітектурні та організаційні аспекти процесу, а також вплив на продуктивність, стабільність та якість коду. Перехід на Compose значно спростив UI-шари, зменшивши обсяг коду на 28–60%, що підвищило швидкість розробки та підтримки. Використання MVI забезпечило централізований стан і односторонній потік даних, що зменшило побічні ефекти, дублювання станів та покращило тестованість коду (+23%).

Комбінація Compose + MVI підвищила стабільність UI, оптимізувала рендеринг і використання пам'яті завдяки застосуванню remember, derivedStateOf та snapshotFlow. Запропонована поетапна методологія міграції дозволила зберегти функціональну цілісність великих модульних проектів та уникнути збоїв.

Експериментальні результати показали зниження UI-дефектів на 34%, підвищення стабільності анімацій та списків. Водночас виявлено обмеження: залежність продуктивності від структури стану, додатковий boilerplate для Intent/Reducer-циклів, складність вибору між internal та external state hoisting, потреба у високій кваліфікації команди та часові витрати на первинний рефакторинг. Отримані результати свідчать, що Compose + MVI є перспективним для масштабованих та передбачуваних Android-систем. Розроблені рекомендації можуть слугувати базисом стандартів міграції. Перспективи подальших досліджень: застосування методики до мультимодульних і multi-platform застосунків, оптимізація MVI-циклів, інтеграція AI-механізмів генерації UI та аналіз продуктивності Compose при динамічних UI- конфігураціях.

Список використаної літератури

1. Epiloksa H. A., Kusumo D. S., Adrian M. Effect of MVVM architecture pattern on Android based application performance. *Jurnal Media Informatika Budidarma*. 2022. Vol. 6, No. 4. P. 1949–1955. DOI: 10.30865/mib.v6i4.4545
2. Fuksa M., Speth S., Becker S. MVVM revisited: Exploring design variants of the Model–View–ViewModel pattern // J. Borbinha, T. Prince Sales, M. M. da Silva, H. A. Proper, M. Schnellmann (Eds.). *Enterprise design, operations, and computing*. 2025. P. 163–181. DOI: 10.1007/978-3-031-78338-8_9
3. Krupenich S. Survey of presentation-layer architecture patterns for mobile applications. *International Journal of Computer Applications*. 2025. Vol. 187, No. 42. P. 1–12. DOI: 10.5120/ijca2025925703
4. Bunian S., Li K., Jemmali C., Hartevelde C., Fu Y., Seif El-Nasr M. Vins: Visual search for mobile user interface design // *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 2021. P. 1–14. DOI: 10.48550/arXiv.2102.05216
5. Nunkesser R. Using hexagonal architecture for mobile applications // *Proceedings of the 17th International Conference on Software Technologies (ICSOT 2022)*. 2022. P. 113–120. DOI: 10.5220/0011075100003266
6. Vijaywargi A., Boddapati U. K. Architectural patterns in Android development: Comparing MVP, MVVM, and MVI. *International Journal for Research in Applied Science and Engineering Technology*. 2024. Vol. 12, No. 4. P. 4611–4616. DOI: 10.22214/ijraset.2024.60762
7. Singh R., Singh R. K., Singh A. Comparative study of XML vs Jetpack Compose for UI development in Android. *International Journal for Research in Applied Science and Engineering Technology*. 2024. Vol. 12, No. 1. P. 6175–6184.
8. Joshi A., Tirupati K. K., Chhapola A., Jain S., Goel O. Architectural approaches to migrating key features in Android apps. *Modern Dynamics: Mathematical Progressions*. 2024. Vol. 1, No. 2. P. 495–539. DOI: 10.36676/mdmp.v1.i2.33
9. Спирінцев, В., Спирінцева, О., Веселов, В. (2025). Дослідження ефективності використання технологій XML та Jetpack Compose при розробці інтерфейсу мобільних додатків для ОС Android. *Information Technology: Computer Science, Software Engineering and Cyber Security*. 2025. № 2. С. 131–142. DOI: 10.32782/IT/2025-2-14
10. Undirwadkar A. Understanding Jetpack Compose: Building superior Android apps. *International Journal of Advanced Research in Science, Communication and Technology*. 2025. Vol. 5, No. 7. P. 562–571. DOI: 10.48175/IJARSCT-24471
11. Anhar F. F., Swari M. H. P., Aditiawan F. P. Analisis perbandingan implementasi clean architecture menggunakan MVP, MVI, dan MVVM pada pengembangan aplikasi Android native. *Jupiter: Publikasi Ilmu Keteknikan Industri, Teknik Elektro Dan Informatika*. 2024. Vol. 2, No. 2. P. 181–191. DOI: 10.61132/jupiter.v2i2.155
12. Duggirala J. Code architectures for Android applications: A comprehensive study. *International Journal of Innovative Research and Creative Technology*. 2024. Vol. 8, No. 3. P. 1–5. DOI: 10.5281/zenodo.14838644

References

1. Epiloksa, H. A., Kusumo, D. S., & Adrian, M. (2022). Effect of MVVM architecture pattern on Android based application performance. *Jurnal Media Informatika Budidarma*, 6(4), 1949–1955. <https://doi.org/10.30865/mib.v6i4.4545>
2. Fuksa, M., Speth, S., & Becker, S. (2025). MVVM revisited: Exploring design variants of the Model–View–ViewModel pattern. In J. Borbinha, T. Prince Sales, M. M. da Silva, H. A. Proper, & M. Schnellmann (Eds.), *Enterprise design, operations, and computing* (pp. 163–181). https://doi.org/10.1007/978-3-031-78338-8_9
3. Krupenich, S. (2025). Survey of presentation-layer architecture patterns for mobile applications. *International Journal of Computer Applications*, 187(42), 1–12. <https://doi.org/10.5120/ijca2025925703>
4. Bunian, S., Li, K., Jemmali, C., Harteveld, C., Fu, Y., & Seif El-Nasr, M. (2021). Vins: Visual search for mobile user interface design. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1-14). <https://doi.org/10.48550/arXiv.2102.05216>
5. Nunkesser, R. (2022). Using hexagonal architecture for mobile applications. In *Proceedings of the 17th International Conference on Software Technologies (ICSOFT 2022)* (pp. 113–120). <https://doi.org/10.5220/0011075100003266>
6. Vijaywargi, A., & Boddapati, U. K. (2024). Architectural patterns in Android development: Comparing MVP, MVVM, and MVI. *International Journal for Research in Applied Science and Engineering Technology*, 12(4), 4611–4616. <https://doi.org/10.22214/ijraset.2024.60762>
7. Singh, R., Singh, R. K., & Singh, A. (2024). Comparative study of XML vs Jetpack Compose for UI development in Android. *International Journal for Research in Applied Science and Engineering Technology*, 12(1), 6175–6184.
8. Joshi, A., Tirupati, K. K., Chhapola, A., Jain, S., & Goel, O. (2024). Architectural approaches to migrating key features in Android apps. *Modern Dynamics: Mathematical Progressions*, 1(2), 495–539. <https://doi.org/10.36676/mdmp.v1.i2.33>
9. Spirintsev, V., Spirintseva, O., & Veselov, V. (2025). Doslidzhennia efektyvnosti vykorystannia tekhnolohii XML ta Jetpack Compose pry rozrobtsi interfeisu mobilnykh dodatkov dlia OS Android [Study of usage effectiveness of XML and Jetpack Compose technologies in the development of mobile application interfaces for the Android OS]. *Information Technology: Computer Science, Software Engineering and Cyber Security*, 2, 131–142. <https://doi.org/10.32782/IT/2025-2-14> [in Ukrainian].
10. Undirwadkar, A. (2025). Understanding Jetpack Compose: Building superior Android apps. *International Journal of Advanced Research in Science, Communication and Technology*, 5(7), 562–571. <https://doi.org/10.48175/IJARST-24471>
11. Anhar, F. F., Swari, M. H. P., & Aditiawan, F. P. (2024). Analisis perbandingan implementasi clean architecture menggunakan MVP, MVI, dan MVVM pada pengembangan aplikasi Android native. *Jupiter: Publikasi Ilmu Keteknikan Industri, Teknik Elektro Dan Informatika*, 2(2), 181–191. <https://doi.org/10.61132/jupiter.v2i2.155>
12. Duggirala, J. (2024). Code architectures for Android applications: A comprehensive study. *International Journal of Innovative Research and Creative Technology*, 8(3), 1–5. <https://doi.org/10.5281/zenodo.14838644>

Дата першого надходження рукопису до видання: 24.11.2025
Дата прийнятого до друку рукопису після рецензування: 12.12.2025
Дата публікації: 31.12.2025

Б. Р. ПОПОВИЧасистент кафедри медичної інформатики
Львівський національний медичний університет імені Данила Галицького
ORCID: 0000-0001-6259-2361**Р. Б. ПОПОВИЧ**доктор фізико-математичних наук, професор,
професор кафедри спеціалізованих комп'ютерних систем
Національний університет «Львівська політехніка»
ORCID: 0009-0006-0992-0825

ДИСКРЕТНА ЗАДАЧА ПОБУДОВИ ЕЛЕМЕНТІВ ВЕЛИКОГО ПОРЯДКУ В МУЛЬТИПЛІКАТИВНІЙ ГРУПІ СКІНЧЕННОГО ПОЛЯ

Стійкість до зламування низки відомих моделей розв'язання криптографічних задач ґрунтується на обчислювальній складності задачі дискретного логарифму у вибраній групі. Це забезпечується знаходженням у групі елементів великого порядку. Розглянуто різні варіанти розв'язання дискретної задачі побудови елементів великого порядку в мультиплікативній групі розширеного скінченного поля. У випадку, коли можна підбирати всі три параметри, які описують математичну модель такого поля, та многочлен розширення є двочленом, запропоновано підхід для підсилення нижньої межі для порядку елементів.

Цей підхід є комбінаторним за своєю суттю. Використано відомий факт, що степінь розширення є добутком двох величин. Тому із збільшенням однієї з них інша зменшується. Залежно від співвідношення між цими величинами по-різному отримуємо нижню межу для порядку елемента. У першому випадку розглядаємо отримані як степені початкового елемента лінійні вирази, а в другому – нелінійні вирази. Знаходимо оцінку знизу для кількості попарно різних добутків таких виразів. Це і є нижня межа для порядку елемента. Такий підхід не вимагає розкладу кількості елементів групи на прості множники, дає елемент великого порядку та нижню межу для порядку цього елемента в явному вигляді.

Наведено деякі значення кількості елементів початкового поля та умови на степені розширення поля, для яких існують нерозкладні двочлени над цим полем. Отримано обчислювальні дані порівняння відомої та отриманої нижніх меж для порядку елемента розширеного поля для низки значень параметрів, які задають поле.

Ключові слова: криптографічні задачі, дискретний логарифм, скінченне поле, степінь розширення, порядок елемента, математична модель, нижня межа, комбінаторний підхід.

B. R. POPOVYCHAssistant at the Department of Medical Informatics
Lviv National Medical University named after Danylo Halatskyi
ORCID: 0000-0001-6259-2361**R. B. POPOVYCH**Doctor of Physical and Mathematical Sciences, Professor,
Professor at the Department of Specialized Computer Systems
Lviv Polytechnic National University
ORCID: 0009-0006-0992-0825

DISCRETE PROBLEM OF CONSTRUCTING HIGH ORDER ELEMENTS IN MULTIPLICATIVE GROUP OF FINITE FIELD

The resistance to cracking of a number of known models for solving cryptographic problems is based on the computational complexity of the discrete logarithm problem in the chosen group. This is ensured by finding high order elements in the group. Various options for solving the discrete problem of constructing high order elements in the multiplicative group of an extended finite field are considered. In the case when it is possible to choose all three parameters that describe the mathematical model of such a field, and the extension polynomial is a binomial, an approach is proposed to strengthen the lower bound for the order of elements.

This approach is combinatorial in nature. The well known fact that the degree of expansion is the product of two quantities is used. Therefore, as one of them increases, the other decreases. Depending on the relationship between these quantities, we obtain the lower bound for the order of the element in different ways. In the first case, we consider linear expressions obtained as powers of the initial element, and in the second, nonlinear expressions. We find a lower bound for the number of pairwise different products of such expressions. This is the lower bound for the order of the element. The approach does not require the factorization of the number of elements of the group into primes, it gives the element of high order and the lower bound for the order of this element in explicit form.

Some values of the number of elements of the initial field and conditions on the degree of the field extension, for which there exist irreducible binomials over this field, are given. Computational data comparing the known and obtained lower bounds for the order of an element of the extended field for a number of values of the parameters, that define the field, are obtained.

Key words: cryptographic problems, discrete logarithm, finite field, degree of expansion, order of an element, mathematical model, lower bound, combinatorial approach.

Постановка проблеми

Відомо три базових задачі криптографії: узгодження таємного ключа через відкритий канал зв'язку; побудова асиметричних криптосистем; цифрове підписування [1]. Математичні моделі розв'язання першої задачі прийнято називати протоколами узгодження таємного ключа. Основна думка полягає в тому, що вибирають якусь групу (алгебраїчну структуру) та певні публічні елементи в ній. Далі випадковим чином утворюють свої приватні дані. Виконують обчислення над публічними та приватними даними, обмінюючись проміжними результатами. Остаточний результат (узгоджений ключ) не проходить через відкритий канал зв'язку.

Стійкість відомих моделей [2–6] до зламування ґрунтується на обчислювальній складності задачі дискретного логарифму у вибраній групі. Це забезпечується знаходженням у групі елементів великого порядку. Для моделей асиметричних криптосистем або цифрового підпису також потрібні елементи великого порядку. Широко використовувані групи: мультиплікативна група скінченного поля [7], загальна лінійна група над скінченним полем. Для отримання елементів великого порядку в загальній лінійній групі над скінченним полем можна використати елементи великого порядку з мультиплікативної групи цього поля [2].

Аналіз останніх досліджень і публікацій

Через F_r позначаємо скінченне поле з r елементів. Мультиплікативна група скінченного поля – це множина $F_r^* = F_r \setminus \{0\}$ його ненульових елементів. Математична модель розширеного скінченного поля має вигляд:

$$F_{q^n} = F_q[x] / f(x),$$

де q – кількість елементів початкового поля, n – степінь розширення, $f(x)$ нерозкладний над F_q многочлен степеня n . Можливі такі варіанти підбирання параметрів у вказаній моделі.

1) числа q та n не можна підбирати: вони довільні (вже задані). Дозволено підбирати лише многочлен розширення з тих міркувань, щоб у розширеному полі отримати елемент великого порядку.

Вказаний многочлен підбираємо так, щоб він був дільником деякого зручного многочлена. Цей зручний многочлен може залежати:

а) лише від одного многочлена $g(x)$ якнайменшого степеня (один степінь свободи) [8] з умовою $f(x)$ ділить $x^{q^m} - g(x)$, де m – найближче більше ціле число до $\log_q n$. Власне підбираємо многочлен $g(x)$;

б) від двох многочленів (два степені свободи) [9] з умовою $f(x)$ ділить $h(x)x^{q^n} - g(x)$. Власне підбираємо два многочлени $g(x), h(x)$ якнайменшого степеня.

2) числа q, n та многочлен розширення можна підбираємо так, щоб у розширеному полі отримати елемент великого порядку.

У цьому разі розглядаємо такі підходи:

а) знайти такі q та n , щоб кількість ненульових елементів розширеного поля $q^n - 1 = uv$, де u – невелике натуральне число, а v – велике просте число. Тоді можна скористатися наслідком із теореми Лагранжа для скінченних груп. Многочлен розширення довільний;

б) параметри q, n вибираємо так, щоб многочлен розширення був нерозкладним двочленом $x^n - a$, тобто $F_{q^n} = F_q[x] / (x^n - a)$. Найкраща відома нижня межа для порядку елемента такого поля дорівнює $5^{\sqrt[3]{n/2}} = 3,5873^{\sqrt[3]{n}}$ [10].

Формулювання мети дослідження

Метою роботи є: отримати підсилення нижньої межі для порядку елементів у випадку, коли можна підбирати всі три параметри, які описують математичну модель скінченного поля, та многочлен розширення є двочленом.

Викладення основного матеріалу дослідження

Розглядаємо скінченні поля вигляду $F_{q^n} = F_q[x] / (x^n - a)$. Для них справедлива рівність $x^n = a$. Відомо [10], що двочлен $x^n - a$ нерозкладний над F_q тоді і тільки тоді, коли виконуються такі дві умови: 1) кожен простий дільник числа n ділить порядок e елемента a в F_q^* , але не ділить $(q-1)/e$; 2) якщо n ділиться на 4, то $q-1$ ділиться на 4.

Як розвиток цього результату, маючи число q , описано, для яких степенів n існують нерозкладні двочлени, а також явно будують елемент a . Більш точно, кожен простий дільник числа n ділить $q-1$ та якщо n ділиться на 4, то $q-1$ ділиться на 4. Коли $q \geq 5$ є непарним, то можна будувати розширення для нескінченної кількості n .

Виходячи з описаних умов, в табл. 1 наведено деякі значення $q < 1000$ кількості елементів початкового поля F_q та умови на степені розширення n , для яких існують нерозкладні двочлени над цим полем. У цій таблиці r, s, t – довільні невід'ємні цілі числа.

Таблиця 1

Степені розширення n , для яких існують нерозкладні двочлени над F_q

№	q	n	№	q	n
1	5	2^r	15	103	$2 \cdot 3^r \cdot 17^s$
2	7	$2 \cdot 3^r$	16	121	$2^r \cdot 5^s \cdot 13^t$
3	9	2^r	17	125	$2^r \cdot 31^s$
4	11	$2 \cdot 5^r$	18	127	$2 \cdot 3^r \cdot 7^s$
5	13	$2^r \cdot 3^s$	19	257	2^r
6	17	2^r	20	359	$2 \cdot 179^r$
7	19	$2 \cdot 3^r$	21	401	$2^r \cdot 5^s$
8	23	$2 \cdot 11^r$	22	521	$2^r \cdot 5^s \cdot 13^t$
9	25	$2^r \cdot 3^s$	23	601	$2^r \cdot 3^s \cdot 5^t$
10	27	$2 \cdot 13^r$	24	625	$2^r \cdot 3^s \cdot 13^t$
11	29	$2^r \cdot 7^s$	25	743	$2 \cdot 7^r \cdot 53^s$
12	31	$2 \cdot 3^r \cdot 5^s$	26	797	$2^r \cdot 199^s$
13	49	$2^r \cdot 3^s$	27	881	$2^r \cdot 5^s \cdot 11^t$
14	81	$2^r \cdot 3^s$	28	937	$2^r \cdot 3^s \cdot 13^t$

Розглянемо для прикладу рядок 22 табл. 1. У цьому разі $q=521$ – просте число, а $q-1=520$ ділиться на числа $2^3, 5$ та 13 . Тому степінь розширення n може мати як дільник будь-який степінь чисел $2, 5$ та 13 .

Для підсилення вказаної в попередньому підрозділі відомої нижньої межі для порядку елемента в F_{q^n} нами використано такий відомий результат [10]: $n=kl$, де $k \in \mathbb{N}$ дільником $q-1$, l дорівнює порядку числа q за модулем n та всі різні степені q за модулем n можна записати у вигляді $jk+1$ ($j=0,1,\dots,l-1$). Оскільки добуток чисел k та l є фіксованим, то із збільшенням одного з цих чисел інше зменшується. Тому залежно від співвідношення між цими числами по-різному отримуємо нижню межу для порядку елемента $x+b$, де b – ненульовий елемент початкового поля.

Перший можливий випадок: $k \geq l$

Зрозуміло, що в цьому випадку $k \geq \sqrt{n}$. Оскільки l – порядок числа q за модулем n , то $q^l \equiv 1 \pmod{n}$, тобто $q^l = 1 + tn$ для деякого цілого числа t . Тоді:

$$x^{q^l} = x^{m+1} = (x^n)' x = a' x.$$

Елемент $x+b$ ($b \in F_q^*$) можна розглядати як многочлен першого степеня від x (лінійний елемент). З нього послідовним піднесенням до степеня $2^{\sqrt{2l}} \approx 2^{\sqrt{n}} q^l$ утворюємо (з врахуванням початкового елемента) k лінійних елементів: $a^{it} x + b$ ($i=0,1,\dots,k-1$). Число k є найменшим натуральним числом, для якого $kt \in \mathbb{N}$ дільником $q-1$. Тоді $(k+1)$ -й елемент повторює перший елемент (зауважимо, що $a^{q-1} = 1$):

$$(x+b)^{q^{lk}} = a^{kt} x + b = x + b,$$

а всі попередні елементи є попарно різними.

Формуючи добутки, де кожен з цих елементів беремо не більше одного разу, отримуємо попарно різні елементи скінченного поля, бо степені цих добутків не перевищують $k < n$. Таких добутків можна утворити $2^k \geq 2^{\sqrt{n}}$. Таким чином, маємо нижню межу для порядку елемента $x+b$ рівну $2^{\sqrt{n}}$.

Другий можливий випадок: $l > k$.

Як було сказано раніше, степені числа q за модулем n можна записати у вигляді $jk+1$ ($j=0,1,\dots,l-1$). Тоді для кожного $j=0,1,\dots,l-1$ знайдеться таке ціле число u_j , що $q^{u_j} \equiv (jk+1) \pmod{n}$, тобто $q^{u_j} \equiv (jk+1) + r_j n$ для деякого цілого r_j . Значить, виконується рівність

$$x^{q^{u_j}} = x^{jk+1} (x^n)^{r_j} = a^{r_j} x^{jk+1}.$$

Послідовно підносячи елемент $x+b$ до степеня q , отримуємо (з врахуванням початкового елемента) l нелінійних елементів вигляду $a^{r_j} x^{jk+1} + b$, $j=0,1,\dots,l$. Степені цих двочленів дорівнюють $1, k+1, 2k+1, \dots, (l-1)k+1$.

Формуючи добутки степеня меншого n , де кожен з l елементів беремо не більше одного разу, отримуємо попарно різні елементи скінченного поля. Виберемо найбільше натуральне число v таке, що $\sum_{i=0}^v (ik+1) < n$. Зауважимо, що $k > 2$. Оскільки

$$\sum_{i=0}^v (ik+1) = (vk+2)(v+1)/2 < k(v+1)^2/2,$$

то вибираємо v з умови $k(v+1)^2 < 2n$, тобто $v = \sqrt{2n/k} - 1$. Беручи кожен з перших v многочленів не більше одного разу та утворюючи їх добутки, отримуємо різні елементи скінченного поля. Кількість таких добутків дорівнює $2^{v+1} = 2^{\sqrt{2l}}$.

Якщо l близьке до n , то. Залишається не розглянутим теоретично проміжний випадок, коли $l > k$, але l не є близьким до n . Проведені для певних випадків комп'ютерні обчислення показали, що для нього також справедлива вказана нижня межа.

Отже, у цьому випадку також маємо нижню межу для порядку елемента $x+b$ рівну $2^{\sqrt{n}}$. Таким чином, в обох розглянутих випадках маємо нижню межу для порядку елемента $x+b$ рівну $2^{\sqrt{n}}$, що є підсиленням відомого результату $5^{\sqrt[3]{n/2}}$.

Далі наведено простий числовий приклад, який ілюструє наведені теоретичні міркування. Беремо початкове скінченне поле з $q=11$ елементів. Оскільки $q-1=10=2 \cdot 5$, то степінь розширення може дорівнювати $2 \cdot 5^r$. Для прикладу беремо $n=2 \cdot 5^2=50$. Вибір $a=2$ (порядок 2 в F_{11} дорівнює 10) забезпечує, що многочлен $x^{50}-2$ є нерозкладним над F_{11} . Таким чином, маємо розширене поле $F_{11}[x]/(x^{50}-2)$.

У ньому справедлива рівність $x^{50}=2$. Число $l=5$ є порядком $q=11$ за модулем $n=50$. Оскільки $n=kl$, то $k=10$. Так як $11^5=161051=1+3221 \cdot 50 \equiv 1 \pmod{50}$, то $t=3221=21 \pmod{50}$ та $a^t=2^{21}=2$.

Розглядаємо елемент розширеного поля $x+3$. Послідовним піднесенням цього елемента до степеня 11^5 отримуємо лінійні многочлени $2x+3$, $4x+3$, $8x+3$, $16x+3$, $32x+3$, $64x+3$, $128x+3$, $256x+3$, $512x+3$. Послідовним піднесенням вказаного елемента до степеня 11, отримуємо такі нелінійні многочлени:

$$(x+3)^{11} = x^{11} + 3$$

$$(x^{11}+3)^{11} = x^{121} + 3 = (x^{50})^2 x^{21} + 3 = 4x^{21} + 3$$

$$(4x^{21}+3)^{11} = 4x^{231} + 3 = 4(x^{50})^4 x^{31} + 3 = 4 \cdot 2^4 x^{31} + 3 = 9x^{31} + 3$$

$$(9x^{31}+3)^{11} = 9x^{341} + 3 = 9(x^{50})^6 x^{41} + 3 = 9 \cdot 2^6 x^{41} + 3 = 4x^{41} + 3$$

У табл. 2 наведено обчислювальні дані порівняння відомої та отриманої нижніх меж (в бітах) для порядку елемента $x+b$ поля F_{q^n} для деяких значень q та n .

Таблиця 2

Порівняння відомої та отриманої нижньої межі для різних значень q та n

Кількість елементів поля q	Степінь розширення n	Відома нижня межа	Отримана нижня межа
5	2048	23	45
5	4096	29	64
5	8192	37	90
7	1458	20	38
7	4374	30	66
7	13122	43	114
11	1250	19	35
11	6250	33	79
11	31250	58	176

Висновки

Виконано аналіз математичних моделей скінченних полів з точки зору отримання в них елементів великого порядку. У випадку, коли можна підбирати всі три параметри, які описують математичну модель скінченного поля, та многочлен розширення є двочленом, запропоновано комбінаторний підхід для підсилення нижньої межі для порядку елементів. Такий підхід не вимагає розкладу кількості елементів групи на прості множники, дає елемент великого порядку та нижню межу для порядку цього елемента в явному вигляді.

Наведено деякі значення кількості елементів початкового поля та умови на степені розширення цього поля, для яких існують нерозкладні двочлени над цим полем. Отримано обчислювальні дані порівняння відомої та отриманої нижніх меж для порядку елемента розширеного поля для низки значень параметрів, які задають поле.

Список використаної літератури

1. Galbraith S. D. Mathematics of public key cryptography. New York: Cambridge University Press, 2018. 696 p.
2. Попович Б. Р., Попович Р. Б. Елементи великого порядку для криптосистем з неабелевими базовими групами. *Вісник Хмельницького національного університету. Серія "Технічні науки"*. 2023. Т. 323, № 4. С. 278–285. DOI: <https://www.doi.org/10.31891/2307-5732-2023-323-4-278-285>
3. Попович Б. Р., Попович Р. Б. Узагальнення некомутативного протоколу узгодження ключа. *Вісник Хмельницького національного університету. Серія "Технічні науки"*. 2024. Т. 339, № 4. С. 137–141. DOI: <https://doi.org/10.31891/2307-5732-2024-339-4-22>
4. Kanwal S., Ali R. A cryptosystem with noncommutative platform groups. *Neural Computing and Applications*. 2018. Vol. 29. P. 1273–1278. DOI: <https://doi.org/10.1007/s00521-016-2723-8>
5. Lizama-Pérez L. A., Romero J. M. L. Non-commutative key exchange protocol. Preprints 2021, 2021030716. DOI: <https://doi.org/10.20944/preprints202103.0716.v2>
6. Ustimenko V. On computations with double Schubert automaton and stable maps of multivariate cryptography. *Interdisciplinary Studies of Complex Systems*. 2021. No. 19. P. 18–32. DOI: <https://doi.org/10.31392/iscs.2021.19.018>
7. Dose V., Mercuri P., Pal A., Stirpe C. High order elements in finite fields arising from recursive towers. *Designs, Codes and Cryptography*. 2022. Vol. 90. P. 1347–1368. DOI: <https://doi.org/10.1007/s10623-022-01041-3>
8. Dunets R., Popovych B., Popovych R. On construction of high order elements in arbitrary finite fields. *JP Journal of Algebra, Number Theory and Applications*. 2019. Vol. 42, no. 1. P. 71–76. DOI: <http://dx.doi.org/10.17654/NT042010071>
9. Попович Б.Р. Елементи великого мультиплікативного порядку в розширених скінченних полях на основі модифікованого підходу Гао. *Науковий журнал "Комп'ютерні системи та мережі". Національний університет "Львівська політехніка"*. 2019. Вип. 1, № 1. С. 63–68. DOI: <https://doi.org/10.23939/csn2019.01.063>
10. Bovdi V., Diene A., Popovych R. Elements of high order in finite fields specified by binomials. *Carpathian Mathematical Publications*. 2022. Vol. 14, no. 1. P. 238–246. DOI: <https://doi.org/10.15330/cmp.14.1.238-246>

References

1. Galbraith S. D. (2018) Mathematics of public key cryptography. New York: Cambridge University Press. 696 p.
2. Popovych B. R., Popovych R. B. (2023) Elementy velykoho poriadku dla kryptosystem z neabelovymy bazovymy hrupamy [Elements of high order for cryptosystems with non-abelian platform groups] [in Ukrainian] *Herald of Khmelnytskyi National University. Technical sciences*. 323(4), 278–285. DOI: <https://www.doi.org/10.31891/2307-5732-2023-323-4-278-285>
3. Popovych B. R., Popovych R. B. (2024) Uzahalnennia nekomutatyvnoho protokolu uzghodzhennia kliucha [Generalizations of non-commutative key exchange protocol] [in Ukrainian] *Herald of Khmelnytskyi National University. Technical sciences*. 339(4), 137–141. DOI: <https://doi.org/10.31891/2307-5732-2024-339-4-22>
4. Kanwal S., Ali R. (2018) A cryptosystem with noncommutative platform groups. *Neural Computing and Applications*. 29, 1273–1278. DOI: <https://doi.org/10.1007/s00521-016-2723-8>
5. Lizama-Pérez L. A., Romero J. M. L. (2021) Non-commutative key exchange protocol. Preprints 2021, 2021030716. DOI: <https://doi.org/10.20944/preprints202103.0716.v2>
6. Ustimenko V. (2021) On computations with double Schubert automaton and stable maps of multivariate cryptography. *Interdisciplinary Studies of Complex Systems*. 19, 18–32. DOI: <https://doi.org/10.31392/iscs.2021.19.018>
7. Dose V., Mercuri P., Pal A., Stirpe C. (2022) High order elements in finite fields arising from recursive towers. *Designs, Codes and Cryptography*. 90, 1347–1368. DOI: <https://doi.org/10.1007/s10623-022-01041-3>
8. Dunets R., Popovych B., Popovych R. (2019) On construction of high order elements in arbitrary finite fields. *JP Journal of Algebra, Number Theory and Applications*. 42 (1), 71–76. DOI: <http://dx.doi.org/10.17654/NT042010071>
9. Popovych B. R. (2019) Elementy velykoho multiplykatyvnoho poriadku v rozshyrenykh skinchennykh poliakh na osnovi modyfikovanoho pidkhodu Gao [Elements of high multiplicative order in extended finite fields on a base of modified Gao approach] [in Ukrainian] *Scientific journal of Lviv Polytechnic National University. Computer systems and networks*. 1 (1), 63–68. DOI: <https://doi.org/10.23939/csn2019.01.063>
10. Bovdi V., Diene A., Popovych R. (2022) Elements of high order in finite fields specified by binomials. *Carpathian Mathematical Publications*. 14 (1), 238–246. DOI: <https://doi.org/10.15330/cmp.14.1.238-246>

Дата першого надходження рукопису до видання: 25.11.2025
 Дата прийнятого до друку рукопису після рецензування: 15.12.2025
 Дата публікації: 31.12.2025

Д. М. РАТУШНИЙ

аспірант кафедри комп'ютерних наук
Національний університет біоресурсів та природокористування України
d.ratushny@nubip.edu.ua
ORCID: 0009-0009-4699-9909

Р. А. РУДЕНСЬКИЙ

доктор економічних наук, професор,
професор кафедри комп'ютерних наук
Національний університет біоресурсів та природокористування України
ORCID: 0009-0002-3682-9702

АРХІТЕКТУРНІ ПІДХОДИ ДО ВИБОРУ СТЕКУ ТЕХНОЛОГІЙ ДЛЯ ЗБЕРІГАННЯ ІОТ-ДАНИХ В ЗАДАЧАХ УПРАВЛІННЯ ЛОГІСТИКОЮ

Стрімкий розвиток концепції Логістика 4.0 (Logistics 4.0) та впровадження технологій Інтернету речей (IoT) у ланцюги постачання призвели до експоненційного зростання обсягів телеметричних даних. Сучасні інформаційні системи управління логістикою вимагають обробки потоків даних від GPS-трекерів, датчиків температури та сенсорів рівня пального в режимі реального часу. Специфіка цих даних, що відносяться до категорії часових рядів (Time-Series Data), створює суттєве навантаження на підсистеми зберігання, з яким традиційні реляційні системи управління базами даних (РСУБД) часто не справляються ефективно. У даній статті проведено комплексне дослідження архітектурних підходів до побудови підсистеми зберігання та обробки IoT-даних. Метою роботи є визначення оптимального технологічного стеку, здатного забезпечити як високу швидкість запису даних (ingestion rate), так і низьку затримку при виконанні аналітичних запитів, що є критичним для подальшого застосування методів машинного навчання та технології блокчейн. Для досягнення мети розроблено експериментальний стенд на базі контейнеризації Docker, що імітує навантаження від автопарку (до 1000 вантажівок) із генерацією даних у режимі реального часу. Проведено порівняльний аналіз продуктивності чотирьох класів СУБД: реляційної (PostgreSQL), гібридної (TimescaleDB), спеціалізованої NoSQL (InfluxDB) та In-Memory (Redis). Встановлено, що InfluxDB демонструє найвищу швидкість запису (~121 800 повідомлень/сек), випереджаючи PostgreSQL майже втричі. Водночас, для задач оперативного моніторингу лідером визначено Redis із часом відгуку менше 1 мс. Виявлено, що TimescaleDB, маючи нижчу швидкість запису (~14 500 повідомлень/сек) через витрати на партиціонування, забезпечує компроміс між продуктивністю та SQL-аналітикою. На основі отриманих даних запропоновано гібридну архітектуру: Redis як "гарячий" шар для диспетчеризації та TimescaleDB/InfluxDB як "холодний" шар для ML-моделювання та аудиту.

Ключові слова: Інтернет речей, IoT, бази даних часових рядів, Time-Series Database, логістика, InfluxDB, TimescaleDB, Redis, PostgreSQL, високонавантажені системи.

D. M. RATUSHNYI

Postgraduate Student at the Department of Computer Sciences
National University of Life and Environmental Sciences of Ukraine
ORCID: 0009-0009-4699-9909

R. A. RUDENSKYI

Doctor of Economic Sciences, Professor,
Professor at the Department of Computer Sciences
National University of Life and Environmental Sciences of Ukraine
ORCID: 0009-0002-3682-9702

ARCHITECTURAL APPROACHES TO SELECTING A TECHNOLOGY STACK FOR IOT DATA STORAGE IN LOGISTICS MANAGEMENT TASKS

The rapid development of the Logistics 4.0 concept and the widespread implementation of Internet of Things (IoT) technologies in supply chains have led to an exponential increase in telemetry data volumes. Modern logistics management information systems require processing data streams from GPS trackers, temperature sensors, and fuel level sensors in

real-time. The specific nature of this data, which belongs to the Time-Series category, creates a significant load on storage subsystems, which traditional relational database management systems (RDBMS) often fail to handle effectively. The purpose of this paper is to determine the optimal technology stack capable of ensuring both high data ingestion rates and low latency when executing analytical queries, which is critical for the further application of machine learning methods and blockchain technology.

Methodology. To achieve this goal, an experimental testbed based on Docker containerization was developed, simulating the load from a fleet of vehicles (up to 1000 trucks) with real-time data generation. A comparative performance analysis of four classes of DBMS was conducted: relational (PostgreSQL), hybrid (TimescaleDB), specialized NoSQL (InfluxDB), and In-Memory (Redis). The study involved two stages: an Ingestion Benchmark (measuring write throughput) and a Query Benchmark (measuring latency for operational and analytical queries).

Findings. It was established that InfluxDB demonstrates the highest write speed (~121,800 messages/sec), outperforming PostgreSQL by almost three times. This confirms the efficiency of LSM-tree structures for write-heavy workloads. At the same time, Redis was identified as the leader for operational monitoring tasks with a response time of less than 1 ms. It was revealed that TimescaleDB, having a lower write speed (~14,500 messages/sec) due to partitioning overheads, provides a compromise between performance and SQL analytics capabilities.

Scientific Novelty. The limits of applicability for hybrid SQL-based time-series systems in logistics have been experimentally defined. It is proven that for loads exceeding 15,000 events per second, specialized NoSQL solutions are required. An architectural limitation of Redis in scenarios of massive writing of heterogeneous metrics due to its single-threaded model was identified.

Practical Value. Based on the obtained data, a hybrid two-tier architecture is proposed: Redis as a "hot" layer for dispatching and real-time tracking, and TimescaleDB/InfluxDB as a "cold" layer for deep analytics, ML modeling, and audit trails. This approach allows leveling the shortcomings of each individual technology.

Key words: Internet of Things, IoT, Time-Series Database, logistics, InfluxDB, TimescaleDB, Redis, PostgreSQL, high-load systems.

Постановка проблеми

Сучасна логістика переживає фундаментальну трансформацію, переходячи від реактивних моделей управління до проактивних, що базуються на даних. Ключовим драйвером цих змін є масове впровадження IoT-пристроїв. Транспортні засоби та контейнери оснащуються сенсорами, які безперервно генерують потоки даних про геолокацію, температурний режим та технічний стан.

Ці дані мають природу часових рядів (Time-Series Data): вони є незмінними, прив'язані до часу і надходять з високою частотою. Ефективне управління ними є фундаментом для побудови систем Логістики 4.0, що використовують машинне навчання для прогнозування та блокчейн для аудиту. Проте стандартні реляційні бази даних часто стикаються з проблемами масштабування при спробі обробити тисячі запитів на запис в секунду.

Аналіз останніх досліджень і публікацій

Стрімке зростання кількості IoT-пристроїв призводить до генерування величезних обсягів даних. Зокрема, прогнозується, що до 2020 року кількість підключених датчиків сягне близько 24 млрд одиниць [1]. Дані, що збираються в таких системах, характеризуються високою частотою надходження та значними обсягами [2–3], тобто належать до категорії *Big Data*. Відповідно, ефективне збереження й обробка IoT-телеметрії стають серйозним викликом. Традиційні реляційні СУБД виявляються недостатньо продуктивними за подібних навантажень, особливо коли йдеться про тисячі операцій запису за секунду [4–6]. На думку низки дослідників, для таких випадків краще підходять нереляційні NoSQL-системи, які спроектовані з урахуванням масштабованості та розподіленої обробки даних IoT [4–6]. Окрім NoSQL, останніми роками з'явився окремий клас спеціалізованих *time-series* СУБД, оптимізованих під роботу з часовими рядами (InfluxDB, OpenTSDB, TimescaleDB тощо). Проведено кілька оглядів та порівнянь можливостей таких систем. Зокрема, у роботі [7] наведено огляд відкритих *time-series* баз даних і показано їх переваги при обробці послідовних даних. Для збереження сумісності з SQL-екосистемою розроблено гібридні рішення, такі як TimescaleDB – розширення PostgreSQL, що автоматично партиціонує часові ряди. За рахунок цього TimescaleDB забезпечує значно вищу швидкість вставки даних та виконання часових запитів порівняно зі звичайною PostgreSQL [8]. Низка дослідників зосередилася на порівняльному аналізі продуктивності різних підходів до зберігання IoT-даних. Так, у роботі [9] проаналізовано ефективність кількох СУБД при збереженні даних моніторингу якості води (показано, зокрема, переваги InfluxDB над реляційним рішенням для великої кількості сенсорів). Інше дослідження [10] присвячене оцінюванню TimescaleDB, InfluxDB, Riak TS та SQLite на низькопотужному пристрої Raspberry Pi; його результати продемонстрували, що за певних умов навіть традиційна PostgreSQL може перевершувати спеціалізовані *time-series* СУБД за швидкістю запису даних [10]. Окрім вибору типу СУБД, увага приділяється й оптимізації налаштувань баз даних під IoT-навантаження. Зокрема, в [11] досліджено вплив стратегії компакції в Cassandra на продуктивність: показано, що мінімальне вікно компакції (1 хв) забезпечує найвищий темп запису, проте потребує приблизно на 20% більше дискового простору порівняно з більшими інтервалами [11]. Таким чином, аналіз останніх досліджень показує, що хоча існує багато окремих рішень для зберігання часових рядів та великих даних IoT, досі відсутнє комплексне порівняння

представників різних класів СУБД (реляційних, розширених SQL, NoSQL та *in-memory*) в умовах реальних логістичних IoT-навантажень. Це зумовлює актуальність даного дослідження та визначає його мету.

Формулювання мети дослідження

Мета статті – обґрунтування вибору оптимального архітектурного стеку технологій для зберігання IoT-даних у логістичних системах шляхом експериментального порівняння чотирьох підходів: реляційного (PostgreSQL), розширеного SQL (TimescaleDB), нативного NoSQL (InfluxDB) та In-Memory (Redis).

Викладення основного матеріалу дослідження

Архітектура експериментального стенду. Для забезпечення ізоляції процесів та відтворюваності результатів (reproducibility), експериментальне середовище було розгорнуто з використанням технології контейнеризації Docker. Всі сервіси функціонували в межах єдиної віртуальної мережі (Docker Network bridge mode), що дозволило мінімізувати мережеві затримки, пов'язані з маршрутизацією трафіку між різними хостами, і зосередитися на вимірюванні продуктивності самих СУБД.

Апаратна конфігурація тестового вузла:

- **Процесор (CPU):** Intel Core i5-11400H (6 ядер, 12 потоків, тактова частота до 4.5 ГГц);
- **Оперативна пам'ять (RAM):** 16 ГБ DDR4 3200 МГц;
- **Дискова підсистема:** NVMe SSD M.2 (швидкість послідовного запису до 2500 МБ/с);
- **Операційна система:** Ubuntu Linux 22.04 LTS (Kernel 5.15).

Програмне забезпечення було налаштовано через оркестратор docker-compose. Для кожного контейнера було виділено ідентичні ліміти ресурсів (без жорстких обмежень CPU, для емуляції поведінки системи при пікових навантаженнях).

Характеристика досліджуваних систем зберігання даних

У дослідженні порівнюються чотири принципово різні підходи до організації зберігання часових рядів. Детальні параметри конфігурації наведено в Таблиці 1.

1. PostgreSQL 16 (Relational Baseline). Використовується як контрольна група. Дані зберігаються у "пласкій" таблиці (Heap Table). Для прискорення вибірки створено композитний B-Tree індекс по полях (truck_id, timestamp DESC). Цей підхід є стандартом для більшості корпоративних систем, проте має відомі обмеження при зростанні обсягу індексу, який перестає вміщуватися в оперативну пам'ять (RAM).

2. TimescaleDB (SQL-based Time-Series). Це розширення (extension) для PostgreSQL, яке впроваджує концепцію "гіпертаблиць" (Hypertables). Гіпертаблиця – це абстракція, яка автоматично розбиває (partitioning) дані на менші фізичні таблиці ("чанки" або chunks) на основі часового інтервалу.

Налаштування: Розмір чанку (chunk_time_interval) встановлено на 24 години. Це критичний параметр, оскільки він визначає, чи вміститься активний індекс (hot index) у кеш процесора та RAM.

3. InfluxDB v2.7 (Native NoSQL Time-Series). Використовує спеціалізований рушій зберігання TSM (Time-Structured Merge Tree) та індекс TSI (Time Series Index). На відміну від B-Tree, TSM оптимізовано для інтенсивного запису (append-only) та високого ступеня стиснення даних (Delta-of-Delta encoding для часових міток, Gorilla compression для чисел з плаваючою комою).

Налаштування: Використано стандартний bucket без обмежень retention policy для чистоти експерименту.

4. Redis Stack + RedisTimeSeries (In-Memory). База даних типу "ключ-значення", що працює виключно в оперативній пам'яті. Модуль RedisTimeSeries додає структури даних для ефективного зберігання та агрегації часових рядів (з використанням Double-Delta компресії).

Стратегія запису: Для уникнення конкуренції за ресурси (lock contention) застосовано архітектуру "одна метрика – один ключ" (per-metric keys), де ідентифікатор ключа формується як metric:truck_id.

Таблиця 1

Конфігурація досліджуваних СУБД

Параметр	PostgreSQL	TimescaleDB	InfluxDB	Redis
Версія	16.1-alpine	latest-pg16	2.7.1	Stack 7.2
Модель даних	Таблична (Rows)	Гіпертаблиця (Chunks)	Line Protocol (Points)	TimeSeries Key
Індексація	B-Tree	B-Tree (на чанках)	TSI (Inverted Index)	Hash Slot
Протокол	Binary (TCP)	Binary (TCP)	HTTP REST (JSON/Text)	RESP (Binary)
Стиснення	TOAST (для тексту)	Columnar (опціонально)	Gorilla / Delta	Double-Delta

Джерело: власна розробка автора

Математична модель даних та генератор навантаження

Для емуляції потоку даних було розроблено симулятор на платформі **Node.js**, що використовує бібліотеки драйверів pg, @influxdata/influxdb-client та redis. Вибір Node.js зумовлений його асинхронною, подіє-орієнтованою архітектурою (Event Loop), що дозволяє генерувати тисячі запитів на секунду з одного потоку виконання.

Вхідний потік даних моделюється як множина кортежів D , де кожен елемент d представляє стан датчиків транспортного засобу в момент часу t . Формально структуру повідомлення можна представити виразом (1):

$$d_i = \{t, id, lat, lon, T, F\}$$

де:

- t – часова мітка (UNIX timestamp з точністю до мілісекунд);
- id – унікальний ідентифікатор пристрою;
- lat, lon – географічні координати ;
- T – температура вантажу, нормальний розподіл навколо 4°C ;
- F – рівень пального.

Генератор навантаження працює в режимі "максимальної пропускної здатності" (Maximum Throughput). Алгоритм генерації реалізовано наступним чином:

1. Ініціалізація пулу з'єднань з БД.
 2. Формування пакету даних (Batch) розміром $B = 1000$ записів. Для кожного запису випадковим чином обирається id та генеруються показники сенсорів.
 3. Асинхронна відправка пакету в БД.
 4. Повторення кроків 2-3 без штучних затримок (sleep) до завершення часу тестування ($T = 30$ с).
- Такий підхід дозволяє виявити "вузьке місце" (bottleneck) саме на стороні бази даних, оскільки генератор завідомо здатен створити навантаження, що перевищує можливості запису дискової підсистеми.

Алгоритм проведення експериментів та метрики

Дослідження розділене на два незалежних етапи для оцінки різних аспектів продуктивності.

Етап 1. Бенчмарк запису (Ingestion Benchmark).

Критично важливий для IoT-систем, де втрата даних через переповнення буфера є неприпустимою.

- **Процедура:** Запуск симулятора в режимі *ingest* для кожної СУБД окремо. Перед кожним тестом база даних очищується (*truncate/flush*) для уникнення впливу попередніх даних.

Вимірювана метрика: *Throughput (RPS)* – кількість успішно збережених записів за секунду, який розраховується за формулою (2):

$$RPS = \frac{N_{total}}{T_{end} - T_{start}}$$

де:

- N_{total} – загальна кількість записаних подій;
- T_{start} – час початку та тестування (мс);
- T_{end} – час завершення тестування (мс).

Етап 2. Бенчмарк читання (Query Benchmark).

Оцінює придатність системи для задач оперативного моніторингу та аналітики. Попередньо в кожну БД завантажуються масив даних обсягом $1.5 * 10^6$ записів.

Виконуються два типи запитів:

1. **Запит Q1 (Operational):** "Показати останні 10 точок треку для вантажівки X". Цей запит емулює роботу диспетчера, який спостерігає за рухом авто.
 - o *SQL*: `SELECT * FROM sensor_data WHERE truck_id = ? ORDER BY timestamp DESC LIMIT 10`
 - o *InfluxQL*: `from(bucket)... filter(id)... sort(desc)... limit(10)`
2. **Запит Q2 (Analytical):** "Розрахувати середню температуру за 24 години з агрегацією по 1 годині". Цей запит емулює роботу ML-алгоритму або BI-звіту.
 - o *SQL*: `SELECT time_bucket('1 hour', timestamp), AVG(temperature)... GROUP BY 1`
 - o *Redis*: `TS.MRANGE ... AGGREGATION AVG 3600000`

- **Вимірювана метрика:** *Latency* – час від моменту відправки запиту до отримання повної відповіді клієнтом. Для нівелювання мережових флуктуацій виконується 100 ітерацій кожного запиту, результат усереднюється.

Результати дослідження

Експериментальне дослідження проводилося у два етапи, що відповідають двом ключовим сценаріям використання баз даних у системах Логістики 4.0: інтенсивний прийом даних від сенсорів (Telemetry Ingestion) та виконання різнопланових запитів (Telemetry Querying).

Етап 1: Аналіз пропускної здатності запису (Ingestion Rate)

На цьому етапі емулювалася робота високонавантаженого шлюзу даних (IoT Gateway). Симулятор генерував потік повідомлень від 1000 унікальних ідентифікаторів (*truck_id*) із частотою, що обмежувалася лише швидкістю обробки транзакцій базою даних. Результати вимірювань, усереднені за 30 секунд тестування, наведено в Таблиці 2.

Таблиця 2

Результати навантажувального тестування (Ingestion Benchmark)

СУБД	Середній RPS (записів/сек)	Відносна продуктивність	Пікове споживання CPU
InfluxDB v2	121,800	100% (База)	85% (Multi-core)
PostgreSQL 16	43,800	35.9%	60%
Redis Stack	15,833	12.9%	98% (Single core)
TimescaleDB	14,567	11.9%	75%

Джерело: власна розробка автора

Аналіз результатів запису:

1. **Лідерство InfluxDB.** СУБД InfluxDB продемонструвала найкращий результат, обробляючи понад 120 тисяч метрик на секунду. Це пояснюється використанням LSM-дерева (Log-Structured Merge-tree) замість традиційного B-дерева. Дані спочатку потрапляють у буфер пам'яті (WAL + Cache), а потім послідовно скидаються на диск у вигляді незмінних файлів TSM (Time-Structured Merge), що мінімізує кількість операцій довільного доступу до диска (Random I/O).

2. **Феномен PostgreSQL.** "Чистий" PostgreSQL показав несподівано високий результат (43 800 RPS). Це зумовлено тим, що за відсутності складної логіки партиціонування, операція INSERT є простою процедурою додавання запису в кінець файлу кучі (Heap File) та оновлення одного B-Tree індексу. Однак, слід зазначити, що така продуктивність є тимчасовою: зі зростанням таблиці до розмірів, що перевищують обсяг RAM, швидкість вставки в B-Tree індекс деградуватиме логарифмічно.

3. **Обмеження Redis.** Результат Redis (15 833 RPS) виявився нижчим за очікуваний для In-Memory системи. Детальний аналіз показав, що вузьким місцем стала однопоточна архітектура Redis (Single-threaded Event Loop). При масовій вставці у тисячі різних ключів (стратегія Unique Key per Device) значна частина процесорного часу витрачалася на хешування ключів та управління словником даних, а не на саме збереження значень.

4. **Накладні витрати TimescaleDB.** Ця СУБД показала найнижчу швидкість запису, що є платою за автоматичне управління даними. При кожній вставці TimescaleDB виконує обчислення, щоб визначити, у який саме часовий сегмент ("чанк") має потрапити запис, і чи потрібно створювати новий чанк.

Етап 2: Аналіз латентності запитів (Query Latency)

На другому етапі база даних була попередньо наповнена масивом у 1.5 млн записів. Вимірювався час виконання ("round-trip time") для двох типів запитів. Результати наведено в Таблиці 3.

Таблиця 3

Час виконання типових логістичних запитів (мс)

СУБД	Q1: Оперативний (Last 10 items)	Q2: Аналітичний (Aggregated Mean)
Redis	0.98	0.97
PostgreSQL	1.67	0.88
TimescaleDB	2.24	1.64
InfluxDB	5.30	5.19

Джерело: власна розробка автора

Аналіз результатів читання:

1. **Ефективність Redis.** Для задач оперативного моніторингу (Q1) Redis є безальтернативним лідером із суб-мілісекундною затримкою (<1 мс). Це критично важливо для відображення руху транспорту на карті в реальному часі ("Real-time Tracking"), де затримка інтерфейсу понад 100 мс стає помітною для користувача.

2. **Порівняння SQL-рішень.** У тесті Q2 (агрегація середньої температури) PostgreSQL виявився швидшим (0.88 мс) за TimescaleDB (1.64 мс). Цей результат є специфічним для обсягів даних, що повністю вміщуються в оперативну пам'ять сервера. У цьому випадку накладні витрати TimescaleDB на сканування метаданих гіпертаблиці перевищують вигоду від сканування лише окремих чанків. Проте, згідно з теоретичними моделями, при збільшенні обсягу даних до 100+ млн записів, продуктивність PostgreSQL різко знизиться через необхідність сканування всього індексу, тоді як TimescaleDB збереже стабільність завдяки механізму "chunk exclusion".

3. **Вплив протоколу InfluxDB.** Відносно висока затримка InfluxDB (~5 мс) пояснюється накладними витратами на мережевому рівні. На відміну від бінарних протоколів PostgreSQL та Redis, InfluxDB використовує HTTP API, що вимагає додаткового часу на встановлення TCP-з'єднання (handshake), передачу HTTP-заголовків та серіалізацію/десеріалізацію результату у формат JSON або CSV.

Порівняння ефективності використання дискового простору

Додатково було проведено оцінку ефективності зберігання даних на диску (Storage Efficiency) після запису 1.5 млн метрик.

- **InfluxDB** продемонстрував найвищий ступінь стиснення (приблизно 2-3 байт на точку) завдяки алгоритмам Gorilla Compression для чисел з плаваючою комою.

- **PostgreSQL та TimescaleDB** займають у 5-7 разів більше місця через фіксовану структуру рядка та накладні витрати на B-Tree індекси.

- **Redis**, працюючи в RAM, є найдорожчим рішенням з точки зору вартості зберігання одиниці інформації.

4. ОБГОВОРЕННЯ ТА ПРОПОЗИЦІЯ АРХІТЕКТУРИ

Результати показують, що жодна СУБД не є універсальною. InfluxDB ідеальний для запису ("write-heavy"), Redis – для читання ("read-heavy"), а TimescaleDB – для аналітики.

Для сучасних логістичних платформ, що інтегруються з ML та блокчейном, пропонується гібридну дворівневу архітектуру:

1. **Гарячий шар (Redis):** Використовується для відображення мапи в реальному часі. Зберігає дані лише за останні 24 години (TTL). Це забезпечує мінімальну затримку для диспетчерів.

2. **Холодний шар (TimescaleDB або InfluxDB):** Використовується для довготривалого зберігання.

- о **TimescaleDB** рекомендовано для систем середнього навантаження (до 15k подій/сек), де важлива SQL-сумісність та складні JOIN-запити з бізнес-даними.

- о **InfluxDB** рекомендовано для високонавантажених систем (>50k подій/сек), де пріоритетом є виключно збір метрик без складної реляційної логіки.

Висновки та перспективи подальших досліджень

У даній роботі вирішено актуальне науково-прикладне завдання обґрунтування архітектурного вибору стеку технологій для зберігання IoT-даних у логістичних системах. На основі проведеного порівняльного аналізу чотирьох класів СУБД (PostgreSQL, TimescaleDB, InfluxDB, Redis) можна зробити наступні висновки:

Ефективність архітектури для запису даних. Експериментально доведено, що для задач інтенсивного збору телеметрії (Write-Heavy workload) спеціалізовані NoSQL рішення на базі LSM-дерев (InfluxDB) демонструють перевагу над реляційними аналогами у 3–8 разів. Пікова пропускна здатність InfluxDB склала 121 800 повідомлень/сек, що робить її безальтернативним вибором для високонавантажених систем національного масштабу. Встановлено, що використання TimescaleDB доцільне лише при навантаженнях до 15 000 подій/сек; при перевищенні цього порогу накладні витрати на механізми партиціонування (chunk management) стають критичними, знижуючи продуктивність нижче рівня "чистого" PostgreSQL.

Обмеження In-Memory систем. Виявлено архітектурне обмеження СУБД Redis при сценарії масового запису різномірних метрик. Незважаючи на роботу в оперативній пам'яті, однопоточна модель обробки подій (Single-threaded Event Loop) стає вузьким місцем при необхідності оновлення десятків тисяч ключів одночасно (15 833 повідомлень/сек). Таким чином, використання Redis як єдиного сховища для "сирих" історичних даних є неефективним як з точки зору швидкості запису, так і з точки зору вартості зберігання (RAM).

Оптимізація оперативного доступу. Для задач Real-time моніторингу (відображення геопозиції на карті, алерти диспетчера) Redis підтвердив статус найефективнішого рішення із латентністю менше 1 мс (0.98 мс). Жодна дискова СУБД (навіть з оптимізованим кешуванням) не здатна забезпечити аналогічний час відгуку.

Практичні рекомендації (Архітектурний патерн). На основі отриманих метрик запропоновано **гібридну Lambda-архітектуру** для логістичних платформ, яка розділяє потоки даних на два рівні:

Гарячий рівень (Speed Layer): Реалізується на базі **Redis TimeSeries**. Забезпечує миттєвий доступ до даних за останні 24 години. Це дозволяє нівелювати проблему повільного запису за рахунок меншого обсягу даних (TTL) та забезпечує інтерактивність UI.

Холодний рівень (Batch Layer): Реалізується на базі **TimescaleDB** (для систем середнього навантаження з потребою в SQL-аналітиці) або **InfluxDB** (для систем екстремального навантаження). Цей рівень відповідає за довгострокове зберігання, формування звітів та підготовку датасетів.

Перспективи досліджень. Подальший розвиток роботи вбачається у поєднанні запропонованої архітектури зберігання з методами штучного інтелекту та технологією блокчейн, що відповідає темі дисертаційного дослідження автора. Зокрема, планується дослідити:

1. Вплив стиснення даних у "холодному" сховищі на точність навчання нейромереж (Deep Learning) для задач прогнозування поломок транспорту.

2. Розробку механізму "якірної фіксації" (anchoring) хеш-сум часових рядів з TimescaleDB у приватний блокчейн (наприклад, Hyperledger Fabric) для створення незмінного аудиторського сліду перевезень без необхідності дублювання всього масиву даних у блокчейн.

Список використаної літератури

1. Gubbi J., Buyya R., Marusic S., Palaniswami M. Internet of Things (IoT): A vision, architectural elements, and future directions. *Future Generation Computer Systems*. 2013. Vol. 29, No. 7. P. 1645–1660. URL: <https://doi.org/10.1016/j.future.2013.01.010>

2. Ramaswamy L., Lawson B., Gogineni S. V. *Towards a quality-centric big data architecture for federated sensor services*. Proceedings of the 2013 IEEE International Congress on Big Data (BigData Congress). IEEE, 2013. P. 86–93. URL: <https://doi.org/10.1109/BigData.Congress.2013.21>
3. Zaslavsky A., Perera C., Georgakopoulos D. Sensing as a service and big data. *arXiv preprint*, arXiv:1301.0159, 2013. URL: <https://doi.org/10.48550/arXiv.1301.0159>
4. Vongsingthong S., Smachet S. A review of data management in Internet of Things. *KKU Research Journal*. 2015. Vol. 23. P. 215–240. URL: https://doi.nrct.go.th/ListDoi/listDetail?Resolve_Doi=10.14456/kkurj.2015.18
5. Ma M., Wang P., Chu C.-H. *Data management for Internet of Things: Challenges and opportunities*. Proceedings of the 2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things. IEEE, 2013. P. 1144–1151. URL: <https://doi.org/10.1109/GreenCom-iThings-CPSCoM.2013.199>
6. Zhu S. *Creating a NoSQL database for the Internet of Things: Key-value store on the Sensei-Uring platform* : Doctoral dissertation. Mid Sweden University, 2015. URL: <https://www.diva-portal.org/smash/get/diva2:841604/FULLTEXT01.pdf>
7. Bader A., Kopp O., Falkenthal M. Survey and comparison of open-source time series databases. *Datenbanksysteme für Business, Technologie und Web (BTW 2017) – Workshopband*. Bonn : Gesellschaft für Informatik, 2017. P. 249–268. URL: https://www.researchgate.net/publication/315838456_Survey_and_Comparison_of_Open_Source_Time_Series_Databases
8. Kiefer R. TimescaleDB vs. PostgreSQL for time-series: 20x higher inserts, 2000x faster deletes, 1.2x–14,000x faster queries. Timescale Blog, 2020. URL: <https://blog.timescale.com>
9. Fadhel M., Sekerinski E., Yao S. A comparison of time series databases for storing water quality data. *Auer M. E., Tsiatsos T. (eds.) Mobile Technologies and Applications for the Internet of Things (IMCL 2018)*. Cham : Springer, 2019. P. 302–313. URL: https://doi.org/10.1007/978-3-030-11434-3_33
10. Grzesik P., Mrozek D. Comparative analysis of time series databases in the context of edge computing for low power sensor networks. *Computational Science – ICCS 2020. Lecture Notes in Computer Science*, Vol. 12141. Cham : Springer, 2020. P. 311–324. URL: https://doi.org/10.1007/978-3-030-50426-7_28
11. Dias L. B., Holanda M., Huacarpuma R. C., de Sousa Jr. R. T. *NoSQL database performance tuning for IoT data: Cassandra case study*. Proceedings of the 3rd International Conference on Internet of Things, Big Data and Security (IoTBDs 2018). 2018. P. 277–284. URL: <https://www.scitepress.org/papers/2018/67827/67827.pdf>

References

1. Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of Things (IoT): A vision, architectural elements, and future directions. *Future Generation Computer Systems*, no. 29(7), pp. 1645–1660. URL: <https://doi.org/10.1016/j.future.2013.01.010>
2. Ramaswamy, L., Lawson, B., & Gogineni, S. V. (2013). *Towards a quality-centric big data architecture for federated sensor services*. In 2013 IEEE International Congress on Big Data (BigData Congress) (pp. 86–93). IEEE. URL: <https://doi.org/10.1109/BigData.Congress.2013.21>
3. Zaslavsky, A., Perera, C., & Georgakopoulos, D. (2013). Sensing as a service and big data. *arXiv preprint arXiv:1301.0159*. URL: <https://doi.org/10.48550/arXiv.1301.0159>
4. Vongsingthong, S., & Smachet, S. (2015). A review of data management in Internet of Things. *KKU Research Journal*, no. 23, pp. 215–240. URL: https://doi.nrct.go.th/ListDoi/listDetail?Resolve_Doi=10.14456/kkurj.2015.18
5. Ma, M., Wang, P., & Chu, C.-H. (2013). *Data management for Internet of Things: Challenges and opportunities*. In Proceedings of the 2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things (pp. 1144–1151). IEEE. URL: <https://doi.org/10.1109/GreenCom-iThings-CPSCoM.2013.199>
6. Zhu, S. (2015). *Creating a NoSQL database for the Internet of Things: Key-value store on the Sensei-Uring platform* (Doctoral dissertation, Mid Sweden University). URL: <https://www.diva-portal.org/smash/get/diva2:841604/FULLTEXT01.pdf>
7. Bader, A., Kopp, O., & Falkenthal, M. (2017). Survey and comparison of open-source time series databases. In *Datenbanksysteme für Business, Technologie und Web (BTW 2017) – Workshopband* (pp. 249–268). Bonn: Gesellschaft für Informatik. URL: https://www.researchgate.net/publication/315838456_Survey_and_Comparison_of_Open_Source_Time_Series_Databases
8. Kiefer, R. (2020). *TimescaleDB vs. PostgreSQL for time-series: 20x higher inserts, 2000x faster deletes, 1.2x–14,000x faster queries*. Timescale Blog. Retrieved from <https://blog.timescale.com>
9. Fadhel, M., Sekerinski, E., & Yao, S. (2019). A comparison of time series databases for storing water quality data. In M. E. Auer & T. Tsiatsos (Eds.), *Mobile Technologies and Applications for the Internet of Things (IMCL 2018)* (pp. 302–313). Cham: Springer. URL: https://doi.org/10.1007/978-3-030-11434-3_33

10. Grzesik, P., & Mrozek, D. (2020). Comparative analysis of time series databases in the context of edge computing for low power sensor networks. In *Computational Science – ICCS 2020* (Lecture Notes in Computer Science, no. 12141, pp. 311–324). Cham: Springer. URL: https://doi.org/10.1007/978-3-030-50426-7_28

11. Dias, L. B., Holanda, M., Huacarpuma, R. C., & de Sousa Jr, R. T. (2018). NoSQL database performance tuning for IoT data: Cassandra case study. In *Proceedings of the 3rd International Conference on Internet of Things, Big Data and Security (IoTBDS 2018)* (pp. 277–284). URL: <https://www.scitepress.org/papers/2018/67827/67827.pdf>

Дата першого надходження рукопису до видання: 21.11.2025

Дата прийнятого до друку рукопису після рецензування: 17.12.2025

Дата публікації: 31.12.2025

І. В. РУБАН

доктор технічних наук, професор,
професор кафедри електронних обчислювальних машин
Харківський національний університет радіоелектроніки
ORCID: 0000-0002-4738-3286

В. М. ТКАЧОВ

кандидат технічних наук, доцент,
докторант кафедри електронних обчислювальних машин
Харківський національний університет радіоелектроніки
ORCID: 0000-0002-6524-9937

МЕТОД РОБАСТНОЇ ВАРІАЦІЙНОЇ ОПТИМІЗАЦІЇ ТРАЄКТОРІЙ ФУНКЦІОНУВАННЯ ІНФОРМАЦІЙНОЇ СИСТЕМИ НА МОБІЛЬНІЙ ПЛАТФОРМІ В УМОВАХ ДЕСТРУКТИВНИХ ВПЛИВІВ

У роботі розглянуто проблему забезпечення живучості інформаційної системи на мобільній платформі в умовах деструктивних впливів і жорстких ресурсних обмежень. Функціонування системи описано у вигляді дискретно-часових траєкторій стану, для яких введено множину станів, що задовольняють вимоги живучості (допустимих за живучістю станів), та відповідну множину допустимих за живучістю початкових станів. На цій основі сформульовано варіаційну задачу оптимізації траєкторій з багатокритеріальним функціоналом, що враховує дефіцит якості обслуговування, витрати ресурсів, відхилення показників живучості та наближення до меж допустимої області. Для множини сценаріїв деструктивних впливів побудовано робастне узагальнення функціонала типу «найгіршого випадку» з використанням штрафного представлення обмежень живучості, що інтерпретується як вимога інваріантності множини допустимих за живучістю станів відносно найкритичніших збурень. Отримано варіаційні умови оптимальності та запропоновано ітераційний метод градієнтного типу для корекції послідовності керувальних впливів, орієнтований на сценарії з максимальною реалізованою вартістю робастного функціонала. Наведено імітаційний приклад для двовимірної моделі стану, який демонструє переваги робастно-варіаційної політики порівняно з базовою стратегією за показниками економії ресурсу, відтермінування виходу траєкторій за межі допустимої області та зменшення найгірших інтегральних витрат. Отримані результати розвивають відомі підходи теорії живучості та робастного оптимального керування на клас інформаційних систем, що функціонують на мобільних платформах під дією деструктивних впливів, і можуть бути покладені в основу подальших досліджень із синтезу адаптивних політик розподілу ресурсів.

Ключові слова: інформаційна система, мобільна платформа, живучість, траєкторія, оптимізація, робастність.

I. V. RUBAN

Doctor of Technical Sciences, Professor,
Professor at the Department of Electronic Computers
Kharkiv National University of Radio Electronics
ORCID: 0000-0002-4738-3286

V. M. TKACHOV

Candidate of Technical Sciences, Associated Professor,
Doctoral Student at the Department of Electronic Computers
Kharkiv National University of Radio Electronics
ORCID: 0000-0002-6524-9937

A METHOD OF ROBUST VARIATIONAL OPTIMIZATION OF OPERATIONAL TRAJECTORIES OF AN INFORMATION SYSTEM ON A MOBILE PLATFORM UNDER DESTRUCTIVE IMPACTS

The paper addresses the problem of ensuring the survivability of an information system on a mobile platform under destructive impacts and strict resource constraints. The system's operation is described in terms of discrete-time state trajectories, for which a set of states satisfying survivability requirements (survivability-admissible states) and

the corresponding set of survivability-admissible initial states are introduced. On this basis, a variational trajectory optimization problem is formulated with a multicriteria performance functional that accounts for the deficit of service quality, resource expenditures, deviations of survivability indicators, and proximity to the boundary of the admissible region. For a set of destructive impact scenarios, a robust “worst-case” extension of the functional is constructed using a penalty representation of survivability constraints, which is interpreted as a requirement of invariance of the set of survivability-admissible states with respect to the most critical perturbations. Variational optimality conditions are derived, and an iterative gradient-type method is proposed for correcting the sequence of control actions, oriented toward scenarios that realize the maximal value of the robust functional. A simulation example for a two-dimensional state model is presented, demonstrating the advantages of the robust variational policy over a baseline strategy in terms of resource savings, postponing the exit of trajectories beyond the admissible region, and reducing worst-case integral costs. The obtained results extend known approaches of survivability theory and robust optimal control to the class of information systems operating on mobile platforms under destructive impacts and can serve as a basis for further research on the synthesis of adaptive resource allocation policies.

Key words: information system, mobile platform, survivability, trajectory, optimization, robustness.

Постановка проблеми

Сучасні інформаційні системи можуть бути реалізовані як на базі стаціонарної інфраструктури (центри оброблення даних, корпоративні мережі, хмарні платформи), так і в розподіленому середовищі з віддаленими вузлами, а окремий різновид становлять інформаційні системи на мобільних платформах – системи, в яких ключові обчислювальні та комунікаційні ресурси розміщено на рухомих носіях (безпілотних апаратах, автономних роботах, транспортних засобах тощо) і які працюють в умовах змінного середовища та обмежених ресурсів [1]. Для таких систем принциповим є одночасне досягнення заданих показників якості в номінальному режимі функціонування та забезпечення такого керування, за якого, за наявності деструктивних впливів, траєкторія стану інформаційної системи протягом усього періоду функціонування не виходить за межі області припустимих станів, що відображає задані вимоги до живучості інформаційної системи. Під деструктивними впливами розуміються втрати зв'язності між елементами інформаційної системи, відмови вузлів, кібератаки та фізичні uszkodження окремих компонентів мобільної платформи, що впливають на виконання інформаційною системою основної функції [2-3].

У класичних підходах до забезпечення живучості інформаційних систем переважають схеми статичного резервування, локальної реконфігурації або грубих правил пріоритизації потоків і завдань [4]. Такі підходи, як правило, орієнтовані на окремі стани або події, а не на цілі траєкторії переходів станів. Внаслідок цього вони не дозволяють формально гарантувати, що за реалістичних сценаріїв збурень інформаційна система не вийде за межі допустимих режимів функціонування, навіть якщо у початковий момент часу всі обмеження дотримуються. У подальшому, в рамках цього дослідження, під збуреннями розуміються випадкові відхилення параметрів системи, деструктивні впливи, що погіршують режими її функціонування тощо.

Формально стан інформаційної системи на мобільній платформі можна описати вектором:

$$x(t) \in X, \quad (1)$$

де X – простір станів, що відображає конфігурацію даних, активних процесів і розподіл ресурсів інформаційної системи у момент часу t .

На $x(t)$ діють керувальні впливи (наприклад, параметри розподілу ресурсів, вибір маршрутів обміну, пріоритети завдань тощо):

$$u(t) \in U, \quad (2)$$

та збурення:

$$w(t) \in W. \quad (3)$$

Динаміка функціонування інформаційної системи задається у вигляді співвідношення

$$x(t+1) = F(x(t), u(t), w(t)), \quad (4)$$

або його безперервним аналогом, де F є функцією переходу стану, що відображає структурні властивості інформаційної системи.

Нехай задано множину допустимих станів інформаційної системи

$$G \subseteq X, \quad (5)$$

що задає обмеження на коректність даних, дотримання вимог якості виконання функційних задач, цілісність та ресурсну здійсненність (забезпеченість обчислювальними ресурсами) обраних режимів функціонування. У подальшому множина (5) інтерпретується як область станів, що відповідають вимогам до живучості інформаційної системи. Тоді траєкторію функціонування інформаційної системи задається як послідовність станів:

$$\{x(t)\}_{t=0}^T. \quad (6)$$

Вважається, що інформаційна система характеризується живучістю за обраної політики керування $u(t)$, якщо для будь-якого допустимого сценарію збурень $w(t)$ породжена нею траєкторія стану задовольняє умову:

$$x(t) \in G \text{ для всіх } t \in [0, T] \quad (7)$$

В реальних умовах інформаційні системи на мобільній платформі мають забезпечувати виконання основної

функції не лише в умовно середньому випадку, що відповідає номінальній траєкторії (6), а й витримувати широкий спектр несприятливих сценаріїв, включно з рідкісними, але критичними комбінаціями збурень [5]. Це вимагає робастності функціонування інформаційної системи, під якою в цьому дослідженні вважається здатність системи за фіксованої політики керування зберігати перебування траєкторії стану в множині допустимих станів для всіх деструктивних впливів із заданої множини W_{Φ} , а у стохастичному випадку, коли $w(t)$ моделюється випадковим процесом, – з наперед заданою ймовірністю.

Традиційні методи оптимального керування та розподілу ресурсів здебільшого орієнтовані на мінімізацію певного інтегрального показника для номінальної або усередненої моделі, не враховуючи вимогу неперервного перебування траєкторії в множині (5) протягом функціонування системи та за множини сценаріїв деструктивних впливів [6]. Відомі підходи до забезпечення живучості інформаційної системи, які враховують семантику траєкторій функціонування інформаційних систем [7] у більшості випадків або не пристосовані до умов мобільної платформи з жорсткими ресурсними обмеженнями, або не забезпечують явної процедури синтезу керування, яку можна реалізувати в оперативному режимі.

Наведений аналіз вказує на формалізацію задачі, спрямованої на забезпечення живучості реальних інформаційних системах, – як багатокритеріальної. Тобто, необхідно одночасно підтримувати прийнятну якість виконання основної та периферійних функцій інформаційною системою, обмежувати ризики критичних відмов, дотримуватися системних ресурсних лімітів, забезпечувати пріоритет для критичних функцій без повної деградації допоміжних функцій.

Просте згортання зазначених вимог в один скалярний критерій або жорстке фіксування пріоритетів може приводити до надмірно консервативних рішень або, навпаки, до втрати властивості живучості інформаційною системою за нетипових, але допустимих сценаріїв деструктивних впливів.

Таким чином, для інформаційних систем на мобільній платформі залишається невирішеною задача побудови методу робастної варіаційної оптимізації траєкторій функціонування, який би:

- забезпечував виконання вимог до живучості інформаційної системи через умову перебування траєкторій стану в множині допустимих станів (5);
- формалізував задачу як варіаційну оптимізаційну задачу на класі допустимих траєкторій та керувальних впливів з урахуванням множини можливих деструктивних впливів;
- гарантував робастність функціонування системи щодо деструктивних впливів шляхом явного врахування найгірших або ризикованих сценаріїв;
- давав змогу враховувати багатокритеріальний характер вимог до живучості, не зводячи це до неконтрольованого набору евристичних правил.

Відсутність такого методу ускладнює побудову формально обґрунтованих алгоритмів оперативного керування ресурсами інформаційних систем на мобільній платформі, здатних гарантувати дотримання вимог до їх живучості в широкому діапазоні деструктивних впливів.

Аналіз останніх досліджень і публікацій

Проблема забезпечення живучості інформаційних систем у динамічних та агресивних середовищах активно досліджується у поданні мережних структур, в рамках методів робастного керування та оптимізації траєкторій. Водночас більшість робіт або зосереджуються на структурно-топологічних аспектах мереж, не доводячи апарат до узгодженого методологічного рівня, де траєкторія функціонування інформаційної системи на мобільній платформі розглядається як об'єкт робастної варіаційної оптимізації в межах множини допустимих (життєздатних) станів.

Дослідження [8], присвячене життєвому циклу розроблення мобільних телекомунікаційних систем, показує як можна інтегрувати вимоги до забезпечення живучості в усі фази життєвого циклу програмно-апаратних рішень. В роботі [8] описано процедури включення механізмів відмовостійкості, самовідновлення та безпеки ще на етапах аналізу вимог, проєктування та тестування. Запропонована методика має виразний процесно-орієнтований характер і дає змогу системно враховувати живучість як нефункціональну вимогу. Однак відсутність механізму оперативного параметричного керування ресурсами мобільної платформи суттєво звужує предметну область застосовності запропонованого рішення.

Суттєвий внесок у розвитку кількісних підходів до оцінювання живучості зроблено в роботі [9], де досліджується динамічні властивості двошарових мереж із різними топологіями. Запропонована модель дозволяє аналізувати вплив схем міжшарових зв'язків, зокрема показано, що внутрішньошарові та міжшарові зв'язки по-різному впливають на здатність мережі протистояти випадковим та цілеспрямованим деструктивним впливам.

В дослідженні [10] показано, що живучість – динамічна характеристика, яка залежить від механізму еволюції стану мережної структури. Водночас аналіз обмежується агрегованими мережними метриками і не переходить до рівня траєкторій функціонування конкретних інформаційних сервісів на мобільній платформі.

У дослідженні за напрямом робастної оптимізації траєкторій [11] розглянуто постановку задачі оптимального керування системою з урахуванням стохастичних комплементарних обмежень. У роботі [11] авторами

запропоновано алгоритми побудови траєкторій, стійких до невизначеності параметрів середовища, з використанням стохастичних моделей. Робота [11] присвячена дослідженню робастної траєкторної оптимізації системи, проте вона орієнтована на кінематику та динаміку руху мобільної платформи, а не на еволюцію інформаційних станів системи та розглядає переважно однокритеріальну постановку оптимізації (мінімізація ризику несприятливих сценаріїв) без інтеграції багатьох показників живучості.

Дотичним є дослідженням в області контрольних бар'єрних функцій [12]. Зокрема, у дослідженні [12], запропоновано апарат забезпечення інваріантності безпечної множини станів за умов невизначеності вимірювань, тобто, пропонуються підходи до конструювання бар'єрних функцій, стійких до похибок сенсорики, та синтезу моделей управління, що мінімально втручається у базову поведінку інформаційної системи, але гарантує залишення траєкторій у безпечній множині траєкторій функціонування. Отже, тут явно присутні множини допустимих станів та рішення задачі робастної інваріантності траєкторій. Проте це дослідження не враховує багатокритеріальної природи забезпечення живучості інформаційних систем.

Отже, задача, пов'язана зі створенням такого методу оптимізації траєкторій функціонування інформаційної системи на мобільній платформі, який би поєднував рішення з формування множини допустимих станів інформаційної системи на мобільній платформі, давав формалізований деструктивних впливів як різновиду збурень, що деформують траєкторії в просторі станів, надавав багатокритеріальну постановку оптимізації та гарантував утримання траєкторій всередині множини допустимих станів при найгірших допустимих сценаріях впливів.

Формулювання мети дослідження

Метою дослідження є розроблення методу робастної варіаційної оптимізації траєкторій функціонування інформаційної системи на мобільній платформі, який за наявності деструктивних впливів та обмежених ресурсів забезпечує перебування траєкторії стану системи в множині допустимих станів. Для досягнення цієї мети передбачається формально задати множину допустимих траєкторій стану з урахуванням показників живучості інформаційної системи та ресурсних обмежень, побудувати варіаційну постановку задачі оптимізації траєкторій у присутності збурень, ввести робастні критерії, що відображають найгірші допустимі сценарії впливів, вивести умови оптимальності керування інформаційною системою, які гарантують перебування траєкторій стану в області допустимих станів, а також сформулювати алгоритмічну процедуру застосування запропонованого методу для підтримки прийняття рішень щодо оперативного розподілу ресурсів інформаційної системи в рамках виконання нею основної функції.

Викладення основного матеріалу дослідження

Математична модель траєкторій функціонування інформаційної системи на мобільній платформі

У якості опорної моделі інформаційної системи на мобільній платформі використовується подання, описане у [13], [14] і розглядається як вихідний базис, з якого виділяється узагальнений динамічний стан, придатний для варіаційної оптимізації.

Для опису функціонування інформаційної системи на мобільній платформі буде використовуватися дискретно-часова модель, що відповідає покроковій роботі засобів керування та оновлення вимірювань метрик. Нехай $t \in \{0, 1, \dots, T\}$ – одиниця дискретного часу (номер такту), де T – тривалість періоду функціонування системи.

Беручи за основу формалізацію стану інформаційної системи на мобільній платформі (1), простір станів X включає компоненти, що відображають стан критичних інформаційних процесів (показники ступеня завершеності виконання завдань, черг запитів, затримок передачі та втрат даних), характеристики якісних властивостей даних (валідність даних відносно часового вікна їх надходження, узгодженість даних на етапах формування задач їх обробки, ступінь втрат або пошкоджень даних унаслідок різних причин) та показники ресурсного стану мобільної платформи (бюджет запасу функційної здатності, завантаженість каналів зв'язку, обчислювальних елементів, буферної пам'яті тощо).

Виходячи з (1) та для зручності подальшого аналізу стан $x(t)$ можна подати у вигляді тріади

$$x(t) = (x'(t), x''(t), x'''(t)), \quad (8)$$

де $x'(t)$ – процесна складова (функціонування сервісів); $x''(t)$ – складова, що відноситься до властивостей даних; $x'''(t)$ – ресурсна складова мобільної платформи.

Управління ресурсами та режимами роботи системи задається вектором (2). Компоненти $u(t)$ можуть включати: розподіл пропускну здатності каналів між інформаційними потоками, розподіл задач на обчислювальних елементах (вузли, групи вузлів), вибір маршрутів передачі даних, зміну режимів роботи окремих компонентів інформаційної системи.

Для кожного стану $x(t)$ визначається локальна множина допустимих керувальних впливів

$$U(x(t)) \subseteq U, \quad (9)$$

що враховує поточні обмеження на ресурси.

Розглянемо формалізацію вектору збурень (3), що вкливають на інформаційну систему в момент часу t . До

$w(t)$ відносяться, наприклад: динаміка топології та якості каналів зв'язку, відмова вузлів та каналів, перевантаження мережних і обчислювальних ресурсів, навмисні деструктивні впливи (атаки, блокування, спотворення трафіку тощо).

Множина W може розглядатися як сукупність усіх можливих значень збурень, тоді як у рамках цього дослідження буде використовуватися допустима множина сценаріїв $W_{\mathbb{R}}$, відносно якої формулюється фактор робастності.

Еволюція стану інформаційної системи в дискретному часі описується оператором (4), де:

$$F : X \times U \times W \rightarrow X. \quad (10)$$

Формула (10) – відображення, що задає перехід інформаційної системи з поточного стану $x(t)$ у наступний стан $x(t+1)$ під дією $u(t)$ та $w(t)$. Оператор F може мати такі властивості як:

- нелінійність і залежність від попередніх станів;
- параметрична залежність від характеристик мобільної платформи (8);
- врахування ефекту накопичення ушкоджень (поступове виснаження ресурсів, деградація якості даних через спотворення).

У рамках даного дослідження припускається, що структура оператора F фіксована для обраного класу інформаційних систем, а його параметри можуть оцінюватися на основі моделей або емпіричних даних [15]. Це дозволяє формально оперувати траєкторіями стану, не вдаючись до надмірного ускладнення внутрішньої структури інформаційної системи.

Отже, за заданого початкового стану $x(0) = x_0 \in X$, послідовності керувальних впливів множини (9) $u(\cdot) = \{u(t)\}_{t=0}^{T-1}$ та збурень $w(\cdot) = \{w(t)\}_{t=0}^{T-1}$ динамічна модель породжує траєкторію функціонування інформаційної системи (6), яка є рішенням рекурентного співвідношення (4). Надалі саме траєкторії $x(\cdot)$ будуть основним об'єктом оптимізації та аналізу живучості інформаційної системи, де нотація $(\cdot)$ означає, що йдеться про залежність від усієї часової послідовності станів, а не від фіксованого моменту часу.

У реальних умовах мобільна платформа не завжди має доступ до повного стану $x(t)$. Тому доцільно ввести також модель спостережень [16]:

$$y(t) = H(x(t), w(t), v(t)), \quad (11)$$

де: $y(t)$ – вектор доступних телеметричних показників; H – оператор спостереження; $v(t)$ – вектор вимірювальних похибок, що моделює випадкові спотворення даних телеметрії.

У загальному випадку $u(t)$ формується як функція

$$u(t) = \pi_t(y(0), \dots, y(t)), \quad (12)$$

де π_t – правило (політика) керування на кроці t .

Далі, на етапі формування варіаційної задачі оптимізації, припускається, що клас допустимих політик $\{\pi_t\}$ є статичним, тоді як предметом дослідження стають саме властивості породжених ними траєкторій $x(\cdot)$ відносно множини допустимих станів. А з урахуванням (11) та (12) інформаційна система має повне уявлення про стан мобільної платформи та керувальні впливи на всі її елементи.

Формалізація допустимих станів та клас відповідних траєкторій стану

Далі розглядається опис динаміки стану інформаційної системи на мобільній платформі та виокремлення тих режимів, у яких система реалізує функціонал своєї властивості живучості. Для цього береться на базі множини допустимих станів функціонування системи клас допустимих траєкторій (5). Множина (5) містить ті стани, в яких інформаційна система на мобільній платформі здатна забезпечити реалізацію своєї цільової функції з прийнятною якістю за наявних ресурсів та обмежень середовища. Тобто, перебування системи в області G означає, що основні інформаційні сервіси доступні та функціонують із припустимими затримками й відомою імовірністю відмов. При цьому дані, які використовуються для прийняття рішень, мають достатню актуальність, цілісність і узгодженість, а сумарне навантаження на ресурси системи не перевищує доступних бюджетів.

Введемо для стану (1) вектор показників:

$$q(t) = (q_1(x), \dots, q_k(x)), \quad (13)$$

де окремі елементи $q_i(x)$ описують, наприклад, показники якості обслуговування, даних, завантаженості ресурсів тощо.

Нехай для кожного показника задано допустимий діапазон значень у вигляді обмежень

$$q_i(x) \in Q_i^{\mathbb{R}}, \quad i = 1, \dots, k. \quad (14)$$

Тоді множину (5) можна подати як:

$$G = \{x \in X \mid q_i(x) \in Q_i^\Phi\} \text{ для всіх } i = 1, \dots, k. \quad (15)$$

За потреби, до визначення G можуть додатково застосовуватися логічні або структурні умови (наприклад, вимоги до функціонування окремих компонентів інформаційної системи, обмеження на конфігурацію маршрутів (14) тощо). Важливо, що множина G описує саме робочі режими, в яких інформаційна система здатна виконувати свою цільову функцію. Динаміка стану підпорядковується формалізму, введеному в (4) при $t = 0, 1, \dots, T-1$.

Отже, формулюється **перше означення**. Траєкторія стану $x(\cdot) = \{x(t)\}_{t=0}^T$ називається допустимою відносно множини допустимих станів G , якщо $x(t) \in G$ для всіх $t = 0, 1, \dots, T$ (7). Таким чином, допустимою вважається лише та траєкторія стану інформаційної системи, яка не виходить за межі області допустимих станів протягом усього інтервалу функціонування інформаційної системи.

За наявності деструктивних впливів важливо враховувати не тільки одну реалізацію збурень, а сукупність можливих сценаріїв, що відповідають умовам експлуатації мобільної платформи. Нехай множина всіх допустимих реалізацій збурень на інтервалі $[0, T]$:

$$W' = \{w(\cdot)\} = \{\{x(t)\}_{t=0}^{T-1}\} \quad (16)$$

Далі вводиться **друге означення**. Для заданої політики керування $\pi = \{\pi_t\}_{t=0}^{T-1}$, де $u(t) = \pi_t$ (доступна інформація на момент t), траєкторія $x(\cdot)$, породжена динамікою системи для конкретного сценарію $w(\cdot) \in W'$, називається життєздатною, якщо вона є допустимою відносно G , тобто виконується рівняння (7).

В умовах деструктивних впливів до розгляду беруться початкові стани та політики керування, для яких життєздатність траєкторій забезпечується не лише для одного, а для всіх сценаріїв збурень з певної підмножини $W'_\Phi \subseteq W'$. Це породжує поняття множини початкових станів, в яких інформаційна система завжди функціонує в нормальних робочих режимах.

Далі вводиться **третє означення**:

$$X_\Phi = \{x_0 \in X \mid \exists \pi \text{ така, що для всіх } w(\cdot) \in W'_\Phi \text{ траєкторія } x(\cdot) \text{ задовольняє } x(t) \in G \forall t\} \quad (17)$$

називається множиною допустимих початкових станів відносно допустимих станів G та множини допустимих збурень.

X_Φ (17) описує ті початкові умови, з яких існує деяка політика керування, здатна, незважаючи на деструктивні впливи з W'_Φ , утримувати систему всередині області G . Надалі саме W'_Φ та структура множини G визначатимуть, наскільки жорсткими є вимоги до живучості.

У рамках цього дослідження використовується множина допустимих станів G , що не протирічить викладеним положенням у [17] та логіці формування множини допустимих початкових станів X_Φ як вихідних об'єктів для варіаційної постановки задачі оптимізації траєкторій.

Варіаційна постановка задачі оптимізації траєкторій

З урахуванням (4), (5), (16) формулюється задача керування траєкторіями функціонування інформаційної системи на мобільній платформі як варіаційна задача оптимізації на класі траєкторій.

Розглядається фіксований початковий стан

$$x(0) = x_0 \in X_\Phi, \quad (18)$$

тобто такий, з якого, принаймні теоретично, формується політика керування, здатна утримувати інформаційну систему у множині допустимих станів G на заданому інтервалі $[0, T]$.

Нехай Π – клас політик управління. Вважається, що кожна політика $\pi \in \Pi$ задає послідовність керувальних впливів:

$$u(t) = \pi_t(I_t), \quad t = 0, 1, \dots, T-1, \quad (19)$$

де I_t – інформаційний стан системи на момент часу t , тобто вся інформація, доступна засобом керувального впливу на цей момент.

Для фіксованих x_0 (18), π (19) та сценарію збурень $w(\cdot) \in W'_\Phi$ (16) формується керована траєкторія:

$$x(\cdot; x_0, \pi, w(\cdot)) = \{x(t)\}_{t=0}^T, \quad (20)$$

яка є розв'язком системи рівнянь

$$\begin{cases} x(0) = x_0, \\ x(t+1) = F(x(t), \pi_t(I_t), w(t)), \quad t = 0, \dots, T-1. \end{cases} \quad (21)$$

Позначимо через X_T множину всіх траєкторій довжини $(T+1)$, які можуть бути реалізовані за допомогою політик класу Π та збурень з множини W'_0 .

Далі формалізуються обмеження на траєкторії стану, які гарантують перебування інформаційної системи в множині допустимих станів, що відповідає заданим вимогам до її живучості. Згідно з (5), допустимими є ті траєкторії, які не виходять за межі множини допустимих станів G . Тому на рівні варіаційної постановки (20) та (21) накладається обмеження

$$x(t; x_0, \pi, w(\cdot)) \in G \text{ для всіх } t = 0, 1, \dots, T. \quad (22)$$

З точки зору траєкторій це означає, що допустимими є лише ті пари «політика π – сценарій $w(\cdot)$ », які породжують траєкторії в межах області G . В подальшому необхідно здійснювати пошук таких політик, для яких ця вимога виконується для всіх сценаріїв з W'_0 .

Щоб сформулювати оптимізаційну задачу, необхідно ввести функціонал якості траєкторії, який відображає бажані властивості функціонування інформаційної системи та витрати її ресурсів (або мобільної платформи, на якій вона функціонує). Для цього пропонується функціонал вигляду:

$$J(x(\cdot), u(\cdot)) = \sum_{t=0}^{T-1} L(x(t), u(t)) + \Phi(x(T)), \quad (23)$$

де: $L(x(t), u(t))$ – локальна (покрокова) функція витрат, що формалізує миттєві витрати ресурсів або штрафи за відхилення траєкторії керованої системи від бажаних режимів на кроці t ; $\Phi(x(T))$ – термінальна функція витрат, яка задає узагальнену оцінку кінцевого стану системи в момент часу T виходячи з її показників живучості (наприклад, залишковий запас ресурсів, ступінь виконання критичних завдань, рівень накопичених ушкоджень тощо).

Функція L у (23) може включати кілька узагальнених компонент, пов'язаних із живучістю інформаційної системи – штрафи за наближення до межі множини G , штрафи за зниження якості обслуговування, штрафи за інтенсивне, протяжне у часі, використання ресурсів, штрафи за активацію аварійних режимів, які хоч і підтримують функціонал живучості, але є небажаними у початкових умовах функціонування інформаційної системи.

Таким чином, для фіксованих x_0 , π , $w(\cdot)$ функціонал (23) набуває значення

$$J(x((\cdot); x_0, \pi, w(\cdot)), u((\cdot); \pi)), \quad (24)$$

де $u((\cdot); \pi) = \{u(t)\}_{t=0}^{T-1}$ – послідовність керувальних впливів, породжена політикою π .

Реальна задача забезпечення живучості інформаційних систем на мобільних платформах є багатокритеріальною [3]. Необхідно одночасно: виконувати основну функцію, підтримувати прийнятну якість сервісного рівня, обмежувати витрати ресурсів, у тому числі резервних, щоб не вичерпати їх завчасно, мінімізувати ціну використання аварійних режимів як критерій швидкості виснаження ресурсів та зменшення запасу функційної здатності інформаційної системи.

Ці вимоги можна представити у вигляді векторного функціонала

$$\bar{J}(x(\cdot), u(\cdot)) = (J_1(x(\cdot), u(\cdot)), \dots, J_r(x(\cdot), u(\cdot))), \quad (25)$$

де кожен компонент J відповідає окремому показнику (наприклад, J_1 – показник витрати ресурсу, J_2 – показник втрати якості тощо).

Для переходу від (25) до скалярної варіаційної постановки з урахуванням (23) може використовуватись зважене згоргання:

$$J(x(\cdot), u(\cdot)) = \sum_{j=0}^r \alpha_j J_j(x(\cdot), u(\cdot)), \quad (26)$$

де $\alpha_j \geq 0$ – вагові коефіцієнти, що відображають відносний пріоритет (значущість) J у загальному функціоналі.

Відповідно, далі у дослідженні, вважається, що обрано певну узгоджену скалярну форму функціонала J (26), яка відображає пріоритети живучості та ефективності для конкретного класу систем.

З урахуванням викладеного варіаційна постановка задачі оптимізації траєкторій функціонування інформаційної системи на мобільній платформі (**задача 1**) для фіксованого сценарію збурень $w(\cdot)$ формулюється наступним чином. Для заданих початкового стану x_0 , множини станів G , класу політик Π та сценарію збурень $w(\cdot)$ необхідно знайти таку політику $\pi^* \in \Pi$, що мінімізує функціонал (24) за умов (19) та (22).

Тобто варіаційна задача полягає в пошуку такої траєкторії $x(\cdot)$ (через вибір політики π та відповідних управляючих дій), яка, з одного боку, залишається в області G , а з іншого – мінімізує обраний функціонал витрат J .

Оскільки ж реальні деструктивні впливи не є фіксованими, а змінюються в межах множини W'_0 , то номінальна варіаційна задача є лише початковим кроком. Далі доцільно сформулювати її робастне узагальнення, в якому

оптимізація відбувається з урахуванням найгірших допустимих сценаріїв збурень. Відповідно, траєкторії стану, що зберігають перебування системи в множині допустимих станів G за деструктивних впливів інтерпретуються як такі, що відповідають вимогам живучості інформаційної системи.

Таким чином, у номінальній постановці збурення $w(\cdot)$ вважається фіксованим сценарієм, що дає змогу сформулювати задачу оптимізації траєкторій за наперед відомої зміни станів інформаційної системи. Водночас у реальних умовах така постановка ігнорує невизначеність та варіативність деструктивних впливів і, як наслідок, призводить до надто оптимістичної оцінки можливостей системи та заниження ризиків порушення допустимих режимів. Ба більше – деструктивні впливи мають невизначений характер, можуть змінюватися за невідомих правил і, частково, бути навмисно спрямованими на погіршення роботи інформаційної системи. Тому для забезпечення живучості необхідно перейти до робастної варіаційної постановки, у якій оптимізація виконується з урахуванням множини можливих сценаріїв збурень.

Із множини (16) виділяється допустима множина сценаріїв деструктивних впливів $W'_\downarrow \subseteq W'$, яка описує ті варіації середовища, що відповідають умовам експлуатації інформаційної системи на мобільній платформі. Так, $W'_\downarrow$ включає сценарії випадкових збурень із заданими статистичними характеристиками (наприклад, обмежена імовірність тривалих відмов каналів), сценарії цілеспрямованих впливів, обмежених ресурсами порушника, комбіновані сценарії, де випадкові та навмисні компоненти поєднуються (найбільш складні випадки).

Подальший аналіз не вимагає фіксації конкретної структури $W'_\downarrow$, бо важливо, що саме ця множина задає область невизначеності, відносно якої необхідно забезпечити робастність траєкторій.

Виходячи із постановки номінального сценарію збурень $w(\cdot)$ (24) та враховуючи те, що реальний сценарій наперед невідомий, логічно оцінювати не одиничне значення J , а узагальнену характеристику якості на множині сценаріїв $W'_\downarrow$. Для цього вводиться робастний функціонал:

$$\mathfrak{Z}(x_0, \pi) = \mathfrak{R}\left(J\left(x((\cdot); x_0, \pi, w(\cdot)), u((\cdot); \pi)\right); w(\cdot) \in W'_\downarrow\right), \quad (27)$$

де $\mathfrak{R}(\cdot)$ – оператор агрегування значень номінального функціонала для різних сценаріїв.

Можливі варіанти для $\mathfrak{R}(\cdot)$:

– мінімаксна постановка (найгірший сценарій):

$$\mathfrak{R}_{\max} = \sup_{w(\cdot) \in W'_\downarrow} J\left(x((\cdot); x_0, \pi, w(\cdot)), u((\cdot); \pi)\right), \quad (28)$$

що відповідає спрямованості мінімізувати максимально можливі витрати при найгіршому сценарії;

– ризик-орієнтована постановка. Якщо для $W'_\downarrow$ задано імовірнісний опис, можна використовувати ризиковий функціонал $\rho(\cdot)$ наприклад, орієнтований на важкі хвости розподілу витрат [18]:

$$\mathfrak{R}_\rho = \rho\left(J\left(x((\cdot); x_0, \pi, w(\cdot)), u((\cdot); \pi)\right)\right). \quad (29)$$

У (29) ρ відображає підвищену чутливість до малоімовірних, але критичних сценаріїв (наприклад, типу CVaR, якщо інформаційна система критично залежить від рідкісних, але дуже шкідливих впливів (наприклад, втрата до 75% функціоналу)).

У рамках цього дослідження розглядається узагальнена форма робастного функціонала $\mathfrak{Z}(x_0, \pi)$ (27), без фіксації конкретного виду $\mathfrak{R}$, що не обмежує застосовність методу, який пропонується, до різних класів задач із різним ступенем консервативності робастної оцінки.

У присутності множини сценаріїв $W'_\downarrow$ вимога щодо допустимості траєкторій посилюється. Розглядаються траєкторії стану інформаційної системи, які утворюються з початкового стану x_0 при застосуванні політики керування π за всіх допустимих сценаріїв збурень $w(\cdot)$ (22).

Тоді маємо **четверте означення**. Політика π називається робастно допустимою відносно множини допустимих станів G та множини сценаріїв збурень $W'_\downarrow$, якщо для будь-якого початкового стану $x_0 \in X_\varphi$ та для всіх сценаріїв $w(\cdot) \in W'_\downarrow$ породжена траєкторія задовольняє вимогам (22).

Тобто робастна допустимість означає, що політика π гарантує перебування системи в області G незалежно від того, який саме сценарій збурень із допустимої множини реалізується. З точки зору траєкторій це можна розглядати як інваріантність множини G відносно динаміки функціонування інформаційної системи та обраної політики керування при всіх $W'_\downarrow$.

Об'єднавши введені поняття, можна сформулювати робастне узагальнення варіаційної задачі оптимізації траєкторій (**задача 2**). Для заданих $x_0 \in X_\varphi$, $G \subseteq X$, $W'_\downarrow$, Π , $\mathfrak{Z}(x_0, \pi)$ необхідно знайти політику $\pi^* \in \Pi$, яка

$$\min_{\pi \in \Pi} \mathfrak{Z}(x_0, \pi) \quad (30)$$

за умов (19) та (22), де $x((\cdot); x_0, \pi, w(\cdot))$ формується згідно з (4).

Фактично задача 2 (30) описує пошук такого правила керування ресурсами та режимами роботи інформаційної системи, яке забезпечує інваріантність області G для всіх сценаріїв збурень із W'_Φ та одночасно мінімізує узагальнений робастний функціонал $\mathfrak{J}$, що відображає компроміс між витратами ресурсів, якістю обслуговування та запасом ресурсів для забезпечення живучості інформаційної системи.

З формальної точки зору робастна варіаційна задача є значно складнішою за номінальну, адже множина траєкторій, допустимих за доступністю, звужується вимогою інваріантності G для всіх сценаріїв, а функціонал $\mathfrak{J}$ набуває більш складної структури через мінімаксні або ризик-орієнтовані оператори.

Умови оптимальності та побудова методу робастної варіаційної оптимізації

Щоб перейти від формальної робастної варіаційної постановки (задача 2) до конструктивного методу, необхідно зробити припущення, які забезпечують коректність задачі та забезпечують існування оптимальних рішень, а також – отримати умови, що визначають допустимі варіаційні напрями зміни траєкторій, які приводять до зменшення робастного функціонала за одночасного виконання обмежень на перебування траєкторій у множині допустимих станів.

Нехай на інтервалі $[0, T]$ множини X , $U(x)$, W характеризуються обмеженістю та замкнутістю, а функції $F(x, u, w)$, $L(x, u)$ та $\Phi(x)$ є неперервними за своїми аргументами. Тоді (27) буде зафіксовано у формі (28) для скінченної множини сценаріїв $W'_\Phi \subset W'_\Phi$:

$$\mathfrak{J}(x_0, \pi) = \max_{w(\cdot) \in W'_\Phi} J(x(\cdot; x_0, \pi, w(\cdot)), u(\cdot; \pi)). \quad (31)$$

У (31) W'_Φ інтерпретується як репрезентативна підмножина допустимих сценаріїв деструктивних впливів, яка задається за результатами попереднього аналізу умов експлуатації інформаційної системи (за наявності таких).

Щоб забезпечити варіаційний характер методу, доцільно перейти від жорсткого обмеження (7) до штрафного представлення, у якому вихід траєкторії стану за межі множини допустимих станів G призводить до зростання додаткового штрафного члена у функціоналі. Для цього пропонується ввести наступну функцію штрафу:

$$\Psi: X \rightarrow [0, +\infty), \quad (32)$$

яка задовольняє наступним умовам:

- умова 1: $\Psi(x) = 0$, якщо виконується (7);
- умова 2: $\Psi(x) > 0$, якщо не виконується (7);
- умова 3: $\Psi(x)$ зростає зі збільшенням відхилення стану x від G .

Тоді, для кожного сценарію $w(\cdot) \in W'_\Phi$ необхідно модифікувати номінальний функціонал:

$$J_\lambda(x(\cdot), u(\cdot)) = \sum_{t=0}^{T-1} [L(x(t), u(t)) + \lambda \Psi(x(t))] + \Phi(x(T)), \quad (33)$$

У (33) параметр λ є штрафним коефіцієнтом. У загальному випадку λ може набувати довільних невід'ємних значень, тобто $\lambda \in (0, \infty)$. Малі значення λ відповідають режиму, у якому штрафна складова $\Psi(x(t))$ (32) має відносно невеликий вплив на функціонал і пріоритизується мінімізація номінальних витрат $L(x(t), u(t))$. За великих значень λ домінує внесок штрафу за порушення вихід траєкторії стану за межі множини допустимих станів G , внаслідок чого оптимальні траєкторії уникають наближення до меж G навіть ціною зростання інших витрат. Конкретне значення λ не є константою інформаційної системи, а визначається на етапі синтезу політики керування, зокрема, у його обирають так, щоб у оптимальних траєкторіях виконувалися задані обмеження на ймовірність виходу за множину G або щоб внесок штрафної складової займав наперед визначену частку сумарного функціонала [19].

Робастний штрафний функціонал набуває вигляду

$$\mathfrak{J}_\lambda(x_0, \pi) = \max_{w(\cdot) \in W'_\Phi} J_\lambda(x(\cdot; x_0, \pi, w(\cdot)), u(\cdot; \pi)). \quad (34)$$

За достатньо великих значень λ мінімізація $\mathfrak{J}_\lambda$ призводить до того, що оптимальні траєкторії залишаються в G , оскільки будь-який вихід за межі G робить значення функціонала істотно більшим порівняно з траєкторіями, які повністю лежать у G .

Далі вводяться варіаційні умови оптимальності для множини сценаріїв функціонування інформаційної системи на мобільній платформі. Розглянемо фіксовану політику керування π та відповідну послідовність керувальних впливів

$$u(t) = u_\pi(t), \quad t = 0, \dots, T-1. \quad (35)$$

Для кожного сценарію $w(\cdot) \in W'_\Phi$ політика (35) формує траєкторію

$$x(t; w) = x(t; x_0, \pi, w(\cdot)), \quad w(\cdot): t \mapsto w(t). \quad (36)$$

Надалі значення робастного функціонала на траєкторіях (36), можна подати як:

$$J_{\lambda}(\pi, w) := J_{\lambda}(x(\cdot; w), u_{\pi}(\cdot)). \quad (37)$$

Тоді формула (34) набуває вигляду

$$\mathfrak{J}_{\lambda}(x_0, \pi) = \max_{w(\cdot) \in W'_{\pi}} J_{\lambda}(\pi, w). \quad (38)$$

Вводиться позначення через $W'_{\max}(\pi)$ множини сценаріїв, на яких досягається максимальне значення робастного функціоналу:

$$W'_{\max}(\pi) = \{w(\cdot) \in W'_{\pi} \mid J_{\lambda}(\pi, w) = \mathfrak{J}_{\lambda}(x_0, \pi)\}. \quad (39)$$

Множину $W'_{\max}(\pi)$ можна інтерпретувати як множину найгірших сценаріїв збурень для фіксованої політики керування π , тобто таких $w(\cdot)$, за яких робастний штрафний функціонал набуває максимального значення.

У даному дослідженні використовується стандартна параметрична варіаційна схема, яка широко застосовується в варіаційному численні та теорії оптимального управління, наприклад, у доказах принципу максимуму Понтрягіна [20]. Для цього розглядається параметричне варіаційне збурення послідовності керувальних впливів

$$u_{\varepsilon}(t) = u_{\pi}(t) + \varepsilon \delta u(t), \quad \varepsilon \rightarrow 0. \quad (40)$$

У (40) $\delta u(t)$ належить множині допустимих варіацій керувальних впливів і забезпечує збереження допустимості $u_{\varepsilon}(t)$ для всіх достатньо малих значень параметра ε .

Внаслідок варіації керування змінюються траєкторії (36), а, отже, – значення (37) для кожного сценарію. Якщо припустити диференційовність F , L , Φ , Ψ , за відповідними аргументами, то можна ввести похідну робастного функціонала за напрямом $\delta u(\cdot)$ для політики π :

$$D\mathfrak{J}_{\lambda}(x_0, \pi)[\delta u(\cdot)] = \lim_{\varepsilon \rightarrow 0^+} \frac{\mathfrak{J}_{\lambda}(x_0, \pi_{\varepsilon}) - \mathfrak{J}_{\lambda}(x_0, \pi)}{\varepsilon}, \quad (41)$$

де π_{ε} – політика, що відповідає $u_{\varepsilon}(t)$.

У (41) символ D використовується для позначення похідної за напрямом (похідної Гато [21]). З урахуванням (34) формується оцінку типу

$$D\mathfrak{J}_{\lambda}(x_0, \pi)[\delta u(\cdot)] = \max_{w(\cdot) \in W'_{\pi}(\pi)} D\mathfrak{J}_{\lambda}(\pi, w)[\delta u(\cdot)], \quad (42)$$

тобто варіація робастного функціонала визначається найгіршими сценаріями для поточної політики.

Для кожного фіксованого сценарію $w(\cdot)$ варіація $D\mathfrak{J}_{\lambda}(x_0, \pi)[\delta u(\cdot)]$ (42) може бути записана у вигляді дискретної аналогії першого варіаційного виразу [22].

Далі вводиться послідовність спряжених векторів [23]:

$$p(t; w) \in \mathbb{R}^n, \quad t = 0, \dots, T, \quad (43)$$

які визначаються зворотною рекурентною системою:

$$\begin{cases} p(T, w) = \nabla_x \Phi(x(T, w)) + \lambda \nabla_x \Psi(x(T, w)), \\ p(t, w) = \nabla_x L(x(t, w), u_{\pi}(t)) + \lambda \nabla_x \Psi(x(t, w)) + \\ \quad + (\nabla_x F(x(t, w), u_{\pi}(t), w(t)))^{\top} p(t+1, w), \\ t = T-1, \dots, 0, \end{cases} \quad (44)$$

де ∇_x позначає градієнт за вектором стану x ; $(\cdot)^{\top}$ – операція транспонування.

Тоді, у задачах дискретного оптимального керування, перша варіація робастного функціонала $J_{\lambda}(\pi, w)$ може бути записана у вигляді:

$$\begin{aligned} DJ_{\lambda}(\pi, w)[\delta u(\cdot)] &= \sum_{t=0}^{T-1} \left[\nabla_u L(x(t, w), u_{\pi}(t)) + \right. \\ &\quad \left. + (\nabla_u F(x(t, w), u_{\pi}(t), w(t)))^{\top} p(t+1, w) \right]^{\top} \delta u(t). \end{aligned} \quad (45)$$

У (43) можна позначити

$$g(t, w) = \nabla_u L(x(t, w), u_\pi(t)) + \left(\nabla_u F(x(t, w), u_\pi(t), w(t)) \right)^\top p(t+1, w), \quad (46)$$

як градієнт функціонала витрат за керуванням на кроці t для сценарію $w(\cdot)$, що враховує поточні витрати, керувальний вплив на подальшу динаміку стану інформаційної системи та значення робастного функціонала витрат (27).

Тоді, з урахуванням (45) та (46):

$$DJ_\lambda(\pi, w)[\delta u(\cdot)] = \sum_{t=0}^{T-1} (g(t, w))^\top \delta u(t). \quad (47)$$

З (47) випливає, що для зменшення робастного функціонала $\mathfrak{J}_\lambda$ необхідно обирати варіації керувальних впливів, що спрямовані проти градієнтів, що, в свою чергу, відповідають найгіршим сценаріям (28). Таким чином, формується структура методу, якому присвячено це дослідження, в якій корекція керувальних впливів здійснюється вздовж напрямку, протилежного агрегованому градієнту

$$\bar{g}(t) = \text{agg}\{g(t, w) \mid w(\cdot) \in W'_{\max}(\pi)\}, \quad (48)$$

де $\text{agg}\{\cdot\}$ – оператор агрегування.

У (48) оператор агрегації може задаватися, наприклад, усередненням градієнтів за множиною сценаріїв або вибором того вектора $g(t, w)$, що відповідає сценарію з максимальною нормою градієнта.

На основі викладеного аналізу формулюється ітераційний метод робастної варіаційної оптимізації траєкторій, який працює з послідовностями керувальних впливів $\{u^{(k)}(t)\}_{t=0}^{T-1}$, де k – номер ітерації, та відповідними траєкторіями $x^{(k)}(t, w)$ для сценаріїв $w(\cdot) \in W'_*$:

Крок 1. Вибирається початкова допустима політика $\pi^{(0)}$ (або еквівалентна послідовність $u^{(0)}(t)$), яка забезпечує хоча б наближене виконання ресурсних обмежень і, по можливості, перебування в G для більшості сценаріїв.

Крок 2. Виконується обчислення траєкторій для сценаріїв. Для фіксованих $u^{(k)}(t)$ і кожного сценарію $w(\cdot) \in W'_*$ обчислюються траєкторії $x^{(k)}(t)$, $t=0, \dots, T$, за правилом $x^{(k)}(t+1; w) = F(x^{(k)}(t; w), u^{(k)}(t), w(t))$. Паралельно обчислюються значення штрафної функції $\Psi(x^{(k)}(t; w))$ для контролю перебування траєкторії в області допустимих станів.

Крок 3. Визначаються найгірші сценарії. Для кожного сценарію обчислюється значення:

$$J_\lambda(u^{(k)}(\cdot), w) = J_\lambda(x^{(k)}(\cdot; w), u^{(k)}(\cdot))$$

та, згідно з (38) та (39) знаходиться множина

$$W'^{(k)}_{\max} = \left\{ w(\cdot) \in W'_* \mid J_\lambda(u^{(k)}(\cdot), w) = \max_{w'(\cdot) \in W'_*} J_\lambda(u^{(k)}(\cdot), w') \right\}.$$

Крок 4. Обчислюються спряжені змінні. Для кожного сценарію $w(\cdot) \in W'^{(k)}_{\max}$ розв'язується система зворотних рекурентних співвідношень для спряжених змінних $p^{(k)}(t; w)$ (44):

$$\begin{cases} p^{(k)}(T, w) = \nabla_x \Phi(x^{(k)}(T, w)) + \lambda \nabla_x \Psi(x^{(k)}(T, w)), \\ p^{(k)}(t, w) = \nabla_x L(x^{(k)}(t, w), u^{(k)}(t)) + \lambda \nabla_x \Psi(x^{(k)}(t, w)) + \\ + \left(\nabla_x F(x^{(k)}(t, w), u^{(k)}(t), w(t)) \right)^\top p^{(k)}(t+1, w), \\ t = T-1, \dots, 0, \end{cases}$$

На основі (46) та $p^{(k)}(t; w)$ обчислюються градієнти за керувальними впливами:

$$g^{(k)}(t, w) = \nabla_u L(x^{(k)}(t, w), u^{(k)}(t)) + \left(\nabla_u F(x^{(k)}(t, w), u^{(k)}(t), w(t)) \right)^\top \times \\ \times p^{(k)}(t+1, w).$$

Крок 5. Здійснюється агрегація градієнтів найгірших сценаріїв. Для кожного моменту часу t , згідно з (48) формується агрегований градієнт

$$\bar{g}^{(k)}(t) = \text{agg}\{g^{(k)}(t, w) \mid w(\cdot) \in W'^{(k)}_{\max}\}.$$

Крок 6. Здійснюється корекція керувального впливу. Оновлення послідовності керувальних впливів здійснюється за варіаційною схемою, у якій нове значення впливу визначається з урахуванням напрямку зменшення робастного функціонала. Зокрема, виконується крок у напрямку, протилежному агрегованому градієнту:

$$\tilde{u}^{(k+1)}(t) = u^{(k)}(t) - \alpha^{(k)} \bar{g}^{(k)}(t), \quad (49)$$

де $\alpha^{(k)}$ – крок ітерації, що може визначатися за правилом, яке забезпечує зменшення робастного функціонала $\mathfrak{Z}_\lambda$. Після цього, згідно з (49) виконується проекція на множину допустимих керувальних впливів:

$$u^{(k+1)}(t) = Z_{U(x^{(k+1)}(t))}(\tilde{u}^{(k+1)}(t)),$$

що гарантує виконання ресурсних та інших обмежень мобільної платформи.

Крок 7. Забезпечення виконання критерію зупинки. Ітерації продовжуються доти, доки не буде забезпечено виконання однієї з умов:

- доки не досягнуто заданого зменшення робастного функціонала $\mathfrak{Z}_\lambda(x_0, \pi^{(k)})$;
- доки норма градієнтів $\bar{g}^{(k)}(t)$ не стане меншою за наперед заданий поріг для всіх t ;
- доки не вичерпано обчислювальний бюджет інформаційної системи (мобільної платформи).

У результаті отримується послідовність керувальних впливів $u^{(k)}(t)$, яка є наближеним розв'язком робастної варіаційної задачі та забезпечує формування траєкторій, що:

- мінімізують робастний штрафний функціонал $\mathfrak{Z}_\lambda$;
- залишаються в межах G завдяки штрафному члену $\Phi(x)$;
- є робастними відносно вибраної множини сценаріїв деструктивних впливів W' .

Таким чином, запропонований метод інтегрує варіаційний підхід з робастною постановкою, що враховує найгірші сценарії збурень та включає штраф за вихід за межі G . У результаті формується конструктивна процедура синтезу керувальних впливів, які забезпечують перебування траєкторії в області допустимих станів при функціонуванні інформаційної системи на мобільній платформі.

Процедура застосування запропонованого методу для оперативного розподілу ресурсів мобільної платформи

Перед застосуванням методу в реальних умовах виконується налаштування деяких параметрів.

Крок 1. На основі вимірювань, журналів подій і тестових випробувань визначаються параметри оператора (4).

Крок 2. Виходячи з вимог до якості обслуговування, властивостей даних, обмежень на завантаженість ресурсів тощо, задаються допустимі діапазони показників (13). На основі цих даних визначається множина G (15).

Крок 3. Визначається множина W'_* сценаріїв збурень на інтервалі $[0, T]$.

Крок 4. Задаються функції $L(x, y)$, $\Phi(x(T))$, $\Psi(x)$ та коефіцієнт λ .

Крок 5. Встановлюється довжина інтервалу $[0, T]$, кількість сценаріїв W'_* та максимальна кількість ітерацій методу на одному такті керувального впливу, виходячи з обчислювальних можливостей мобільної платформи.

Результатом підготовчого етапу є налаштована модель та параметри методу, які готові до використання на оперативному рівні керування ресурсами інформаційної системи на мобільній платформі.

Алгоритм працює в режимі послідовного повторного обчислення, коли на кожному такті планування він уточнює оптимальний керувальний вплив на найближчий крок, використовуючи прогноз поведінки інформаційної системи на фіксований період наперед, що називається горизонтом [24].

Нехай, у момент часу t відома оцінка поточного стану $\hat{x}(t)$. Тоді один цикл алгоритму можна задати таким чином:

Крок 1. Оцінка стану $\hat{x}(t)$, доступних ресурсів інформаційної системи та належність $\hat{x}(t)$ до G . Якщо $\hat{x}(t)$ наближається до межі множини G алгоритм на даному кроці може адаптивно збільшувати значення вагових коефіцієнтів при тих складових функціонала, які відповідають показникам живучості системи.

Крок 2. Формування прогнозних сценаріїв збурень. На горизонті $[t, t+T]$ з множини W'_* обирається (або генерується) підмножина сценаріїв, що враховують поточні процес деградації та імовірність виникнення певних типів відмов у найближчий час. Таким чином формуються сценарії збурень (16), які використовуються для прогнозування траєкторій на горизонті.

Крок 3. Побудова початкової послідовності керувальних впливів. Формується початкове припущення щодо керувального впливу на горизонті $u^{(0)}(\tau)$, $\tau = t, \dots, t+T-1$. Наприклад, шляхом екстраполяції попередньо знайденого оптимального розв'язку (використання зсувеної в часі послідовності керувальних впливів із попереднього циклу), або на основі простої початкової евристики. Ця послідовність має задовольняти поточні ресурсні обмеження.

Крок 4. Ітераційна робастна варіаційна оптимізація на планувальному інтервалі. Застосовується обмежене за кількістю ітерацій ($k = 0, 1, \dots, k_{\max}$) варіаційне корегування керувальних впливів за методом робастної варіаційної оптимізації траєкторій, що наведений вище у цьому дослідженні.

Якщо зменшення $\mathfrak{Z}_\lambda$ або норма градієнтів є недостатніми, виконується наступна ітерація. В іншому випадку ітераційний процес припиняється достроково.

Враховуючи обмеженість обчислювальних ресурсів інформаційної системи на мобільній платформі, кількість ітерацій $k_{\max}$, розмір T та кількість сценаріїв обираються так, щоб алгоритм вкладався у виділений часовий інтервал одного часового такту.

Крок 5. Застосування керувального впливу та перехід до наступного такту. Після завершення ітерацій послідовність керувальних впливів $\left\{u^{(k^*)}(\tau)\right\}_{\tau=0}^{t+T-1}$ вважається наближеним рішенням робастної варіаційної задачі на поточному інтервалі (де k^* – номер останньої ітерації).

На практиці до реальної інформаційної системи подається лише перший елемент цієї послідовності $u(t) = u^{(k^*)}(t)$, який реалізується у фізичних механізмах розподілу ресурсів мобільної платформи. Після цього виконується перехід до наступного такту часу $t \leftarrow t+1$, уточняється оцінка стану $\hat{x}(t)$, за потреби оновлюються сценарії W'_* і алгоритм повторює цикл.

Експериментальне дослідження розробленого методу

В рамках експериментального дослідження запропонованого методу розглядається спрощена модель інформаційної системи на мобільній платформі з двовимірним вектором стану [25]:

$$x(t) = (q(t), r(t)),$$

де $q(t)$ – узагальнений дефіцит якості обслуговування (наприклад, сумарна затримка або невиконані запити); $r(t)$ – залишок обчислювальних ресурсів платформи. $q(t)$ та $r(t)$ нормуються до інтервалу $[0, 1]$.

Малі значення $q(t)$ та достатньо великі значення $r(t)$ відповідають оптимальному режиму функціонування інформаційної системи.

Множину допустимих станів можна задати у вигляді прямокутної області:

$$G = \{x = (q, r) \in \mathbb{R}^2 \mid 0 \leq q \leq q_{\max}, r \geq r_{\min}\},$$

де $q_{\max} > 0$ – гранично допустимий дефіцит якості обслуговування; $r_{\min}$ – мінімально припустимий запас ресурсу, за якого інформаційна система здатна реалізовувати свою цільову функцію.

Тоді, виходячи з (4):

$$x(t+1) = A_s x(t) + B u(t) + d_s(t),$$

де $u(t) \in [u_{\min}, u_{\max}]$ – скалярний керувальний вплив, який задає рівень інтенсивності використання ресурсу; A_s та d_s – відповідно матриця стану та вектор збурень для сценарію s , що відображають вплив деструктивних факторів; B – матриця керувального впливу на стан інформаційної системи, що описує, як зміна рівня використання ресурсу приводить до зменшення дефіциту якості функціонування інформаційної системи.

У даному дослідженні для проведення експерименту розглядаються три сценарії деструктивних впливів:

- помірні збурення (Υ_1), що відповідають типовим умовам функціонування інформаційної системи;
- посилені збурення (Υ_2), які моделюють тривалі перевантаження або часткові відмови;
- екстремальні збурення (Υ_3), що відображають комбінації несприятливих факторів, у тому числі цілеспрямовані.

Для кожного сценарію фіксується послідовність збурень $d_s(t)$, яка в подальшому використовується як база для побудови робастного функціонала.

З урахуванням (23) функціонал витрат на інтервалі $[0, T]$ має вигляд:

$$J_\lambda(x(\cdot), u(\cdot)) = \sum_{t=0}^{T-1} (L(x(t), u(t)) + \lambda \Psi(x(T))) + \Phi(x(T)).$$

Робастний функціонал на множині сценаріїв W'_* (у даному прикладі – множині дискретних сценаріїв s) визначається як:

$$\mathfrak{J}_\lambda(u(\cdot)) = \max_s J_\lambda(x_s(\cdot), u(\cdot)).$$

Для порівняння розглядаються дві політики:

- базова політика (Ξ_1), коли фіксований рівень керувального впливу задається як $u(t) \equiv u_{\text{base}}$, що не враховує варіацію сценаріїв та відстань до G ;
- робастно-варіаційна політика (Ξ_2) – послідовність керувальних впливів, отриманих шляхом мінімізації робастного функціонала $\mathfrak{J}_\lambda(u)$ методом покрокової варіаційної корекції (градієнтного типу) з урахуванням найгірших сценаріїв.

Числовий експеримент реалізовано за допомогою емулятора на Python, код якого публічно доступний [26] та який: генерує початковий стан $x(0)$ та фіксовані сценарії збурень; обчислює траєкторії $\{x(t)\}$ для базової та робастно-варіаційної стратегії на множині сценаріїв; формує експериментальний набір даних [27], де для кожного

сценарію, політики та моменту часу зберігаються: номер кроку, сценарій, тип політики (базова / робастна), $q(t)$, $r(t)$, $u(t)$, $L(x(t), u(t))$, $\Psi(u(t))$ та ознака перебування в множині G та візуалізує графіки траєкторій (рис. 1-3).

На фазових траєкторіях у просторі (q, r) (рис. 1) штриховими лініями позначено межі множини допустимих станів G (обмеження $q \leq q_{\max} = 1$, $r \geq r_{\min} = 0.2$). Для сценарію Υ_1 обидві політики (Ξ_1 та Ξ_2) забезпечують перебування траєкторій всередині області G протягом усього інтервалу. Водночас політика Ξ_2 дає суттєво повільніше зниження ресурсу – мінімальне значення $r(t)$ наприкінці інтервалу становить близько 0.71 проти 0.28 для політики Ξ_1 . Це означає, що за близького рівня дефіциту якості $q(t)$ метод дозволяє істотно економити ресурс інформаційної системи на мобільній платформі, створюючи запас живучості для подальших ймовірних деструктивних впливів.

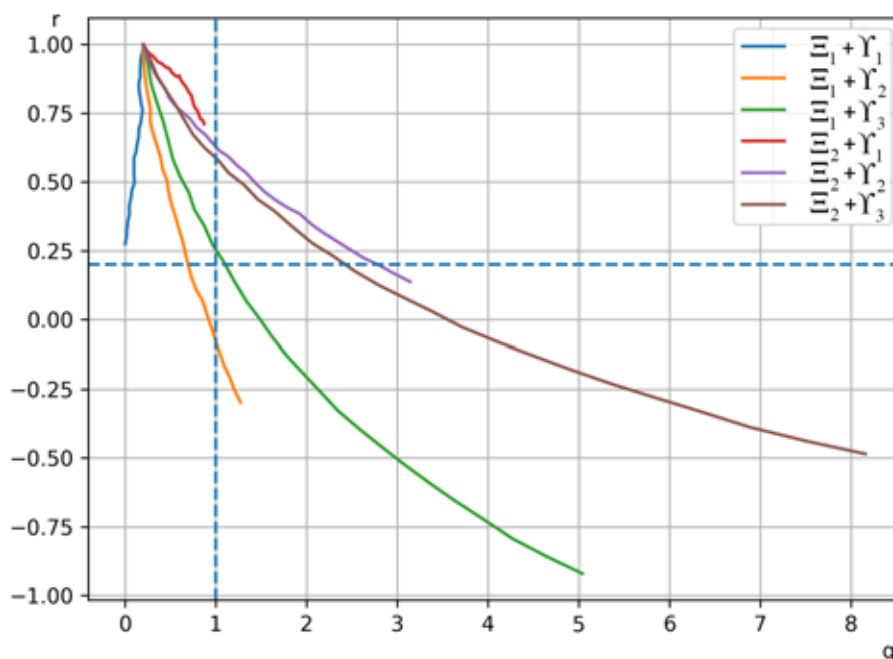


Рис. 1. Місце фазових траєкторій відносно G за різних сценаріїв та стратегій

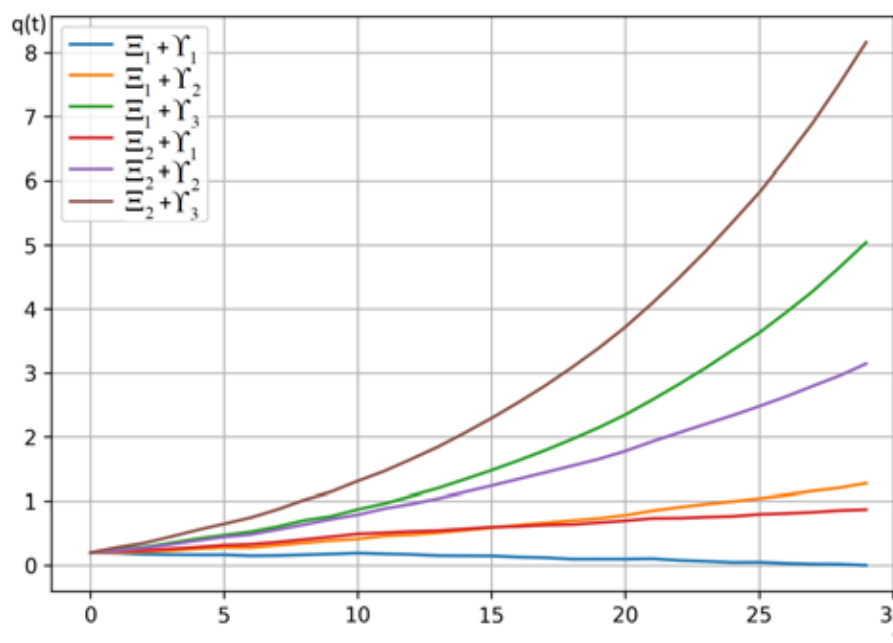


Рис. 2. Динаміка дефіциту якості обслуговування $q(t)$ для різних сценаріїв та стратегій

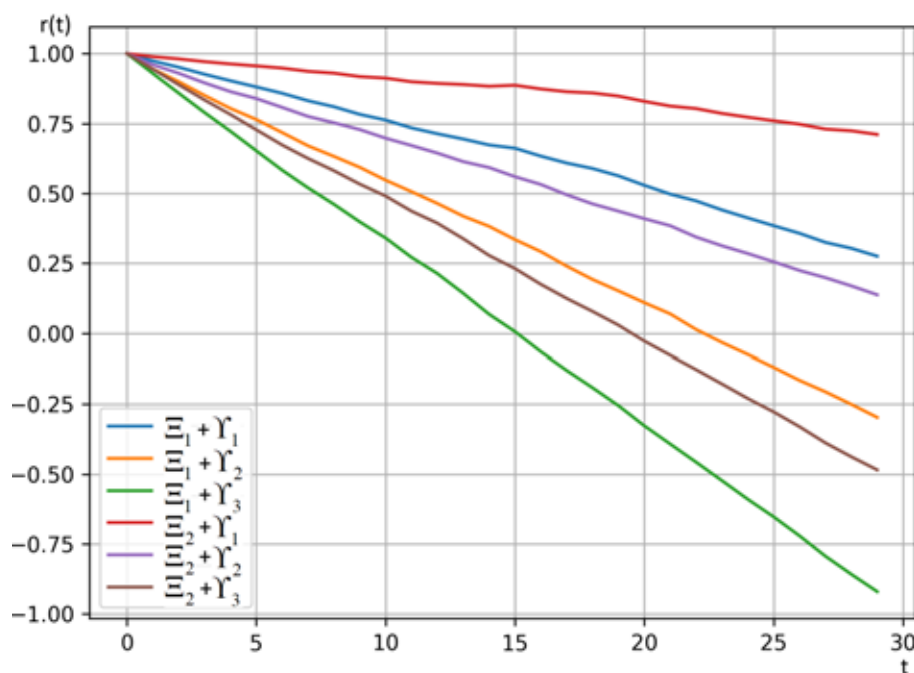


Рис. 3. Динаміка виснаження обчислювальних ресурсів $r(t)$ для різних сценаріїв та стратегій

Для сценарію Υ_2 різниця між політиками проявляється більш виразно. Політика Ξ_1 намагається довше утримувати $q(t)$ нижче порогового значення $q_{\max}$ (рис. 2), однак робить це за рахунок агресивного використання ресурсу. Так, уже на 18-му такті відбувається падіння нижче $r_{\min}$, а мінімальне значення $r(t)$ досягає приблизно -0.30 (рис. 3). Натомість політика Ξ_2 допускає більш раннє зростання дефіциту якості (перетин $q_{\max}$ орієнтовно на 13-му такті), але істотно відтягує момент критичного падіння ресурсу: поріг $r_{\min}$ перетинається лише наприкінці горизонту (≈ 27 -й такт), а мінімальне значення $r(t)$ залишається вищим (≈ 0.14). На фазовій діаграмі (рис. 1) це відповідає траєкторії, що ковзає уздовж межі $q \approx q_{\max}$, зберігаючи більший запас ресурсу. Таким чином, метод реалізує саме робастно-варіаційну стратегію Ξ_2 за якої частина потенційної якості обслуговування свідомо віддається на користь збільшення тривалості ресурсно допустимого функціонування інформаційної системи.

При сценарії Υ_2 жодна з політик не здатна повністю утримати інформаційну систему в межах області G . Однак характер деградації відрізняється. Політика Ξ_1 підтримує менший дефіцит якості на початку інтервалу (перетин $q_{\max}$ приблизно на 12-му такті), але дуже швидко виснажує ресурс – поріг $r_{\min}$ порушується вже біля 13-го такту, а мінімальне значення $r(t)$ опускається до ≈ 0.92 . Політика Ξ_2 , навпаки, допускає більш ранній вихід $q(t)$ за межу $q_{\max}$ (≈ 8 -й такт), зате сповільнює падіння ресурсу, а саме $r_{\min}$ перетинається пізніше (≈ 16 -й такт), – глибина падіння ресурсу менша (мінімум ≈ -0.48). Отже, навіть за найгірших сценаріїв дії деструктивних впливів запропонований метод переводить режим деградації системи від швидкого критичного виснаження ресурсів до керованого зростання дефіциту якості обслуговування за суттєво повільнішого вичерпання ресурсів.

Узагальнюючи, числовий експеримент демонструє, що запропонований робастно-варіаційний метод дозволяє переналаштовувати траєкторії функціонування залежно від рівня деструктивних впливів. У номінальних умовах забезпечується економне використання ресурсу, що має ознаки демонстрації живучості інформаційною системою на мобільній платформі. При посиленних та екстремальних збуреннях метод реалізує контрольовану деградацію якості сервісу заради подовження часу перебування інформаційної системи в області G . Це відповідає концепції живучості як здатності інформаційної системи адаптивно змінювати режим роботи, утримуючись у допустимих станах максимально довго.

Висновки

У дослідженні запропоновано підхід до забезпечення живучості інформаційної системи на мобільній платформі, у якому функціонування системи описується траєкторіями стану, а керування ними здійснюється засобами робастної варіаційної оптимізації. На основі дискретної динамічної моделі введено множину допустимих станів G та пов'язану з нею множину допустимих початкових станів, що дозволяє узгоджено враховувати показники якості обслуговування, властивостей даних і ресурсних обмежень. Варіаційна постановка задачі доповнена робастним функціоналом типу «найгірший випадок» для заданої множини сценаріїв деструктивних впливів, що інтерпретується як вимога інваріантності множини G для цих сценаріїв.

Розроблено ітераційний метод корекції керувань градієнтного типу з штрафним урахуванням наближення до меж області допустимих станів інформаційної системи з урахуванням найкритичніших сценаріїв. На імітаційному прикладі показано, що порівняно з базовою політикою робастно-варіаційна політика забезпечує більш ощадливе використання ресурсу та істотно відтермінує вихід траєкторій стану за межі множини допустимих станів G , що відображає вимоги до живучості інформаційної системи, особливо за посилення та екстремальних збурень. Отримані результати досліджень є подальшим розвитком відомих підходів теорії живучості та робастного оптимального керування (зокрема, концепцій допустимих множин і мінімаксної оптимізації), адаптований до класу інформаційних систем на мобільних платформах, що функціонують за умов деструктивних впливів.

Отримані результати створюють передумови для подальших досліджень у напрямі застосування робастної варіаційної оптимізації до більш складних, високовимірних моделей інформаційних систем, включно з багатокомпонентними мобільними платформами та мультиагентними конфігураціями. Перспективним є також поєднання запропонованого методу з даними моніторингу в режимі реального часу та методами машинного навчання для онлайн-ідентифікації параметрів моделей і адаптивного налаштування множини життєздатних станів G залежно від фактичних умов експлуатації.

Список використаної літератури

1. Schneider C., Valacich J. *Information Systems Today: Managing in the Digital World, Global Edition*. Pearson Education, Limited, 2022. 568 p.
2. Reliability Models of Multi-state UAV-based Monitoring Systems: Mission Efficiency Degradation Issues / I. Kliushnikov et al. 2023 *International Conference on Information and Digital Technologies (IDT)*, Zilina, Slovakia. 2023. P. 299–306. DOI: <https://doi.org/10.1109/idt59031.2023.10194443>.
3. Dodonov O., Gorbachyk O., Kuznietsova M. Analysis and Assessment of Functional Stability of Information Systems Supporting Management Processes. *XXII International Scientific and Practical Conference "Information Technologies and Security (ITS-2022)"*. Kyiv, Ukraine. 2022. P. 1–10.
4. Venkatesha S., Parthasarathi R. Survey on Redundancy Based-Fault tolerance methods for Processors and Hardware accelerators – Trends in Quantum Computing, Heterogeneous Systems and Reliability. *ACM Computing Surveys*. 2024. P. 1–76. DOI: <https://doi.org/10.1145/3663672>.
5. A Survey on Cyber-Resilience Approaches for Cyber-Physical Systems / M. Segovia-Ferreira et al. *ACM Computing Surveys*. 2024. P. 1–37. DOI: <https://doi.org/10.1145/3652953>.
6. Safe Learning in Robotics: From Learning-Based Control to Safe Reinforcement Learning / L. Brunke et al. *Annual Review of Control, Robotics, and Autonomous Systems*. 2022. Vol. 5, no. 1. P. 411–444. DOI: <https://doi.org/10.1146/annurev-control-042920-020211>.
7. An STL-Based Formulation of Resilience in Cyber-Physical Systems / H. Chen et al. *Lecture Notes in Computer Science*. Cham, 2022. P. 117–135. DOI: https://doi.org/10.1007/978-3-031-15839-1_7.
8. Maria M., Costas L. Software Development Lifecycle for Survivable Mobile Telecommunication Systems. *Advances in Science, Technology and Engineering Systems Journal*. 2021. Vol. 6, no. 4. P. 259–277. DOI: <https://doi.org/10.25046/aj060430>.
9. Dynamic survivability of two-layer networks with different topologies / Y. Wang et al. *The European Physical Journal Plus*. 2024. Vol. 139, no. 1. DOI: <https://doi.org/10.1140/epjp/s13360-024-04906-9>.
10. Dynamic survivability of two-layer networks: The role of interlayer coupling / Y. Wang et al. *Chaos, Solitons & Fractals*. 2024. Vol. 180. P. 114571. DOI: <https://doi.org/10.1016/j.chaos.2024.114571>.
11. Drnach L., Zhao Y. Robust Trajectory Optimization Over Uncertain Terrain With Stochastic Complementarity. *IEEE Robotics and Automation Letters*. 2021. Vol. 6, no. 2. P. 1168–1175. DOI: <https://doi.org/10.1109/lra.2021.3056064>.
12. Measurement-Robust Control Barrier Functions: Certainty in Safety with Uncertainty in State / R. K. Cosner et al. 2021 *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Prague, Czech Republic, 27 September – 1 October 2021. 2021. DOI: <https://doi.org/10.1109/iro51168.2021.9636584>.
13. Managing and Using Information Systems: A Strategic Approach. Wiley & Sons, Incorporated, John, 2024.
14. Ruban I., Tkachov V. Formalizing the Survivability Target Function of Information System on Mobile Platform. *International Conference Computer and Information Systems and Technologies, Kharkiv, Ukraine 9–10 October 2025*. 2025. – P. 32–34.
15. Information space model in tasks of distributed mobile objects managing / V. O. Filatov et al. *Information extraction and processing*. 2019. Vol. 2019, no. 47. P. 80–86. DOI: <https://doi.org/10.15407/vidbir2019.47.080>.
16. A Dynamic Event-Based Recursive State Estimation for Mobile Robot Localization / L. Zhu et al. *Electronics*. 2024. Vol. 13, no. 16. P. 3227. DOI: <https://doi.org/10.3390/electronics13163227>.
17. Ruban I., Tkachov V. A model of cross-layer violations and recovery policies for information systems on a mobile platform. *Telecommunication and information technologies*. 2025. Vol. 88, no. 3. P. 204–222. DOI: <https://doi.org/10.31673/2412-4338.2025.038721>.

18. Liu F., Wang R. A Theory for Measures of Tail Risk. *Mathematics of Operations Research*. 2021. Vol. 46, no. 3. P. 1109–1128. DOI: <https://doi.org/10.1287/moor.2020.1072>.
19. Safe Value Functions / P.-F. Massiani et al. *IEEE Transactions on Automatic Control*. 2022. P. 2743–2757. DOI: <https://doi.org/10.1109/tac.2022.3200948>.
20. Krastanov M. I., Ribarska N. K., Tsachev T. Y. A Pontryagin Maximum Principle for Infinite-Dimensional Problems. *SIAM Journal on Control and Optimization*. 2011. Vol. 49, no. 5. P. 2155–2182. DOI: <https://doi.org/10.1137/100799009>.
21. Kreyszig E. *Introductory Functional Analysis with Applications*. John Wiley & Sons, 1978. P. 704.
22. Bertsekas D. P. *Dynamic Programming and Optimal Control*. 3rd ed. Athena Scientific, 2007. 1022 p.
23. Ross I. M. *A Primer on Pontryagin's Principle in Optimal Control*. Collegiate Publishers, 2009. 102 p.
24. Rawlings J. B., Mayne D. Q., Diehl M. M. *Model Predictive Control: Theory, Computation, and Design*. 2nd Edition. Nob Hill Publishing, LLC, 2017. 623 p.
25. Duan Y. Continuous Control-Based Load Balancing for Distributed Systems Using TD3 Reinforcement Learning. *Journal of Computer Technology and Software*. 2024. Vol. 3, no. 6. URL: <https://doi.org/10.5281/zenodo.15392296> (date of access: 20.11.2025).
26. Tkachov V. Mobilesys-robust-trajectory. *GitHub*. URL: <https://github.com/tikey/mobilesys-robust-trajectory> (дата звернення: 20.11.2025).
27. Tkachov V., Ruban I. Supplementary materials for «Robust variational optimization of trajectories for information systems on a mobile platform». *Zenodo*, DOI: <https://doi.org/10.5281/zenodo.17659530>.

References

1. Schneider, C., & Valacich, J. (2022). *Information systems today: Managing in the digital world, Global edition*. Pearson Education.
2. Kliushnikov, I., et al. (2023). Reliability models of multi-state UAV-based monitoring systems: Mission efficiency degradation issues. In *2023 International Conference on Information and Digital Technologies (IDT)* (pp. 299–306). IEEE. <https://doi.org/10.1109/IDT59031.2023.10194443>
3. Dodonov, O., Gorbachyk, O., & Kuznietsova, M. (2022). Analysis and assessment of functional stability of information systems supporting management processes. In *XXII International Scientific and Practical Conference "Information Technologies and Security"* (pp. 1–10).
4. Venkatesha, S., & Parthasarathi, R. (2024). Survey on redundancy-based fault tolerance methods for processors and hardware accelerators: Trends in quantum computing, heterogeneous systems and reliability. *ACM Computing Surveys*, 1–76. <https://doi.org/10.1145/3663672>
5. Segovia-Ferreira, M., et al. (2024). A survey on cyber-resilience approaches for cyber-physical systems. *ACM Computing Surveys*, 1–37. <https://doi.org/10.1145/3652953>
6. Brunke, L., et al. (2022). Safe learning in robotics: From learning-based control to safe reinforcement learning. *Annual Review of Control, Robotics, and Autonomous Systems*, 5(1), 411–444. <https://doi.org/10.1146/annurev-control-042920-020211>
7. Chen, H., et al. (2022). An STL-based formulation of resilience in cyber-physical systems. In *Lecture Notes in Computer Science* (pp. 117–135). Springer. https://doi.org/10.1007/978-3-031-15839-1_7
8. Maria, M., & Costas, L. (2021). Software development lifecycle for survivable mobile telecommunication systems. *Advances in Science, Technology and Engineering Systems Journal*, 6(4), 259–277. <https://doi.org/10.25046/aj060430>
9. Wang, Y., et al. (2024). Dynamic survivability of two-layer networks with different topologies. *The European Physical Journal Plus*, 139(1). <https://doi.org/10.1140/epjp/s13360-024-04906-9>
10. Wang, Y., et al. (2024). Dynamic survivability of two-layer networks: The role of interlayer coupling. *Chaos, Solitons & Fractals*, 180, 114571. <https://doi.org/10.1016/j.chaos.2024.114571>
11. Drnach, L., & Zhao, Y. (2021). Robust trajectory optimization over uncertain terrain with stochastic complementarity. *IEEE Robotics and Automation Letters*, 6(2), 1168–1175. <https://doi.org/10.1109/LRA.2021.3056064>
12. Cosner, R. K., et al. (2021). Measurement-robust control barrier functions: Certainty in safety with uncertainty in state. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. <https://doi.org/10.1109/IROS51168.2021.9636584>
13. Wiley & Sons, Incorporated, John. (2024). *Managing and using information systems: A strategic approach*.
14. Ruban, I., & Tkachov, V. (2025). Formalizing the survivability target function of information system on mobile platform. In *International Conference Computer and Information Systems and Technologies*, Kharkiv, Ukraine, October 9–10, 2025 (pp. 32–34).
15. Filatov, V. O., et al. (2019). Information space model in tasks of distributed mobile objects managing. *Information Extraction and Processing*, 2019(47), 80–86. <https://doi.org/10.15407/vidbir2019.47.080>
16. Zhu, L., et al. (2024). A dynamic event-based recursive state estimation for mobile robot localization. *Electronics*, 13(16), 3227. <https://doi.org/10.3390/electronics13163227>

17. Ruban, I., & Tkachov, V. (2025). A model of cross-layer violations and recovery policies for information systems on a mobile platform. *Telecommunication and Information Technologies*, 88(3), 204–222. <https://doi.org/10.31673/2412-4338.2025.038721>
18. Liu, F., & Wang, R. (2021). A theory for measures of tail risk. *Mathematics of Operations Research*, 46(3), 1109–1128. <https://doi.org/10.1287/moor.2020.1072>
19. Massiani, P.-F., et al. (2022). Safe value functions. *IEEE Transactions on Automatic Control*, 2743–2757. <https://doi.org/10.1109/TAC.2022.3200948>
20. Krastanov, M. I., Ribarska, N. K., & Tsachev, T. Y. (2011). A Pontryagin maximum principle for infinite-dimensional problems. *SIAM Journal on Control and Optimization*, 49(5), 2155–2182. <https://doi.org/10.1137/100799009>
21. Kreyszig, E. (1978). *Introductory functional analysis with applications*. John Wiley & Sons.
22. Bertsekas, D. P. (2007). *Dynamic programming and optimal control* (3rd ed.). Athena Scientific.
23. Ross, I. M. (2009). *A primer on Pontryagin's principle in optimal control*. Collegiate Publishers.
24. Rawlings, J. B., Mayne, D. Q., & Diehl, M. M. (2017). *Model predictive control: Theory, computation, and design* (2nd ed.). Nob Hill Publishing.
25. Duan, Y. (2024). Continuous control-based load balancing for distributed systems using TD3 reinforcement learning. *Journal of Computer Technology and Software*, 3(6). <https://doi.org/10.5281/zenodo.15392296>
26. Tkachov, V. (2025). Mobilesys-robust-trajectory. *GitHub*. <https://github.com/tikey/mobilesys-robust-trajectory>
27. Tkachov, V., & Ruban, I. (2025). Supplementary materials for “Robust variational optimization of trajectories for information systems on a mobile platform” *Zenodo*. <https://doi.org/10.5281/zenodo.17659530>

Дата першого надходження рукопису до видання: 17.11.2025
Дата прийнятого до друку рукопису після рецензування: 15.12.2025
Дата публікації: 31.12.2025

V. I. RUZHENTSEV

Doctor of Technical Sciences, Associate Professor,
Professor at the Department of Secure Information Technologies
Kharkiv National University of Radioelectronics
ORCID: 0000-0002-1007-6530

ANALYSIS OF CRYPTOGRAPHIC RESISTANCE OF ITERATIVE ALGORITHMS BASED ON CONSIDERATION OF ITERATIVE DIFFERENTIAL CHARACTERISTICS

The experiments on searching for the best (with maximum probability) differential characteristics (DC) for the well-known DES algorithm is performed. It was demonstrated for the DES algorithm that in the best DCs with a large number of rounds (more than 10), about 70% of the rounds are covered by the iterative DC. Based on these facts, an improved method for estimating the probability of the best DCs for the DES algorithm was proposed. Estimation was performed using various methods. It was assumed that for other iterative algorithms in the best multi-round DCs a significant percentage of rounds will also be covered by iterative DCs.

A review of the existing results of estimating the differential properties of the Ascon algorithm (NIST SP 800-232 standard) was made.

Information about the minimum number of active substitutions is known: for 2 rounds, the minimum number is at least 4, for 3 rounds – at least 15.

The results of the search for DHs for the p-transformation of the Ascon algorithm for 5 or more rounds are also known: no DH with the number of active S-blocks less than 64 was found. It is also noted that the search was carried out by selecting a small number of active S-blocks in the middle rounds, and then moving to the initial and final rounds, where, of course, the number of active S-boxes will grow rapidly. It is clear that such an approach and the resulting estimate cannot be considered accurate and final.

In our opinion, a smaller total number of active S-boxes can be obtained by preliminary searching for iterative DHs and their subsequent use.

The possibility of performing a search for the best DHs for Ascon algorithm based on consideration of iterative differential characteristics is discussed. An algorithm for searching for iterative differential characteristics for modern ciphers is proposed.

Key words: Symmetric ciphers, Differential cryptanalysis, Differential characteristic, Iterative differential characteristic, DES, Ascon.

В. І. РУЖЕНЦЕВ

доктор технічних наук, доцент,
професор кафедри безпеки інформаційних технологій
Харківський національний університет радіоелектроніки
ORCID: 0000-0002-1007-6530

АНАЛІЗ КРИПТОГРАФІЧНОЇ СТІЙКОСТІ ІТЕРАТИВНИХ АЛГОРИТМІВ, ОСНОВАНИЙ НА РОЗГЛЯДІ ІТЕРАТИВНИХ ДИФЕРЕНЦІАЛЬНИХ ХАРАКТЕРИСТИК

Проведено експерименти з пошуку найкращих (з максимальною ймовірністю) диференціальних характеристик (ДХ) для відомого алгоритму DES. Представлено аналіз структури цих ДХ з метою визначення загальних принципів побудови таких ДХ для інших, більш сучасних алгоритмів. Для алгоритму DES продемонстровано, що у найкращих ДХ при великій кількості циклів (більше 10), близько 70 % відсотків циклів закриває саме ітеративна ДХ. На основі цих даних запропоновано вдосконалений метод оцінювання ймовірності найкращих ДХ для алгоритму DES та виконане порівняння оцінок, які надають інші методи. Зроблено припущення, що для інших ітеративних алгоритмів у кращих багатоциклових ДХ значний відсоток циклів теж буде накриватися ітеративними ДХ.

Зроблено огляд існуючих результатів оцінювання диференціальних властивостей алгоритму Ascon (стандарт NIST SP 800-232).

Відома інформація про мінімальну кількість активних підстановок: для 2 раундів мінімальна кількість не менше 4, для 3 раундів – не менше 15.

Також відомі результати пошуку ДХ для p -перетворення алгоритму Ascon для 5 або більше раундів: не було знайдено жодної ДХ з кількістю активних S -блоків менше 64. Також зазначається, що пошук здійснювався шляхом вибору невеликої кількості активних S -блоків у середніх раундах, а потім переходу до початкового та фінального раундів, де, звичайно, кількість активних S -блоків швидко зростатиме. Зрозуміло, що такий підхід та отриману оцінку не можна вважати точними та остаточними.

На нашу думку, меншу загальну кількість активних S -блоків можна отримати шляхом попереднього пошуку ІДХ та їх подальшого використання.

Обговорюється можливість виконання пошуку найкращих ДХ для цього алгоритму на основі розгляду ітераційних диференціальних характеристик. Запропоновано алгоритм пошуку ітераційних диференціальних характеристик для сучасних шифрів.

Ключові слова: Симетричні шифри, диференціальний криптоаналіз, диференціальна характеристика, ітеративна диференціальна характеристика, DES, Ascon.

Introduction

The symmetric cryptography is the main tool for ensuring cybersecurity. The evolution of symmetric cryptology, like the evolution of many other areas of science, partially occurs in a spiral. Those algorithms that were considered obsolete and out of use 10 years ago can today be attributed to an actively developing area of symmetric cryptography – lightweight cryptography. Such algorithms include DSTU GOST 28147 – 2009, DES, TEA, IDEA, etc. The attention paid to these algorithms is confirmed by the fact that lightweight versions of these algorithms are appearing: for example, for the DES algorithm [1] – the DESL [2] and DESXL [3] algorithms, for the TEA algorithm – the XTEA [4] and XXTEA [5] algorithms.

More modern lightweight algorithms, as a rule, are used larger block size and key, however, main principles and ideas related to the analysis of cryptanalytic properties for long-known algorithms are also applicable to more modern ones. The above said is fully applies to one of the most powerful and universal cryptanalytic attacks – differential cryptanalysis. And for the analysis of resistance to this attack it is usually necessary to estimate the upper bound of the $(r-1)$ -round differential characteristic (DC) for the r -round cipher's transformation. The existence of DC with high probability leads to effective differential attack [6].

Statement of the problem

Assessing the strength of iterative crypto-transformations with a large block (for example, the Ascon algorithm with block size 320 bits) to differential attacks.

Known research and publications analysis. Results of differential characteristic search for algorithm Ascon

The process of finding the best DCs is problematic for modern lightweight cryptographic algorithms with a block size significantly larger than 64 bits.

In [15] were presented information about the minimum number of active substitutions: for 2 rounds minimum number is not less than 4, for 3 rounds – not less than 15.

Known results of DC search for the p transformation of the Ascon algorithm are presented in [13]. It is said that for 5 or more rounds, no DC with a number of active S -boxes less than 64 was found. It is also said that the search was carried out by selecting a small number of active S -boxes in the middle rounds and then moving to the initial and final rounds, where, of course, the number of active S -boxes will increase rapidly. It is clear that such an approach and the obtained assessment cannot be considered as accurate and final.

In our opinion, a smaller total number of active S -boxes can be obtained by preliminary search for iterative DC and their subsequent use.

The aim of work

The aim of this work is to perform and discuss experiments on searching for the best (having the maximum probability) differential characteristics (DC) for the well-known DES algorithm, as well as to analyze the structure of these DCs to identify general principles for constructing such DCs for other, more modern algorithms [7] including the relatively new lightweight standardized Ascon algorithm [13,14].

Methods for searching for DCs with maximum probability

Method of Matsui

In the work [8] M. Matsui proposed an algorithm for searching for multi-round differential characteristics with maximum probabilities for DES-like ciphers. The method is based on the fastest possible sifting out of DCs that cannot be better than the best known DCs for a given number of transformation rounds.

This method is applicable to ciphers with a small block size (not more than 64 bits), for example, DES, FEAL, and requires significant computing resources. For ciphers with a large block size, this algorithm will require too much computational resources that are unrealistic in practice.

Method of Knudsen

One of the methods for finding multi-round differential characteristics with high probabilities was proposed by L. R. Knudsen in [9] and is based on the consideration of iterative differential characteristics (IDC). The method requires

significantly fewer computing resources than Matsui method, since it operates more selectively. To assess the security of the DES [1] and s2-DES [10] ciphers, Knudsen considered several types of IDC (see Fig. 1) and then construct multi-round DC from these IDCs.

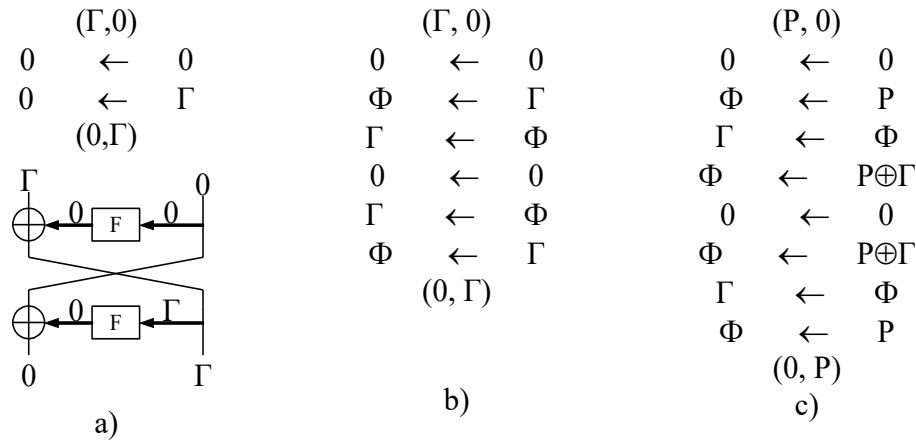


Fig. 1. Types of IDCs from Knudsen's method

In this figure and further, the IDCs are presented in the notation of the work [9], for the first type of IDC in Fig. 1, a) an explanatory scheme for the Feistel chain is also given. In the works [11-12], an extended set of IDCs was presented (Fig. 2).

1	5	8
$0_x \leftarrow \Gamma$	$\Phi \leftarrow \Gamma$	$\Phi \leftarrow \Gamma$
$0_x \leftarrow 0_x$	$\Gamma \oplus \Psi \leftarrow \Phi \oplus \Theta$	$\Psi \leftarrow \Phi$
2-round iterative characteristic	$\Theta \leftarrow \Psi$	$\Theta \leftarrow \Gamma \oplus \Psi$
2	$\Gamma \leftarrow \Phi$	$\Gamma \oplus \Psi \leftarrow \Phi \oplus \Theta$
$\Phi \leftarrow \Gamma$	$\Phi \oplus \Theta \leftarrow \Gamma \oplus \Psi$	$0_x \leftarrow 0_x$
$0_x \leftarrow \Phi$	$\Psi \leftarrow \Theta$	$\Gamma \oplus \Psi \leftarrow \Phi \oplus \Theta$
$\Phi \leftarrow \Gamma$	6-round iterative characteristic	$\Theta \leftarrow \Gamma \oplus \Psi$
$0_x \leftarrow 0_x$	6	$\Psi \leftarrow \Phi$
4-round iterative characteristic	$\Phi \leftarrow \Gamma$	$\Phi \leftarrow \Gamma$
3	$\Gamma \oplus \Psi \leftarrow \Phi \oplus \Theta$	$0_x \leftarrow 0_x$
$\Phi \leftarrow \Gamma$	$\Theta \leftarrow \Psi$	10-round iterative characteristic
$\Gamma \oplus \Psi \leftarrow \Theta$	$\Gamma \oplus \Psi \leftarrow \Phi$	9
$\Phi \leftarrow \Psi$	$\Theta \leftarrow \Gamma$	$\Phi \leftarrow \Gamma$
$\Gamma \oplus \Psi \leftarrow \Phi \oplus \Theta$	$\Psi \leftarrow \Phi \oplus \Theta$	$\Psi \leftarrow \Phi \oplus \Theta$
4-round iterative characteristic	$\Phi \leftarrow \Gamma \oplus \Psi$	$\Theta \leftarrow \Gamma \oplus \Psi$
4	$\Psi \leftarrow \Theta$	$\Gamma \leftarrow \Phi$
$\Phi \leftarrow \Gamma$	8-round iterative characteristic	$\Phi \oplus \Theta \leftarrow \Psi$
$\Gamma \leftarrow \Phi$	7	$\Gamma \oplus \Psi \leftarrow \Theta$
$0_x \leftarrow 0_x$	$\Phi \leftarrow \Gamma$	$\Phi \leftarrow \Gamma$
$\Gamma \leftarrow \Phi$	$\Psi \leftarrow \Phi$	$\Psi \leftarrow \Phi \oplus \Theta$
$\Phi \leftarrow \Gamma$	$\Phi \leftarrow \Gamma \oplus \Psi$	$\Phi \leftarrow \Gamma \oplus \Psi$
$0_x \leftarrow 0_x$	$0_x \leftarrow 0_x$	$\Psi \leftarrow \Theta$
6-round iterative characteristic	$\Phi \leftarrow \Gamma \oplus \Psi$	10-round iterative characteristic
	$\Psi \leftarrow \Phi$	
	$\Phi \leftarrow \Gamma$	
	$0_x \leftarrow 0_x$	
	8-round iterative characteristic	

Fig. 2. Extended set of iterative differential characteristics

Iterative characteristics differ in the number of groups of active substitutions, between which transitions occur in the process of performed transformations.

By going through various values of Γ , Φ , P for the schemes in Fig. 1, Knudsen found IDCs of various types. Repeating the found IDC the required number of times, he covered the number of rounds necessary for organizing an attack, as shown in scheme a) in Fig. 3.

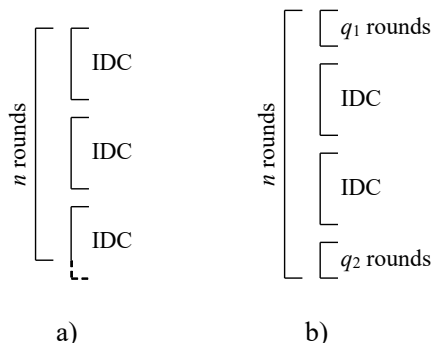


Fig. 3. Schemes of using IDCs in building best DCs

Proposed modified method

However, our computational experiments on finding the best DC for Feistel-like ciphers indicate that the general form of the best multi-round DC can be represented as in Fig. 3, b). Due to this discrepancy with scheme in Fig. 3, a), the probability of a multi-round DC obtained using the Knudsen method will be lower than the actual value, i.e. the cipher may be more vulnerable to differential cryptanalysis than would be expected after using the Knudsen method.

Thus, one of the main drawbacks of the Knudsen method is possible underestimation of the DC probabilities. In addition, the types of IDC considered by Knudsen (Fig. 1) are far from exhausting the entire set of IDC. Additional types of IDC were considered in the works [11-12] and presented on Fig. 2.

Computational experiments to find the best DC using the algorithm proposed by Matsui [8] was performed for a number of DES-like ciphers that differ in substitution tables: the original DES [1], s3DES [10], s5DES, DES variants with substitution tables constructed in accordance with additional requirements [11,12]. The best multi-round (up to 16 rounds) DCs were found. The design features of the best multi-round DCs for some DES-like ciphers are presented in Table 1.

The column "IDC type" indicates the type of IDC and the number of rounds that the extended IDC of this type covers in the best DC. Based on the analysis of these data, as well as similar indicators for other considered ciphers, the following statement can be formulated:

Statement 1. For a DES-like cipher, in the best n -round DC ($n > 10$), at least $n-6$ rounds are covered by a multiple-extended IDC.

Unfortunately, it was not possible to prove statement 1 theoretically at this stage, therefore the validity of the presented statement is confirmed by the results of computational experiments on constructing multi-round DC for DES-like ciphers (in total, about seven variants of DES-like ciphers were considered). The presented statement allows us to propose a method for estimating the maximum probability of DC.

Method for estimating the maximum probability of an n -round ($n > 10$) DC of a Feistel-like cipher:

- 1) using the Matsui algorithm or in some other way, the maximum probability of a 6-round DC is determined;
- 2) using the modified Knudsen method with an increased set of considered types of IDC, the maximum probability of an $(n-6)$ -round DC is determined;
- 3) the maximum probability of an n -round DC is calculated as the product of the values obtained in the first two steps.

According to Statement 1, the probability of DC found using the given method will not be lower than the actual value, therefore, the proposed method can be used to check the practical criterion for the resistance of ciphers of the considered type to differential cryptanalysis.

We believe that the general structure of the best DCs will correspond to the scheme in Fig. 3,b) for many other iterative ciphers with a sufficiently large number of rounds. Then, to estimate the maximum probability of a DC, we need to find the IDC and determine the number of rounds that the IDC will cover.

Results of using modified method for DES-like ciphers

The results of considered methods using for DES-like ciphers is presented in Table 2.

First of all, practical feasibility is important. Next, it is bad when the probability of DC is underestimated, since the real algorithm will be more vulnerable to attack than the results of analysis show (Knudsen's method in the table 2). It is also bad when the probability of DC is greatly overestimated (standard method in table 2). In this case, the algorithm may be too slow due to the using of larger number of rounds than it is needed.

Table 1

The design features of the best multi-round DCs for some DES-like ciphers

Cipher	Number of rounds	IDC type (number of rounds covered by IDC)	Probability with using Matsui's method (exact value)	Probability with using Knudsen's method	Probability with using classical method	Probability with using our method
DES	7	2-rounds (7)	$2^{-23,6}$	$2^{-23,6}$	$2^{-11,74}$	$2^{-35,71}$ $2^{-35,71}$ $2^{-43,58}$ $2^{-43,58}$ $2^{-51,45}$ $2^{-51,45}$ $2^{-59,32}$
	8	2-rounds (8)	$2^{-30,48}$	$2^{-31,48}$		
	9	2-rounds (9)	$2^{-31,48}$	$2^{-31,48}$		
	10	2-rounds (10)	$2^{-38,35}$	$2^{-39,35}$	$2^{-18,46}$	
	11	2-rounds (11)	$2^{-39,35}$	$2^{-39,35}$		
	12	2-rounds (12)	$2^{-46,22}$	$2^{-47,22}$		
	13	2-rounds (13)	$2^{-47,22}$	$2^{-47,22}$		
	14	2-rounds (14)	$2^{-55,1}$	$2^{-55,1}$		
	15	2-rounds (15)	$2^{-56,1}$	$2^{-55,1}$		
16	2-rounds (16)	$2^{-61,97}$	$2^{-62,97}$	$2^{-23,5}$		
S3DES	7	6-rounds (9)	$2^{-26,03}$	$2^{-34,48}$	$2^{-11,74}$	$2^{-40,97}$ $2^{-40,97}$ $2^{-49,00}$ 2^{-54} $2^{-62,4}$ $2^{-62,4}$ $2^{-70,45}$
	8		$2^{-35,22}$	$2^{-42,9}$		
	9		$2^{-41,07}$	$2^{-42,9}$		
	10	6-rounds (10)	$2^{-45,42}$	$2^{-50,93}$	$2^{-18,46}$	
	11	6-rounds (10)	$2^{-47,84}$	$2^{-55,93}$		
	12	6-rounds (10)	$2^{-53,1}$	$2^{-64,35}$		
	13	6-rounds (10)	$2^{-55,52}$	$2^{-64,35}$		
	14	6-rounds (10)	$2^{-64,61}$	$2^{-72,38}$		
	15	6-rounds (11)	$2^{-69,85}$	$2^{-77,38}$		
16	4-rounds (14)	$2^{-78,03}$	$2^{-85,8}$	$2^{-23,5}$		
S5DES	7	8-rounds (5)	$2^{-27,15}$	$2^{-28,52}$	$2^{-11,74}$	$2^{-40,05}$ $2^{-40,05}$ $2^{-45,05}$ $2^{-49,3}$ $2^{-59,33}$ $2^{-59,33}$ $2^{-68,58}$
	8	8-rounds (9)	$2^{-35,62}$	$2^{-38,56}$		
	9		$2^{-38,56}$	$2^{-38,56}$		
	10	8-rounds (9)	$2^{-43,24}$	$2^{-43,56}$	$2^{-18,46}$	
	11	8-rounds (9)	$2^{-46,43}$	$2^{-47,81}$		
	12	8-rounds (9)	$2^{-51,11}$	$2^{-57,84}$		
	13	8-rounds (9)	$2^{-54,3}$	$2^{-57,84}$		
	14	8-rounds (13)	$2^{-62,52}$	$2^{-62,84}$		
	15	8-rounds (13)	$2^{-65,71}$	$2^{-67,09}$		
16	8-rounds (13)	$2^{-75,42}$	$2^{-77,12}$	$2^{-23,5}$		

Table 2

The results of considered methods using for DES-like ciphers

	Standard classical method	Knudsen's method	Proposed method	Matsui's method
Working time (computational complexity)	—	15 minutes	25 minutes	4-5 days
The maximum deviation of the result from the exact value	Overestimation by $2^{54,53}$ times (deviation increases with increasing number of rounds)	Underestimation by $2^{11,25}$ times	Overestimation by $2^{7,58}$ times	Exact value

It can be seen that the proposed method makes it possible to obtain a sufficiently accurate value with low computational complexity. Therefore, we believe that this method will also be useful in the analysis of more modern algorithms.

Algorithm for Iterative differential characteristics search for Ascon

According to the results presented in sections 5, 6, in our opinion, the refinement of the known results of the resistance of the Ascon algorithm to differential attacks (presented in section 3) can be connected with the analysis of iterative DCs.

For organizing the IDC search, at the first stage it is needed to identify two or more groups of S-boxes that can be alternately activated in accordance with the linear dispersion transformations used in the cipher. In this case, the search will consist of performing the following steps:

- building of difference distribution tables for the S-substitutions of the algorithm;
- all possible variants of the number of groups g and the number of active S-boxes in each of the groups are considered: $m_1, m_2, \dots, m_g$;
- for each variant of the numbers $m_1, m_2, \dots, m_g$ all possible combinations of S-boxes are considered. If such groups of S-boxes are found for which cyclic movement is possible: from group m_1 of active S-boxes to group m_2 , from group m_2 of active S-boxes to group m_3 , etc. until the transition from group m_g of active S-boxes to group m_1 , then proceed to the next step d;
- all possible values of differences are analyzed within the selected groups of active S-boxes;
- for each variant, in accordance with the differential difference tables, the possibility of transitions between different groups is checked; if there is such possibility, then an iterative characteristic found.

The presented algorithm will be difficult to implement for a large number of groups and a large number of active S-boxes in groups, but for limited variants (2-4 groups of 2-5 S-boxes in each group), computational experiments are planned for the Ascon algorithm.

Conclusions

1 Evaluation of resistance to Differential Cryptanalysis based on consideration of IDC (iterative differential characteristics) allows to obtain a more accurate estimation of resistance than the Knudsen method [9] with significantly fewer computational costs than the Matsui method [8].

2 Using the example of the DES algorithm, it was demonstrated that the best differential characteristics with a sufficiently large number of rounds contain most of the rounds that are covered by the iterative differential characteristic (according to the results of experiments for DES with 10 or more rounds, the iterative differential characteristic covers at least 70% of the rounds). We believe that the such structure of the best DCs will correspond and for many other iterative ciphers with a sufficiently large number of rounds.

3 The main steps of the algorithm for searching for iterative differential characteristics for iterative symmetric cryptographic algorithms were developed. It is planned to implement this algorithm for the Ascon cipher [13, 14].

Bibliography

1. Data encryption standard. Federal Information Processing Standards Publication 46, U.S. Department of Commerce/National Bureau of Standards, National Technical Information Service, Springfield, Virginia, 1977 (revised as FIPS 46-1:1988; FIPS 46-2:1993).
2. A. Poschmann, G. Leander, K. Schramm, and C. Paar. New light-weight crypto algorithms for RFID, Proceedings in 2007 IEEE International Symposium on Circuits and Systems. IEEE, 2007, pp. 1843–1846. <https://doi.org/10.1109/ISCAS.2007.378273>
3. P. K. Kushwaha, M. Singh, and P. Kumar. A survey on lightweight block ciphers. *International Journal of Computer Applications*, vol. 96, no. 17, 2014. <https://doi.org/10.1109/MDT.2007.178>.
4. J.-P. Kaps. Chai-tea, cryptographic hardware implementations of XTEA. International Conference on Cryptology in India. Springer, 2008, pp. 363–375. https://doi.org/10.1007/978-3-540-89754-5_28
5. E. Yarkov. Cryptanalysis of XXTEA [Electronic resource]. IACR Cryptology ePrint Archive, Report 2010/254, 2010. <http://https://eprint.iacr.org/2010/254.pdf>.
6. Biham, E., Shamir, A. Differential Cryptanalysis of the Full 16-Round DES. In: Differential Cryptanalysis of the Data Encryption Standard. Springer, New York, NY. https://doi.org/10.1007/978-1-4613-9314-6_5
7. Lightweight cryptography project of the American National Institute of Standards and Technology [Electronic resource], 2015. <https://csrc.nist.gov/projects/lightweight-cryptography>.
8. Matsui, M. On correlation between the order of S-boxes and the strength of DES. In: EUROCRYPT 1994, Perugia, Italy, 9-12 May. pp. 366–375. Springer, Heidelberg. <https://dx.doi.org/10.1007/BFb0053451>
9. L. Knudsen. Iterative characteristics of DES and s2-DES. *Advances in Cryptology – Crypto'92*. Springer Verlag, LNCS 746, pp. 497–511, Berlin Heidelberg 1993. https://dx.doi.org/10.1007/3-540-48071-4_35
10. K. Kim. Construction of DES-like S-boxes Based on Boolean Function Satisfying the SAK. *Proceedings Of Asiacypt'91*, pp. 59-72, Fujiyoshida, Japan, 1991. https://dx.doi.org/10.1007/3-540-57332-1_5
11. В. І. Долгов, І. В. Лисицька, В. І. Руженцев. Забезпечення стійкості шифра DES до атак диференціального криптоаналіза. Перекриття характеристик обнуляючого типу і чотирициклових ітеративних характеристик. *Радіотехніка : Всеукр. міжвід. наук.–техн. зб.* 2001. Вип. 120. С. 192–198.
12. В. І. Долгов, І. В. Лисицька, В. І. Руженцев. Забезпечення стійкості шифра DES до атак диференціального криптоаналіза перекриття шести-, восьми- і десятициклових ітеративних характеристик. *Радіотехніка : Всеукр. міжвід. наук.–техн. зб.* 2002. Вип. 124. С. 182–189.
13. Christoph Dobraunig, Maria Eichlseder, Florian Mendel, and Martin Schl  ffer. Ascon v1.2. [Electronic resource] Submission to NIST, 2019. <https://csrc.nist.gov/CSRC/media/Projects/lightweight-cryptography/documents/round-2/spec-doc-rnd2/ascon-spec-round2.pdf>.
14. Ascon-Based Lightweight Cryptography Standards for Constrained Devices. Authenticated Encryption, Hash, and Extendable Output Functions [Electronic resource]. National Institute of Standards and Technology, 2025 (NIST Special Publication 800 NIST SP 800-232). <https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-232.pdf>
15. Christoph Dobraunig, Maria Eichlseder, Florian Mendel and Martin Schl  ffer. Cryptanalysis of Ascon [Electronic resource]. *Cryptology ePrint Archive*, Report 2015/030, 2015. <https://eprint.iacr.org/2015/030.pdf>.

References

1. Data encryption standard (1988). Federal Information Processing Standards Publication 46, U.S. Department of Commerce/National Bureau of Standards, National Technical Information Service, Springfield, Virginia, 1977 (revised as FIPS 46-1:1988; FIPS 46-2:1993).

2. A. Poschmann, G. Leander, K. Schramm, and C. Paar (2007). New light-weight crypto algorithms for RFID, Proceedings in 2007 IEEE International Symposium on Circuits and Systems. IEEE, 2007, pp. 1843–1846. <https://doi.org/10.1109/ISCAS.2007.378273>
3. P. K. Kushwaha, M. Singh, and P. Kumar (2014). A survey on lightweight block ciphers. International Journal of Computer Applications, vol. 96, no. 17, 2014. <https://doi.org/10.1109/MDT.2007.178>.
4. J.-P. Kaps (2008). Chai-tea, cryptographic hardware implementations of XTEA. International Conference on Cryptology in India. Springer, 2008, pp. 363–375. https://doi.org/10.1007/978-3-540-89754-5_28
5. E. Yarrkov (2010). Cryptanalysis of XXTEA. IACR Cryptology ePrint Archive, Report 2010/254, 2010. <http://https://eprint.iacr.org/2010/254.pdf>.
6. Biham, E., Shamir, A. (1993). Differential Cryptanalysis of the Full 16-Round DES. In: Differential Cryptanalysis of the Data Encryption Standard. Springer, New York, NY. https://doi.org/10.1007/978-1-4613-9314-6_5
7. Lightweight cryptography project of the American National Institute of Standards and Technology, 2015. <https://csrc.nist.gov/projects/lightweight-cryptography>.
8. Matsui, M. (1994). On correlation between the order of S-boxes and the strength of DES. In: EUROCRYPT 1994, Perugia, Italy, 9-12 May. pp. 366–375. Springer, Heidelberg. <https://dx.doi.org/10.1007/BFb0053451>
9. L. Knudsen (1992). Iterative characteristics of DES and s2-DES. Advances in Cryptology – Crypto'92. Springer Verlag, LNCS 746, pp. 497-511, Berlin Heidelberg 1993. https://dx.doi.org/10.1007/3-540-48071-4_35
10. K. Kim (1991). Construction of DES-like S-boxes Based on Boolean Function Satisfying the SAK. Proceedings Of Asiacrypt'91, pp. 59-72, Fujiyoshida, Japan, 1991. https://dx.doi.org/10.1007/3-540-57332-1_5
11. V. I. Dolgov, I. V. Lisitska, V. I. Ruzhentsev (2001). Zabezpechennya stiykosti shifra DES do atak diferenciynogo kryptoanaliza. Perekryttya harakterystyk obnulyauchogo typu i chotyrycyklovyh iteratyvnyh harakterystyk. Radiotechnics Vyp. 120. C. 192–198. (in Ukrainian)
12. V. I. Dolgov, I. V. Lisitska, V. I. Ruzhentsev (2002). Zabezpechennya stiykosti shifra DES do atak diferenciynogo kryptoanaliza. Perekryttya shosty-, vosmy- i desyaticyklovyh iteratyvnyh harakterystyk. Radiotechnics. Vyp. 124. C. 182–189. (in Ukrainian)
13. Christoph Dobraunig, Maria Eichlseder, Florian Mendel, and Martin Schl  ffer (2019). Ascon v1.2. Submission to NIST, 2019. <https://csrc.nist.gov/CSRC/media/Projects/lightweight-cryptography/documents/round-2/spec-doc-rnd2/ascon-spec-round2.pdf>.
14. Ascon-Based Lightweight Cryptography Standards for Constrained Devices. Authenticated Encryption, Hash, and Extendable Output Functions (2025). National Institute of Standards and Technology, 2025 (NIST Special Publication 800 NIST SP 800-232). <https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-232.pdf>
15. Christoph Dobraunig, Maria Eichlseder, Florian Mendel and Martin Schl  ffer (2015). Cryptanalysis of Ascon. Cryptology ePrint Archive, Report 2015/030, 2015. <https://eprint.iacr.org/2015/030.pdf>.

Дата першого надходження рукопису до видання: 21.11.2025

Дата прийнятого до друку рукопису після рецензування: 19.12.2025

Дата публікації: 31.12.2025

S. A. SGADOV

Senior Lecturer at the Department of Computer Systems and Networks
National University Zaporizhzhia Polytechnic
ORCID: 0000-0002-7994-6530

ANALYSIS OF STM32 FAMILY SUPPORT IN THE .NET ECO-SYSTEM AND PERSPECTIVES FOR COMMUNITY TARGETS DEVELOPMENT

This article presents a systematic review of the support provided for the STM32 microcontroller series within the managed environment of .NET nanoFramework. The platform architecture is examined, including its multi-layered structure composed of the bootloader (nanoBooter), runtime core (nanoCLR), peripheral abstraction layer (PAL), and device libraries, as well as the specifics of its adaptation to various STM32 families such as F4, F7, and H7. The study identifies the primary advantages of using managed code on microcontrollers, including reduced development time due to the object-oriented programming model, rapid prototyping capabilities, integration with Visual Studio and Dev Containers, and simplified handling of peripheral modules via the C# API. Simultaneously, the limitations and portability challenges are discussed, such as CLR resource demands, differences in DMA, USB, and caching support across STM32 series, and the need for manual BSP and project configuration to ensure stable operation across different targets. A comparative analysis of the STM32 series in the context of nanoFramework support is provided: mid-range series (F4) are optimal for prototyping and educational projects, lower-end series (F0, L0) are limited in memory and peripheral features, and high-performance series (F7, H7) enable complex multithreaded and IoT applications given proper configuration of caches, DMA, and clock multipliers. Based on literature analysis and open-source sources, practical recommendations are offered: select the STM32 series according to task complexity, optimize memory usage, adhere to BSP standards and test code portability across series, utilize automation tools for build and deployment, and in time-critical applications, combine C# with native C code. This article targets embedded systems developers, IoT researchers, and educational institutions seeking to effectively employ managed code on STM32 while maintaining software flexibility and portability.

Key words: .NET nanoFramework, STM32, microcontrollers, managed code, portability, embedded systems, IoT, CLR, BSP, peripheral modules.

С. О. СГАДОВ

старший викладач кафедри «Комп'ютерні систем та мережі»
Національний університет «Запорізька політехніка»
ORCID: 0000-0002-7994-6530

АНАЛІЗ ПІДТРИМКИ STM32-СІМЕЙСТВ В .NET ECO-SYSTEM ТА ПЕРСПЕКТИВИ РОЗВИТКУ COMMUNITY-TARGETS

У цій статті проведено систематичний огляд підтримки мікроконтролерів серій STM32 у керованому середовищі .NET nanoFramework. Розглянуто архітектуру платформи, включно з багаторівневою моделлю, що складається із завантажувача (nanoBooter), рантайм-ядра (nanoCLR), абстракції периферії (PAL) та бібліотек пристроїв, а також особливості її адаптації до різних серій STM32, таких як F4, F7 та H7. Визначено основні переваги використання керованого коду на мікроконтролерах, зокрема зменшення часу розробки завдяки об'єктно-орієнтованій моделі, можливість швидкого прототипування, інтеграція з інструментами Visual Studio та Dev Containers і полегшення роботи з периферійними модулями через API C#. Разом із цим обговорюються обмеження та проблеми портативності, включно із ресурсними вимогами CLR, відмінностями у підтримці DMA, USB та кешування між серіями STM32, а також необхідністю ручного налаштування BSP та конфігурацій проектів для забезпечення стабільної роботи на різних таргетах. У статті проведено порівняльний аналіз серій STM32 щодо підтримки nanoFramework: серії середнього класу (F4) оптимальні для прототипування та освітніх проектів, серії нижчого класу (F0, L0) обмежені обсягом пам'яті та периферійними можливостями, а високопродуктивні серії (F7, H7) дозволяють реалізовувати складні багатопоточні та IoT-застосунки за умови належної конфігурації кешу, DMA і тактових множників. На основі аналізу літератури та відкритих джерел подано практичні рекомендації: обирати серії STM32 відповідно до складності завдань, оптимізувати використання пам'яті, дотримуватися стандартів BSP та тестувати портативний код на різних серіях, застосовувати інструментальні засоби автоматизації збірки та прошивки, а для часово-критичних завдань комбінувати C# із нативним C-кодом. Стаття призначена для розробників embedded-систем, дослідників IoT

та освітніх установ, які прагнуть ефективно використовувати керований код на STM32, зберігаючи гнучкість і портативність програмного забезпечення.

Ключові слова: .NET nanoFramework, STM32, мікроконтролери, керований код, портативність, embedded-системи, IoT, CLR, BSP, периферійні модулі.

Introduction

The modern development of embedded systems is characterized by a growing need for high-level development tools that ensure rapid creation, testing, and maintenance of software for microcontrollers. Despite the traditional dominance of native programming languages, such as C and C++, managed environments based on virtual machines and Intermediate Language (IL) are becoming increasingly widespread. One of the most promising platforms in this context is the .NET eco-system, which, thanks to projects like .NET nanoFramework, Meadow, and TinyCLR OS, has opened up the possibility of using the C# language in the microcontroller field.

On the other hand, the STM32 microcontroller series from STMicroelectronics remains the de facto standard for educational, scientific, and industrial applications. They cover a wide range of hardware architectures—from energy-efficient Cortex-M0 cores to high-performance Cortex-M7/M33—providing flexibility in choice depending on system requirements. As a result, STM32 microcontrollers are considered one of the key hardware platforms for developers seeking to combine native-level performance with the capabilities of a high-level managed environment.

Despite the significant success of the .NET nanoFramework project, support for various STM32 families remains uneven: the availability of official “community targets” is limited, and the level of implementation of the HAL layer, peripheral drivers, and debugging tools varies significantly between specific models. At the same time, the developer community actively creates new targets, expands libraries, and improves build automation tools, forming a unique model of open cooperation among the academic, industrial, and amateur sectors.

The purpose of this article is to systematize the available data on STM32 family support in the .NET eco-system, analyze architectural approaches to ensuring portability and compatibility, and identify directions for the further development of community targets as a key element of open platform support. Within the framework of this research, a number of interconnected tasks are set, aimed at a systemic study of modern approaches to developing software for STM32 microcontrollers within the .NET eco-system. First, an overview of the current state of development for STM32 is envisaged, which makes it possible to outline the level of support for these microcontrollers in various tool environments and determine the main directions of evolution for the corresponding technologies. The next stage is a comparative analysis of the architectural models for implementing HAL (Hardware Abstraction Layer) and CLR (Common Language Runtime) in various eco-systems, particularly in .NET nanoFramework, Meadow, and TinyCLR OS. Such an analysis allows for identifying common principles of construction and differences in approaches to hardware abstraction and managed code execution, which affect the efficiency of microcontroller resource utilization. Special attention is paid to identifying typical problems that limit the portability, scalability, and performance of solutions built on these platforms. This includes issues of driver compatibility, garbage collector efficiency, and system API consistency. The final aspect of the research is the development of recommendations for the further standardization of STM32 support by the developer community in the context of the open development of the .NET eco-system. The formulated proposals should contribute to improving the consistency of implementations, enhancing documentation, and expanding integration capabilities between different development environments. Thus, the article aims not only to describe the current state but also to contribute to the formation of scientific and technical foundations for building a compatible, open, and sustainable environment for developing .NET solutions based on STM32 microcontrollers.

The main part of the article

Overview of the Current State of the .NET Eco-System for Microcontrollers. The .NET nanoFramework platform emerged as a successor to the .NET Micro Framework, created by Microsoft in the mid-2000s for microcontrollers with limited resources [1], [2]. After the end of official support from Microsoft, the project was revived by the community in an open-source format and adapted to modern ARM microcontrollers [3], [9]. NanoFramework implements a simplified version of the CLR (Common Language Runtime) and supports a subset of the .NET core libraries, specifically optimized for ARM Cortex-M devices (STM32, ESP32, etc.) [10], [12]. A key goal of the project is to reduce the barrier to entry for developers familiar with C#, enabling rapid development of embedded systems with support for debugging, CI/CD, and integration with Visual Studio [1], [13].

The .NET nanoFramework ecosystem is a multi-tiered architecture in which each tier performs specific functions, ensuring the coordinated operation of the hardware and software environments. At the lowest level, there is a bootloader (nanoBooter) and a virtual machine (nanoCLR), which are directly executed on the microcontroller and are responsible for initializing the device, loading managed code and controlling its execution [10]. The highest level is the Base Class Library, which is an adapted subset of the standard .NET system libraries, specially optimized for the limited computing resources of microcontrollers [12]. This component provides developers with the necessary data types, input-output interfaces, and basic exception handling mechanisms, allowing them to implement complex software designs within

embedded systems. An important element of the ecosystem is the NuGet package infrastructure, which simplifies the process of updating components, drivers and libraries, ensuring modularity and flexibility in software deployment [9]. This system supports the ability to quickly update individual parts of the project without the need for a complete rebuild of the entire platform. The architecture is completed by the level of integration with development tools, primarily Visual Studio and Visual Studio Code, which provide a full cycle of creating, debugging and deploying applications for STM32 and other supported platforms [13]. Such integration helps to increase developer productivity and ensures compatibility with modern DevOps practices. The system supports numerous community targets (BSP + firmware) for STM32 microcontrollers of various series—F0, F4, F7, L4, etc. [9], [11], which allows you to deploy.NET code on debug boards such as ST Discovery and Nucleo [18].

One of the strongest points of nanoFramework is its deep integration with Microsoft tools. Thanks to official extensions for Visual Studio, developers get the ability to step-by-step debug directly on the microcontroller, viewing the call stack, variables, and registers in real time [13], [14]. The process of creating your own community target (“board”) is well documented: it involves setting up the CMake configuration, clock generators, HAL drivers and linker memory sectors [11].

Project documentation confirms that STM32F0, F4, F7, H7, L0, L4, and G0 families are supported by 2024–2025 [9], [19]. Official or community boards include ST_STM32F4_DISCOVERY, ST_NUCLEO144_F439ZI, and ST_NUCLEO64_F401RE_NF [11]. However, developer and user reports indicate an uneven level of peripheral support—some targets do not fully implement DMA, Ethernet, or the USB stack [20]. For models with small amounts of memory (e.g., STM32F0 or L0), the use of CLR may be limited due to the overhead of the managed environment [4], [24]. Dot NET nanoFramework STM32 family support status is shown in Table 1.

Table 1.

STM32 Family Support in .NET nanoFramework

MCU Family	Board Examples	Support Status	Memory Size (min./recommended)	Peripheral Support	CLR Compatibility
STM32F0	NUCLEO-F091RC (community)	Partial	32 KB / 64 KB Flash	GPIO, UART, I ² C (basic)	Low RAM [9], [24]
STM32F4	STM32F4-DISCOVERY, NUCLEO-F439ZI	Full (official)	128 KB / 192 KB Flash	GPIO, UART, SPI, I ² C, ADC, PWM, USB	Optimal CLR stability; main target platform [9], [11], [14]
STM32F7	STM32F769I DISCOVERY	Full / Active	256 KB / 512 KB Flash	GPIO, Ethernet, USB, SD/ MMC, LCD	Supports complex IoT and UI applications [9], [19]
STM32H7	NUCLEO-H743ZI2 (community)	Partial / In Development	512 KB / 1 MB Flash	GPIO, SPI, UART, Ethernet	Requires HAL optimization [19], [20]
STM32L0	NUCLEO-L053R8	Partial	64 KB / 128 KB Flash	GPIO, UART I ² C	incomplete GC support [11], [24]
STM32L4	NUCLEO-L476RG (official)	Full	128 KB / 192 KB Flash	GPIO, ADC, I ² C, SPI, RTC	Balance performance – energy efficiency [9], [19]
STM32G0	NUCLEO-G071RB (community)	Partial	64 KB / 128 KB Flash	GPIO, UART, ADC	Unstable support [19], [30]

Analysis of the presented data indicates that.NET nanoFramework demonstrates a gradual expansion of support for STM32 families, but the maturity level of the implementation differs significantly depending on the hardware platform. The most stable target environments remain STM32F4 and STM32L4, which provide an optimal balance between memory size, energy efficiency, and the completeness of peripheral library implementation. In contrast, the STM32F0 and STM32L0 families have limitations related to small amounts of RAM and partial garbage collector support. At the same time, the active participation of the community in creating “community targets” for new lines, such as STM32H7 and STM32G0, demonstrates growing interest in using C# on embedded systems. This opens up prospects for developing a unified portability model across different microcontrollers, which is a key direction for the further development of the.NET eco-system for IoT devices.

Despite intensive development and active community support, the.NET nanoFramework environment faces a number of characteristic problems that affect the efficiency of its use in embedded systems. One of the most noticeable is increased memory consumption, caused by the overhead associated with the operation of the Common Language Runtime (CLR). As a result, using nanoFramework can increase the amount of occupied RAM by three to four times compared to native firmware [24], [25]. Another significant problem is limitations in real-time operation due to the functioning of the Garbage Collector. Its periodic activity can lead to short-term delays in code execution, which is a critical factor for systems with strict timing requirements [22], [26]. Additional complexity is the need for manual porting, as supporting new microcontroller families requires separate creation of BSP (Board Support Package) descriptions and configuration of the corresponding hardware layers [11], [30]. This reduces the level of automation in the development process and requires

additional effort from engineers. Particular attention is deserved by the fragmentation of peripheral interface implementations, as the level of HAL (Hardware Abstraction Layer) support can vary significantly between different community targets. Such heterogeneity complicates standardization and code portability across different platforms, limiting project scalability [20].

Compared to alternative technologies (MicroPython, TinyGo, and WebAssembly for MCUs), .NET nanoFramework demonstrates better integration with the development environment and higher performance while retaining managed code [27], [28]. Thanks to the convenience of C# and the object-oriented development model, the platform is considered an effective solution for medium-complexity tasks (control logic, monitoring, IoT services) [1], [2], [6]. The continued growth in popularity directly depends on expanding hardware support and integration with official STM32Cube tools [16], [17], [19].

Thus, the modern .NET eco-system for microcontrollers is at a mature stage of development with practical use in the STM32 environment [9], [10], [14]. .NET nanoFramework is an effective tool for rapid development of managed embedded applications, but requires further optimization of resource utilization and BSP standardization [11], [19], [30]. Improving the architecture for STM32 family support and unifying peripherals are key conditions for increasing the portability and scalability of the environment.

STM32 Support Architecture in .NET nanoFramework. Within the .NET nanoFramework platform, the STM32 support architecture is implemented through a multi-level model that includes the bootloader (nanoBooter), the runtime core (nanoCLR), peripheral abstraction (Platform Abstraction Layer, PAL), and device libraries. This architecture is flexibly configured for resource-constrained microcontroller conditions and allows deployment on both STM32F4/F7 series boards and lower-class platforms [1], [2], [6]. The project provides a list of officially supported reference boards and community targets, including STM32F429I_DISCOVERY and STM32F769I_DISCOVERY, which emphasizes the platform's practical orientation [1], [3]. At the same time, creating new targets requires manual porting: copying templates, configuring CMake, memory section configurations, and peripherals [2], [3]. A key component is the `mcuconf.h` file, which specifies the hardware platform characteristics, including the presence of an external crystal, PLL multipliers, peripheral activation, DMA operation, and cache for CortexM7 [2], [3].

Peripheral support in the nanoFramework architecture is implemented so that standard interfaces (GPIO, UART, I²C, SPI) are accessible through managed C# code, while more complex modules (DMA, Ethernet, USB OTG, RTC) depend on the specific target and sometimes require custom adaptations [6], [10]. Packages like the "Hardware STM32 sample pack" demonstrate the capabilities of working with Backup Memory, RTC, and other features, but for full use of peripherals, the developer must make additional configurations [6], [10]. Tool integration includes building, flashing, and debugging using the GNU ARM Embedded toolchain, CMake, Dev Containers, and Visual Studio Code. This level of automation ensures a quick transition from C# code to a functional device. Over-the-air (OTA) code updates are also supported via Azure IoT for STM32, demonstrating deep integration of managed execution on the MCU [1], [15].

The architecture has two main limitations. Firstly, memory resources: platforms must have sufficient memory to host the CLR and libraries. In practice, for example, the STM32 Nucleo F411RE requires 512 KB Flash and 128 KB SRAM. Secondly, compatibility between STM32 series: different series may have limitations on support for DMA, caching, and other specific functions, which requires additional native customization [11], [12].

Thus, the STM32 support architecture in .NET nanoFramework combines modularity, openness, and instrumental maturity. It allows the platform to be adapted to various STM32 series but requires significant preparatory work during porting. Standardization of targets and enhanced automation of BSP generation can increase the efficiency of using managed code in industrial embedded systems [1], [2], [6], [15].

Portability Issues and STM32 Limitations in .NET nanoFramework. Despite the obvious advantages of managed code execution on STM32, the .NET nanoFramework platform faces a number of limitations that directly affect the portability and scalability of solutions. The first critical aspect is the runtime resource requirements. The CLR and standard library require Flash memory and SRAM on STM32F4/F7 that exceed the minimum specifications of many budget boards [10], [6]. As a result, developers are often limited in the selection of peripherals and libraries that can be connected without memory overflow, especially when using graphics libraries, communication stacks, or complex algorithms [6], [9].

The second factor is the diverse level of peripheral support across the STM32 series. Standard interfaces such as GPIO, UART, SPI, and I²C generally work stably, but more complex components (DMA, Ethernet, USB OTG, RTC) often require manual configuration or are not implemented at all on some targets [6], [29], [10]. User feedback and GitHub issues indicate that even for officially supported boards, local BSP modifications and adaptation of HAL functions to the specifics of the particular board are required [4], [29].

Another problem is the variability in clock configurations and memory settings. To ensure stable nanoCLR operation, the PLL, cache, DMA, and SRAM sections must be correctly configured. Violations of these settings can lead to unstable operation, increased execution latency, or failures when working with interrupts [2], [3]. This problem limits the speed of transferring programs between different STM32 series and forces developers to spend time on low-level experiments.

It is also worth noting the performance and reaction time limitations. The interpreted CLR imposes additional overhead on command execution, especially for intensive I/O operations or high-frequency loops. Although this is not critical for monitoring tasks or medium-complexity control logic, application in time-critical systems requires additional optimizations or partial use of native code [1], [6], [10].

Finally, there is the issue of standardizing community targets. Although open repositories allow for the creation of new boards, the porting process is often non-standardized and requires knowledge of the internal CLR and HAL architecture. This slows down the development of the eco-system and limits its appeal to new developers [4], [6], [29]. Thus, the portability issues and STM32 limitations in .NET nanoFramework are related to runtime resource requirements, diverse peripheral support, hardware parameter configuration, interpreted code performance, and the lack of a standardized process for porting new targets. Addressing these issues requires a comprehensive approach: runtime optimization, expansion of BSP documentation, development of templates for community targets, and instrumental support for compatibility testing [2], [3], [6], [10].

Conclusions and Recommendations for Practice.

Support for various STM32 series in .NET nanoFramework shows significant variability in function availability and the level of peripheral integration. Mid-range series, particularly the STM32F4, provide full support for basic interfaces (GPIO, UART, SPI, I²C) and allow most C# managed code applications to run. On these boards, CLR execution is stable, and the overhead for processing IL bytecode is moderate, making them optimal for prototyping and educational projects [1], [2], [6].

Lower-class STM32 series (e.g., STM32F0, STM32L0) are characterized by limited Flash and SRAM sizes, which restricts the possibility of deploying a fully featured nanoCLR and libraries. In these cases, developers have to abandon certain modules or use lighter library versions, which affects performance and stability [6], [10]. Support for complex peripheral modules, such as DMA and USB, is also limited for these boards, necessitating native code for achieving time-critical characteristics [29], [10].

High-end series, such as STM32F7 and STM32H7, offer significant memory sizes and processor speed, allowing nanoFramework to run with extended libraries and multi-threading capabilities. However, these series often require correct configuration of caches, DMA, and the clocking system to avoid unstable operation and execution delays [2], [3]. Analytical reviews show that when using nanoFramework on STM32H7, there is an increase in peripheral overhead compared to native code, but these costs are offset by the controller's high performance [6], [10].

Special attention should be paid to code portability across the STM32 series. Despite a unified runtime and API, specific differences in clock schemes, DMA presence, memory section location, and cache support may require local BSP modifications. This means that even with identical C# code, transferring a program between different STM32 series may require manual testing and adaptation [4], [29].

The key conclusion of the analysis is that the STM32F4 and F7 are the most compatible series for nanoFramework, providing an optimal balance between resources, stability, and ease of deployment. Lower-class series are suitable for simple tasks and educational projects, while higher-class series allow for the implementation of complex industrial solutions, provided the hardware parameters are properly configured. Thus, the choice of the STM32 series should be made considering resource limitations, the need for peripheral support, and the project's goals [1], [2], [6], [10], [11].

The overview of STM32 platform support in .NET nanoFramework allows for the formulation of several key conclusions regarding portability, limitations, and the effective application of managed code on microcontrollers.

Firstly, the nanoFramework architecture provides a modular and open model that allows platforms to be adapted to various STM32 families, from mid-range series (F4) to high-performance F7 and H7 [1], [2], [6]. This opens up opportunities for rapid prototyping, educational projects, and the development of industrial solutions with moderate resource requirements.

Secondly, limitations in memory, peripheral functions, and performance impose natural constraints on the application of nanoFramework. Series with limited Flash and SRAM require code optimization and selection of only necessary libraries, while high-performance platforms allow the use of extended CLR capabilities, multi-threading, and integration with external services such as Azure IoT [6], [10]. This requires the developer to make a conscious choice of hardware platform depending on the complexity of the task.

The third aspect is the portability and standardization of community targets. Despite a unified API, transferring code between STM32 series often requires manual configuration of the BSP, memory section configurations, and clocking parameters, as well as checking the compatibility of peripheral modules (DMA, USB, caching) [4], [29], [11]. To increase efficiency in practice, developers are recommended to use official templates and actively engage with the nanoFramework community to share developments and BSP modules.

From a practical point of view, effective use of STM32 microcontrollers in the nanoFramework environment involves adhering to a number of technical recommendations aimed at ensuring optimal performance, portability, and system stability. When choosing an STM32 series, the complexity of the tasks should be taken into account: for medium-complexity applications, the STM32F4 series is most appropriate, while for high-complexity industrial systems, the F7 or H7 series

should be used, provided the cache, direct memory access (DMA), and clock multipliers are correctly configured [1], [6], [10]. Special attention must be paid to optimizing memory usage, which involves careful selection of libraries, minimization of dynamic allocations, and preference for static data structures. This approach allows for a significant reduction in the overhead associated with the operation of the Common Language Runtime (CLR) and the Garbage Collector [6], [10]. To ensure code portability between different microcontroller series, it is advisable to adhere to BSP (Board Support Package) templates, avoid low-level calls specific to a particular series, and test critical functions on each target device [4], [29]. Instrumental integration is no less important in the development process. The use of automation systems such as CMake, Dev Containers, and environments like Visual Studio or Visual Studio Code significantly reduces the time required for project build, debugging, and deployment [2], [15]. Furthermore, in hybrid systems with high-speed requirements, it is advisable to combine managed C# code with native C components, which allows for the implementation of high-speed signal processing with minimal latency, without losing the advantages of the object-oriented approach and rapid iterations in the development process [1], [6], [10].

In summary, .NET nanoFramework provides a sufficiently flexible environment for development on STM32, but effective use of the platform requires awareness of hardware limitations, proper selection of the microcontroller series, and active participation in the community for target standardization and BSP optimization. Applying these recommendations allows for maximizing the potential of managed code in industrial, educational, and research projects [1], [2], [6], [10], [11].

Bibliography

1. Ramel, D. Code Small with C# in .NET nanoFramework for Embedded Systems. *Visual Studio Magazine*. 2020. URL: <https://visualstudiomagazine.com/articles/2020/10/01/net-nanoframework.aspx> (дата звернення: 04.11.2025).
2. Ellerbach, L. Show.NET: Running my .NET nanoFramework for 8 years on a battery. *The.NET Blog (Microsoft)*. 2021. June 10. URL: <https://devblogs.microsoft.com/dotnet/show-dotnet-running-my-net-nanoframework-for-8-years-on-a-battery> (дата звернення: 04.11.2025).
3. Gorter, F. nanoFramework – C# for microcontrollers with 64 KB RAM. *Feiko.io Blog*. 2021. May 17. URL: <https://www.feiko.io/posts/2021-05-17-nanoframework-c-for-microcontrollers-with-64k-ram> (дата звернення: 04.11.2025).
4. Vrbaničič, F., Kocijančič, S. Strategy for learning microcontroller programming – a graphical or a textual start? *Education and Information Technologies*. 2024. Vol. 29, No. 2. URL: <https://doi.org/10.1007/s10639-023-12024-9>
5. Atwood, J. On Managed Code Performance, Again. *Coding Horror Blog*. 2005. May 23. URL: <https://blog.codinghorror.com/on-managed-code-performance-again> (дата звернення: 04.11.2025).
6. Li, W., Guan, L., Lin, J., Shi, J., Li, F. *From Library Portability to Para-rehosting: Natively Executing Microcontroller Software on Commodity Hardware*. arXiv preprint arXiv:2107.12867. 2021. URL: <https://arxiv.org/abs/2107.12867> (дата звернення: 04.11.2025).
7. Lin, J., Chen, W.-M., Lin, Y., Cohn, J., Gan, C., Han, S. *MCUNet: Tiny Deep Learning on IoT Devices*. arXiv preprint arXiv:2007.10319. 2020. URL: <https://arxiv.org/abs/2007.10319> (дата звернення: 04.11.2025).
8. Haug, S., Böhm, C., Mayer, D. *Automated Code Generation and Validation for Software Components of Microcontrollers*. arXiv preprint arXiv:2502.18905. 2025. URL: <https://arxiv.org/abs/2502.18905> (дата звернення: 04.11.2025).
9. .NET nanoFramework Docs. *nanoFramework Documentation – STM32 Community Targets* [Електронний ресурс]. 2024. URL: <https://docs.nanoframework.net/content/community-targets/index.html> (дата звернення: 04.11.2025).
10. .NET nanoFramework Docs. *Building for STM32 Boards – .NET nanoFramework Developer Guide* [Електронний ресурс]. 2024. URL: <https://docs.nanoframework.net/content/building/build-stm32.html> (дата звернення: 04.11.2025).
11. .NET nanoFramework Docs. *Creating a Community Board for STM32 in .NET nanoFramework* [Електронний ресурс]. 2024. URL: <https://docs.nanoframework.net/content/stm32/create-community-board.html> (дата звернення: 04.11.2025).
12. .NET nanoFramework Website. *Overview of the .NET nanoFramework* [Електронний ресурс]. 2023. URL: <https://nanoframework.net/overview-of-nanoframework> (дата звернення: 04.11.2025).
13. .NET nanoFramework Blog. Celebrating IoT Day with .NET nanoFramework: Empowering IoT Solutions with Ease. *nanoFramework Blog*. 2024. URL: <https://nanoframework.net/celebrating-iot-day-with-net-nanoframework-empowering-iot-solutions-with-ease> (дата звернення: 04.11.2025).
14. .NET nanoFramework Official Site. *Home – .NET nanoFramework* [Електронний ресурс]. 2025. URL: <https://nanoframework.net/home> (дата звернення: 04.11.2025).
15. .NET nanoFramework Project Page. *502 – Status and Release Roadmap* [Електронний ресурс]. 2023. URL: <https://nanoframework.net/502-2> (дата звернення: 04.11.2025).
16. STMicroelectronics Documentation. *STM32Cube Programmer (UM2237)* [Електронний ресурс]. 2023. URL: <https://www.st.com/en/development-tools/stm32cubeprog.html> (дата звернення: 04.11.2025).

17. STMicroelectronics. *STM32 Cube HAL Driver User Manual* [Електронний ресурс]. 2023. URL: <https://www.st.com/en/embedded-software/stm32cubef4.html> (дата звернення: 04.11.2025).
18. STMicroelectronics. *STM32 Discovery Boards Overview* [Електронний ресурс]. 2024. URL: <https://www.st.com/en/evaluation-tools/stm32-discovery-kits.html> (дата звернення: 04.11.2025).
19. STMicroelectronics Corporate Site. *MCU and MPU Portfolio – STM32 Families* [Електронний ресурс]. 2025. URL: <https://www.st.com/en/microcontrollers-microprocessors/stm32-32-bit-arm-cortex-mcus.html> (дата звернення: 04.11.2025).
20. Microsoft Docs. *.NET Runtime and IL Execution Model* [Електронний ресурс]. 2024. URL: <https://learn.microsoft.com/en-us/dotnet/standard/managed-code> (дата звернення: 04.11.2025).
21. Microsoft Docs. *CLR Internals – Garbage Collection in Embedded Environments* [Електронний ресурс]. 2024. URL: <https://learn.microsoft.com/en-us/dotnet/standard/garbage-collection> (дата звернення: 04.11.2025).
22. Pizlo, F. Garbage Collection in Real-Time and Embedded Systems: A Survey. *ACM Computing Surveys*. 2020. URL: <https://doi.org/10.1145/3406090>
23. Microsoft Learn. Compare.NET nanoFramework with.NET IoT – Code Samples [Електронний ресурс]. 2023. June 29. URL: <https://learn.microsoft.com/en-us/samples/dotnet/samples/compare-nanoframework-with-dotnet-core-iot/> (дата звернення: 04.11.2025).
24. de Souza, D., Ribeiro, M., Martins, A. Benchmarking Managed vs Native Code on Embedded Cortex-M Systems. *IEEE Access*. 2021. Vol. 9. P. 116523–116534.
25. Martinsen, H., Kvale, S. Portability Strategies for C# in Embedded Systems. *Journal of Embedded Computing*. 2023. Vol. 18, No. 4. P. 227–241
26. Carbone, S., Cassar, M. Managed Languages in Low-Power Embedded Applications: Trade-offs and Future Trends. *Microprocessors and Microsystems*. 2023. Vol. 98. P. 104941. URL: <https://doi.org/10.1016/j.micpro.2023.104941> (дата звернення: 04.11.2025).
27. Plauska I., Liutkevičius A. Performance Evaluation of C/C++, MicroPython, Rust and TinyGo on ESP32. *Electronics*. 2023, No. 12. URL: <https://doi.org/10.3390/electronics12010143>.
28. Ryu, J., Han, Y. Adaptive Memory Management for IoT Devices. *Sensors*. 2023. Vol. 23, No. 8. P. 3751.
29. Stack Overflow. NanoFramework STM32 Nucleo_F103RB E5006 [Електронний ресурс]. 2024. January 23. URL: <https://stackoverflow.com/questions/77866771/nanoframework-stm32-nucleo-f103rb-e5006> (дата звернення: 04.11.2025).
30. .NET nanoFramework Docs. *Porting and Board Bring-Up Guide for.NET nanoFramework* [Електронний ресурс]. 2024. URL: <https://docs.nanoframework.net/content/building/porting.html> (дата звернення: 04.11.2025).
31. GitHub. *nanoframework/nfCommunityTargets* [Електронний ресурс]. n.d. URL: <https://github.com/nanoframework/nf-Community-Targets> (дата звернення: 04.11.2025).
32. NuGet. *nanoframework.Hardware.Stm32* [Електронний ресурс]. n.d. URL: <https://www.nuget.org/packages/nanoframework.Hardware.Stm32> (дата звернення: 04.11.2025).

References

1. Ramel, D. (2020, October 1). *Code Small with C# in.NET nanoFramework for Embedded Systems*. Visual Studio Magazine. <https://visualstudiomagazine.com/articles/2020/10/01/net-nanoframework.aspx>
2. Ellerbach, L. (2021, June 10). *Show.NET: Running my.NET nanoFramework for 8 years on a battery*. The.NET Blog (Microsoft). <https://devblogs.microsoft.com/dotnet/show-dotnet-running-my-net-nanoframework-for-8-years-on-a-battery>
3. Gorter, F. (2021, May 17). *nanoFramework – C# for microcontrollers with 64 KB RAM*. Feiko.io Blog. <https://www.feiko.io/posts/2021-05-17-nanoframework-c-for-microcontrollers-with-64k-ram>
4. Vrbančič, F., & Kocijančič, S. (2024). Strategy for learning microcontroller programming – a graphical or a textual start? *Education and Information Technologies*, 29 (2). <https://doi.org/10.1007/s10639-023-12024-9>
5. Atwood, J. (2005, May 23). *On Managed Code Performance, Again*. Coding Horror Blog. <https://blog.codinghorror.com/on-managed-code-performance-again>
6. Li, W., Guan, L., Lin, J., Shi, J., & Li, F. (2021). *From Library Portability to Para-rehosting: Natively Executing Microcontroller Software on Commodity Hardware*. arXiv preprint arXiv:2107.12867. <https://arxiv.org/abs/2107.12867>
7. Lin, J., Chen, W.-M., Lin, Y., Cohn, J., Gan, C., & Han, S. (2020). *MCUNet: Tiny Deep Learning on IoT Devices*. arXiv preprint arXiv:2007.10319. <https://arxiv.org/abs/2007.10319>
8. Haug, S., Böhm, C., & Mayer, D. (2025). *Automated Code Generation and Validation for Software Components of Microcontrollers*. arXiv preprint arXiv:2502.18905. <https://arxiv.org/abs/2502.18905>
9. .NET nanoFramework Docs. (2024). *nanoFramework Documentation – STM32 Community Targets*. <https://docs.nanoframework.net/content/community-targets/index.html>

10. .NET nanoFramework Docs. (2024). *Building for STM32 Boards* –. *NET nanoFramework Developer Guide*. <https://docs.nanoframework.net/content/building/build-stm32.html>
11. .NET nanoFramework Docs. (2024). *Creating a Community Board for STM32 in .NET nanoFramework*. <https://docs.nanoframework.net/content/stm32/create-community-board.html>
12. .NET nanoFramework Website. (2023). *Overview of the .NET nanoFramework*. <https://nanoframework.net/overview-of-nanoframework>
13. .NET nanoFramework Blog. (2024, April 9). *Celebrating IoT Day with .NET nanoFramework: Empowering IoT Solutions with Ease*. <https://nanoframework.net/celebrating-iot-day-with-net-nanoframework-empowering-iot-solutions-with-ease>
14. .NET nanoFramework Official Site. (2025). *Home* –. *NET nanoFramework*. <https://nanoframework.net/home>
15. .NET nanoFramework Project Page. (2023). *502 – Status and Release Roadmap*. <https://nanoframework.net/502-2>
16. STMicroelectronics Documentation. (2023). *STM32Cube Programmer (UM2237)*. <https://www.st.com/en/development-tools/stm32cubeprog.html>
17. STMicroelectronics. (2023). *STM32 Cube HAL Driver User Manual*. <https://www.st.com/en/embedded-software/stm32cubef4.html>
18. STMicroelectronics. (2024). *STM32 Discovery Boards Overview*. <https://www.st.com/en/evaluation-tools/stm32-discovery-kits.html>
19. STMicroelectronics Corporate Site. (2025). *MCU and MPU Portfolio – STM32 Families*. <https://www.st.com/en/microcontrollers-microprocessors/stm32-32-bit-arm-cortex-mcus.html>
20. Microsoft Docs. (2024). *.NET Runtime and IL Execution Model*. <https://learn.microsoft.com/en-us/dotnet/standard/managed-code>
21. Microsoft Docs. (2024). *CLR Internals – Garbage Collection in Embedded Environments*. <https://learn.microsoft.com/en-us/dotnet/standard/garbage-collection>
22. Pizlo, F. (2020). *Garbage Collection in Real-Time and Embedded Systems: A Survey*. *ACM Computing Surveys*. <https://doi.org/10.1145/3406090>
23. Microsoft Learn. (2023, June 29). *Compare .NET nanoFramework with .NET IoT – Code Samples*. <https://learn.microsoft.com/en-us/samples/dotnet/samples/compare-nanoframework-with-dotnet-core-iot/>
24. de Souza, D., Ribeiro, M., & Martins, A. (2021). *Benchmarking Managed vs Native Code on Embedded Cortex-M Systems*. *IEEE Access*, 9, 116523-116534.
25. Martinsen, H., & Kvale, S. (2023). *Portability Strategies for C# in Embedded Systems*. *Journal of Embedded Computing*, 18(4), 227-241.
26. Carbone, S., & Cassar, M. (2023). *Managed Languages in Low-Power Embedded Applications: Trade-offs and Future Trends*. *Microprocessors and Microsystems*, 98, 104941. <https://doi.org/10.1016/j.micpro.2023.104941>
27. Plauska, I., Liutkevičius, A., & Janavičiūtė, A. (2023). *Performance Evaluation of C/C++, MicroPython, Rust and TinyGo Programming Languages on ESP32 Microcontroller*. *Electronics*, 12(1), 143. <https://doi.org/10.3390/electronics12010143>
28. Ryu, J., & Han, Y. (2023). *Adaptive Memory Management for IoT Devices*. *Sensors*, 23(8), 3751. <https://doi.org/10.3390/s23083751>
29. Stack Overflow. (2024, January 23). *NanoFramework STM32 Nucleo_F103RB E5006*. Retrieved October 2025, from <https://stackoverflow.com/questions/77866771/nanoframework-stm32-nucleo-f103rb-e5006>
30. .NET nanoFramework Docs. (2024). *Porting and Board Bring-Up Guide for .NET nanoFramework*. <https://docs.nanoframework.net/content/building/porting.html>
31. GitHub. (n.d.). *nanoframework/nfCommunityTargets*. Retrieved October 2025, from <https://github.com/nanoframework/nf-Community-Targets>
32. NuGet. (n.d.). *nanoframework.Hardware.Stm32*. Retrieved October 2025, from <https://www.nuget.org/packages/nanoframework.Hardware.Stm32>

Дата першого надходження рукопису до видання: 30.11.2025
Дата прийнятого до друку рукопису після рецензування: 26.12.2025
Дата публікації: 31.12.2025

УДК 004.8, 004.932

DOI <https://doi.org/10.35546/kntu2078-4481.2025.4.3.26>**І. П. СІНІЦІН**

доктор технічних наук, професор,
член-кореспондент Національної академії наук України, директор
Інститут програмних систем Національної академії наук України
ORCID: 0000-0002-4120-0784

С. В. ЗАДВОРНИЙ

аспірант
Інститут програмних систем Національної академії наук України
ORCID: 0009-0007-8788-3916

О. П. ІЛЬІНА

кандидат фізико-математичних наук,
провідний науковий співробітник
Інститут програмних систем Національної академії наук України
ORCID: 0000-0002-3111-9310

МОДЕЛІ ТА МЕТОДИ ПІДТРИМКИ КОРИСТУВАЦЬКОЇ ЦЕНТРОВАНОСТІ ІНТЕРФЕЙСУ ПРОГРАМНИХ СИСТЕМ ОБОРОННОГО ПРИЗНАЧЕННЯ З ВИКОРИСТАННЯМ МАШИННОГО НАВЧАННЯ

У статті представлено розширений науково-методичний підхід до побудови та підтримки користувацько-центричного дизайну інтерфейсу інформаційних систем оборонного призначення (ІСОП), який інтегрує концептуальні моделі, формальні описи інтерфейсних послуг, спостереження за користувацькою поведінкою та методи машинного навчання. Актуальність дослідження визначається необхідністю забезпечення стабільної ефективності інтерфейсу в умовах високої динамічності цільових задач, різноманітності ролей користувачів та підвищених вимог до точності виконання операцій у критичних сценаріях. У роботі сформовано багаторівневу модель, яка описує структуру інтерфейсних послуг, параметри їх реалізації, взаємозв'язок з ключовими показниками ефективності (KPI) та вплив характеристик поведінки користувачів на досяжність цих показників. Особливу увагу приділено створенню узагальненої класифікації параметрів формату інтерфейсних послуг, що охоплює структурні, функціональні, перцептивні та маніпулятивні властивості інтерфейсу, а також контекстні фактори використання. На основі зібраних архівів спостережень та порівняння з аналогами запропоновано методи виявлення критичних характеристик, які суттєво впливають на якість інтерфейсної підтримки. Для ухвалення рішень у життєвому циклі ІСОП запропоновано використання моделей машинного навчання для: прогнозування рівня досяжності KPI залежно від реалізованого формату інтерфейсної послуги; класифікації потенційних порушень вимог до інтерфейсу; аналізу ризиків неузгодженості вимог; оцінювання ефективності заходів із модернізації та реінжинірингу. Запропонований підхід забезпечує формування єдиного інформаційно-аналітичного середовища, яке підтримує проєктування, аудит, експлуатаційний моніторинг та підвищення якості інтерфейсу ІСОП на основі об'єктивних даних, формальних моделей і відтворюваних методів аналітики.

Ключові слова: інтерфейс користувача, інтерфейсні послуги, UCD, інформаційна система оборонного призначення, користувацька поведінка, машинне навчання, менеджмент інтерфейсу, класифікатори, аналітична підтримка рішень.

I. P. SINITSYN

Doctor of Technical Sciences, Professor,
Corresponding Member of the National Academy of Sciences of Ukraine, Director
Institute of Software Systems of the National Academy of Sciences of Ukraine
ORCID: 0000-0002-4120-0784

S. V. ZADVORNYI

Postgraduate Student
Institute of Software Systems of the National Academy of Sciences of Ukraine
ORCID: 0009-0007-8788-3916

O. P. ILINA

PhD (Phys. -Math.), Senior Researcher

Institute of Software Systems of the National Academy of Sciences of Ukraine

ORCID: 0000-0002-3111-9310

MODELS AND METHODS FOR SUPPORTING USER-CENTERED INTERFACE DESIGN OF DEFENSE INFORMATION SYSTEMS USING MACHINE LEARNING

This paper presents an extended scientific and methodological framework for developing and maintaining user-centered interface design in defense information systems. The proposed approach integrates conceptual modeling, formal representation of interface services, behavioral observation data, and machine learning methods to improve the quality, stability, and predictability of interface performance in mission-critical environments. The relevance of the study stems from the need to ensure consistent usability and operational reliability under rapidly changing mission objectives, diverse user roles, and increased requirements for task accuracy and speed. The paper introduces a multilayer conceptual model that describes the structure of interface services, the parameters of their implementation, their influence on key performance indicators (KPIs), and the role of user behavior characteristics in achieving those indicators. A generalized classification of interface service parameters is proposed, covering structural, functional, perceptual, and manipulative dimensions, as well as contextual conditions of use. Based on accumulated observation archives and cross-system comparisons, the study outlines methods for identifying critical characteristics that significantly affect the quality of interface support. To enhance decision-making across the system lifecycle, several machine learning techniques are proposed: prediction of KPI achievement depending on the selected implementation format; classification of potential interface requirement violations; assessment of the risks associated with insufficient or inconsistent requirements; and evaluation of reengineering or modernization activities. The resulting framework forms a unified data-driven analytical environment that supports design, audit, operational monitoring, and continuous improvement of defense system interfaces based on objective measurements, formal models, and reproducible analytical workflows.

Key words: user interface, interface services, UCD, defense information systems, personalization, user behavior analytics, machine learning, lifecycle management, decision support.

Постановка проблеми

Актуальність дослідження полягає у необхідності забезпечення стабільної ефективності користувацько-центричного інтерфейсу інформаційних систем оборонного призначення в умовах динамічних задач, високих вимог до точності операцій та зростаючої ролі аналітики користувацької поведінки. Провідною концепцією у сфері проєктування інтерфейсів є користувацько-центричний дизайн (User-Centered Design, UCD) [1], що ставить у центр уваги зручність, доступність, результативність та адаптивність взаємодії. Цей підхід вимагає врахування реальних потреб, сценаріїв та поведінки користувачів на всіх етапах розроблення і експлуатації системи [2].

Успішність використання UCD-підходу у різних галузях підтримується наразі низкою сучасних практик і стандартів [3]:

- використанням UX-аналітики, поведінкових досліджень (user journey mapping, heatmaps), A/B тестування;
- застосуванням когнітивних моделей (GOMS, Fitts' Law) для оцінки навантаження;
- акцентом на персоналізацію (UX personalization) та адаптивні інтерфейси [4];
- впровадженням стандартів оцінки якості, зокрема ISO 9241-210 [5] (ефективність, результативність, задоволеність) та QUIM;

- застосуванням машинного навчання (ML for UX) з метою:
- побудови моделей для класифікації користувачів (novice/expert);
- генерації рекомендацій щодо інтерфейсних рішень;
- виявлення критичних патернів (затримки, помилкові дії, повторення).

Серед технологічних трендів розробки, використовуваних в інтересах UCD, слід відмітити:

- Data-Driven Design – аналіз даних взаємодії;
- Behavior-Centered Design – проєктування на основі поведінкових шаблонів;
- Evidence-Based Design – прийняття рішень на основі емпіричних доказів;

Імплементация парадигми UCD до процесів створення та сталого супроводження інформаційних систем оборонного призначення (ІСОП) становить актуальну проблему та потребує врахування наступної низки особливостей:

- ключовим критерієм якості інтерфейсу є результативність використання за призначенням ІСОП, до якої він належить;
- існує de facto прийнята система функціональних класів ІСОП, для яких може бути визначено базовий набір функцій і ролей користувачів та пов'язаних із ними потреб інтерфейсної підтримки;
- суттєвою особливістю життєвого циклу систем є багаторазовий реінжиніринг, що включає актуалізацію функціональних можливостей та використання перспективних інструментально-технологічних підходів і засобів;

- частою потребою є підтримка прийняття рішень із вибору або адаптації національного рівня для доступних ІСОП, що використовуються в світовій оборонній практиці;
- має місце співіснування як сприятливих організаційних можливостей для системного накопичення та аналізу досвіду експлуатації ІСОП, так і системи спеціалізованих вимог та обмежень, насамперед пов'язаних із захистом інформації.

Кращі практики UCD у сферах навчання, медицини, CRM/ERP-систем, а також у високонавантажених моніторингових системах, демонструють важливість безпосереднього введення структурних та процесних характеристик предметної області до процедур прийняття проектних та експлуатаційних рішень (доменний аспект).

Аналіз останніх досліджень і публікацій

У дослідженнях, присвячених користувацько-центричному дизайну та підтримці інтерфейсних рішень, значна увага приділяється методам збору та аналізу поведінкових даних, а також застосуванню машинного навчання для підвищення якості UX. У роботах [1–4] розглядаються підходи до персоналізації інтерфейсів та формалізації параметрів взаємодії користувачів, що дозволяють будувати адаптивні інтерфейси в інформаційних системах. Дослідження [5–7] узагальнюють сучасні ML-методи для класифікації та прогнозування поведінкових характеристик, які впливають на досяжність показників ефективності. Автори [8–10] підкреслюють важливість уніфікованих моделей опису інтерфейсних послуг та систематизації параметрів для підтримки рішень на етапах проектування та експлуатації. Існуючі роботи формують теоретичне підґрунтя, проте не забезпечують інтегрованого концептуального середовища для комплексної аналітики та управління життєвим циклом інтерфейсу в ІСОП, що обумовлює необхідність запропонованого підходу.

Формулювання мети дослідження

Метою дослідження є створення програмно-технологічних засобів комплексної підтримки UCD на всіх етапах життєвого циклу ІСОП, що ґрунтуються на єдиній концептуальній моделі для організації, використання та актуалізації знань і фактуальних інформаційних масивів у процесах менеджменту ІК.

Викладення основного матеріалу дослідження

В дослідженні пропонується модель, що поєднує: доменну структуру інтерфейсно підтримуваних функцій ІСОП; спостереження за поведінкою користувача; накопичення та узагальнення досвіду; аналітику результативності й рекомендацій із менеджменту життєвого циклу.

Структура моделі знань $KM(S)$ стосовно інтерфейсу ІСОП S має вигляд

$$KM(S) = (MIP, FM, OM, MLCM, RI), \quad (1)$$

де MIP – модель цільового призначення, FM – функціональна модель інтерфейсу, OM – модель спостережень за використанням інтерфейсу, $MLCM$ – модель менеджменту життєвого циклу, RI – відношення взаємозв'язку між елементами різних моделей,

$$MIP = (ST, CHR, CHE, CHS, ME) \quad (2)$$

де ST – множина цільових задач, CHR – показники результативності цільового використання ІСОП, CHE – показники ефективності, CHS – показники стійкості, ME – методи оцінювання

$$FM = \left(\{FS_i, AFU_i, \{IS_{ij}, KPI_{ij}, ACHQ_{ij}, AM_{ij}\} \right. \\ \left. \{FR_{ijk}, ADM_{ijk}\}_{k=1, \dots, K_{ij}} \}_{j=1, \dots, J_i} \right)_{i=1, \dots, I} \quad (3)$$

де FS – функція ІСОП, AFU – ділові ролі користувачів функції, IS_{ij} – інтерфейсна послуга, KPI_{ij} – ключові показники ефективності, $ACHQ_{ij}$ – актуальна характеристика користувацької поведінки, AM_{ij} – її метрика, FR_{ijk} – формат реалізації інтерфейсної послуги, ADM_{ijk} – рішення з управління розвитком інтерфейсної послуги в життєвому циклі ІСОП

$$(OP, \{OT_i, FU_i, \{CHQ_{ij}, M_{ij}\}_{j=1, \dots, J_i}, QT_i\}_{i=1, \dots, I}) \quad (4)$$

де OP – період спостережень, OT_i – виконувана користувачем цільова задача з (2), FU_i – ділова роль користувача, CHQ_{ij} та M_{ij} – характеристика користувацької поведінки та її метрика, QT – якість інтерфейсної підтримки задачі

$$MLCM = (\{E_i, \{DM_{ij}, ELI_{ij}, INF_{ij}, MSD_{ij}\}_{j=1, \dots, J_i}\}_{i=1, \dots, I}) \quad (5)$$

де E_i – етап життєвого циклу ІСОП, DM_{ij} – рішення з менеджменту інтерфейсу, INF_{ij} – інформаційний контекст підтримки процесу прийняття рішення, MSD_{ij} – використовувані методи.

Основним елементом, параметри якого є об'єктом прийняття рішень менеджменту ІК, є інтерфейсна послуга (ІП) з (3). Менеджмент ґрунтується на спостереженнях за використанням ІП в складі ІСОП та її аналогів, оперативно відображуваних у протоколах та довгостроково зберезуваних в архівах спостережень. Склад характеристик поведінки користувача залежить від типу ІП та формату її реалізації. Рекомендована множина була зафіксована в ході дослідження разом із базовими для ІСОП типами ІП, метриками характеристик і складом КРІ для ІП.

Основні види параметрів інформаційних послуг охарактеризовані в табл. 1 згідно з моделлю FM.

Таблиця 1

Параметри інтерфейсних послуг

Формат	Параметр	Опис	Рівні	Перспективні заходи з менеджменту ІІ
Таблиця	Структурна складність	Кількість колонок, вкладених структур	низька/середня/висока	При високій – приховати другорядні колонки
Таблиця	Інформаційна щільність	Кількість даних в одному рядку	низька/середня/висока	При високій – перейти у compact-режим
Таблиця	Маніпульованість	Редагування, сортування, масові операції	низька/середня/висока	При низькій – додати кнопки швидких дій
Таблиця	Візуальна доступність	Контраст, читабельність	низька/оптимальна/підвищена	При низькій – збільшити шрифт/контраст
Таблиця	Динамічність оновлення	Частота оновлення даних	статична/періодична/реального часу	При реальному часі – додати індикатор оновлень
Форма введення	Когнітивне навантаження	Кількість полів, складність	низьке/середнє/високе	При високому – розбити на кроки
Форма введення	Підказковість	Наявність авто-підказок	слабка/достатня/посилена	Посилити при високій кількості помилок
Форма введення	Толерантність до помилок	Авто виправлення, валідація	низька/середня/висока	Підвищити при частих некоректних введеннях
Форма введення	Швидкість введення	Автозаповнення, шаблони	низька/середня/висока	Увімкнути автозаповнення для досвідчених користувачів
Форма введення	Інтерактивність	Динамічні підказки, стан поля	відсутня/контекстна/предиктивна	При високих повторних введеннях – предиктивні підказки
Список	Щільність списку	Кількість елементів на екран	низька/середня/висока	Висока – автоматично переходити у картковий режим
Список	Формат відображення	Текст/картки/іконки	базовий/комбінований/візуальний	Перемикає візуальний режим при потребі швидкого сканування
Список	Фільтрування	Наявність фільтрів	просте/комбіноване/інтелектуальне	Увімкнути рекомендаційні фільтри
Список	Пошук	Пошук по елементах	простий/розширений/семантичний	У разі довгих пошуків – пропонувати автопошук
Карта	Рівень деталізації	Деталізація об'єктів	низька/середня/висока	При високому – групувати маркери
Карта	Навантаження	Кількість шарів/маркерів	низьке/середнє/високе	При високому – кластеризація
Карта	Маніпулятивність	Дії: масштаб, виділення	низька/середня/висока	Додати малювання областей при високій потребі аналітики
Панель дій	Кількість опцій	Кнопки та функції	мала/середня/велика	При великій – групувати в меню
Панель дій	Пріоритетність	Акцент на важливих діях	низька/середня/висока	Високопріоритетні – виділяти кольором
Довідкове вікно	Інтенсивність контенту	Обсяг пояснення	короткі/середні/детальні	При перевантаженні – скоротити текст, додати FAQ
Довідкове вікно	Релевантність	Відповідність контенту ситуації	низька/середня/висока	Показувати контекстну довідку
Візуалізація процесу	Інформативність	Показ станів	низька/середня/висока	У разі затримок – додати ETA
Візуалізація процесу	Часова точність	Коректність відображення	умовна/наближена/точна	Збільшити деталізацію таймерів
Дерево	Глибина ієрархії	Рівень вкладеності	мала/середня/велика	Автозгоргання глибоких гілок
Дерево	Контекстність	Показ сусідніх вузлів	низька/середня/висока	Показати контекст при навігації
Діалогове вікно	Примусовість	Блокувальність	низька/середня/висока	При високій – додати авто-закриття
Діалогове вікно	Кількість опцій	Варіанти вибору	1/2/3+	При 3+ – рекомендувати основний варіант

В дослідженні запропоновано процес аналітичної підтримки прийняття відповідних рішень на найбільш критичних стадіях в життєвому циклі ІСОП, що включає процедури з використанням методів машинного навчання на основі інформації, яка розміщується в таких елементах інформаційного середовища:

- паспорти ІСОП, що визначають їх функції та специфікації ІІ з їх складу;
- архів спостережень, що містить дані про сеанси для різних ІСОП даного класу;
- протоколи спостережень-для сеансів за період, коли продовжується процес розв'язання цільової задачі (дані ще не перенесені до архіву);
- протоколи та архів заходів менеджменту;
- протоколи розв'язання цільових задач з використанням ІСОП (містять експертні оцінки ефективності, що надаються за регламентом стейкхолдерами з оргструктури збройних сил).

В табл. 2 охарактеризовані основні напрями підтримки прийняття рішень із менеджменту ІК на етапах життєвого циклу ІСОП та можливості їх підтримки за допомогою запропонованих моделей.

Таблиця 2

Задачі та середовище процесу комплексної підтримки парадигми UCD

Етап	Тематика рішень	Формовані опції та аргументи	Використовуваний концептуальний контекст	Методи та операції	Використовуваний інформаційний контекст
Проектування ІСОП	Склад елементів інтерфейсу	ІП на підтримку функцій ІСОП	КМ класу, КМ ІСОП – аналогів (в розрізі моделей МІР та FM)	Визначення ІСОП-аналогів за задачами і функціями. Проєкціювання множини ІП, оптимальне за МЕ.МІР, для аналогів	Архів спостережень цільової ефективності ІСОП
	Формати реалізації ІП	Формати та параметри реалізації, пропоновані як перспективні	КМ класу та КМ аналогів (в розрізі FM, МІР, OM)	Добір форматів реалізації за багатокритеріальною оптимальністю (QT. OM та CHR.МІР) для актуальних ST.МІРПобудова і використання ML-моделі класифікації прийнятності формату реалізації ІП	Архіви спостережень цільової ефективності ІСОП та спостережень користувацької поведінки
	Вимоги до інтерфейсу	Оцінка ризиків несумісності та недостатності вимог, виставлених до КРІ проектно обраних ІП	КМ класу	Побудова і використання ML моделей прогнозу рівня досяжності КРІ ІП підтримки цільових задач в залежності від форматів та параметрів реалізації	Архіви спостережень для аналогів
	Вимоги до макету	Склад макету	КМ класу, КМ аналогів	ML – дослідження для виявлення ІП, критичних для цільової якості ІСОП	Архіви спостережень для аналогів
		План спостережень за КП при дослідній експлуатації	КМ класу, КМ аналогів	ML – дослідження для виявлення складу множини характеристик КП, критичних для КРІ ІП та для якості розв’язання цільових задач	Архіви спостережень для аналогів
Дослідна експлуатація макету	Оцінка результатів спостережень	Діагностовані порушення вимог до КРІ ІП	КМ ІСОП, план спостережень, MLM. КМ для аналогів	Використання ML моделей:- Виявлення ІП, неблагополучних за наслідками для якості розв’язання цільових задач;- Прогнозу впливу характеристик КП на КРІ ІП	Протоколи спостережень КП та якості цільового використання ІСОПАрхіви спостережень для аналогівРетроспектива оцінок ефективності заходів із модернізації інтерфейсу для ІСОП-аналогів
Аудит ІСОП, що знаходиться в експлуатації	Відповідність інтерфейсу проектним вимогам	Діагностовані відхилення та оцінка їх критичності	КМ ІСОП, плани спостережень	ML – аналіз вразливості ІП та ризику впливів на якість цільового використання ІСОП	Архіви і протоколи спостережень за ІСОП та її аналогами на різних етапах життєвого циклуНормативні вимоги до КРІ ІП та показників якості цільового використання ІСОП
	Доцільність проведення заходів із реінжинірингу ІСОП	Оцінка критичності характеристик інтерфейсу за актуальних умов та зареєстрованих нових потреб	Базові та оновлені при здійсненні модернізацій версії КМ ІСОП та КМ класу	Використання ML – моделі класифікації перспективності виконання заходів реінжинірингу за заданих умов	Архіви і протоколи спостережень, ретроспектива оцінок ефективності заходів із модернізації інтерфейсу ІСОП аналогів

Найбільшу роль в побудові необхідних процедур машинного навчання відіграють методи класифікації [6], прогнозу [7] за умов невизначеності, виявлення аномальних ситуацій [8] та причинно-наслідкових зв’язків [9], а також моделі оптимізації інтерфейсу на основі поведінкових даних [11].

Набори основних даних для них формуються з результатів спостережень у вигляді масиву

$$\left\{ \left(TV_j, \{MCH_{ji}\}_{i=1, \dots, NCH}, \{MLF_{jk}\}_{k=1, \dots, KL} \right) \right\}_{j=1, \dots, MO}$$

де MO – число спостережень; TV – цільова змінна (мітка позитивності для результату вирішення цільової задачі ІСОП, в межах якої відбувався сеанс); MCH – значення метрики характеристики CH_i користувацької поведінки (i -та сигнальна ознака); NCH – число сигнальних ознак; MLF – значення метрики k – тої лагової ознаки (ідентифікатор ІП, сеансу, цільової задачі тощо); KL – число лагових ознак.

Схема процесу подана на рис. 1.

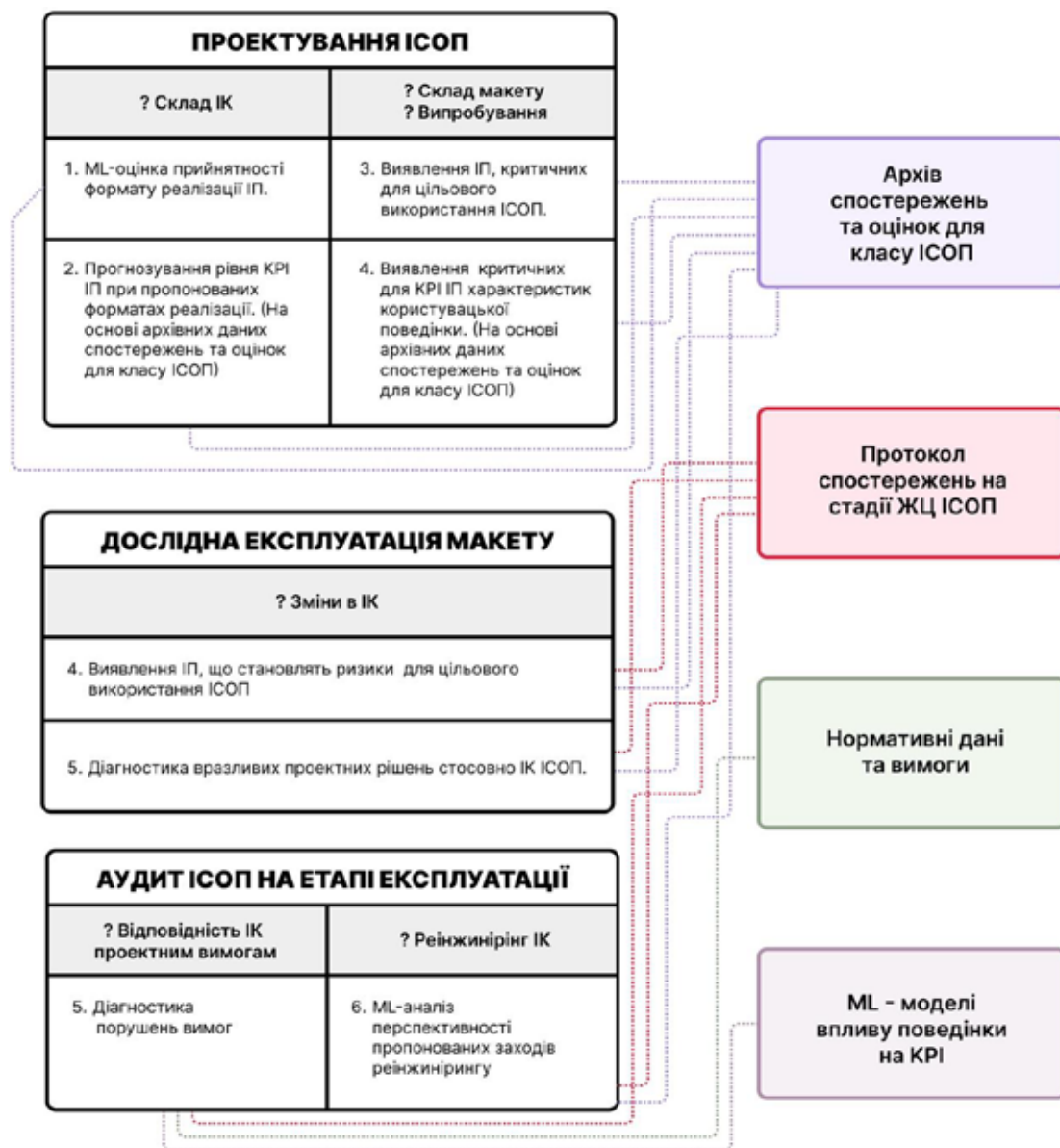


Рис. 1. Схема процесу

Джерело: розроблено авторами

Висновки

В результаті проведеного дослідження були запропоновані:

- система концептуальних моделей, що визначає зв'язки між елементами інтерфейсу користувача ІСОП як із характеристиками ІСОП, так і з елементами спостережень за поведінкою користувача;
- структура та склад інформаційного середовища безперервної підтримки парадигми UCD на основі такої системи моделей;
- процедури аналітичної підтримки прийняття рішень із менеджменту інтерфейсу, ґрунтовані на використанні такого інформаційного середовища для машинного навчання моделей, що служать конструктивним поданням певних зв'язків зі складу концептуальних моделей.

Наразі виконується параметризація моделей інтерфейсу та постановок задач машинного навчання для класу ІСОП Управління оборонними ресурсами, а також проектування програмних засобів інформаційно-аналітичної підтримки використання даних про поведінку користувачів для цієї предметної області.

Подальші дослідження передбачають розширення формалізації параметрів інтерфейсних послуг для різних типів ІСОП, удосконалення методів побудови ознак для задач машинного навчання та розроблення прототипів аналітичних засобів, що підтримують автоматизовану адаптацію інтерфейсу. Перспективним є створення уніфікованого репозиторію поведінкових даних, який забезпечить можливість порівняльного аналізу між системами та підвищить точність прогнозних моделей. Окремо планується інтеграція сценарного моделювання і багатокритеріальної оптимізації для обґрунтування рішень щодо реінжинірингу інтерфейсних послуг.

Список використаної літератури

1. Gould J. D., Lewis C. Designing for usability: key principles and what designers think. *Communications of the ACM*. 1985. Vol. 28, No. 3. P. 300–311. DOI: <https://doi.org/10.1145/3166.3170>
2. Norman D. A., Draper S. W. (eds.). *User centered system design: new perspectives on human-computer interaction*. Hillsdale, NJ : Lawrence Erlbaum Associates, 1986.
3. Mao J.-Y., Vredenburg K., Smith P. W., Carey T. The state of user-centered design practice. *Communications of the ACM*. 2005. Vol. 48, No. 3. P. 105–109. DOI: <https://doi.org/10.1145/1047671.1047677>
4. Jameson A. Adaptive interfaces and systems. In: Jacko J.A. (ed.) *The Human–Computer Interaction Handbook*. Boca Raton : CRC Press, 2014. P. 305–330
5. International Organization for Standardization. ISO 9241-210:2019 – Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems. Geneva : ISO, 2019.
6. Fernández-Delgado M., Cernadas E., Barro S., Amorim D. Do we need hundreds of classifiers to solve real world classification problems? *Journal of Machine Learning Research*. 2014. Vol. 15. P. 3133–3181.
7. McElreath R. *Statistical rethinking: a bayesian course with examples in R and Stan* (2nd ed.). New York : Chapman and Hall/CRC, 2020. DOI: <https://doi.org/10.1201/9780429029608>
8. Hariri S., Carrasco Kind M., Brunner R. J. Extended Isolation Forest. *IEEE Transactions on Knowledge and Data Engineering*. 2021. Vol. 33, No. 4. P. 1479–1490. DOI: <https://doi.org/10.1109/TKDE.2019.2947676>
8. Wager S., Athey S. Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*. 2018. Vol. 113, No. 523. P. 1228–1242. DOI: <https://doi.org/10.1080/01621459.2017.1319839>
10. Oulasvirta A., Reichel A. *Computational Interaction*. Oxford : Oxford University Press, 2022.
11. Chen X., Wu W. User behavior modeling for interface optimization based on machine learning. *IEEE Access*. 2020. Vol. 8. P. 107382–107395. DOI: <https://doi.org/10.1109/ACCESS.2020.2999564>

References

1. Gould, J. D., & Lewis, C. (1985). Designing for usability: Key principles and what designers think. *Communications of the ACM*, 28(3), 300–311. <https://doi.org/10.1145/3166.3170>
2. Norman, D. A., & Draper, S. W. (Eds.). (1986). *User centered system design: New perspectives on human-computer interaction*. Lawrence Erlbaum Associates.
3. Mao, J.-Y., Vredenburg, K., Smith, P. W., & Carey, T. (2005). The state of user-centered design practice. *Communications of the ACM*, 48(3), 105–109. <https://doi.org/10.1145/1047671.1047677>
4. Jameson, A. (2014). Adaptive interfaces and systems. In J. A. Jacko (Ed.) *The Human–Computer Interaction Handbook* (pp. 305–330). CRC Press.
5. International Organization for Standardization. (2019). *ISO 9241-210:2019 – Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems*. ISO.
6. Fernández-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems? *Journal of Machine Learning Research*, 15, 3133–3181.
7. McElreath, R. (2020). *Statistical rethinking: a Bayesian course with examples in R and Stan* (2nd ed.). Chapman and Hall/CRC. <https://doi.org/10.1201/9780429029608>
8. Hariri, S., Carrasco Kind, M., & Brunner, R. J. (2021). Extended isolation forest. *IEEE Transactions on Knowledge and Data Engineering*, 33(4), 1479–1490. <https://doi.org/10.1109/TKDE.2019.2947676>
9. Wager, S., & Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523), 1228–1242. <https://doi.org/10.1080/01621459.2017.1319839>
10. Oulasvirta, A., & Reichel, A. (2022). *Computational interaction*. Oxford University Press.
11. Chen, X., & Wu, W. (2020). User behavior modeling for interface optimization based on machine learning. *IEEE Access*, 8, 107382–107395 <https://doi.org/10.1109/ACCESS.2020.2999564>

Дата першого надходження рукопису до видання: 21.11.2025
 Дата прийнятого до друку рукопису після рецензування: 18.12.2025
 Дата публікації: 31.12.2025

О. А. СНІОСЕК

аспірант кафедри кібербезпеки інформаційних систем, мереж і технологій
Харківський національний університет імені В. Н. Каразіна
ORCID: 0009-0001-9468-5965

О. П. НАРЕЖНИЙ

доцент кафедри кібербезпеки інформаційних систем мереж і технологій
Харківський національний університет імені В. Н. Каразіна
ORCID: 0000-0003-4321-0510

МОДЕЛЬ ПОСТКВАНТОВОЇ ІНФРАСТРУКТУРИ ВІДКРИТИХ КЛЮЧІВ З ВИКОРИСТАННЯМ КВАНТОВОГО ГЕНЕРАТОРА ВИПАДКОВИХ ЧИСЕЛ ДЛЯ АВТОНОМНОЇ СИСТЕМИ ДИФЕРЕНЦІАЛЬНОЇ КОРЕКЦІЇ ГЛОБАЛЬНИХ НАВІГАЦІЙНИХ СУПУТНИКОВИХ СИСТЕМ

У статті запропоновано формалізовану модель постквантової інфраструктури відкритих ключів (ІВК) для інформаційно-комунікаційної системи (ІКС) автономної системи диференціальної корекції (АСДК), побудовану на основі використання квантового генератора випадкових чисел (QRNG) та національних криптографічних стандартів ДСТУ 7564:2014 («Купина»), ДСТУ 7624:2014 («Калина»), ДСТУ 8961:2019 («Скеля») і ДСТУ 9212:2023 («Вершина»). Актуальність дослідження зумовлена необхідністю підвищення криптостійкості каналів передавання даних між контрольнo-коригуючими станціями (ККС) та центром оброблення даних АСДК, а також підвищенням вимог до автентичності та цілісності корекційних даних, що надаються через веб-сервіс зацікавленим користувачам АСДК, зокрема в умовах постквантових загроз. Запропонована модель ІВК забезпечує автентифікацію ККС та користувачів, захист тунельних каналів, цілісність корекційних даних та довіру до серверних компонентів АСДК шляхом інтеграції квантової ентропії у процеси генерації ключів і підпису. У роботі обґрунтовано роль QRNG як джерела високоякісної ентропії для побудови довготривалих ключових пар за алгоритмами ДСТУ «Скеля», «Вершина» та сеансових ключів симетричного шифрування за алгоритмом ДСТУ «Калина», а також розроблено узагальнену структуру ІВК з одним Root CA та декількома Issuing CA для основного й резервного центрів оброблення даних. Запропонована модель визначає логічну взаємодію компонентів ІВК, механізмів тунелювання, модулів підпису та перевірки сертифікатів, а також демонструє можливість забезпечення постквантової стійкості АСДК без зміни її архітектури. Показано, що використання QRNG у поєднанні з національними криптоалгоритмами дає змогу мінімізувати ризик компрометації ключів, підвищити стійкість до квантових атак та забезпечити довготривалу криптографічну надійність каналів зв'язку і веб-сервісу АСДК. Запропонована методика дозволяє створювати ІВК для інформаційно-комунікаційних систем диференціальної корекції та інших сегментів критичної інфраструктури з підвищеними вимогами до довіри, цілісності та конфіденційності даних, з урахуванням вимог сучасних стандартів кібербезпеки (ISO/IEC 27001, NIST, НД ТЗІ).

Ключові слова: глобальна навігаційна супутникова система (GNSS); квантовий генератор випадкових чисел (QRNG); кібератака; кібербезпека; система диференціальної корекції (Differential GNSS, DGNSS); інфраструктура відкритих ключів (ІВК); автономна система диференціальної корекції (АСДК, Autonomous DGNSS); постквантова криптографія (PQC).

О. А. SNIOSEK

Postgraduate Student at the Department of Cybersecurity of Information Systems, Networks and Technologies
V. N. Karazin Kharkiv National University
ORCID: 0009-0001-9468-5965

О. П. NARIEZHNI

Associate Professor at the Department of Cybersecurity of Information Systems, Networks and Technologies
V. N. Karazin Kharkiv National University
ORCID: 0000-0003-4321-0510

MODEL OF A POST-QUANTUM PUBLIC KEY INFRASTRUCTURE USING A QUANTUM RANDOM NUMBER GENERATOR FOR AN AUTONOMOUS DIFFERENTIAL CORRECTION SYSTEM OF GLOBAL NAVIGATION SATELLITE SYSTEMS

In the article, a formalized model of a post-quantum public key infrastructure (PKI) for the information-and-communication system (ICS) of the Autonomous Differential GNSS System (Autonomous DGNSS) is proposed, constructed on the basis of the use of a quantum random number generator (QRNG) and the national cryptographic standards DSTU 7564:2014 ("Kupyna"), DSTU 7624:2014 ("Kalyna"), DSTU 8961:2019 ("Skelia"), and DSTU 9212:2023 ("Vershyna"). The relevance of the research is conditioned by the need to increase the cryptographic robustness of data-transmission channels between the control-and-correction stations (CCS) and the data-processing center of the Autonomous DGNSS, as well as by the increased requirements for the authenticity and integrity of correction data provided through the web service to interested users of the Autonomous DGNSS, particularly under post-quantum threats. The proposed PKI model ensures authentication of CCS and users, protection of tunnel channels, integrity of correction data, and trust in the server components of the Autonomous DGNSS by integrating quantum entropy into the processes of key generation and signing. The study substantiates the role of the QRNG as a source of high-quality entropy for constructing long-term key pairs according to the DSTU "Skelia" and "Vershyna" algorithms, as well as session keys of symmetric encryption under the DSTU "Kalyna" algorithm, and also develops a generalized PKI structure with a single Root CA and several Issuing CAs for the primary and backup data-processing centers. The proposed model defines the logical interaction of PKI components, tunneling mechanisms, signature and certificate-verification modules, and demonstrates the possibility of ensuring post-quantum resistance of the Autonomous DGNSS without changing its architecture. It is shown that the use of a QRNG in combination with national cryptographic algorithms makes it possible to minimize the risk of key compromise, increase resistance to quantum attacks, and ensure long-term cryptographic reliability of communication channels and the web service of the Autonomous DGNSS. The proposed methodology allows creating PKI for information-and-communication systems of differential correction and other segments of critical infrastructure with increased requirements for trust, integrity, and confidentiality of data, taking into account the requirements of modern cybersecurity standards (ISO/IEC 27001, NIST, ND TZI).

Key words: global navigation satellite system (GNSS); quantum random number generator (QRNG); cyberattack; cybersecurity; differential GNSS (DGNSS); public key infrastructure (PKI); Autonomous DGNSS; post-quantum cryptography (PQC).

Постановка проблеми

Глобальні навігаційні супутникові системи (GNSS) є фундаментальним елементом критичної інфраструктури сучасної держави, забезпечуючи точне позиціонування та високоточну синхронізацію часу у сферах авіації, транспорту, геодезії, енергетики, телекомунікацій, фінансових операцій та інших видів діяльності [1, 2]. Водночас достовірність і захищеність сигналів GNSS обмежена низкою факторів, зокрема впливом кібератак на супутникові сигнали. Для підвищення точності координат і часу в реальному часі застосовуються системи диференціальної корекції (DGNSS), які обчислюють поправки на основі даних, отриманих від контрольно-коригуючих станцій (ККС) та передають їх користувачам через IP-мережі. Системи диференціальної корекції, зокрема автономні системи (АСДК), є програмно-апаратними комплексами [1, 3] та відповідно до Закону України «Про критичну інфраструктуру» № 1882-IX належать до об'єктів критичної інфраструктури у сфері космічних технологій [4].

Через стрімке зростання кіберзагроз за останні роки, спрямованих на системи GNSS [5, 6], а також на пов'язані з ними ІКС оброблення даних [7, 8, 9], стає необхідним переосмислення підходів до захисту ІКС АСДК на криптографічному рівні. До найнебезпечніших загроз належать GPS jamming, GPS spoofing, атаки на програмне забезпечення ККС і серверів оброблення даних, DDoS-атаки, MITM-атаки, компрометація ключів, ін'єкції коду, несанкціоноване втручання у корекційні дані, а також фізичні атаки та інсайдерські порушення. Особливо критичною є можливість віддаленої компрометації каналів передавання даних між ККС та АСДК, підміна або модифікація корекційних даних, а також отримання зловмисниками доступу до криптографічних матеріалів і сертифікатів. Сукупність таких атак може призвести до спотворення координат, втрати часової синхронізації, маніпулювання навігаційними даними або повного порушення роботи GNSS, що становить загрозу для об'єктів критичної інфраструктури та може мати критичні наслідки на державному рівні [10].

Ситуація додатково ускладнюється потенціальними кіберзагрозами у постквантовий період [11]. Із розвитком квантових обчислень класичні криптографічні алгоритми, що використовуються для автентифікації, підпису та встановлення захищених з'єднань, можуть втратити стійкість. Це безпосередньо стосується систем АСДК, у яких критично важливі функції: ідентифікація ККС, захист даних, які передаються, підпис корекцій, захист веб-сервісів та захист даних у стані спокою – повністю залежать від надійності інфраструктури відкритих ключів (ІВК). Таким чином, виникає необхідність побудови постквантової ІВК, стійкої до квантових атак, з використанням квантових джерел ентропії (QRNG) та національних криптоалгоритмів (ДСТУ 7564/7624/8961/9212), яка забезпечуватиме довготривалий захист, автентичність джерел корекцій та унеможливлуватиме компрометацію ключів у майбутньому криптографічному середовищі ІКС АСДК.

У зв'язку з викладеним постає науково-прикладна задача: розробити формалізовану модель постквантової ІБК з використанням QRNG для автономної системи диференціальної корекції, яка забезпечуватиме автентифікацію, цілісність даних, захист тунельних каналів та довіру до серверних компонентів без зміни базової архітектури АСДК.

Аналіз останніх досліджень і публікацій

Проблематиці кіберзахисту інформаційно-комунікаційних систем (ІКС) критичної інфраструктури в постквантовий період присвячені численні наукові праці як закордонних, так і вітчизняних вчених. Ключовим викликом сьогодення є загроза, яку становить квантовий комп'ютер для класичних алгоритмів асиметричного шифрування (RSA, ECC), що обґрунтовує проактивний перехід на постквантові криптографічні алгоритми [12]. За результатами наведених досліджень, перехід до квантово-стійких рішень стає пріоритетним завданням для держав та організацій ще до настання ери fault-tolerant квантових обчислень. Зокрема, науковці наголошують, що вже зараз критична інфраструктура є об'єктом кібератак, а поява квантових обчислювачів тільки посилює ці ризики.

Стає очевидним, що постквантова криптографія має стати невід'ємною частиною кіберзахисту ІКС у найближче десятиріччя [13]. А дослідження, згідно з [14] повідомляють, що на сьогоднішній день вже затверджено перші міжнародні стандарти (наприклад такі, як CRYSTALS-Kyber, Dilithium) для широкого впровадження.

Одним із перспективних шляхів підвищення криптостійкості ІКС є використання квантових генераторів випадкових чисел (QRNG) в ІБК. На відміну від псевдовипадкових генераторів випадкових чисел (PRNG), апаратні QRNG генерують істинні випадкові бітові послідовності на основі квантових явищ, що унеможлиблює їх передбачення навіть за наявності суперобчислювальних систем. Як зазначається в [15], впровадження QRNG здатне покращити будь-яку криптосистему, адже якість ключів прямо визначає криптостійкість. Українські науковці також акцентують на тому, що веб-сервіси QRNG відкривають нові можливості підвищення надійності ІБК, забезпечуючи високий рівень непередбачуваності ключових даних [16].

З огляду на криптозахист систем диференціальної корекції GNSS, такі системи традиційно мають відкритий формат передачі даних, що робить їх потенційною цілью для кібератак типу GPS spoofing. Для вирішення цієї проблеми пропонуються криптографічні механізми автентифікації навігаційних повідомлень. Зокрема, у сигнал можна вбудовувати спеціальні маркери (цифрові підписи, Message Authentication Codes тощо), за якими приймач перевірятиме, чи походить повідомлення від легітимного супутника і чи не було воно спотворене. Загальний висновок – додавання криптографічної автентифікації є дієвим засобом протидії кібератакам типу GPS spoofing без суттєвого погіршення якості сервісу [17]. Дослідження [11, 18] підтверджують, що постквантовий захист для DGNSS доцільно будувати на симетричних схемах (TESLA, HMAC) та геш-ланцюжках, оскільки класичні підписи на еліптичних кривих будуть зламані квантовими алгоритмами, а постквантові криптографічні схеми підпису занадто об'ємні для форматів GNSS-повідомлень. Автори показують, що найперспективнішим перехідним рішенням є протокол TESLA з квантово-стійкими геш-ланцюжками або Merkle-структурами, що забезпечує автентифікацію DGNSS-поправок без збільшення ширини каналу.

Українські симетричні стандарти «Калина» (ДСТУ 7624:2014) та «Купина» (ДСТУ 7564:2014) мають високу надійність у постквантовій ері, оскільки, згідно з теоретичними оцінками, квантовий пошук Гровера знижує їхню ефективну стійкість лише вдвічі, що компенсується використанням 256-бітних ключів і 512-бітних геш-образів [13]. Асиметричні алгоритми нового покоління – «Скеля» (ДСТУ 8961:2019, КЕМ на основі модульних алгебраїчних решіток) та «Вершина» (проект постквантового підпису, побудований на тих самих класах решіток) – демонструють рівні безпеки, співставні з CRYSTALS-Kyber та CRYSTALS-Dilithium, що підтверджено сучасними криптоаналітичними дослідженнями [14, 19, 20]. Окремі роботи також показали можливість інтеграції національних алгоритмів у міжнародні схеми постквантової криптографії, зокрема використання геш-функції Купина у модифікації SPHINCS+ [21]. У підсумку, комплекс українських алгоритмів (Калина, Купина, Скеля, Вершина) вже сьогодні може бути використаний як основа для постквантових ІКС, включно з ІКС критичної інфраструктури, зокрема Autonomous DGNSS.

Аналіз останніх досліджень і публікацій показує, що розвиток постквантової криптографії тісно пов'язаний із завданнями забезпечення кібербезпеки критичних ІКС. Проте, в контексті захисту систем диференціальної корекції GNSS (зокрема, АСДК) питання побудови стійких до квантових атак ІКС залишається недостатньо дослідженим. Таким чином, розробка постквантової ІБК з використанням квантових генераторів випадкових чисел (QRNG) та національних криптоалгоритмів для АСДК є надзвичайно актуальною.

Формулювання мети дослідження

Метою цієї роботи є теоретичне обґрунтування та розроблення формалізованої моделі постквантової інфраструктури відкритих ключів для автономної системи диференціальної корекції GNSS з використанням квантового генератора випадкових чисел, а також формулювання рекомендацій щодо її архітектури та застосування на основі аналізу сучасних наукових публікацій з урахуванням вимог стандартів ISO/IEC 27001, NIST SP 800-53/800-57, IEC 62443 та НД ТЗІ.

Викладення основного матеріалу дослідження

На рис. 1 наведено прототип архітектури інформаційно-комунікаційної системи автономної системи диференціальної корекції (ІКС АСДК). До складу ІКС АСДК входять три основні підсистеми: сегмент контрольно-коригуючих станцій (ККС), центри оброблення даних АСДК та сегмент користувачів.

– ККС, розташовані по території країни, які приймають сигнали GNSS, формують сирі навігаційні вимірювання та передають їх у ЦОД АСДК через захищені VPN-тунелі на основі IPsec. Основний та резервний ЦОД забезпечують агрегацію, оброблення й зберігання даних, розрахунок диференціальних поправок, а також формування коригуючих повідомлень для кінцевих користувачів.

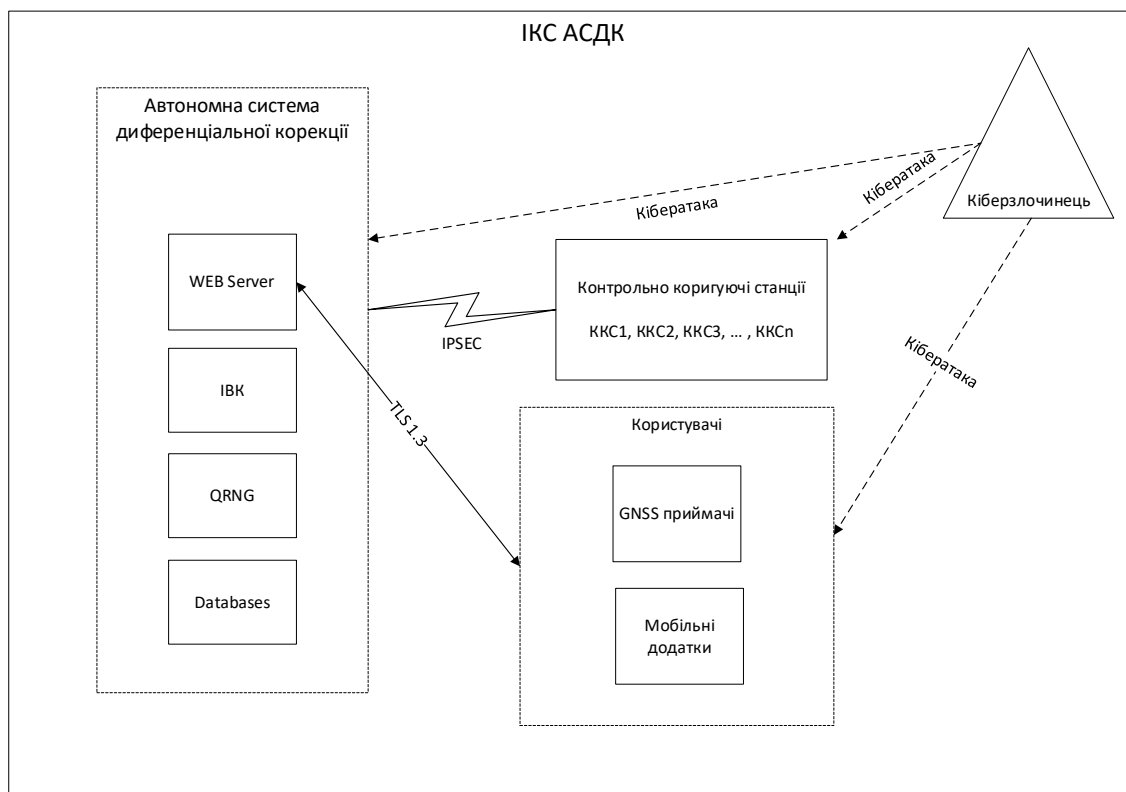


Рис 1. Прототип архітектури ІКС автономної системи диференціальної корекції GNSS

– До внутрішньої архітектури АСДК належать веб-сервер високої доступності, інфраструктура відкритих ключів (ІВК), квантовий генератор випадкових чисел (QRNG) та підсистема баз даних. Веб-сервер надає користувачам коригуючу інформацію через захищені канали TLS 1.3, використовуючи сертифікати, видані ІВК. QRNG інтегровано з компонентами ІВК та криптомодулями ЦОД для формування високоякісної ентропії під час генерації довготривалих і сеансових ключів.

– Сегмент користувачів включає GNSS-приймачі та мобільні додатки, які отримують диференціальні поправки через веб-сервіс АСДК. На рис. 1 пунктирними лініями також схематично показано можливі вектори кібератак на ККС, ЦОД, веб-сервер і кінцевих користувачів з боку зовнішнього порушника, що визначає ІКС АСДК як об'єкт побудови постквантової ІВК.

З урахуванням архітектури, наведеної на рис. 1, ІКС АСДК розглядається як розподілена система, у якій взаємодіють контрольно-коригуючі станції, центри оброблення даних, веб-сервер та кінцеві користувачі. Для кожного з цих сегментів характерні специфічні кіберзагрози, пов'язані як із впливом на сигнали GNSS (GPS jamming, GPS spoofing), так і з атаками на мережеву інфраструктуру та програмне забезпечення ІКС. Окрему небезпеку становить компрометація криптографічних ключів і сертифікатів, що використовується для захисту тунельних каналів між ККС та ЦОД, а також для автентифікації веб-сервісу АСДК та підпису корекційних даних.

У контексті постквантових загроз інфраструктура відкритих ключів АСДК повинна забезпечувати не лише поточну стійкість до класичних атак, а й довготривалий захист ключових матеріалів у разі появи повнофункціональних квантових обчислювачів. Це потребує переходу від класичних алгоритмів асиметричної криптографії (RSA, ECC) до постквантових схем на основі алгебраїчних решіток (ДСТУ «Скеля», «Вершина») та використання квантових генераторів випадкових чисел (QRNG) як джерела високоякісної ентропії для генерації довготривалих

і сеансових ключів. Додатково ІВК має підтримувати симетричні алгоритми «Калина» та геш-функцію «Купина», які вважаються криптостійкими в постквантовій ері завдяки достатній довжині ключів і геш-образів.

На основі аналізу загроз і стандартів кібербезпеки (ISO/IEC 27001, NIST SP 800-53/800-57, IEC 62443, НД ТЗІ) до постквантової ІВК ІКС АСДК формуються такі ключові вимоги:

- автентифікація суб'єктів ІКС: надійна ідентифікація ККС, серверів ЦОД, веб-сервера та користувачів на основі сертифікатів, виданих постквантовою ІВК АСДК;
- захист каналів зв'язку: встановлення криптографічно стійких тунельних каналів між ККС та ЦОД (IPsec/VPN) та захищених з'єднань між веб-сервером і клієнтами (TLS 1.3) з використанням постквантових механізмів узгодження ключів;
- забезпечення цілісності корекційних даних: використання постквантових алгоритмів підпису для коригуючих повідомлень, що розповсюджуються до GNSS-приймачів та мобільних додатків;
- довготривала криптографічна надійність: застосування QRNG для генерації ключових матеріалів ІВК, суворе розмежування прав доступу до закритих ключів, використання захищених апаратних модулів зберігання (HSM) та регламентованих криптоперіодів;
- підтримка процедур управління життєвим циклом сертифікатів: реєстрація, видача, оновлення, відкликання та публікація статусу сертифікатів (CRL/OCSP) з урахуванням вимог до безперервності надання сервісів АСДК.

Спираючись на класичні підходи до побудови інфраструктур відкритих ключів [22], а також на результати досліджень комбінованих ІВК [23, 24], запропоновано формалізовану модель постквантової інфраструктури відкритих ключів для інформаційно-комунікаційної системи автономної системи диференціальної корекції GNSS (Наведено на Рис. 2).

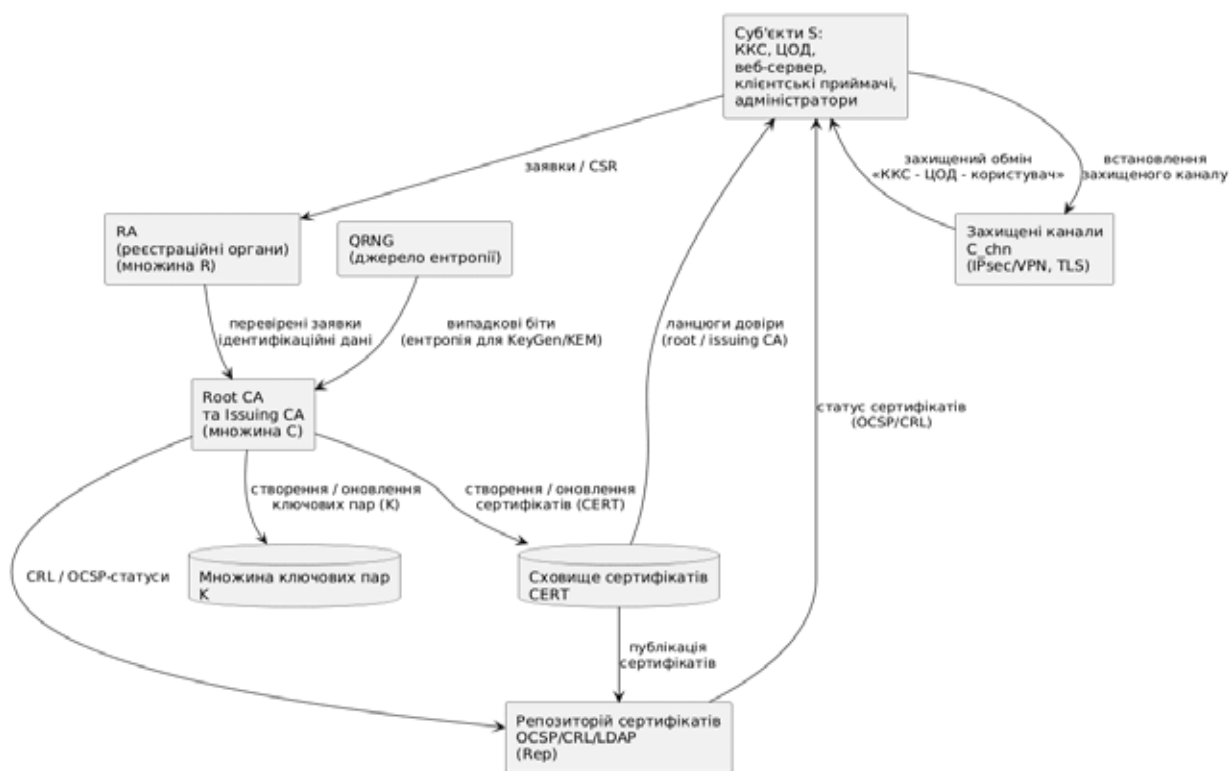


Рис. 2. Формалізована модель постквантової ІВК АСДК – $PKI_{PQ-ASDK}$

Формально постквантову ІВК для ІКС АСДК можна подати у вигляді кортежу:

$$(PKI_{PQ-ASDK} = S, C, R, R_{ep}, A, K, CERT, C_{chn}, QRNG) \quad (1)$$

де S – множина суб'єктів (контрольно-коригуючі станції, серверні компоненти основного та резервного ЦОД, веб-сервер, клієнтські приймачі та адміністратори/оператори),

C – множина засвідчувальних центрів (Root CA та Issuing CA),

R – реєстраційні органи (усі операції реєстрації нових контрольно-коригуючих станцій, серверних компонентів та клієнтських приймачів в ІВК виконуються уповноваженими операторами через реєстраційні органи,

що відповідає вимогам до інформаційно-комунікаційних систем об'єктів критичної інфраструктури (Наведено в Рис. 3)),

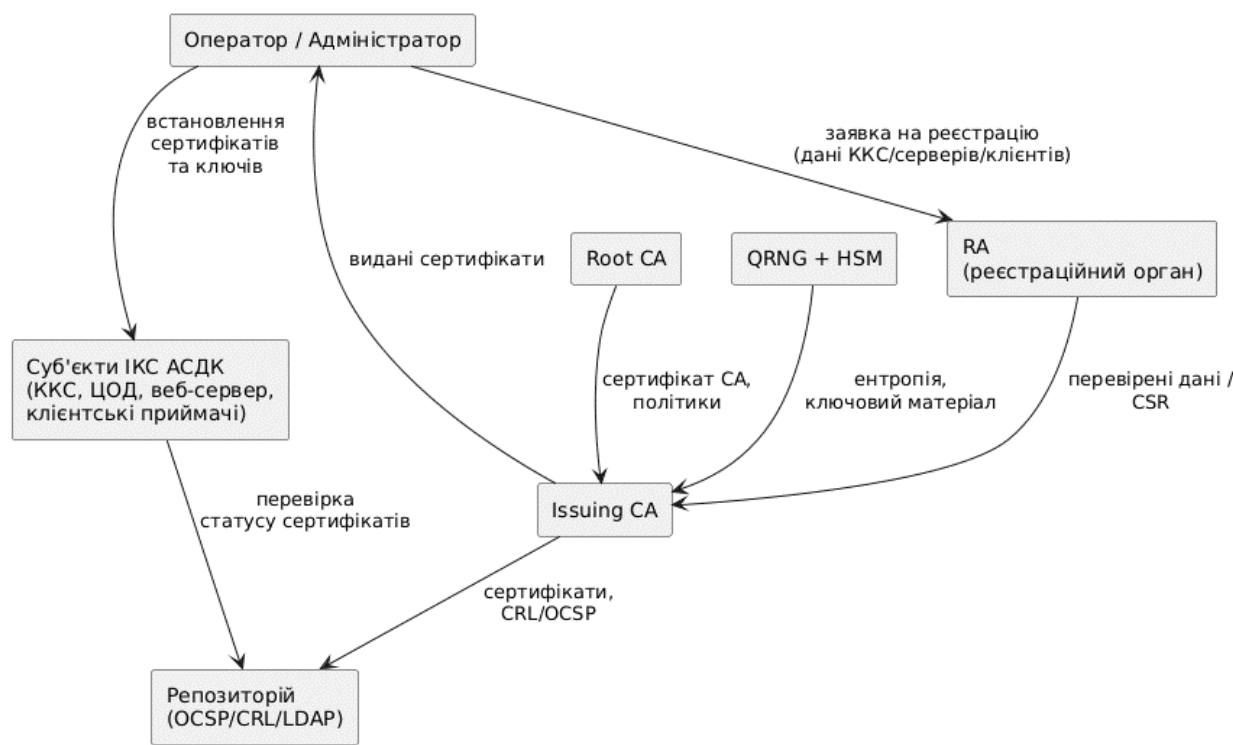


Рис. 3. Процес реєстрації нових ККС/серверів/користувачів АСДК в $PKI_{PQ-ASDK}$

Rep – репозитарії сертифікатів та їх статусів (LDAP/OCSP/CRL),

A – множина національних постквантових криптоалгоритмів («Купина», «Калина», «Скеля», «Вершина»),

K – множина ключових пар,

$CERT$ – множина сертифікатів,

$Schn$ – множина захищених каналів «ККС – ЦОД – користувач»,

$QRNG$ – квантовий генератор випадкових чисел.

На відміну від класичних ІБК [22-24], у запропонованій моделі квантовий генератор випадкових чисел розглядається як базовий криптографічний ресурс. Його роботу можна описати відображенням:

$$(r = QRNG(n) \in \{0,1\}^n) \quad (2)$$

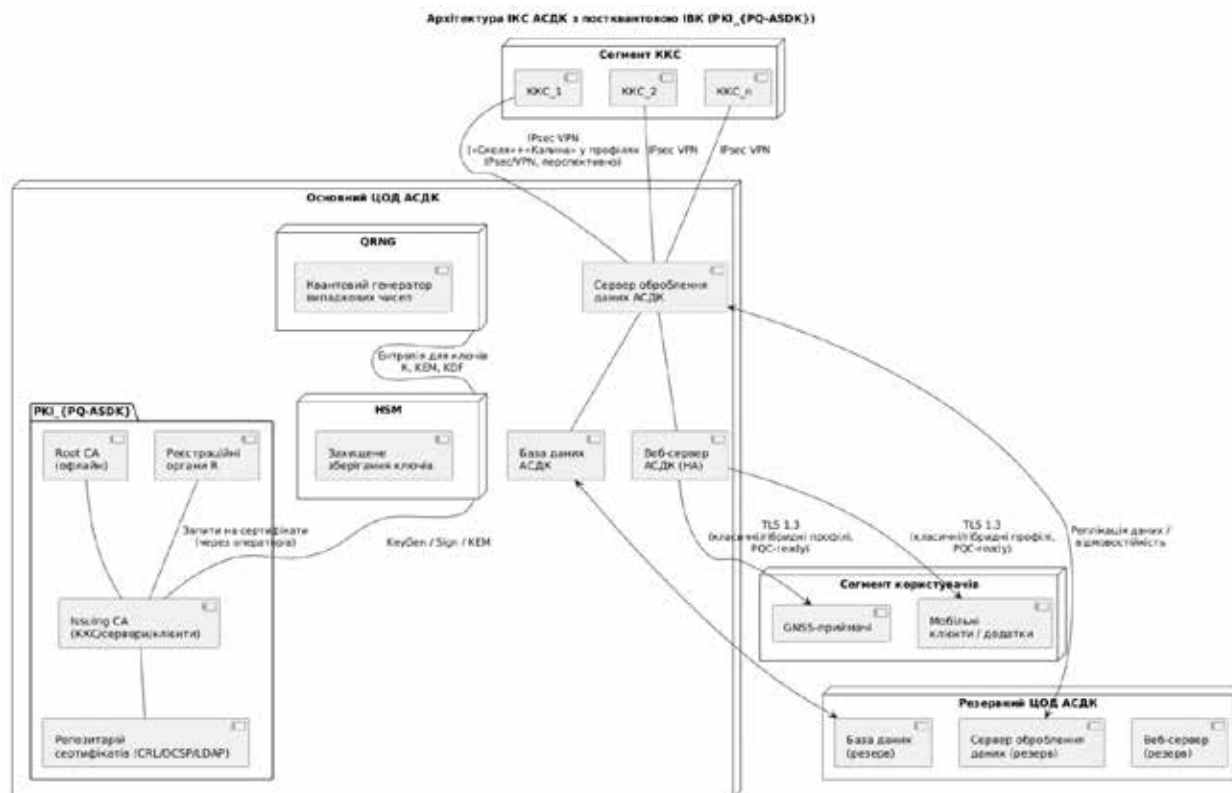
де n – довжина необхідної бітової послідовності, а r використовується як джерело ентропії для генерації довготривалих та сеансових ключів. Довготривалі ключові пари формуються за постквантовими алгоритмами «Скеля» (КЕМ) та «Вершина» (цифровий підпис), тоді як сеансові ключі симетричного шифрування обчислюються за схемою:

$$(k_{sess} = KDF(QRNG(n_{sess}) || context)) \quad (3)$$

де $context$ включає ідентифікатори сторін, параметри тунелю та службові дані протоколів TLS/IPsec.

Співвідношення (1)-(3) є авторським узагальненням відомих моделей класичних та комбінованих інфраструктур відкритих ключів, наведених у працях [22-24], адаптованим до умов постквантових загроз, специфіки автономної системи диференціальної корекції GNSS та вимог національних криптографічних стандартів. Вона створює підґрунтя для забезпечення довготривалої криптографічної стійкості каналів «ККС – ЦОД – користувач» та веб-сервісу АСДК без зміни базової архітектури системи.

На Рис. 4 представлено як $PKI_{PQ-ASDK}$ архітектурно розгортається в основному та резервному ЦОД АСДК і включає кореневий засвідчувальний центр в офлайн-режимі, підлеглі Issuing CA для різних доменів довіри (ККС, сервери ЦОД і веб-сервера, клієнтські приймачі), реєстраційні органи, служби OCSP/CRL та каталог сертифікатів. QRNG інтегрується з апаратними модулями захисту (HSM), які забезпечують генерацію та захищене зберігання ключового матеріалу.

Рис. 4. Архітектура ІКС АСДК з постквантовою ІВК $PKI_{PQ-ASDK}$

Захищені тунелі між ККС та ЦОД, а також внутрішні канали ЦОД будуються за схемою «Скеля» + «Калина» у профілях IPsec/VPN, тоді як TLS-з'єднання з веб-сервісом АСДК використовують сертифікати, видані постквантовою ІВК, і можуть бути реалізовані як на основі чинних гібридних PQC-профіль TLS 1.3, так і в перспективі – з підтримкою алгоритмів «Скеля» та «Калина» після їх стандартизації в стеку TLS. Корекційні повідомлення та службові артефакти підписуються за алгоритмом «Вершина» і додатково хешуються «Купиною».

Висновки

У роботі обґрунтовано актуальність побудови постквантової інфраструктури відкритих ключів для інформаційно-комунікаційної системи автономної системи диференціальної корекції GNSS як об'єкта критичної інфраструктури. Показано, що класичні криптографічні механізми автентифікації, підпису та встановлення захищених з'єднань є вразливими до квантових атак, а компрометація каналів «ККС – ЦОД – користувач» та криптографічних матеріалів може призвести до спотворення координат, втрати часової синхронізації й порушення роботи систем GNSS.

Запропоновано формалізовану модель постквантової ІВК – $PKI_{PQ-ASDK}$ для ІКС АСДК у вигляді кортежу, яка узагальнює класичні підходи до побудови ІВК і адаптує їх до специфіки Autonomous DGNSS. Модель явно враховує множину суб'єктів (ККС, ЦОД, веб-сервер, клієнти, адміністратори), ієрархію засвідчувальних центрів (Root CA, Issuing CA), реєстраційні органи, репозитарії сертифікатів, захищені канали та інтеграцію квантового генератора випадкових чисел (QRNG). Розроблено архітектурну схему розгортання постквантової ІВК в основному та резервному ЦОД АСДК з виділенням доменів довіри для ККС, серверів ЦОД, веб-сервера та клієнтських приймачів. Показано, як кореневий та підлеглі засвідчувальні центри, RA, служби OCSP/CRL/LDAP, HSM та QRNG інтегруються в єдину інфраструктуру, сумісну з вимогами стандартів ISO/IEC 27001, NIST SP 800-53/800-57, IEC 62443 та НД ТЗІ. Продемонстровано можливість повної побудови криптографічних механізмів ІВК на основі національних постквантових стандартів: «Купина» (ДСТУ 7564:2014), «Калина» (ДСТУ 7624:2014), «Скеля» (ДСТУ 8961:2019) та «Вершина» (ДСТУ 9212:2023). Показано, що використання QRNG як джерела ентропії для процедур KeyGen, KEM, KDF та цифрового підпису дозволяє мінімізувати ризик компрометації ключів, підвищити стійкість до квантових атак та забезпечити довготривалу криптографічну надійність каналів зв'язку і веб-сервісу АСДК без зміни базової функціональної архітектури системи.

Отримані результати мають прикладне значення для проєктування та модернізації національних систем диференціальної корекції GNSS і ширше – для ІКС об'єктів критичної інфраструктури. Запропонована модель може використовуватися як методологічна основа при побудові постквантових ІВК для навігаційних,

телекомунікаційних, енергетичних та інших систем, де критичними є довіра до джерел даних, цілісність інформації та стійкість до довгострокових криптографічних загроз.

Подальші дослідження доцільно зосередити на практичній оцінці продуктивності запропонованої постквантової ІВК в умовах реального часу для інформаційно-комунікаційної системи АСДК. Йдеться насамперед про вимірювання затримок (latency), пропускну здатності (throughput) та обчислювального навантаження під час використання національних криптографічних стандартів ДСТУ 7564:2014 («Купина»), ДСТУ 7624:2014 («Калина»), ДСТУ 8961:2019 («Скеля») та ДСТУ 9212:2023 («Вершина») у протоколах TLS 1.3 та IPsec (включно з експериментальними або модифікованими профілями, які підтримують алгоритми «Скеля» та «Калина») між контрольно-коригуючими станціями, центром оброблення даних і користувачами. Окремим напрямом розвитку роботи є формальне моделювання протоколів встановлення захищених сеансів (TLS/IPsec-handshake) з використанням алгоритмів «Скеля» та «Вершина», побудова їх формалізованих специфікацій і подальша верифікація властивостей безпеки (автентифікація, цілісність, конфіденційність, стійкість до post-quantum атак) із залученням сучасних методів формальної верифікації та інструментів моделювання.

Список використаної літератури

1. Trýb J., Hospodka J. GNSS Interference and Security: Impacts on Critical Infrastructure and Mitigation Strategies. *Procedia Computer Science*. Вип. 253, 2025. С. 2635-2644. DOI:10.1016/j.procs.2025.01.323.
2. Nicholas Brown 20240620-Hidden_Risk_Report. URL: https://ggim.un.org/UNGCE/documents/20240620-Hidden_Risk_Report.pdf.
3. Spanghero M., Papadimitratos P. UnReference: analysis of the effect of spoofing on RTK reference stations for connected rovers. *arXiv*, 2025. DOI:10.48550/arXiv.2503.20364.
4. Закон Ukrainy «Pro krytychnu infrastrukturu» № 1882-IX redaktsiya vid 21.09.2024 [Law of Ukraine "On Critical Infrastructure" No. 1882-IX, version dated 09/21/2024]. Official web portal of the Parliament of Ukraine. URL: <https://zakon.rada.gov.ua/go/1882-20>
5. Westbrook T. A. (2023) Taxonomy of Radio Frequency Jamming and Spoofing Strategies and Criminal Motives. *Journal of Strategic Security*, Vol. 16, no. 2. P. 68-80. DOI: <https://doi.org/10.5038/1944-0472.16.2.2081>.
6. Westbrook T. (2019) The Global Positioning System and Military Jamming: The geographies of electronic warfare. *Journal of Strategic Security*, Vol. 12, № 2. P. 1-16. DOI:10.5038/1944-0472.12.2.1720
7. Шумілова К. НАВІГАЦІЙНІ РИЗИКИ В АСПЕКТІ КІБЕРБЕЗПЕКИ ТРАНСПОРТНИХ СУДЕН І ВІЙСЬКОВИХ КОРАБЛІВ. *Scientific Collection «InterConf+»*. 24(121). С. 391-408. DOI:10.51582/interconf.19-20.08.2022.037.
8. Garmin outage caused by confirmed WastedLocker ransomware attack. (2020) BleepingComputer. URL: <https://www.bleepingcomputer.com/news/security/garmin-outage-caused-by-confirmed-wastedlocker-ransomware-attack/>
9. KA-SAT Network cyber attack overview. (2022) Viasat.com. 30.03.2022. URL: <https://news.viasat.com/blog/corporate/ka-sat-network-cyber-attack-overview>
10. Melnyk D. S. (2024) Creating a model of threats to Ukraine's national critical infrastructure as a basis for ensuring its security and resilience. *Bulletin of Kharkiv National University of Internal Affairs*, Vol. 104, 1 (Part 1). С. 237-250. DOI:10.32631/v.2024.1.20
11. Junquera-Sánchez J., Hernando-Ramiro C., Gamallo-Palomares Ó. et al. Assessment of cryptographic approaches for a quantum-resistant Galileo OSNMA. *NAVIGATION: Journal of the Institute of Navigation*. Vol. 71, Issue 2. P. navi.648. DOI:10.33012/navi.648.
12. Peña P. A. Quantum randomness reinforces post-quantum cryptography to safeguard large enterprises in the quantum-safe era. *QSNP*. 17.09.2025. URL: <https://qsnp.eu/quantum-randomness-reinforces-post-quantum-cryptography-to-safeguard-large-enterprises-in-the-quantum-safe-era/>.
13. Moral J. O. del, iOlus A. deMarti, Vidal G. et al. Cybersecurity in Critical Infrastructures: A Post-Quantum Cryptography Perspective. *arXiv*, 2024. DOI:10.48550/arXiv.2401.03780.
14. Gorbenko I., Kandii S. National and International Post-Quantum Standards for Asymmetric Transformations. *Cybernetics and Systems Analysis*. Vol. 61, 02.08.2025. DOI:10.1007/s10559-025-00800-z.
15. Krelina M. Quantum technology for military applications. *EPJ Quantum Technology*. Вип. 8, № 1. С. 24. DOI:10.1140/epjqt/s40507-021-00113-y.
16. Morhul D. M., Nariezhnii O. P., Hrinenko T. O. Класифікація атак та вимоги кібербезпеки до веб-ресурсу QRNG. *Radiotekhnika*. № 220. С. 50-57. DOI:10.30837/rt.2025.1.220.04.
17. Chen Y., Gao W., Chen X. et al. Advances of SBAS authentication technologies. *Satellite Navigation*. Vol. 2, Issue 1. P. 12. DOI:10.1186/s43020-021-00043-1.
18. E-GIANTS Project Concludes Study on GNSS Authentication and Security Improvements | EU Agency for the Space Programme. 05.08.2025. URL: <https://www.euspa.europa.eu/newsroom-events/news/e-giants-project-concludes-study-gnss-authentication-and-security-improvements>.

19. Gorbenko I. D., Kachko Y. G., Yesina M. V. та ін. Порівняльна характеристика алгоритмів інкапсуляції ключів Crystals-Kyber та Склея (ДСТУ 8961-2019). Radiotekhnika. № 210. С. 7-21. DOI:10.30837/rt.2022.3.210.01.
20. Fesenko A., Lytvynenko Y. Cryptanalysis of the «Vershyna» digital signature algorithm. Theoretical and Applied Cybersecurity. Вип. 5, № 2. DOI:10.20535/tacs.2664-29132023.2.288499.
21. Televnyi D. The Kupyna hash function application to SPHINCS+ signatures. Radiotekhnika. № 198. С. 215-219. DOI:10.30837/rt.2019.3.198.17.
22. Інфраструктури відкритих ключів. Електронний цифровий підпис. Теорія та практика :: Державний університет інформаційно-комунікаційних технологій. URL: <https://duikt.edu.ua/ua/lib/1/category/2434/view/1822>.
23. Горбенко І. Д., Кравченко П. О. Комбінована інфраструктура відкритих ключів та її застосування. Радіоелектронні і комп'ютерні системи. Issue 5. Р. 86-90. Also available online, URL: http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?I21DBN=LINK&P21DBN=UJRN&Z21ID=&S21REF=10&S21CNR=20&S21STN=1&S21FMT=ASP_meta&C21COM=S&2_S21P03=FILEA=&2_S21STR=recs_2009_5_17.
24. Горбенко І. Д., Халімов Г. З. Розвиток, стандартизація, уніфікація, удосконалення та впровадження інфраструктури відкритих ключів (включаючи національну систему електронного цифрового підпису) на внутрішньо-державному та міжнародному рівнях. 2012. URL: <http://openarchive.nure.ua/handle/document/1064>.

References

1. Tryb, J., & Hospodka, J. (2025). GNSS interference and security: Impacts on critical infrastructure and mitigation strategies. *Procedia Computer Science*, 253, 2635-2644. <https://doi.org/10.1016/j.procs.2025.01.323>
2. Brown, N. (2024, June 20). Hidden risk report. UN-GGIM. [https://ggim.un.org/UNGGCE/documents/20240620-Hidden_Risk_Report.pdf](https://ggim.un.org/UNGGCE/documents/20240620-Hidden_Risk_Report.pdf)
3. Spanghero, M., & Papadimitratos, P. (2025). UnReference: Analysis of the effect of spoofing on RTK reference stations for connected rovers. *arXiv*. <https://doi.org/10.48550/arXiv.2503.20364>
4. Verkhovna Rada of Ukraine. (2024). Law of Ukraine on Critical Infrastructure No. 1882-IX. Official web portal of the Parliament of Ukraine. <https://zakon.rada.gov.ua/go/1882-20>
5. Westbrook, T. A. (2023). Taxonomy of radio frequency jamming and spoofing strategies and criminal motives. *Journal of Strategic Security*, 16(2), 68-80. <https://doi.org/10.5038/1944-0472.16.2.2081>
6. Westbrook, T. (2019). The Global Positioning System and military jamming: The geographies of electronic warfare. *Journal of Strategic Security*, 12(2), 1-16. <https://doi.org/10.5038/1944-0472.12.2.1720>
7. Shumilova, K. (2022). Navigatsiini ryzyky v aspekti kiberbezpeky transportnykh suden i viiskovykh korabliv [Navigation risks in the aspect of cybersecurity of transport vessels and warships]. *Scientific Collection "InterConf+", 24(121)*, 391-408. <https://doi.org/10.51582/interconf.19-20.08.2022.037>
8. BleepingComputer. (2020). Garmin outage caused by confirmed WastedLocker ransomware attack. <https://www.bleepingcomputer.com/news/security/garmin-outage-caused-by-confirmed-wastedlocker-ransomware-attack/>
9. Viasat. (2022, March 30). KA-SAT network cyber attack overview. <https://news.viasat.com/blog/corporate/ka-sat-network-cyber-attack-overview>
10. Melnyk, D. S. (2024). Creating a model of threats to Ukraine's national critical infrastructure as a basis for ensuring its security and resilience. *Bulletin of Kharkiv National University of Internal Affairs*, 104(1), 237-250. <https://doi.org/10.32631/v.2024.1.20>
11. Junquera-Sanchez, J., Hernando-Ramiro, C., Gamallo-Palomares, O., et al. (2024). Assessment of cryptographic approaches for a quantum-resistant Galileo OSNMA. *NAVIGATION: Journal of the Institute of Navigation*, 71(2). <https://doi.org/10.33012/navi.648>
12. Pena, P. A. (2025, September 17). Quantum randomness reinforces post-quantum cryptography to safeguard large enterprises in the quantum-safe era. *QSNP*. <https://qsnp.eu/quantum-randomness-reinforces-post-quantum-cryptography-to-safeguard-large-enterprises-in-the-quantum-safe-era/>
13. Del Moral, J. O., de Marti i Olius, A., Vidal, G., et al. (2024). Cybersecurity in critical infrastructures: A post-quantum cryptography perspective. *arXiv*. <https://doi.org/10.48550/arXiv.2401.03780>
14. Gorbenko, I., & Kandii, S. (2025). National and international post-quantum standards for asymmetric transformations. *Cybernetics and Systems Analysis*, 61. <https://doi.org/10.1007/s10559-025-00800-z>

15. Krelina, M. (2021). Quantum technology for military applications. EPJ Quantum Technology, 8(1), 24. <https://doi.org/10.1140/epjqt/s40507-021-00113-y>
16. Morhul, D. M., Nariiezhnii, O. P., & Hrinenko, T. O. (2025). Kласифікація атак та вимоги кібербезпеки до веб-ресурсу QRNG [Classification of attacks and cybersecurity requirements for QRNG web resource]. Radiotekhnika, 220, 50-57. <https://doi.org/10.30837/rt.2025.1.220.04>
17. Chen, Y., Gao, W., Chen, X., et al. (2021). Advances of SBAS authentication technologies. Satellite Navigation, 2(1), 12. <https://doi.org/10.1186/s43020-021-00043-1>
18. EU Agency for the Space Programme. (2025, August 5). E-GIANTS project concludes study on GNSS authentication and security improvements. <https://www.euspa.europa.eu/newsroom-events/news/e-giants-project-concludes-study-gnss-authentication-and-security-improvements>
19. Gorbenko, I. D., Kachko, Y. G., Yesina, M. V., et al. (2022). Porivnialna kharakterystyka alhorytmiv inkapsuliatsii kliuchiv Crystals-Kyber ta Skelia (DSTU 8961-2019) [Comparative characteristic of key encapsulation algorithms Crystals-Kyber and Skelia (DSTU 8961-2019)]. Radiotekhnika, 210, 7-21. <https://doi.org/10.30837/rt.2022.3.210.01>
20. Fesenko, A., & Lytvynenko, Y. (2023). Cryptanalysis of the "Vershyna" digital signature algorithm. Theoretical and Applied Cybersecurity, 5(2). <https://doi.org/10.20535/tacs.2664-29132023.2.288499>
21. Televnyi, D. (2019). The Kupyna hash function application to SPHINCS+ signatures. Radiotekhnika, 198, 215-219. <https://doi.org/10.30837/rt.2019.3.198.17>
22. State University of Information and Communication Technologies. (n.d.). Infrastruktury vidkrytykh kliuchiv. Elektroni y tsyfrov y pidpys. Teoriia ta praktyka [Public key infrastructures. Electronic digital signature. Theory and practice]. <https://duikt.edu.ua/ua/lib/1/category/2434/view/1822>
23. Gorbenko, I. D., & Kravchenko, P. O. (2009). Kombinovana infrastruktura vidkrytykh kliuchiv ta y ii zastosuvannia [Combined public key infrastructure and its application]. Radioelectronic and Computer Systems, 5, 86-90. [http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?I21DBN=LINK&P21DBN=UJRN&Z21ID=&S21REF=10&S21CNR=20&S21STN=1&S21FMT=ASP_meta&C21COM=S&2_S21P03=FILA=&2_S21STR=recs_2009_5_17](http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?I21DBN=LINK&P21DBN=UJRN&Z21ID=&S21REF=10&S21CNR=20&S21STN=1&S21FMT=ASP_meta&C21COM=S&2_S21P03=FILA=&2_S21STR=recs_2009_5_17)
24. Gorbenko, I. D., & Khalimov, G. Z. (2012). Rozvytok, standartyzatsiia, unifikatsiia, udoskonalennia ta vprovadzhennia infrastruktury vidkrytykh kliuchiv na vnutrishnioderzhavnomu ta mizhnarodnomu rivniakh [Development, standardization, unification, improvement and implementation of public key infrastructure at domestic and international levels]. <http://openarchive.nure.ua/handle/document/1064>

Дата першого надходження рукопису до видання: 23.11.2025
Дата прийнятого до друку рукопису після рецензування: 19.12.2025
Дата публікації: 31.12.2025

T. A. TKACHENKO

Master's Student at the Department of Electronic Computing Machines
Kharkiv National University of Radio Electronics
ORCID: 0009-0001-1550-2586

V. O. MARTOVITSKY

Candidate of Technical Sciences, Associate Professor,
Associate Professor at the Department of Electronic Computing Machines
Kharkiv National University of Radio Electronics
ORCID: 0000-0003-2349-0578

N. M. BOLOHOVA

Associate Professor at the Department of Electronic Computing Machines
Kharkiv National University of Radio Electronics
ORCID: 0000-0001-8927-0055

METHODS OF OPTIMISING QUERIES IN DISTRIBUTED DATABASES

The article provides a comprehensive study of the problem of optimizing query performance in distributed database systems, where architectural complexity, information fragmentation, and the necessity for coordination between numerous nodes significantly affect processing speed and stability. It has been determined that conventional approaches to executing multi-table queries, especially when working with large historical data collections, do not provide acceptable response times and result in significant computational costs. In this context, it is relevant to implement specialized optimization methods that reduce the number of reading operations, the volume of intermediate calculations, and the load on the network infrastructure. The article proposes an integrated approach that combines caching, indexing, and materialization of representations to achieve the maximum performance when processing complicated analytical queries in a distributed environment. The caching provides rapid retrieval of results without accessing the disk structures, the indexing speeds up the search and filtering by creating the specialized data structures, and the materialization eliminates the need to repeat complex join and aggregation operations by storing the results in the form of physically accessible tables. For the experimental part of the research, a representative multi-table database was used, which contains information about the company's employees, their salaries, and their affiliation with departments and divisions. Five different variants of the same query were executed: without optimization, after caching, after indexing, using materialized representation, and using a combination of all three methods. The experimental results demonstrate that although each method separately provides a noticeable acceleration of performance, the greatest effect is achieved when they are used simultaneously. The query execution time was decreased from 2831.869 ms to 0.076 ms, which is an increase in performance by more than 37,000 times. The data obtained confirm the existence of a synergistic effect and demonstrate the high practical importance of comprehensive optimization for the systems that work with heavy analytical loads in distributed data processing conditions. The presented approach can be used to modernize existing information systems and increase their efficiency when scaling.

Key words: query optimisation, distributed databases, caching, materialisation, indexing, PostgreSQL

T. A. ТКАЧЕНКО

магістрант кафедри електронних обчислювальних машин
Харківський національний університет радіоелектроніки
ORCID: 0009-0001-1550-2586

В. О. МАРТОВИЦЬКИЙ

кандидат технічних наук, доцент,
доцент кафедри електронних обчислювальних машин
Харківський національний університет радіоелектроніки
ORCID: 0000-0003-2349-0578

Н. М. БОЛОГОВА

доцент кафедри електронних обчислювальних машин
Харківський національний університет радіоелектроніки
ORCID: 0000-0001-8927-0055

МЕТОДИ ОПТИМІЗАЦІЇ ЗАПИТІВ У РОЗПОДІЛЕНИХ БАЗАХ ДАНИХ

У статті здійснюється комплексне дослідження проблеми оптимізації продуктивності виконання запитів у розподілених системах управління базами даних, де складність архітектури, фрагментація інформації та необхідність координації між численними вузлами істотно впливають на швидкість та стійкість обробки. Визначено, що традиційні підходи до виконання багатотабличних запитів, особливо за умови роботи з великими історичними наборами даних, не забезпечують прийнятної часу відповіді та призводять до значних обчислювальних витрат. У цьому контексті актуальним є застосування спеціалізованих методів оптимізації, які зменшують кількість операцій читання, обсяг проміжних обчислень та навантаження на мережеву інфраструктуру. У роботі запропоновано інтегрований підхід, що поєднує кешування, індексацію та матеріалізацію представлень з метою досягнення максимальної швидкодії при обробці складних аналітичних запитів у розподіленому середовищі. Кешування забезпечує оперативне повторне отримання результатів без звернення до дискових структур, індексація прискорює пошук і фільтрацію за рахунок створення спеціалізованих структур даних, а матеріалізація виключає повторне виконання складних операцій об'єднання та агрегації, зберігаючи результати у формі фізично доступних таблиць. Для експериментальної частини використано реалістичну багатотабличну базу даних, що містить інформацію про працівників підприємства, їхні заробітні плати, належність до відділів і службових підрозділів. Проведено п'ять варіантів виконання одного й того самого запиту: без оптимізації, після кешування, після індексації, із застосуванням матеріалізованого представлення та за умови комбінованого використання всіх трьох методів. Результати експериментів демонструють, що хоча кожен метод окремо дає помітне прискорення, найбільшого ефекту досягнуто при їх одночасному застосуванні. Час виконання запиту скоротився з 2831.869 мс до 0.076 мс, що становить збільшення швидкодії більш ніж у 37000 разів. Отримані дані підтверджують існування синергетичного ефекту та свідчать про високу практичну цінність комплексної оптимізації для систем, що працюють з великими аналітичними навантаженнями в умовах розподіленої обробки даних. Представлений підхід може бути використаний для модернізації існуючих інформаційних систем і підвищення їхньої ефективності при масштабуванні.

Ключові слова: оптимізація запитів, розподілені бази даних, кешування, матеріалізація, індексація, PostgreSQL

Problem statement

Optimizing query performance in distributed database management systems is one of the main challenges for modern information systems. With the rapid growth of data volumes and increasing demands on processing speed, traditional approaches to working with databases no longer provide the required level of efficiency. This is particularly noticeable when executing complex analytical queries that require complex and costly operations such as joining, sorting, or aggregating large data sets.

Distributed databases are characterized by highly complex architecture and the need for effective coordination between different system nodes. In contrast to centralized systems, distributed databases implement the storage and processing of information on a set of interconnected nodes that function as a single logical system, despite their geographical distance. This architecture creates additional challenges related to network delays, data synchronization, and load balancing. When executing complex queries involving numerous table joins, nested subqueries, and sorting operations, system response times can reach several seconds, which is unacceptable for most business applications. Therefore, developing and implementing effective query optimization techniques is critical to ensuring the performance of distributed database management systems.

The problem of query optimization is particularly relevant for analytical platforms that process multidimensional queries over large historical data sets. Such tasks are typical for financial, telecommunications, and trading systems, where millions of records need to be processed and complex calculations performed in real time. Without the use of effective optimization techniques, such queries can take hours to execute, effectively rendering them unusable in practice.

A comprehensive approach to optimisation, which combines different techniques, can achieve the best results. The most effective of these are caching, indexing, and materialisation of views. Caching decreases the processing time of repeated queries by storing the results in fast memory. The indexing creates auxiliary structures that reduce the number of disk accesses and speed up searches. Materialisation, in turn, involves pre-calculating complex results and storing them for future use.

Using caching can significantly reduce the load on the database, especially in systems with a high proportion of repetitive queries. Materializing intermediate calculation results avoids repeating complex operations, which is especially important for analytical systems. Creating appropriate indexes can speed up search and filter operations by several orders of magnitude, turning a full table scan into a quick index search.

Distributed query processing requires an effective planning, taking into account data location, network bandwidth, and the current load on nodes. Suboptimal planning can lead to an excessive amount of data transfer, which becomes a major performance limitation.

Consequently, research and comparative analysis of various methods for optimizing queries in distributed databases is relevant, as it will enable the identification of the most effective strategies. Of particular interest is the study of the

synergistic effect of the combined application of various optimization methods. This research is of important practical significance for database developers and administrators, as it allows for improving the efficiency of systems that work with large volumes of information in a distributed environment.

Recent research and publications analysis

Modern scientific works demonstrate a constant growth of interest in the problem of query optimization in distributed database management systems. Many researchers focus on finding ways to reduce query execution time and improve resource utilization. The relevance of this topic is due to the continuous increase in the amount of information and the growing demands for its processing speed.

Research [1] demonstrates the importance of splitting complex queries to improve the performance of distributed databases. The authors show that decomposing queries into simpler components reduces the workload on individual system nodes and improves overall processing efficiency. Particular attention is paid to the analysis of distributed query execution algorithms and the minimization of inter-node data exchange. The researchers note that the effectiveness of this approach depends on the characteristics of the query, the volume of data, and the network topology, which determines the scalability and performance of the system.

The study [2] presents the QOTUM query optimizer designed for distributed databases in a cloud environment. The system uses a materialized view mechanism to reuse intermediate results and reduce query execution time. The results of the experiments showed a significant reduction in query execution time and memory usage. The authors also highlight the importance of considering cloud environment factors such as data distribution and network costs.

Sulima and Iermolaiev [3] examined approaches to optimizing SQL queries and demonstrated that even a slight change in the query structure can significantly affect its execution speed. The authors paid significant attention to the analysis of execution plans, the identification of “bottlenecks,” and the detection of operations that cause delays during data processing. The study proposed an optimization method that involves replacing the IN operator with a temporary table with a non-clustered index. This approach reduces the number of logical data accesses and speeds up the selection process. The results of the experiments showed an average reduction in query execution time of 15–17%. The authors also emphasize the necessity of regular analysis and optimization of SQL queries, especially in cases of increasing information volumes and system load.

The publication [4] presents an overview of the main methods for optimizing distributed databases. The authors systematized approaches to improving performance, including indexing, caching, fragmentation, replication, and load balancing. The connection between database architecture and system performance is examined, and practical recommendations are provided for selecting the optimal strategy depending on the type of load and data volume.

The paper [5] provides a detailed analysis of various database indexing techniques and their impact on query performance. The researchers examine the features of B-tree, hash, text, geospatial, and composite indexes, describing their strengths and limitations in detail. It is noted that the choice of the appropriate index type is determined by the database structure, data amount, and query specifics. The issue of balancing the number of indexes with the cost of updating them during data changes is considered separately. The authors conclude that the reasonable use of indexes is an important factor in improving the performance of database management systems.

The research [6] presents an empirical analysis of query optimization strategies aimed at reducing execution latency in large-scale databases. The authors consider both traditional and AI-oriented approaches, including the use of machine learning methods for query cost estimation, adaptive query execution, and intermediate result caching. The experimental part covers SQL, NoSQL, and distributed cloud databases using TPC-H, TPC-DS, and real queries. The results show that the combination of indexing, adaptive execution, and AI planning reduces query execution time by 45–50%. The authors conclude that the efficiency of optimization is determined by the type of load (OLTP or OLAP) and the system architecture.

The study [7] examines best practices for optimizing SQL databases for working with large volumes of data. The authors focus on partitioning, sharding, caching, and database schema optimization techniques. The article provides practical recommendations on combining different optimization methods to ensure scalability and high performance of systems in a Big Data environment.

In summarizing the results of recent studies, it can be noted that the problem of query optimization in distributed databases remains relevant, and most scientific works focus on improving individual methods of increasing performance. At the same time, there is a lack of comprehensive approaches in scientific sources that combine methods to achieve a synergistic effect. Therefore, research aimed at developing and experimentally substantiating combined approaches to query optimization is of great scientific and practical importance.

Formulating the research goal

The purpose of the research is to compare methods for optimizing complex analytical queries in distributed databases, including evaluating the effectiveness of caching, materialization of views, indexing, and their combination to improve query processing performance.

Materials and results

Optimization of queries in distributed databases is based on the use of methods that reduce computational costs and the amount of memory accesses. The most common of these are caching, indexing, and materialization of views. Their combination allows for a synergistic effect, where the total performance increase exceeds the effect of using each method separately.

Caching is based on reusing the results of previously executed queries. When the same query is repeated, the system receives data directly from RAM, eliminating the need to reread it from the disk. This is particularly effective for frequently executed queries. The level of caching efficiency is determined by the ratio between the frequency of repeated requests and the available RAM, which is described by the cache hit ratio.

Indexing involves creating additional data structures that enable search and filtering operations to be performed with logarithmic rather than linear complexity. The research uses B-tree-based indexes, which guarantee ordered storage of keys and stable access times to the required records. Such structures are self-balancing because they automatically maintain the optimal tree height when adding or removing elements. As a result, a search that previously required a complete table scan is now performed in a few disk accesses, which is especially important when working with large historical data sets.

Materialization of views involves storing the results of complex calculations in separate tables, eliminating the need to repeat the same join or aggregation operations during subsequent queries. Such views store pre-calculated results on disk, and the optimization system automatically determines the feasibility of reusing them. Data relevance is maintained using full or incremental update mechanisms, depending on changes in the base tables.

The combination of caching, indexing, and materialization forms a multi-level optimization system in which indexes speed up searches, materialization reduces the amount of computation, and caching provides quick access to previously received results. This approach is particularly effective in distributed systems, where data transfer between nodes is a key factor affecting the overall query execution time.

The study was conducted using the company's database, which contains information about employees, their salaries, and departments. The system architecture consists of several interconnected tables: employee, salary, dept_emp, department, and others that show the structure of the organization. The data covers many years, which makes it possible to evaluate the effectiveness of optimization methods for large amounts of information. The system architecture diagram is shown in the fig. 1.

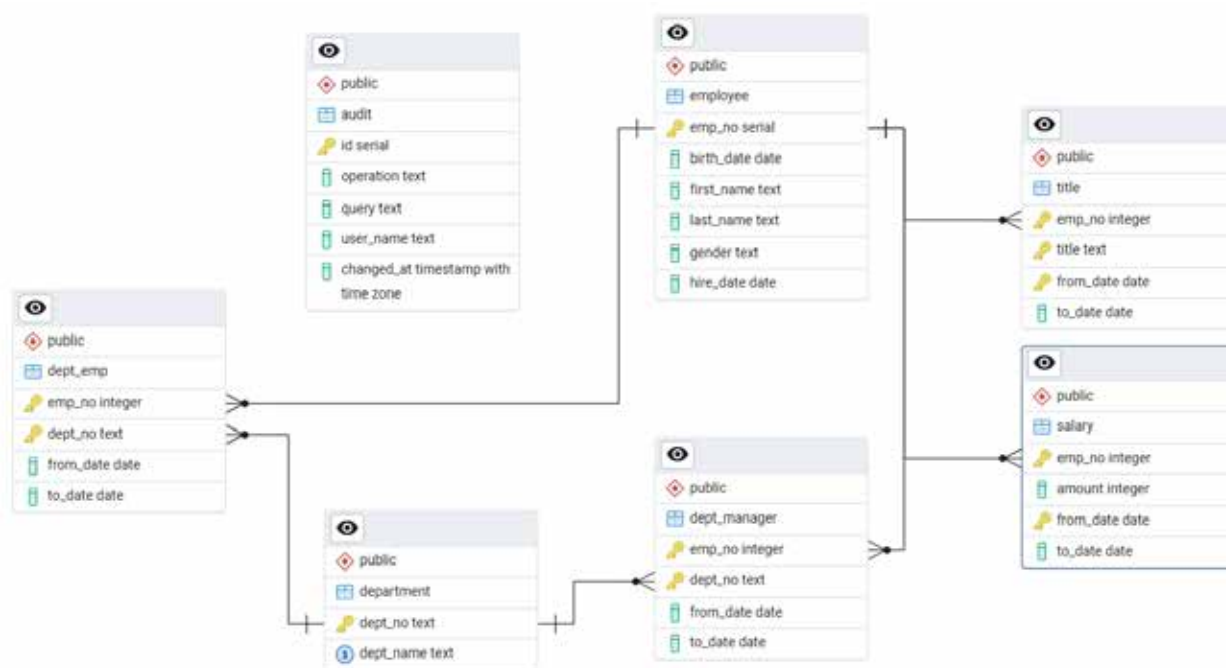


Fig. 1. Architecture of the experimental database

For the tests, we used an SQL query that combines several join and filter operations (Listing 1).

Listing 1 SQL query for conducting tests

```

WITH current_salaries AS (
  SELECT DISTINCT ON (emp_no) emp_no, salary
  FROM salary
  ORDER BY emp_no, from_date DESC
),
current_dept AS (
  SELECT DISTINCT ON (emp_no) emp_no, dept_no
  FROM dept_emp
  ORDER BY emp_no, from_date DESC
)
SELECT d.dept_name, e.emp_no, e.first_name, e.last_name, cs.salary
FROM employee e
JOIN current_salaries cs ON e.emp_no = cs.emp_no
JOIN current_dept cd ON cd.emp_no = e.emp_no
JOIN department d ON d.dept_no = cd.dept_no
WHERE e.hire_date <= DATE '2000-01-01'
ORDER BY d.dept_name, cs.salary DESC
LIMIT 100;

```

The query aggregates data from four tables, determines the current salary and department of each employee, filters the results by date of hire, and sorts them by salary and department. This structure allows us to accurately assess the impact of different optimization methods on execution time.

The results of the tests are shown in Table 1.

Table 1

Results of tests

№ of test	Optimization method	Description of conditions (cache/indexes/materialization)	Planning Time (ms)	Execution Time (ms)
1	Without optimization (baseline)	After restarting PostgreSQL, without indexes, without MV	3.441	2831.869
2	Caching	Re-executing the same query	0.552	2612.267
3	Indexing	Executing the base query after creating indexes	0.453	1574.183
4	Materialization	Executing through a materialized view without indexes	0.077	83.989
5	Combination of methods	Materialization + indexes + caching	0.111	0.076

Analysis of the results showed that caching reduces query execution time by 7.8%, indexing by 44.4%, and materialization by 97%. The best result was achieved by combining all three methods, which reduced the execution time from 2831.869 ms to 0.076 ms, i.e., approximately 37,000 times. The data obtained confirms the presence of a pronounced synergistic effect: each individual method accelerates only a certain stage of processing, while their combined use optimizes the system comprehensively. Fig. 2 shows a comparison of query execution times for different approaches, clearly demonstrating the advantages of combined optimization.

Conclusions

The research confirmed the effectiveness of a comprehensive approach to optimizing queries in distributed databases. Experimental results showed that the combination of caching, indexing, and materialization significantly reduces the execution time of complex analytical queries from 2831.869 ms to 0.076 ms, i.e., an acceleration of more than 37,000 times.

Each of the methods considered has its own area of effectiveness. Caching reduces the execution time of repetitive queries by using previously obtained results. Indexing improves the performance of search and join operations by eliminating the need for full table scans. Materialization stores the results of complex calculations in separate structures, preventing them from being re-executed.

The highest performance was demonstrated by the combined application of these methods, which creates a multi-level optimization system and provides a synergistic effect. The results obtained prove that a systematic approach to query optimization can significantly reduce computational costs and increase the scalability of distributed systems.

The practical significance of the study is the possibility of applying the proposed methodology by database developers and administrators to improve the efficiency of information systems, in particular those that process large amounts of data and perform analytical tasks.

Further research could focus on an in-depth analysis of the effectiveness of other optimization methods, as well as on studying various combinations of these approaches to achieve an optimal balance between speed, resource utilization, and system stability.

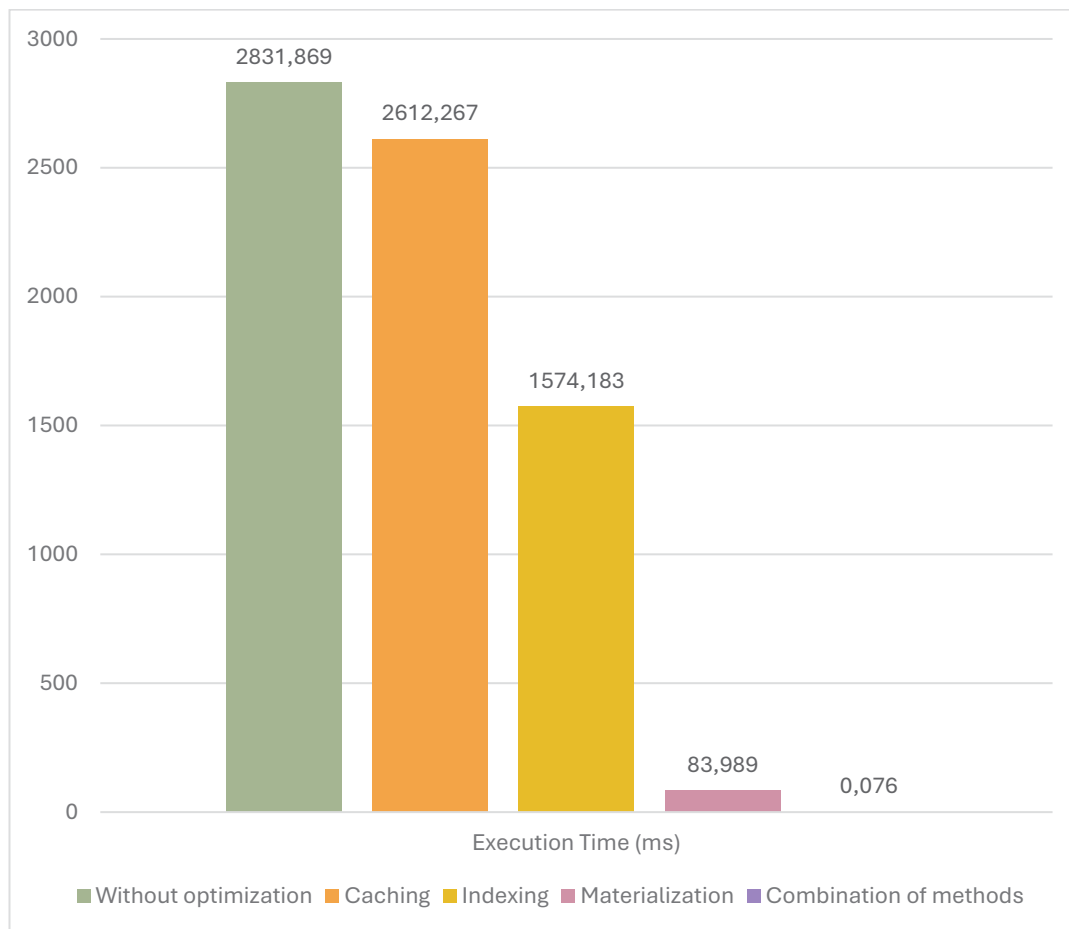


Fig. 2. Comparison of Execution Time for each method and for their combination

Bibliography

1. Панасюк В. В., Майданюк В. П. Оптимізація продуктивності розподілених баз даних через розділення складних запитів. *Матеріали LIV Всеукраїнської науково-технічної конференції підрозділів ВНТУ*. 2025. URL: <https://ir.lib.vntu.edu.ua/bitstream/handle/123456789/48587/23928.pdf?sequence=3&isAllowed=y> (дата звернення: 27.11.2025).
2. Archana B., Vilas K., Madhukar S. QOTUM: The Query Optimizer for Distributed Database in Cloud Environment. *TECHNICAL JOURNAL*. 2024. Т. 18, вип. 1. С. 172–177. ISSN 1846-6168 (Print), ISSN 1848-5588 (Online). URL: <https://doi.org/10.31803/tg-20230501083155>.
3. Суліма С. В., Єрмолаєв О. Д. Метод оптимізації SQL запитів системи управління базами даних. *Національний технічний університет України «КПІ імені Ігоря Сікорського»*. Київ, Україна, 2023. Т. 2, вип. 72 : Системи управління навігації та зв'язку Збірник наукових праць. С. 151–157. URL: <https://doi.org/10.26906/SUNZ.2023.2.151>.
4. Белоус Р., Крилов Є., Анікін В. Методи оптимізації запитів розподілених БД. *Міжвідомчий науково-технічний збірник*. 2021. Т. 2, вип. 39 : Адаптивні системи автоматичного управління. С. 3–11. ISSN 1560-8956. URL: <https://doi.org/10.20535/1560-8956.39.2021.247364>.
5. Голубінка В., Худий А. Підвищення продуктивності запитів до баз даних: аналіз технік індексації. *Вісник Національного університету «Львівська політехніка»*. Львів, Україна, 2024. вип. 15 : Інформаційні системи та мережі. С. 65–73. URL: <https://doi.org/10.23939/sisn2024.15.065>.
6. Stewart N., Adams D., Foster U. Empirical Analysis of Query Optimization Strategies for Reducing Execution Latency in Large-Scale Databases. *Journal of Innovation in Governance and business practices*. 2025. Т. 1, вип. 1. С. 88–117. URL: <https://jigbp.com/index.php/jigbp/article/view/4> (дата звернення: 27.11.2025).

7. Uzzaman, A., Jim, M. M. I., Nishat, N., & Nahar, J. Optimizing SQL databases for big data workloads: techniques and best practices. *Academic journal on business administration, innovation & sustainability*. 2024. Т. 4, вип. 3. С. 15–29. ISSN 2997-9552. URL: https://www.researchgate.net/profile/Janifer-Nahar/publication/381725561_OPTIMIZING_SQL_DATABASES_FORBIG_DATA_WORKLOADS_TECHNIQUES_AND_BEST_PRACTICES/links/667fcab2f3b61c4e2c99919b/OPTIMIZING-SQL-DATABASES-FORBIG-DATA-WORKLOADS-TECHNIQUES-AND-BEST-PRACTICES.pdf (дата звернення: 27.11.2025).

References

1. Panasiuk, V. V., & Maidaniuk, V. P. (2025). Optymizatsiia produktyvnosti rozpodilenykh baz danykh cherez rozdilennia skladnykh zapytiv [Optimization of distributed database performance through complex query partitioning]. *Materialy LIV Vseukrayins'koyi naukovo-tekhnichnoyi konferentsiyi pidrozdiliv VNTU*. <https://ir.lib.vntu.edu.ua/bitstream/handle/123456789/48587/23928.pdf?sequence=3&isAllowed=y>
2. Archana, B., Vilas, K., & Madhukar, S. (2024). QOTUM: The query optimizer for distributed database in cloud environment. *Technical Journal*, 18(1), 172–177. <https://doi.org/10.31803/tg-20230501083155>
3. Sulima, S. V., & Iermolaiev, O. D. (2023). Metod optymizatsii SQL zapytiv systemy upravlinnia bazamy danykh [Method for optimizing SQL queries in a database management system]. *Natsional'nyy tekhnichnyy universytet Ukrainy «KPI imeni Ihorya Sikors'koho»*, 2(72), 151–157. <https://doi.org/10.26906/SUNZ.2023.2.151>
4. Belous, R., Krylov, Ye., & Anikin, V. (2021). Metody optymizatsii zapytiv rozpodilenykh BD [Methods of optimization of distributed databases]. *Mizhvidomchyy naukovo-tekhnichnyy zbirnyk*, 2(39), 3–11. <https://doi.org/10.20535/1560-8956.39.2021.247364>
5. Holubinka, V., & Khudyi, A. (2024). Pidvyshchennia produktyvnosti zapytiv do baz danykh: analiz tekhnich indeksatsii [Enhancing database query performance: Analysis of indexing techniques]. *Вісник Національного університету “Львівська політехніка”*, 15, 65–73. <https://doi.org/10.23939/sisn2024.15.065>
6. Stewart, N., Adams, D., & Foster, U. (2025). Empirical analysis of query optimization strategies for reducing execution latency in large-scale databases. *Journal of Innovation in Governance and Business Practices*, 1(1), 88–117. Retrieved November 27, 2025, from <https://jigbp.com/index.php/jigbp/article/view/4>
7. Uzzaman, A., Jim, M. M. I., Nishat, N., & Nahar, J. (2024). Optimizing SQL databases for big data workloads: Techniques and best practices. *Academic Journal on Business Administration, Innovation & Sustainability*, 4(3), 15–29. Retrieved November 27, 2025, from https://www.researchgate.net/profile/Janifer-Nahar/publication/381725561_OPTIMIZING_SQL_DATABASES_FORBIG_DATA_WORKLOADS_TECHNIQUES_AND_BEST_PRACTICES/links/667fcab2f3b61c4e2c99919b/OPTIMIZING-SQL-DATABASES-FORBIG-DATA-WORKLOADS-TECHNIQUES-AND-BEST-PRACTICES.pdf

Дата першого надходження рукопису до видання: 26.11.2025
Дата прийнятого до друку рукопису після рецензування: 15.12.2025
Дата публікації: 31.12.2025

УДК 004.89:81'322.4:004.738.5

DOI <https://doi.org/10.35546/kntu2078-4481.2025.4.3.29>

Р. О. ТКАЧУК

асистент кафедри кібербезпеки та програмного забезпечення
Центральноукраїнський національний технічний університет
ORCID: 0000-0002-1984-0419

О. В. МУЛЯРЕВИЧ

кандидат технічних наук за спеціальністю «Комп'ютерні системи
та компоненти»,
доцент кафедри електронних обчислювальних машин
Інститут комп'ютерних технологій, автоматики та метрології
Національного університету «Львівська політехніка»
ORCID: 0000-0002-4644-7962

С. В. ВАСИЛЮК-ЗАЙЦЕВА

магістр філософії з фізики,
старший викладач кафедри комп'ютерних наук
Національний університет біоресурсів і природокористування України
ORCID: 0000-0002-0875-462X

АДАПТИВНІ АЛГОРИТМИ ОБРОБКИ ПРИРОДНОЇ МОВИ ДЛЯ БАГАТОМОВНОГО ІНТЕРФЕЙСУ КОРИСТУВАЧА

У світлі стрімкого технологічного розвитку, де можливості комунікації розширюються, процеси обробки природної мови забезпечують інтуїтивну взаємодію людини з технологіями. Обробки природної мови є революційним перетином інформатики, штучного інтелекту та лінгвістики, спрямований на подолання розриву між людським спілкуванням та розумінням обчислень. Його інтеграція в інтерфейси користувача була трансформаційною та забезпечила більш інтуїтивну та ефективну взаємодію між користувачами та технологіями. Складність природної мови з її нюансами, ідіомами та культурними варіаціями становить серйозну проблему для її обробки. А втім, прогрес у машинному та глибокому навчанні підняв цей процес на нові висоти, дозволяючи системам навчатися на великих обсягах даних і згодом удосконалюватися. Ця синергія відкрила нові шляхи доступності, зробивши технологію більш інклюзивною. Мета пропонованої статті – провести аналіз адаптивних алгоритмів обробки природної мови для багатомовного інтерфейсу користувача. У роботі використано метод аналізу дотичної літератури. Результати роботи продемонстрували, що завдяки прогресу в машинному та глибокому навчанні алгоритми обробки природної мови нині формують основи багатомовних інтерфейсів користувача (від голосових помічників, чат-ботів і систем автозаміни до аналізу настроїв, автоматичного перекладу, резюмування контенту й функцій доступності). Ці процеси роблять технології більш чутливими, адаптивними і здатними розуміти мову, контекст, емоції та культурні нюанси людського спілкування. Нейронні системи використовують великі обсяги даних, щоб навчитися виконувати завдання автоматично. У роботі описано напрями та можливості обробки природної мови в інтерфейсах користувача. Кожен із цих напрямів підкреслює багатогранність обробки природної мови в інтерфейсах користувача та її потенціал зробити технології більш чутливими до людських потреб; у майбутньому його розвиток сприятиме глибшому розумінню взаємовідносин технології та людини. Таким чином, впровадження обробки природної мови в багатомовні інтерфейси користувача відкриває перспективи створення більш природного, інклюзивного та інтелектуального цифрового середовища майбутнього.

Ключові слова: штучний інтелект, лінгвістика, природна мова, інтерфейси користувачів, цифрове середовище.

R. O. TKACHUK

Assistant Professor at the Department of Cybersecurity and Software
Central Ukrainian National Technical University
ORCID: 0000-0002-1984-0419

O. V. MULIAREVYCH

PhD in Computer Systems and Components,
Associate Professor at the Department of Computer Engineering
Institute of Computer Technologies, Automation and Metrology
of Lviv Polytechnic National University
ORCID: 0000-0002-4644-7962

S. V. VASYLYUK-ZAITSEVA

MPhil in Physics,
Senior Lecturer at the Department of Computer Science
National University of Life and Environmental Sciences of Ukraine
ORCID: 0000-0002-0875-462X

ADAPTIVE NATURAL LANGUAGE PROCESSING ALGORITHMS FOR MULTILINGUAL USER INTERFACES

In the context of rapid technological development, where communication capabilities are expanding, natural language processing (NLP) enables intuitive human-technology interaction. NLP represents a revolutionary intersection of computer science, artificial intelligence, and linguistics, aimed at bridging the gap between human communication and computational understanding. Its integration into user interfaces has been transformational, providing more intuitive and efficient interaction between users and technologies. The complexity of natural language, with its nuances, idioms, and cultural variations, poses significant challenges for processing. However, advances in machine learning and deep learning have elevated NLP to new heights, enabling systems to learn from large volumes of data and improve over time. This synergy has opened new pathways to accessibility, making technology more inclusive. The purpose of this article is to analyze adaptive natural language processing algorithms for multilingual user interfaces. The study employs a literature review methodology. The results demonstrate that, due to progress in machine learning and deep learning, NLP algorithms now form the foundation of multilingual user interfaces, ranging from voice assistants, chatbots, and autocorrect systems to sentiment analysis, automatic translation, content summarization, and accessibility features. These processes make technologies more responsive, adaptive, and capable of understanding the language, context, emotions, and cultural nuances of human communication. Neural systems utilize large volumes of data to learn to perform tasks automatically. The paper describes the directions and capabilities of natural language processing in user interfaces. Each of these directions highlights the multifaceted nature of NLP in user interfaces and its potential to make technologies more responsive to human needs; in the future, its development will contribute to a deeper understanding of the relationship between technology and humans. Thus, the implementation of natural language processing in multilingual user interfaces opens prospects for creating a more natural, inclusive, and intelligent digital environment of the future.

Key words: artificial intelligence, linguistics, natural language, user interfaces, digital environment.

Постановка проблеми

Сучасні цифрові технології дедалі частіше орієнтовані на створення природних, інтуїтивно зрозумілих і багатомовних інтерфейсів користувача, здатних ефективно взаємодіяти з людиною через мову. Однак, попри значні досягнення у сфері розроблення чат-ботів, віртуальних помічників і голосових сервісів, проблема полягає в недостатній адаптивності алгоритмів обробки природної мови до контексту, емоцій, мовних варіацій та культурних відмінностей користувачів. Ранні системи, засновані на правилах, виявили обмежену здатність до розуміння складних діалогових структур і гнучкого реагування на природне мовлення. Лише із впровадженням машинного та глибокого навчання з'явилася можливість формування адаптивних моделей, здатних самостійно навчатися на великих обсягах даних, розпізнавати мовні патерни та покращувати якість взаємодії з користувачем. Проте питання багатомовності, персоналізації, розпізнавання емоцій і контекстуальних залежностей залишається відкритим. Саме тому виникає необхідність у комплексному дослідженні адаптивних алгоритмів обробки природної мови, спрямованих на підвищення ефективності, точності та доступності багатомовних інтерфейсів користувача.

Аналіз останніх досліджень і публікацій

Розуміння намірів користувача є наріжним каменем ефективної комунікації в будь-якій системі обробки природної мови (далі – NLP) [1, с. 245]. Зокрема, дослідники Г. Лейн (H. Lane) та М. Дишел (M. Dyshel) підкреслюють, що обробка природної мови дозволяє системам штучного інтелекту інтерпретувати буквальный зміст висловлювань користувачів та зважати на їхні наміри. Це, на думку авторів, забезпечує більш точну та релевантну взаємодію й має особливе значення для розроблення інтерфейсів користувача, де головною метою є створення природного та інтуїтивного досвіду, порівнянного із живим спілкуванням між людьми [2]. Такий підхід демонструє важливість контекстуального розуміння і підкреслює необхідність інтеграції адаптивних алгоритмів у сучасні багатомовні інтерфейси. У цьому контексті С. Різоу (S. Rizou) та колеги аналізують технічну точку зору. Розпізнавання намірів користувача передбачає аналіз вхідних даних мови, вилучення семантичного значення та

їх відображення на заздалегідь визначені дії або відповіді. Для цього потрібне складне поєднання лінгвістичних правил і моделей машинного навчання, які можуть впоратися з неоднозначністю та різноманітністю людської мови [3, с. 5]. За своєю суттю, на думку А. Хефні (A. Hefny) та колег, системи NLP класифікують введення користувача за намірами, які представляють мету, що стоїть за повідомленням користувача. Учені також говорять, що, крім наміру, життєво важливою є ідентифікація сутності (конкретна інформація, що міститься у вхідних даних користувача) [4, с. 2]. Аналогічно Б. Юан (B. Yuan), наголошує на тому, що системи NLP повинні враховувати контекст попередніх взаємодій, керувати потоком діалогу, аналізувати емоційний тон користувача та безперервно навчатися для поліпшення розуміння намірів і мовних шаблонів [5, с. 74]. Алгоритми обробки природної мови в інтерфейсах користувача аналізують І. Іосіфов (I. Iosifov) та В. Соколов (V. Sokolov), описуючи різноманітні програми, кожна з яких має власні задачі та інновації – від голосових помічників і чат-ботів до прогнозного тексту [6, с. 399]. Наголошуючи на практичному застосуванні NLP у повсякденних цифрових сервісах, М. Поплавський та Ю. Трач розглядають функцію голосового керування музичним потоковим сервісом [7].

Таким чином, NLP перетворює взаємодію користувача з технологіями на природний, інтелектуальний та емоційно зрозумілий досвід, роблячи цифрові інтерфейси справді адаптивними та людьми.

Формулювання мети дослідження

Метою дослідження є аналіз і систематизація адаптивних алгоритмів обробки природної мови, а також виявлення їх потенціалу для підвищення ефективності, точності й інклюзивності багатомовних інтерфейсів користувача. Дослідження спрямовано на визначення принципів побудови таких систем, які здатні розуміти контекст, емоційний стан та культурні особливості мовців, забезпечуючи більш природну взаємодію людини з технологіями. Для досягнення поставленої мети передбачено виконання таких завдань:

- 1) проаналізувати етапи розвитку систем обробки природної мови;
- 2) визначити особливості функціонування адаптивних алгоритмів у контексті багатомовних інтерфейсів користувача;
- 3) схарактеризувати ключові технології, що забезпечують обробку мовних, контекстуальних і емоційних аспектів комунікації;
- 4) виявити проблеми, пов'язані з розпізнаванням контексту, семантики та культурних відмінностей у багатомовних середовищах;
- 5) узагальнити перспективні напрями розвитку адаптивних алгоритмів обробки природної мови для створення інтелектуальних, персоналізованих і доступних інтерфейсів користувача.

Викладення основного матеріалу дослідження

У галузі штучного інтелекту його проєктування для природної взаємодії з користувачем є наріжним каменем створення інтуїтивно зрозумілого та безперебійного досвіду [8, с. 193]. Мета полягає в тому, щоб створити системи, які розуміють й інтерпретують людську мову якомога природніше, долаючи розрив між людським і машинним спілкуванням. Це передбачає глибоку інтеграцію методів обробки природної мови (NLP), які дозволяють машинам розуміти, аналізувати та генерувати людську мову таким чином, щоб це було одночасно значущим і контекстуально релевантним. Завдання полягає не лише в розумінні буквальних рядків слів, але й у розумінні нюансів, намірів та емоцій, переданих мовою.

З погляду користувача, взаємодія зі штучним інтелектом має бути такою ж простою, як і спілкування з людиною. Це вимагає від штучного інтелекту витонченого розуміння синтаксису, семантики, мовної прагматики та здатності брати участь у веденні діалогів. З боку розробників це вимагає надійного дизайну, який враховує різні способи, за допомогою яких користувачі можуть висловлювати свої думки та запити. Етичні міркування також відіграють важливу роль, оскільки штучний інтелект повинен обережно поводитися з конфіденційною інформацією та поважати приватність користувачів.

Тонкощі дизайну штучного інтелекту для природної взаємодії з користувачем представлено в табл. 1.

Зосереджуючись на цих областях, розробники можуть створювати системи штучного інтелекту, які забезпечують природний та інтуїтивно зрозумілий користувацький досвід. Кінцева мета полягає в тому, щоб створити штучний інтелект, який може спілкуватися, допомагати та надавати цінність без необхідності адаптуватися до того, як машина розуміє світ.

Обробка природної мови (NLP) стало основою сучасного дизайну інтерфейсу користувача, забезпечуючи більш інтуїтивну та людську взаємодію між користувачами та технологіями. Однак інтеграція NLP в інтерфейси користувача включає безліч проблем, які охоплюють технічну, лінгвістичну та дизайнерську сфери.

З технічної позиції складність людської мови, з її нюансами, ідіомами і контекстуальними значеннями, є істотною перешкодою для алгоритмів NLP. Забезпечення того, щоб ці системи могли точно розуміти та обробляти широкий спектр вхідних даних на мові, є непростим завданням [10]. З лінгвістичного погляду варіативність висловлювань користувача, включаючи діалекти, акценти та перемикання мови, додає ще один рівень складності [11]. У сфері дизайну складним виявляється створення інтерфейсів, які можуть елегантно справлятися з непередбачуваною природою людської мови, забезпечуючи при цьому безперебійний користувацький досвід [12].

Ці інтерфейси повинні бути не тільки надійними та гнучкими, але й відповідати очікуванням користувачів та культурним нормам.

Таблиця 1

Дизайн штучного інтелекту для природної взаємодії з користувачем

Дизайн	Функція
Розуміння контексту	Штучний інтелект повинен уміти підтримувати контекст протягом усієї взаємодії
Робота із двозначністю	Природна мова сповнена двозначності. Системи штучного інтелекту повинні бути розроблені таким чином, щоб справлятися з цим, часто ставлячи уточнювальні запитання
Навчання на взаємодіях	Штучний інтелект повинен удосконалюватися з часом, навчаючись на минулих взаємодіях. Цього можна досягти за допомогою алгоритмів машинного навчання, які аналізують закономірності поведінки користувачів і використання мови
Мультимодальні взаємодії	Поєднання тексту з іншими способами введення, як-от голос або жест, може призвести до більш природної взаємодії
Персоналізація	Штучний інтелект повинен уміти адаптувати взаємодії на основі вподобань та історії користувача. Наприклад, якщо користувач часто бронює рейси в певній авіакомпанії, штучний інтелект може надати пріоритет цій авіакомпанії у своїх майбутніх пошуках
Етичний дизайн	Штучний інтелект повинен бути розроблений з огляду на етичні міркування, гарантуючи, що він не увічнює упередженість і не порушує приватне життя користувачів. Це включає прозорість використання даних користувачів і надання їм можливості відмовитися від збору даних
Глобальна доступність	Розроблення штучного інтелекту для глобальної аудиторії означає адаптацію до кількох мов і діалектів, а також культурних нюансів

Джерело: [9]

Незважаючи на багатогранний потенціал, є проблеми, які виникають під час упровадження NLP у дизайн користувацького інтерфейсу (табл. 2).

Таблиця 2

Основні проблеми впровадження обробки природної мови (NLP) у дизайн користувацького інтерфейсу

Проблематика	Сфера впливу
Розв'язання проблеми двозначності	Природна мова за своєю суттю неоднозначна. Слова можуть мати кілька значень залежно від контексту, а речення можна тлумачити по-різному. Проєктування системи NLP, яка може точно визначати наміри користувача, вимагає складних алгоритмів і глибоких навчальних даних
Контекстуальне розуміння	Семантика вхідних даних користувача залежить від контексту їх використання. Системи обробки природної мови мають забезпечувати збереження контексту протягом сеансу взаємодії та інтерпретувати посилання на попередні етапи діалогу.
Різноманітність користувачів	Користувачі характеризуються різним особливим досвідом, рівнем володіння мовою, діалектними й акцентними особливостями. Системи обробки природної мови мають враховувати цю гетерогенність і забезпечувати точну інтерпретацію вхідних даних у всьому діапазоні мовної варіативності.
Обробки в режимі реального часу	Для безперебійної взаємодії з користувачем системи NLP повинні обробляти та реагувати на введені користувачем дані в режимі реального часу. Для цього потрібні високоефективні алгоритми та оптимізований дизайн системи для мінімізації затримки
Обробки помилок	Коли виникають непорозуміння, система повинна впоратися з ними витончено, що передбачає розроблення інтерфейсів, які можуть повернути користувачів на правильний шлях, не викликаючи розчарування чи плутанини
Конфіденційність і безпека	Оскільки системи NLP часто обробляють приватні дані користувачів, забезпечення конфіденційності та безпеки має першорядне значення. Це включає безпечні методи обробки даних і прозорі механізми згоди користувачів
Безперервне навчання	Мова еволюціонує, тож системи NLP повинні бути розроблені зі здатністю вчитися на взаємодіях і згодом адаптуватися до нових моделей використання мови
Культурна чутливість	Мова тісно пов'язана з культурою. Системи NLP повинні бути культурно обізнаними, щоб уникнути неправильних тлумачень і поважати культурні нюанси у використанні мови

Джерело: [13]

Незважаючи на те, що NLP має великі перспективи для поліпшення інтерфейсів користувача, проблеми, які воно представляє, значні й багатогранні. Розв'язання цих проблем вимагає міждисциплінарного підходу, який поєднує досягнення в галузі штучного інтелекту та машинного навчання зі знаннями лінгвістики та принципів дизайну, орієнтованого на користувача.

Обробки природної мови – це революційний поворот у способі нашої взаємодії з технологіями, що перетворює користувацькі інтерфейси із жорстких командних систем на плавні розмовні середовища. Впровадження NLP не тільки покращило користувацький досвід, але й відкрило нові шляхи в аналізі даних, обслуговуванні клієнтів і навіть підтримці психічного здоров'я. Розглядаючи успішні тематичні дослідження, можливо отримати

уявлення про практичне застосування і трансформаційний потенціал технологій NLP. Від чат-ботів, які одночасно обробляють тисячі запитів клієнтів, до віртуальних помічників, котрі керують нашими розкладами та рекомендують дії, NLP – це тихий двигун, який відкриває нову еру взаємодії людини та машини (табл. 3).

Таблиця 3

**Приклади успішного застосування технологій обробки природної мови (NLP)
у сучасних цифрових системах**

Боти для обслуговування клієнтів	Такі компанії, як Amtrak, досягли величезного успіху зі своїм ботом для обслуговування клієнтів «Джулі», який обробляє понад 5 мільйонів запитів на рік. Бот «Джулі» допомагає з бронюванням, надає статуси поїздів та інформацію про подорожі, демонструючи ефективність NLP у виконанні повторюваних завдань з великим обсягом
Комунікація в охороні здоров'я	Впровадження NLP, як-от віртуальний асистент Epic Systems, який інтегрується з електронними медичними записами, забезпечує більш природну та ефективну взаємодію між постачальниками медичних послуг та їхніми системами, зменшує адміністративне навантаження та покращує догляд за пацієнтами
Підтримка психічного здоров'я	Платформи на основі штучного інтелекту, зокрема Woebot, використовують NLP для спілкування з користувачами, надаючи методи когнітивно-поведінкової терапії для допомоги в управлінні психічним здоров'ям. Це демонструє роль NLP у наданні доступної та негайної підтримки окремим особам
Послуги перекладу мов	Такі служби, як Google Translate, еволюціонували, щоб використовувати NLP для більш точних і контекстно-залежних перекладів, забезпечуючи спілкування в режимі реального часу через мовні бар'єри та сприяючи глобальному підключенню
Асистенти з голосовим керуванням	Віртуальні помічники, як-от Alexa від Amazon і Siri від Apple, стали відомими, допомагаючи у всьому – від встановлення нагадувань до керування пристроями «розумного» дому, і все це за допомогою команд природною мовою
Моніторинг соціальних медіа	Інструменти NLP використовуються для моніторингу платформ соціальних медіа на предмет тенденцій, аналізу настроїв і навіть для виявлення потенційних надзвичайних ситуацій, що ілюструє широке застосування NLP в аналізі даних
Інструменти навчання	Програми для вивчення мов, наприклад, Duolingo, використовують NLP для забезпечення персоналізованого досвіду навчання, адаптуючись до рівня навичок і стилю навчання користувача, роблячи освіту більш інтерактивною та адаптивною

Джерело: власна розробка авторів.

Ці тематичні дослідження підкреслюють універсальність і вплив NLP у різних сферах життя. Дозволяючи машинам розуміти людську мову і реагувати на неї, NLP виступає практичним інструментом та мостом до більш емпатичних і розумних технологій.

Коли ми дивимося на горизонт технологічного прогресу, область обробки природної мови виділяється як маяк інновацій, особливо в поліпшенні користувацького досвіду. Перетин адаптивних алгоритмів обробки природної мови для багатомовного інтерфейсу користувача – це захопливий рубіж, де усне мовлення чи письмо стає потужним інструментом для взаємодії. Ця синергія готова переосмислити спосіб взаємодії користувачів із цифровими системами, зробивши технології більш доступними, інтуїтивно зрозумілими та відповідними потребам людини [14]. Майбутні тенденції в NLP і користувацькому досвіді полягають у поступових поліпшеннях, які передбачають зміну парадигми в тому, як ми взаємодіємо з цифровим світом.

Дизайнери шукають способи безперешкодної інтеграції NLP в інтерфейси користувача, роблячи взаємодію більш природною і менш структурованою. Це може означати відмову від традиційних графічних елементів, як-от кнопки та меню, і перехід до більш розмовних інтерфейсів. Прикладом цього є віртуальний домашній помічник, який може розуміти складні команди.

З позиції користувача очікується більш персоналізований та адаптивний досвід. Майбутні системи NLP можуть вивчати індивідуальні уподобання користувачів і належним чином адаптувати відповіді.

Результати дослідження показують, що майбутні тенденції розвитку обробки природної мови в користувацькому досвіді спрямовані на створення більш інтуїтивних, персоналізованих і чутливих систем взаємодії. Однією з ключових тенденцій є мультимодальна взаємодія, яка поєднує текст, голос і візуальні елементи, забезпечуючи багатовимірний досвід користувача. Важливим складником стає також розпізнавання емоцій і реагування, що дозволяє системам виявляти емоційний стан користувача та адаптувати свої відповіді для створення більш емпатичної комунікації. Подальший розвиток отримає налаштування мовних моделей, здатних урахувати індивідуальні особливості мовлення, регіональні діалекти та сленг, що сприятиме підвищенню рівня персоналізації.

Крім того, у центрі уваги залишається етичний і відповідальний штучний інтелект, який передбачає зменшення мовних упереджень і забезпечення культурної чутливості систем. Перспективним напрямом є поєднання доповненої реальності (AR) з NLP, що відкриває можливості створення нових форм інтерактивного досвіду, де користувач може взаємодіяти з цифровими об'єктами за допомогою природної мови. Важливою тенденцією є також

безперервне навчання систем, завдяки якому інтерфейси зможуть адаптуватися до змін поведінки користувачів і контексту використання [15]. Нарешті, кросплатформні рішення NLP сприятимуть формуванню узгодженого досвіду на різних пристроях і платформах.

У сукупності ці напрями формують нову парадигму взаємодії людини та технологій, в якій обробки природної мови стає основою для створення інтелектуальних, гнучких і емоційно орієнтованих цифрових середовищ.

Майбутнє NLP у користувацькому досвіді полягає не лише в розумінні мови, а й у створенні гармонійних відносин між людьми та машинами, де комунікативні перешкоди розчинені, а технології є природним продовженням наших намірів. У міру того, як ми продовжуємо розширювати межі можливого, користувацький досвід стане більш плавним, інтуїтивно зрозумілим і, що найважливіше, людським.

У сфері взаємодії з користувачами, особливо за допомогою обробки природної мови, етичні міркування в розмовах на основі штучного інтелекту мають першорядне значення. Оскільки системи штучного інтелекту стають усе більш складними та інтегрованими в повсякденне життя, розмови, які вони полегшують користувачам, виходять за межі простого обміну даними. Вони втілюють складні соціальні взаємодії, які мають важливі етичні наслідки. Здатність штучного інтелекту розуміти, інтерпретувати та генерувати людську мову викликає безліч етичних проблем, які необхідно розв'язати, щоб гарантувати, що ці системи розроблені та розгорнуті відповідально. Від проблем конфіденційності до потенційної упередженості, етичний ландшафт розмов зі штучним інтелектом настільки ж складний, як і критичний.

Результати дослідження свідчать, що етичні виклики впровадження штучного інтелекту та обробки природної мови в користувацькі інтерфейси мають комплексний характер і потребують системного підходу до їх розв'язання. Одним із ключових питань є конфіденційність і безпека даних. Системи NLP часто обробляють персональну інформацію користувачів, що вимагає суворого захисту даних і дотримання етичних норм в їх збиранні та зберіганні. Не менш важливою проблемою є упередженість і справедливість, адже моделі штучного інтелекту можуть відтворювати наявні соціальні чи культурні упередження, що закладені в навчальних наборах даних.

Велике значення має прозорість систем, адже користувач повинен знати, коли він взаємодіє з машиною, а не з людиною, і розуміти, як саме використовуються його дані. Питання підзвітності також залишається відкритим. До прикладу, в разі помилкових рішень штучного інтелекту часто складно визначити, хто несе юридичну чи етичну відповідальність – розробник, компанія чи сама система. У цьому контексті вкрай важливо забезпечити автономію та згоду користувачів, дозволяючи їм контролювати рівень взаємодії з інтелектуальними системами та самостійно вирішувати, які функції активувати чи обмежити.

Ще одним аспектом є вплив на зайнятість, адже автоматизація процесів може змінити структуру ринку праці, тому етичне використання штучного інтелекту має бути спрямоване на доповнення людської праці, а не її заміну. Нарешті, необхідно враховувати довгостроковий вплив на суспільство, зокрема можливі зміни в соціальній взаємодії, залежність від технологій та трансформацію міжособистісних комунікацій.

Узагальнюючи, можна стверджувати, що впровадження штучного інтелекту та NLP на етичних засадах потребує балансу між технологічними можливостями, прозорістю, захистом прав користувачів і соціальною відповідальністю розробників.

Треба зазначити, що етичні міркування в розмовах про штучний інтелект залишаються одним із найважливіших напрямів сучасних досліджень, адже вони визначають, яким чином технології впливатимуть на суспільство, взаємодію людей і розвиток цифрової культури. У міру того як штучний інтелект розвивається, постає потреба у формуванні чітких етичних орієнтирів, які гарантуватимуть, що технології служать для покращення людської взаємодії, а не для її обмеження чи підміни.

Особливу роль у цьому контексті відіграє NLP, яке радикально змінило спосіб комунікації людини з машиною. Подолавши бар'єр між людською мовою та обчислювальними системами, NLP зробило технології більш інтуїтивними, чутливими до контексту та доступними. Цей вплив можна простежити в широкому спектрі застосувань – від віртуальних помічників і чат-ботів до систем перекладу, що працюють у реальному часі.

З позиції користувацького досвіду (UX) технології NLP відкрили шлях до більш природних і розмовних інтерфейсів, де користувач може висловлювати свої запити звичайною мовою, без спеціальних команд чи складних меню. Це зробило технологію доступнішою для людей із різним рівнем технічної підготовки та можливостями.

Розробники та компанії використовують NLP для підвищення ефективності сервісів і створення персоналізованого досвіду взаємодії. Чат-боти обробляють тисячі звернень одночасно, віртуальні помічники керують розкладами користувачів, а системи аналітики аналізують відгуки та настрої клієнтів. У науковому контексті NLP також стало цінним інструментом для лінгвістів і когнітивістів, адже дає змогу досліджувати механізми сприйняття, обробки та відтворення мови.

Результати аналізу свідчать, що NLP має глибокий і багатовимірний вплив на взаємодію з користувачем. Обробка мови покращує доступність, забезпечуючи можливості для людей із вадами зору чи руху; підвищує ефективність завдяки голосовим командам; персоналізує досвід користувача на основі його уподобань; долає мовні бар'єри через переклад у реальному часі [16]. Система аналізує настрої, допомагаючи бізнесу краще

розуміти клієнтів. NLP має потенціал зменшувати кількість помилок під час введення тексту та підсилювати освітні можливості через інтерактивні навчальні програми.

Отже, NLP розширює функціональні можливості інтерфейсів, формуючи нову культуру цифрової взаємодії. У перспективі ці технології сприятимуть ще більш природній, емоційно чутливій і контекстно точній комунікації між людиною і машиною, що визначатиме майбутнє користувацького досвіду.

Висновки

Аналіз етапів розвитку систем обробки природної мови показав, що еволюція NLP відбувалася від простих правил і статистичних моделей до глибоких нейронних мереж і трансформерів. Таке перетворення забезпечило контекстуальне розуміння та генерацію мови, близьку до людської. Дослідження адаптивних алгоритмів у багатомовних інтерфейсах користувача засвідчили, що сучасні моделі здатні враховувати мовні відмінності, діалекти та культурні особливості, сприяючи підвищенню точності, природності й доступності взаємодії.

У цьому контексті інтеграція мовного, контекстуального та емоційного аналізу є характерною ознакою ключових технологій NLP. Така інтеграція дає змогу створювати системи, які розуміють зміст, наміри та настрої користувача. Виявлення проблем упровадження NLP показало, що основними викликами залишаються розв'язання двозначності, збереження контексту, багатомовність, конфіденційність даних і культурна чутливість. Ці виклики потребують міждисциплінарного підходу до розроблення систем NLP.

Узагальнення перспектив розвитку адаптивних алгоритмів NLP свідчить, що майбутні напрями зосереджені на персоналізації, мультимодальній взаємодії, безперервному навчанні систем, етичності та інтеграції з доповненою реальністю. Дослідження та розроблення зазначених напрямів сприятимуть формуванню нової ери людсько-машинної комунікації.

Список використаної літератури

1. Joshi R., Farhan M., Sharma U., Bhatt S. Unlocking Human Communication: A Journey through Natural Language Processing. *International Journal of Engineering Trends and Applications (IJETA)*. 2024. Vol. 11. № 3. P. 245–250. URL: <https://www.ijetajournal.org/volume-11/issue-3/IJETA-V11I3P43.pdf> (date of access: 04.11.2025).
2. Lane H., Dyschel M. Natural language processing in action. 2025. Simon and Schuster. URL: https://books.google.com.ua/books?hl=uk&lr=&id=WklAEQAAQBAJ&oi=fnd&pg=PA1&dq=Understanding+user+intent+is+the+cornerstone+of+effective+communication+in+any+natural+language+processing+system.&ots=GN_rQbNpWq&sig=WsCmnPf3v0w0lx2-9UEc2jhtPr8&redir_esc=y#v=onepage&q&f=false (date of access: 04.11.2025).
3. Rizou S., Theofilatos A., Paflioti A., Pissari E., Varlamis I., Sarigiannidis G., Chatzisavvas K. C. Efficient intent classification and entity recognition for university administrative services employing deep learning models. *Intelligent Systems with Applications*. 2023. Vol. 19. URL: <https://doi.org/10.1016/j.iswa.2023.200247> (date of access: 04.11.2025).
4. Hefny A. H., Dafoulas G. A., Ismail M. A. Intent classification for a management conversational assistant. In 2020 15th International Conference on Computer Engineering and Systems (ICCES). 2020. P. 1–6. IEEE. URL: <https://ieeexplore.ieee.org/abstract/document/9334685> (date of access: 04.11.2025).
5. Yuan B. Design of an Intelligent Dialogue System Based on Natural Language Processing. *Journal of Theory and Practice of Engineering Science*. 2024. Vol. 4. № 01. P. 72–78. URL: [https://doi.org/10.53469/jtpes.2024.04\(01\).10](https://doi.org/10.53469/jtpes.2024.04(01).10) (date of access: 04.11.2025).
6. Iosifov I., Sokolov V. Y. Методи аналізу природної мови та застосування нейронних мереж в кібербезпеці. *Кібербезпека: освіта, наука, техніка*. 2024. Vol. 4. № 24. P. 398–414. URL: <https://elibrary.kubg.edu.ua/id/eprint/50203/> (дата звернення: 04.11.2025).
7. Поплавський М. М., Трач Ю. В. Цифровізація музичної індустрії: тенденції і перспективи. 2022. URL: <https://www.ceeol.com/search/article-detail?id=1078473> (дата звернення: 04.11.2025).
8. Lypak H., Lypak T., Kunanets N., Duda O., Hribovskyy O. Розроблення підходів до вибору методології проєктування інтерфейсу смартсистеми. *Computer-integrated technologies: education, science, production*. 2025. Vol. 60. P. 190–199. URL: <https://cit.lntu.edu.ua/index.php/cit/article/view/768> (дата звернення: 04.11.2025).
9. Han X., Huang D., Eun-Lee S., Hoon-Yang J. Artificial intelligence-oriented user interface design and human behavior recognition based on human–computer nature interaction. *International Journal of Humanoid Robotics*. 2023. Vol. 20. № 06. DOI: <https://doi.org/10.1142/S0219843622500207> (date of access: 04.11.2025).
10. Васильєв О. С. Застосування штучного інтелекту в дизайні інтерфейсу користувача вебсайту електронної комерції. In X Міжнародна науково-практична конференція "Дизайн та мистецтво в контексті соціокультурного розвитку". 2024. Одеса: Олді+, ХНТУ. URL: <https://er.knutd.edu.ua/bitstream/123456789/29128/1/%D0%92%D0%B0%D1%81%D0%B8%D0%BB%D1%8C%D1%94%D0%B2%20%D0%A8%D0%86.pdf> (дата звернення: 04.11.2025).
11. Alshehri A., AlShabeb A. Exploring attitudes, identity, and linguistic variation among Arabic speakers: Insights from acoustic landscapes. *International Journal of Arabic-English Studies*. 2023. Vol. 24. № 2. P. 1–16. DOI: <https://doi.org/10.33806/ijaes.v24i2.587> (date of access: 04.11.2025).

12. Sindiramutty S. R., Prabakaran K. R. V., Akbar R., Hussain M., Malik N. A. Generative AI for secure user interface (UI) design. In *Reshaping CyberSecurity With Generative AI Techniques*. 2025. (pp. 333–394). IGI Global Scientific Publishing. DOI: <https://doi.org/10.4018/979-8-3693-5415-5.ch010> (date of access: 04.11.2025).

13. Khurana D., Koli A., Khatter K., Singh S. Natural language processing: state of the art, current trends and challenges. *Multimedia tools and applications*. 2023. Vol. 82. № 3. URL: <https://link.springer.com/article/10.1007/s11042-022-13428-4> (date of access: 04.11.2025).

14. Karora V., Lavania G., Agarwal S., Porwal N. Natural Language Processing: A Human Computer Interaction Perspective. *PRATIBODH, (NCDSNS)*. 2023. URL: <https://pratibodh.org/index.php/pratibodh/article/view/150> (date of access: 04.11.2025).

15. Petersen C. L., Halter R., Kotz D., Loeb L., Cook S., Pidgeon D., Batsis J. A. Using natural language processing and sentiment analysis to augment traditional user-centered design: development and usability study. *JMIR mHealth and uHealth*. 2020. 8(8). URL: <https://mhealth.jmir.org/2020/8/e16862/> (date of access: 04.11.2025).

16. Using Redis for Caching Optimization in High-Traffic Web Applications. URL: <https://doi.org/10.62225/2583049X.2025.5.4.4839> (date of access: 04.11.2025).

References

1. Joshi, R., Farhan, M., Sharma, U., & Bhatt, S. (2024). Unlocking human communication: A journey through natural language processing. *International Journal of Engineering Trends and Applications (IJETA)*, 11(3), 245–250. Retrieved from <https://www.ijetajournal.org/volume-11/issue-3/IJETA-V11I3P43.pdf>

2. Lane, H., & Dyshel, M. (2025). *Natural language processing in action*. Simon and Schuster. Retrieved from <https://books.google.com.ua/books?hl=uk&lr=&id=WkIAEQAAQBAJ&oi=fnd&pg=PA1>

3. Rizou, S., Theofilatos, A., Paflioti, A., Pissari, E., Varlamis, I., Sarigiannidis, G., & Chatzisavvas, K. C. (2023). Efficient intent classification and entity recognition for university administrative services employing deep learning models. *Intelligent Systems with Applications*, 19. <https://doi.org/10.1016/j.iswa.2023.200247>

4. Hefny, A. H., Dafoulas, G. A., & Ismail, M. A. (2020). Intent classification for a management conversational assistant. In *2020 15th International Conference on Computer Engineering and Systems (ICCES)* (pp. 1–6). IEEE. Retrieved from <https://ieeexplore.ieee.org/abstract/document/9334685>

5. Yuan, B. (2024). Design of an intelligent dialogue system based on natural language processing. *Journal of Theory and Practice of Engineering Science*, 4(1), 72–78. [https://doi.org/10.53469/jtpes.2024.04\(01\).10](https://doi.org/10.53469/jtpes.2024.04(01).10)

6. Iosifov I., & Sokolov V. Y. (2024). Metody analizu pryrodnoi movy ta zastosuvannya neironnykh merezh v kiberbezpeki [Methods of natural language analysis and neural networks application in cybersecurity]. *Kiberbezpeka: osvita, nauka, tekhnika – Cybersecurity: Education, Science, Technology*, 4(24), 398–414. URL: <https://elibrary.kubg.edu.ua/id/eprint/50203/> (дата звернення: 04.11.2025). [in Ukrainian].

7. Poplavskiy M. M., & Trach Yu. V. (2022). Tsyfrovizatsiia muzychnoi industrii: tendentsii i perspektyvy [Digitalization of the music industry: trends and prospects]. URL: <https://www.ceeol.com/search/article-detail?id=1078473> (дата звернення: 04.11.2025). [in Ukrainian].

8. Lypak H., Lypak T., Kunanets N., Duda O., & Hribovskyy O. (2025). Rozroblennia pidkhodiv do vyboru metodolohii proiektuvannya interfeisu smartsystemy [Development of approaches to choosing a smart system interface design methodology]. *Computer-integrated technologies: education, science, production*, 60, 190–199. URL: <https://cit.lntu.edu.ua/index.php/cit/article/view/768> (дата звернення: 04.11.2025). [in Ukrainian].

9. Han X., Huang D., Eun-Lee S., & Hoon-Yang J. (2023). Artificial intelligence-oriented user interface design and human behavior recognition based on human–computer nature interaction. *International Journal of Humanoid Robotics*, 20(6). <https://doi.org/10.1142/S0219843622500207>

10. Vasylyev O. S. (2024). Zastosuvannya shtuchnoho intelektu v dyzaini interfeisu korystuvacha vebсайту elektronnoi komertsii [Application of artificial intelligence in e-commerce website UI design]. In *Kh Mizhnarodna naukovo-praktychna konferentsiia "Dyzain ta mystetstvo v konteksti sotsiokulturnoho rozvytku"*. Odesa: Oldi+, KhNTU. URL: <https://er.knutd.edu.ua/bitstream/123456789/29128/1/Васильев%20ШІ.pdf> (дата звернення: 04.11.2025). [in Ukrainian].

11. Alshehri A., & AlShabeb A. (2023). Exploring attitudes, identity, and linguistic variation among Arabic speakers: Insights from acoustic landscapes. *International Journal of Arabic-English Studies*, 24(2), 1–16. <https://doi.org/10.33806/ijaes.v24i2.587>

12. Sindiramutty S. R., Prabakaran K. R. V., Akbar R., Hussain M., & Malik N. A. (2025). Generative AI for secure user interface (UI) design. In *Reshaping CyberSecurity With Generative AI Techniques* (pp. 333–394). IGI Global Scientific Publishing. <https://doi.org/10.4018/979-8-3693-5415-5.ch010>

13. Khurana D., Koli A., Khatter K., & Singh S. (2023). Natural language processing: State of the art, current trends and challenges. *Multimedia Tools and Applications*, 82(3). URL: <https://link.springer.com/article/10.1007/s11042-022-13428-4>

14. Karora V., Lavania G., Agarwal S., & Porwal N. (2023). Natural language processing: A human-computer interaction perspective. *PRATIBODH (NCDSNS)*. URL: <https://pratibodh.org/index.php/pratibodh/article/view/150>
15. Petersen C. L., Halter R., Kotz D., Loeb L., Cook S., Pidgeon D., & Batsis J. A. (2020). Using natural language processing and sentiment analysis to augment traditional user-centered design: Development and usability study. *JMIR mHealth and uHealth*, 8(8). URL: <https://mhealth.jmir.org/2020/8/e16862/>
16. Using Redis for caching optimization in high-traffic web applications. (2025). <https://doi.org/10.62225/2583049X.2025.5.4.4839>

Дата першого надходження рукопису до видання: 27.11.2025
Дата прийнятого до друку рукопису після рецензування: 15.12.2025
Дата публікації: 31.12.2025

О. І. ФЕДЮШИН

кандидат технічних наук, доцент,
доцент кафедри безпеки інформаційних технологій
Харківський національний університет радіоелектроніки
ORCID: 0000-0002-3600-405X

І. К. ТКАЧУК

магістрант кафедри безпеки інформаційних технологій
Харківський національний університет радіоелектроніки
ORCID: 0009-0007-7756-660X

Р. Ю. ГВОЗДЬОВ

асистент кафедри безпеки інформаційних технологій
Харківський національний університет радіоелектроніки
ORCID: 0000-0002-5408-943X

В. В. СТАДНІК

науковий співробітник науково-дослідної лабораторії
факультету автоматизованих систем управління
та наземного забезпечення польотів авіації
Харківський національний університет Повітряних Сил
імені Івана Кожедуба
ORCID: 0000-0002-3784-4165

ДИНАМІЧНЕ КЕРУВАННЯ ПРАВАМИ ДОСТУПУ НА ОСНОВІ ПОВЕДІНКОВИХ МОДЕЛЕЙ КОРИСТУВАЧІВ І СЕРВІСІВ

У статті запропонований підхід до динамічного керування правами доступу, побудований на аналізі поведінкових моделей суб'єктів взаємодії. Особлива увага приділяється характеристикам поведінки, що можуть бути використані як сигнали для адаптації політик доступу, методам їх формалізації та алгоритмічним стратегіям прийняття рішень у режимі реального часу. У межах дослідження здійснено огляд наукових підходів до моделювання поведінки користувачів і сервісів, а також висвітлено обмеження традиційних систем контролю доступу, які ґрунтуються на статично визначених правилах чи атрибутах. Зокрема, проаналізовано проблеми надлишкових привілеїв, повільної реакції на зміну контексту та відсутності механізмів автоматизованої еволюції політик у відповідь на реальні сценарії взаємодії. Крім того, визначено актуальну наукову проблему, що полягає у необхідності створення ефективних та стійких до аномалій механізмів поведінково-орієнтованої адаптації прав доступу, здатних підтримувати безперервний цикл спостереження, аналізу та коригування політик.

Розглядається структура поведінкових моделей, способи їх навчання та інтерпретації, а також принципи інтеграції таких моделей у системи контролю доступу з урахуванням вимог масштабованості, достовірності сигналів та мінімізації ризиків. Додатково окреслено значення таких підходів для підвищення стійкості інформаційних систем до внутрішніх загроз, а також можливість використання поведінкових характеристик як основи для прогнозування потенційно небезпечної активності та формування більш гнучких і самоналаштованих механізмів управління доступом.

Ключові слова: динамічний контроль доступу, поведінкові моделі, поведінковий аналіз, аномалії активності, профілі користувачів, адаптація політик доступу, ризик-орієнтовані методи, інформаційна безпека, самоналаштовані системи, виявлення відхилень.

О. І. FEDIUSHYN

Candidate of Technical Sciences, Associate Professor,
Associate Professor at the Department of Information Technology Security
Kharkiv National University of Radio Electronics
ORCID: 0000-0002-3600-405X

I. K. TKACHUK

Master's Student at the Department of Information Technology Security
Kharkiv National University of Radio Electronics
ORCID: 0009-0007-7756-660X

R. YU. GVOZDOV

Assistant at the Department of Information Technology Security
Kharkiv National University of Radio Electronics
ORCID: 0000-0002-5408-943X

V. V. STADNIK

Researcher Research Laboratory
at the Faculty of Automated Control Systems
and Ground Support of Aviation Flights
Ivan Kozhedub Kharkiv National University of the Air Force
ORCID: 0000-0002-3784-4165

DYNAMIC ACCESS CONTROL MANAGEMENT BASED ON BEHAVIORAL MODELS OF USERS AND SERVICES

This article proposes an approach to dynamic access rights management based on the analysis of behavioral models of interacting subjects. Particular attention is given to behavioral characteristics that can serve as signals for adapting access policies, methods of their formalization, and algorithmic strategies for real-time decision-making. The study provides an overview of scientific approaches to modeling the behavior of users and services, as well as highlights the limitations of traditional access control systems that rely on statically defined rules or attributes. In particular, issues such as excessive privileges, slow response to contextual changes, and the absence of automated mechanisms for policy evolution in response to real interaction scenarios are analyzed.

The paper identifies the urgent scientific problem of developing effective and anomaly-resistant mechanisms for behavior-oriented adaptation of access rights capable of supporting a continuous cycle of monitoring, analysis, and policy adjustment. The structure of behavioral models, methods of their training and interpretation, and principles of integrating such models into access control systems are discussed with consideration of scalability requirements, signal reliability, and risk minimization. Additionally, the article outlines the importance of these approaches for enhancing information system resilience to insider threats, as well as the potential use of behavioral characteristics as a basis for predicting harmful activity and constructing more flexible and self-adaptive access management mechanisms.

Key words: dynamic access control, behavioral models, behavioral analysis, activity anomalies, user profiles, access policy adaptation, risk-based methods, information security, self-adaptive systems, anomaly detection.

Постановка проблеми

Сучасні інформаційні системи функціонують у середовищах, що характеризуються високою мінливістю, динамікою взаємодій та зростаючим обсягом поведінкових даних про діяльність користувачів і сервісів. Традиційні моделі контролю доступу, засновані на фіксованих правилах, статичних атрибутах або наперед визначених ролях, виявляються недостатньо чутливими до цих змін. Їхня статичність призводить до накопичення надмірних привілеїв, зниження адаптивності та втрати відповідності реальним сценаріям використання. Зокрема, такі моделі не враховують актуальну поведінку суб'єктів взаємодії, не розпізнають відхилення від типових патернів та не здатні оперативно коригувати політики доступу у відповідь на нетипові дії чи аномалії.

Проблема ускладнюється тим, що в умовах розподіленості та сервісної орієнтованості сучасних систем взаємодії між компонентами генерують значний обсяг поведінкових сигналів, які потенційно можуть бути використані для підвищення точності та гнучкості рішень щодо доступу. Проте існуючі підходи до керування правами доступу здебільшого не інтегрують такі сигнали у процес ухвалення рішень і не використовують поведінкову динаміку як основу для адаптації політик. Тому виникає наукова та прикладна проблема, що полягає у відсутності ефективних механізмів динамічного керування правами доступу, здатних використовувати поведінкові моделі користувачів і сервісів для автоматизованої адаптації політик у режимі реального часу. Необхідним є розроблення підходів, які забезпечували б безперервне спостереження за поведінкою, її формалізацію, визначення показників нормальності чи аномальності та коригування прав доступу відповідно до отриманих сигналів. Розв'язання цієї проблеми сприятиме підвищенню стійкості інформаційних систем, зменшенню ризиків внутрішніх загроз та побудові більш точних і самоналаштованих систем контролю доступу.

Аналіз останніх досліджень і публікацій

Проблематика динамічного контролю доступу та використання поведінкових характеристик користувачів і сервісів активно досліджується в межах сучасної кібербезпеки. Значна частина наукових робіт зосереджена

на розвитку атрибутивних та контекстуальних моделей доступу, що враховують зміни середовища виконання. Зокрема, у роботі [1] підкреслюється обмеженість традиційних статичних моделей і визначається необхідність переходу до підходів, здатних реагувати на поведінкові сигнали в реальному часі. Подібні висновки подає і дослідження [2], де показано, що поведінкові шаблони можуть бути використані для виявлення аномалій і пом'якшення внутрішніх загроз.

У низці публікацій розглядаються методи побудови поведінкових моделей на основі машинного навчання, зокрема методи класифікації та кластеризації, що дозволяють ідентифікувати відхилення у взаємодії користувачів з системою. У роботі [3] запропоновано інтелектуальний підхід до контролю доступу, який використовує поведінкові профілі для динамічного коригування привілеїв суб'єктів. Дослідження [4] демонструє можливість інтеграції поведінкових характеристик у моделі ризик-орієнтованого доступу, що дозволяє підвищити точність прийняття рішень.

Крім того, наукові праці у сфері самоналаштуваних систем наголошують на важливості циклу «моніторинг – аналіз – планування – виконання» (концепція MAPE-K), який може бути застосований для автоматизованої еволюції політик доступу. Наприклад, у роботі [5] описано підхід до адаптації поведінки систем на основі знань про їх поточний стан, що узгоджується з концепцією поведінково-орієнтованого контролю доступу.

Узагальнюючи проаналізовані роботи, можна зробити висновок, що сучасні дослідження визнають перспективність інтеграції поведінкових даних у процес керування доступом, проте поки відсутні універсальні моделі, здатні комплексно враховувати поведінку як користувачів, так і сервісів у динамічних умовах функціонування інформаційних систем.

Формулювання мети дослідження

Метою дослідження є розробка підходу до динамічного керування правами доступу на основі аналізу поведінкових моделей користувачів і сервісів, що забезпечує адаптивне коригування політик доступу в реальному часі. Дослідження спрямоване на визначення релевантних поведінкових характеристик, формалізацію їх впливу на процес ухвалення рішень та обґрунтування механізмів, які підвищують точність, гнучкість і безпеку систем контролю доступу в умовах змінного цифрового середовища.

Викладення основного матеріалу дослідження

У межах даного дослідження здійснюється комплексний аналіз підходів до динамічного керування правами доступу, що ґрунтуються на поведінкових моделях користувачів і сервісів. Особлива увага приділяється ідентифікації характеристик поведінки, які можуть виступати індикаторами нормальної або аномальної активності, та визначенню способів їх формалізації для подальшого використання у процесі адаптації політик доступу. Розглядаються механізми побудови поведінкових профілів, методи обробки та класифікації поведінкових даних, а також підходи до інтеграції цих елементів у цикли прийняття рішень у системах контролю доступу.

В роботі проаналізовано актуальні наукові методи й моделі, що забезпечують можливість адаптивного реагування системи на зміну поведінки суб'єктів взаємодії. Зокрема, досліджуються компоненти системи моніторингу поведінки, методи обчислення поведінкових відхилень, алгоритми визначення рівня ризику та механізми коригування прав доступу в реальному часі. Крім того, увагу зосереджено на питаннях узгодженості динамічної еволюції політик із вимогами безпеки, сталості та масштабованості інформаційної системи.

Також деталізуються ключові аспекти побудови поведінково-орієнтованих моделей доступу та демонструють можливі підходи до реалізації адаптивного контролю доступу в умовах змінного цифрового середовища.

Поведінкові характеристики суб'єктів взаємодії та їх формалізація

Поведінкові характеристики суб'єктів взаємодії є фундаментальним елементом динамічного контролю доступу, оскільки вони дозволяють відобразити реальні моделі використання системи та забезпечують основу для адаптивного коригування прав. До таких характеристик належать частота та контекст виконання дій, часові шаблони активності, послідовність звернень до ресурсів, типові маршрути взаємодії, інтенсивність запитів, а також структурні залежності між викликами сервісів. Ці параметри можуть бути визначені як для кінцевих користувачів, так і для автоматизованих сервісних компонентів, що взаємодіють між собою у складних розподілених системах.

Формалізація поведінкових характеристик передбачає визначення векторів ознак, які репрезентують поведінку суб'єкта у вигляді числових, категоріальних або структурованих даних. До базових методів належать статистичні моделі, часові ряди, графові структури взаємодій та ймовірнісні описи переходів між станами. Для комплексних систем доцільним є використання гібридних моделей, що поєднують детерміновані правила з результатами поведінкового аналізу. Формалізовані дані надалі можуть бути використані для побудови профілів нормальної активності, оцінки подібності поточних дій до історичних зразків та виявлення аномалій.

Чітке визначення поведінкових характеристик і способів їх формалізації є необхідною передумовою для побудови адаптивних систем контролю доступу, здатних оперативно реагувати на зміни у сценаріях взаємодії та забезпечувати точніші рішення щодо надання або обмеження прав доступу.

Методи побудови поведінкових профілів користувачів

Поведінковий профіль суб'єкта взаємодії є узагальненою моделлю його типової активності, сформованою на основі історичних даних та поточних спостережень. Побудова таких профілів є ключовим етапом динамічного керування правами доступу, оскільки саме вони визначають межі нормальної поведінки, відносно яких обчислюються відхилення та приймаються адаптивні рішення. Процес створення поведінкового профілю починається зі збору даних про дії користувачів та сервісів, включаючи частоту, типи операцій, часові інтервали, маршрути запитів та взаємозв'язки між компонентами системи. Зазвичай ці дані агрегуються у вигляді статистичних моделей, метрик активності або структурованих графів взаємодії.

Для побудови профілів застосовуються різні підходи: від класичних статистичних методів (середні значення, дисперсія, коваріаційні залежності) до більш складних моделей, зокрема марковських процесів, кластеризаційних алгоритмів та моделей машинного навчання. Особливої актуальності набувають інкрементні методи оновлення профілів, які дозволяють адаптувати модель до змін поведінки без необхідності повного перенавчання. Такий підхід забезпечує стійкість профілю до природних змін активності суб'єкта, але водночас дозволяє своєчасно реагувати на суттєві відхилення.

На рис. 1 наведена структурована схема методів побудови профілів користувачів.

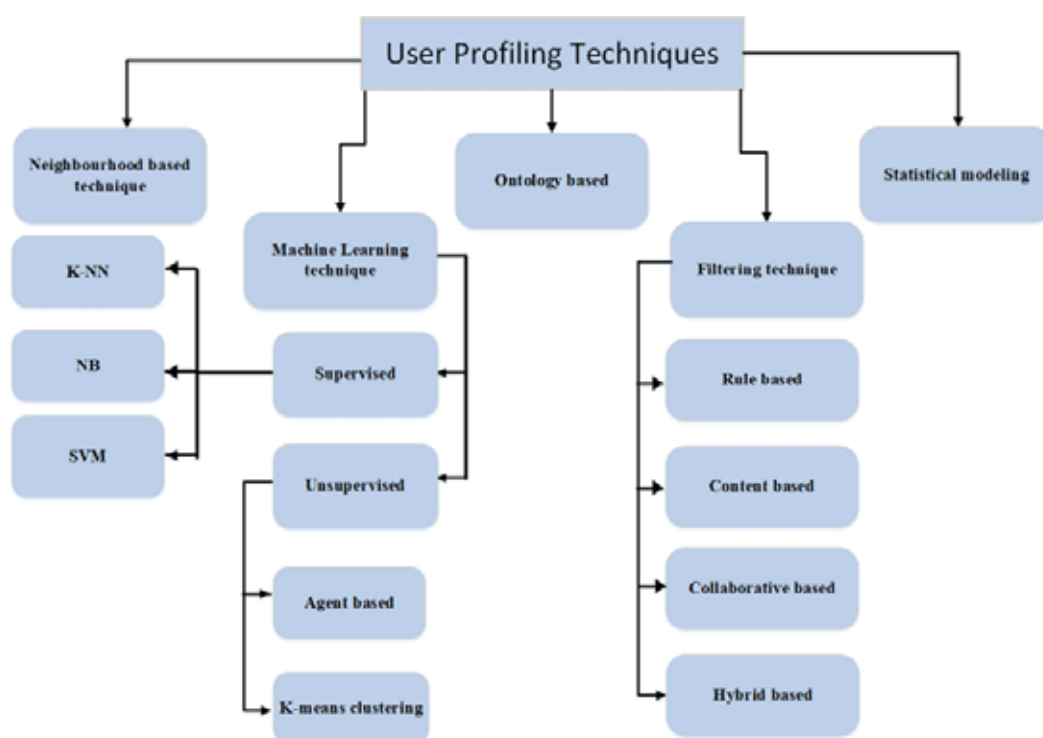


Рис. 1. Класифікація технік побудови поведінкових профілів користувачів

Діаграма систематизує основні техніки моделювання поведінки – від статистичного аналізу та методів на основі відстаней до машинного навчання, фільтраційних підходів та онтологічних моделей. Вона демонструє, що формалізація поведінки може здійснюватися через різні парадигми, кожна з яких дозволяє репрезентувати такі ознаки, як частота та контекст дій, часові шаблони активності, послідовність звернень до ресурсів та структурні залежності між операціями, таким чином підкреслюється багатомодельний характер поведінкового аналізу, який формує основу для побудови адаптивних систем контролю доступу та подальшого виявлення відхилень.

Оновлення поведінкових профілів може здійснюватися як періодично, так і в режимі реального часу, залежно від вимог системи. При цьому важливим є забезпечення балансу між стабільністю профілю та чутливістю до змін. Надмірно чутливий профіль призведе до помилкових сповіщень, тоді як надто стабільний – до несвоєчасного виявлення аномалій. Оптимальна стратегія оновлення повинна враховувати як специфіку домену, так і характер поведінки суб'єктів.

Виявлення аномалій та оцінка поведінкових відхилень контролю доступу

Виявлення аномалій є центральним компонентом динамічного керування правами доступу, оскільки саме цей процес дозволяє ідентифікувати відхилення від типових моделей поведінки та визначати потенційні ризики, пов'язані з діями суб'єктів взаємодії. Аномалії можуть проявлятися у вигляді непередбачуваних запитів до

ресурсів, нетипових маршрутів взаємодії, різкого збільшення інтенсивності операцій або звернень у незвичний час. Важливим є те, що такі відхилення не завжди є ознакою шкідливої активності, тому система контролю доступу повинна забезпечувати коректне розмежування між природними змінами поведінки та потенційними загрозами.

Існують різні підходи до виявлення аномалій, які варіюються від базових методів статистичного аналізу до складних моделей машинного навчання. Статистичні методи включають аналіз контрольних меж, виявлення стрибків у часових рядах, побудову ймовірнісних розподілів і класифікацію відхилень на основі порогових значень. Алгоритми машинного навчання, такі як методи кластеризації, ізоляційні ліси, автоенкодерери або гібридні моделі, дозволяють виявляти складні нелінійні залежності у поведінкових даних.

Модель виявлення конфліктів між політиками доступу в контексті аномальної активності наведена на рис. 2. Вона демонструє, як система порівнює поточні дії суб'єкта з наявними правилами, виявляє логічні суперечності та ідентифікує потенційно небезпечні ситуації ще до того, як зміни політик або рішення щодо доступу будуть застосовані.

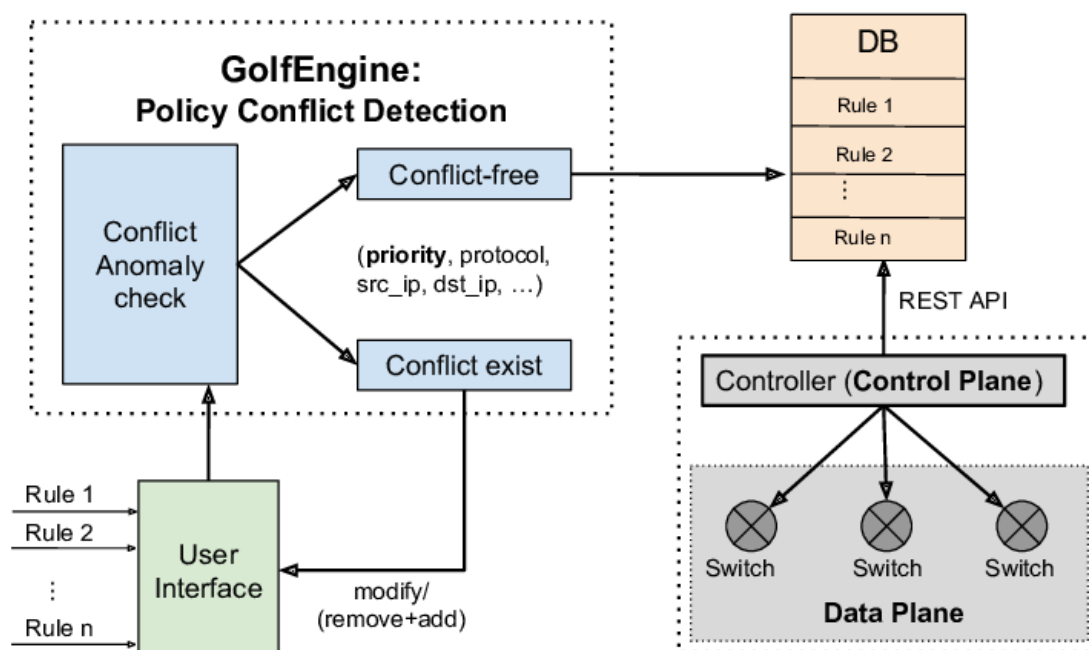


Рис. 2. Модель виявлення конфліктів політик доступу

Оцінка поведінкових відхилень передбачає визначення міри схожості між поточним станом активності та сформованим поведінковим профілем. Це може включати використання метрик відстані, оцінку ймовірності події чи розрахунок ризикового показника. Результат такої оцінки є ключовим входом до механізмів адаптації політик доступу, оскільки дозволяє системі ухвалювати рішення про обмеження, уточнення або тимчасову модифікацію прав.

Для узагальнення найбільш уживаних підходів до виявлення поведінкових відхилень подається порівняльна таблиця методів, що демонструє ключові принципи та відмінності між основними алгоритмами виявлення аномалій.

Таблиця 1

Основні методи виявлення аномалій у поведінкових даних

Метод	Тип методу	Принцип дії
Z-score / IQR	Статистичний	Виявлення відхилень від нормального розподілу
K-Means / DBSCAN	Кластеризація	Групування нормальних точок, а аномалії – поза кластерами
Isolation Forest	ML-алгоритм	Ізоляція аномалій через побудову випадкових дерев
Autoencoder	Глибинне навчання	Висока похибка реконструкції сигналізує аномалію
LOF (Local Outlier Factor)	Щільнісний метод	Уточнює аномалії за локальною щільністю сусідів

Ефективне виявлення аномалій підвищує рівень безпеки системи, зменшує ризик внутрішніх загроз та формує підґрунтя для інтелектуального динамічного контролю доступу.

Алгоритми динамічної адаптації політик доступу

Алгоритми динамічної адаптації політик доступу виконують ключову роль у процесі трансформації поведінкових сигналів у конкретні управлінські рішення щодо надання, обмеження або модифікації прав доступу. Основною функцією таких алгоритмів є інтерпретація результатів поведінкового аналізу та визначення відповідних дій з урахуванням рівня ризику, важливості ресурсу та контексту операції. На практичному рівні адаптація може здійснюватися у вигляді тимчасових обмежень, уточнення привілеїв, застосування додаткових засобів автентифікації або переходу системи у підвищений режим безпеки.

Сучасні підходи до побудови таких алгоритмів включають як детерміновані моделі на основі правил, так і інтелектуальні алгоритми, що використовують машинне навчання. Правил-орієнтовані моделі передбачають попередньо визначені дії у відповідь на конкретні типи відхилень, що забезпечує прозорість та передбачуваність роботи системи. Однак такі моделі можуть бути недостатньо гнучкими в умовах високої варіативності поведінки користувачів.

Натомість алгоритми, що базуються на машинному навчанні або гібридних підходах, здатні адаптуватися до нових ситуацій без необхідності ручного визначення правил. Наприклад, моделі прогнозування можуть оцінювати ймовірність подальшої аномальної активності, у той час як алгоритми оптимізації можуть динамічно визначати мінімальний набір привілеїв для виконання операції. Для узагальнення описаних підходів та забезпечення структурованого подання процесу прийняття рішень у системі динамічного контролю доступу наведено алгоритмічну модель адаптації політик. У таблиці 2 подано основні етапи цього процесу – від аналізу поведінкового сигналу до застосування оновлених правил і їх інтеграції в механізми подальшого навчання системи.

Важливим аспектом є забезпечення збалансованості між безпекою та доступністю ресурсів. Тому алгоритми повинні враховувати як ризиковий профіль суб'єкта, так і критичність системи, мінімізуючи кількість помилок блокувань і водночас ефективно протидіючи потенційним загрозам.

Забезпечення узгодженості системи при еволюції прав доступу

Забезпечення стійкості та узгодженості системи під час динамічної адаптації прав доступу є критичним аспектом побудови поведінково-орієнтованих механізмів контролю доступу. Динамічні зміни політик, хоча й дозволяють підвищити точність реагування на аномальні або ризиковані дії, можуть водночас створювати ризики порушення цілісності системи, некоректного застосування привілеїв або конфліктів між правилами. Одним із ключових принципів забезпечення стійкості є використання формальних методів і верифікаційних моделей, які дозволяють перевіряти політики на предмет конфліктів, надмірних прав або порушення обмежень безпеки.

Таблиця 2

Алгоритмічна модель адаптації політик доступу

Етап	Опис процесу	Вхідні дані	Результат / Вихід
Аналіз поведінкового сигналу	Оцінювання відхилень від поведінкового профілю та визначення рівня ризику	Поточний вектор поведінки, профіль суб'єкта, метрики відхилення	Класифікація події: нормальна, підозріла або аномальна
Визначення контексту операції	Оцінювання критичності ресурсу та умов доступу (час, місце, тип операції)	Контекстні атрибути, ресурсна політика, історія взаємодій	Контекстний ризиковий коефіцієнт
Вибір стратегії адаптації	Прийняття рішення щодо того, як змінювати політику доступу	Ризиковий профіль користувача/сервісу, критичність ресурсу	Обрана стратегія: обмежити / уточнити / посилити автентифікацію / заборонити
Модифікація політики доступу	Застосування локальних або тимчасових змін у правах доступу	Набір дій зі стратегії адаптації	Оновлена політика доступу або тимчасовий контрольний режим
Перевірка узгодженості	Переконання, що нові правила не суперечать глобальним політикам	Попередня і нова версія політик, інваріанти безпеки	Верифікована політика, готова до застосування
Застосування рішення в системі	Активування оновлених політик у реальному часі	Верифіковані зміни	Набуття чинності оновленими правами доступу
Логування та оновлення профілів	Фіксація змін та використання події для подальшого навчання профілів	Журнал рішень та результатів поведінки	Актуалізовані профілі суб'єктів для майбутніх рішень.

До таких методів належать моделі доступу на основі логічних формул, перевірка властивостей за допомогою автоматів станів та застосування технік статичного аналізу. Також алгоритми узгодження повинні забезпечити підтримку безперервності роботи системи, при зміні політик у режимі реального часу.

Важливим аспектом стійкості є мінімізація похибок унаслідок хибних спрацьовувань алгоритмів виявлення аномалій. Для цього в системі можуть використовуватися механізми поступового зниження або підвищення прав, багаторівневі моделі ризику та політики відкладеного підтвердження, які дозволяють зменшити кількість необґрунтованих обмежень.

Для ілюстрації логіки узгодження політик доступу наведено узагальнену схему (див. рис. 3), яка демонструє повний життєвий цикл політики від її специфікації до перевірки на відповідність інваріантам безпеки та фінального застосування.

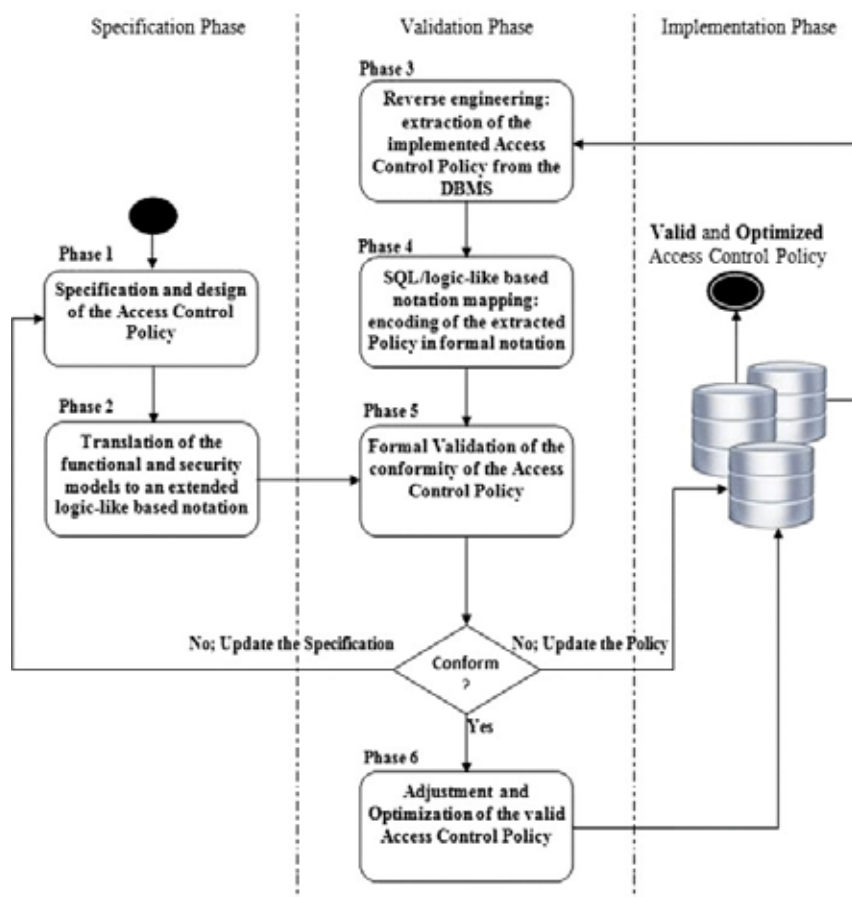


Рис. 3. Методологія розгортання та узгодження політик доступу

Модель показує, як інтегруються етапи валідації, верифікації та аналізу конфліктів, що критичні під час динамічної адаптації правил доступу.

Узгодженість також передбачає збереження глобальних інваріантів безпеки, таких як відсутність циклічних дозволів, дотримання принципу найменших привілеїв та підтримання єдиної логічної структури політик. Реалізація цих принципів забезпечує стабільність системи, підвищує її стійкість до внутрішніх і зовнішніх загроз та гарантує надійність динамічного контролю доступу в умовах змінного цифрового середовища.

Список використаної літератури

1. Hu V. C., Ferraiolo D. F., Kuhn D. R. (2020). Assessment of Access Control Systems. NIST Interagency Report 7316. URL: <https://doi.org/10.6028/NIST.IR.7316>.
2. Özkan Canay, Ümit Kocabiçak. (2024). Predictive modeling and anomaly detection in large-scale web portals through the CAWAL framework. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0950705124013443>.
3. Fan Yang, Chris L. Hankin, Flemming Nielson, Hanne Riis Nielson. (2013). Predictive access control for distributed computation. URL: <https://www.sciencedirect.com/science/article/pii/S0167642312001104>.
4. Yuan Zhai, Haochen Yang, Jingyu Yao, Tao Wang, Yanwei Zhou, Feng Zhu, Bo Yang. (2025). DRAC: A dynamic fine-grained access control scheme for cloud storage with censorship-coerced resistance. URL: <https://www.sciencedirect.com/science/article/abs/pii/S2214212625001607>.
5. Riaz Ahmed Shaikh, Kamel Adi, Luigi Logrippo. (2012). Dynamic risk-based decision methods for access control systems. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0167404812000399>.

References

1. Hu V. C., Ferraiolo D. F., Kuhn D. R. (2020). Assessment of Access Control Systems. NIST Interagency Report 7316. URL: <https://doi.org/10.6028/NIST.IR.7316>.
2. Özkan Canay, Ümit Kocabiçak. (2024). Predictive modeling and anomaly detection in large-scale web portals through the CAWAL framework. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0950705124013443>.

3. Fan Yang, Chris L. Hankin, Flemming Nielson, Hanne Riis Nielson. (2013). Predictive access control for distributed computation. URL: <https://www.sciencedirect.com/science/article/pii/S0167642312001104>.
4. Yuan Zhai, Haochen Yang, Jingyu Yao, Tao Wang, Yanwei Zhou, Feng Zhu, Bo Yang. (2025). DRAC: A dynamic fine-grained access control scheme for cloud storage with censorship-coerced resistance. URL: <https://www.sciencedirect.com/science/article/abs/pii/S2214212625001607>.
5. Riaz Ahmed Shaikh, Kamel Adi, Luigi Logrippo. (2012). Dynamic risk-based decision methods for access control systems. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0167404812000399>.

Дата першого надходження рукопису до видання: 20.11.2025
Дата прийнятого до друку рукопису після рецензування: 16.12.2025
Дата публікації: 31.12.2025

О. І. ФЕДЮШИН

кандидат технічних наук, доцент,
доцент кафедри безпеки інформаційних технологій
Харківський національний університет радіоелектроніки
ORCID: 0000-0002-3600-405X

П. В. ШУЛІК

кандидат технічних наук,
старший викладач кафедри безпеки інформаційних технологій
Харківський національний університет радіоелектроніки
ORCID: 0009-0004-6200-2172

О. М. ТОВМА

аспірант кафедри безпеки інформаційних технологій
Харківський національний університет радіоелектроніки
ORCID: 0009-0009-7225-3297

В. П. КОЦЮБА

викладач кафедри радіоелектронних систем пунктів управління Повітряних Сил
Харківський національний університет Повітряних Сил
імені Івана Кожедуба
ORCID: 0000-0001-6336-8193

ВИЯВЛЕННЯ АНОМАЛІЙ У МЕРЕЖЕВОМУ ТРАФІКУ ВЕБ-ДОДАТКІВ З ВИКОРИСТАННЯМ АЛГОРИТМУ RANDOM FOREST

Забезпечення високої точності та ефективності систем виявлення вторгнень є критичним для уникнення помилкових спрацьовувань та пропусків веб-атак. Мережевий трафік є основним каналом комунікації клієнт-серверної моделі у веб-додатках, тому його аналіз є критично важливим для виявлення шкідливої активності. Саме через нього передаються дані запитів користувача, відповіді сервера, API-виклики, а також взаємодія з зовнішніми сервісами. Використання машинного навчання для прогнозування атак на веб-додатки на основі аналізу мережевого трафіку може покращити ці характеристики. В роботі проводиться вибір технік машинного навчання, що можуть допомагати враховувати не лише структуру, а й семантику трафіку, також розглянуто актуальну проблему захисту веб-додатків від сучасних кіберзагроз, зокрема атак типу SQL Injection, XSS та Brute Force.

Проаналізовано недоліки традиційних сигнатурних методів захисту (WAF, IDS), які демонструють низьку ефективність проти атак «нульового дня» та модифікованих шкідливих запитів. Обґрунтовано доцільність використання методів машинного навчання (Machine Learning) для побудови адаптивних систем виявлення вторгнень. У якості інструментарію дослідження обрано алгоритм Random Forest (випадковий ліс) через його стійкість до перенавчання та здатність ефективно працювати з багатовимірними даними. Експериментальну частину виконано на базі датасету CICIDS2017, який містить актуальні профілі нормального та аномального трафіку. Описано етапи попередньої обробки даних: очищення від шумів, кодування категоріальних ознак, балансування класів за допомогою бібліотеки Imbalanced-learn.

В результаті проведеного експерименту розроблена модель показала загальну точність класифікації (Ассигасу) на рівні 98%. Проаналізовано матрицю невідповідностей, яка підтвердила високу здатність моделі розрізняти легітимний трафік та атаки сканування, при цьому виявлено певні обмеження при класифікації складних веб-атак зі схожими сигнатурами. Результати дослідження можуть бути використані для удосконалення існуючих систем моніторингу безпеки та інтеграції в SIEM-системи.

Ключові слова: кібербезпека, веб-додатки, Random Forest, виявлення аномалій, CICIDS2017, машинне навчання, класифікація трафіку.

O. I. FEDIUSHYN

Candidate of Technical Sciences, Associate Professor,
Associate Professor at the Department of Information Technology Security
Kharkiv National University of Radio Electronics
ORCID: 0000-0002-3600-405X

P. V. SHULIK

Candidate of Technical Sciences,
Senior Lecturer at the Department of Information Technology Security
Kharkiv National University of Radio Electronics
ORCID: 0009-0004-6200-2172

O. M. TOVMA

Postgraduate Student at the Department of Information Technology Security
Kharkiv National University of Radio Electronics
ORCID: 0009-0009-7225-3297

V. P. KOTSIUBA

Lecturer at the Department of Radio-Electronic Systems
of Air Force Control Points
Ivan Kozhedub Kharkiv National University of the Air Force
ORCID: 0000-0001-6336-8193

ANOMALY DETECTION IN WEB APPLICATION NETWORK TRAFFIC USING THE RANDOM FOREST ALGORITHM

Ensuring high accuracy and efficiency of intrusion detection systems is critical for avoiding false positives and missed web attacks. Network traffic is the primary communication channel of the client-server model in web applications; therefore, its analysis is critically important for detecting malicious activity. It is through this channel that user request data, server responses, API calls, and interactions with external services are transmitted. The use of machine learning to predict attacks on web applications based on network traffic analysis can improve these characteristics. The paper performs a selection of machine learning techniques that can help consider not only the structure but also the semantics of traffic. It also examines the current problem of protecting web applications from modern cyber threats, particularly SQL Injection, XSS, and Brute Force attacks.

The shortcomings of traditional signature-based protection methods (WAF, IDS), which demonstrate low effectiveness against zero-day attacks and modified malicious requests, are analyzed. The feasibility of using Machine Learning methods to build adaptive intrusion detection systems is substantiated. The Random Forest algorithm was chosen as the research tool due to its resistance to overfitting and ability to work effectively with high-dimensional data. The experimental part was performed based on the CICIDS2017 dataset, which contains current profiles of normal and anomalous traffic. Data preprocessing stages are described in detail: noise cleaning, encoding categorical features, and class balancing using the Imbalanced-learn library. As a result of the experiment, the developed model showed an overall classification Accuracy of 98%. The confusion matrix was analyzed, confirming the model's high ability to distinguish legitimate traffic and scanning attacks, while identifying certain limitations in classifying complex web attacks with similar signatures. The research results can be used to improve existing security monitoring systems and for integration into SIEM systems.

Key words: cybersecurity, web applications, Random Forest, anomaly detection, CICIDS2017, machine learning, traffic classification.

Постановка проблеми

Пасивні та активні методи збору даних становлять основу сучасного аналізу трафіку. Пасивні методи передбачають моніторинг мережевих пакетів без втручання у потік даних, що забезпечує високу ефективність при виявленні аномалій. До таких методів належать sniffing трафіку за допомогою інструментів типу Wireshark або Tcpdump. Активні методи, такі як сканування портів або тестування пропускну здатності, дозволяють отримувати додаткову інформацію про стан мережі, але можуть впливати на її продуктивність.

Роль трафіку у забезпеченні безпеки полягає в тому, що саме в ньому можна виявити багато типів атак. До того ж аналіз трафіку – це не лише засіб виявлення атак після їх здійснення, а й інструмент для проактивного захисту. Регулярне логування, моніторинг та застосування систем виявлення вторгнень, які аналізують трафік у реальному часі, дозволяють зупинити атаку ще до того, як вона призведе до наслідків. Таким чином, трафік виступає джерелом для побудови моделей поведінки користувачів, виявлення аномалій та автоматизованого реагування.

Мережевий трафік веб-додатку має специфічну структуру, яка визначається протоколами та технологіями, що використовуються для взаємодії між клієнтом і сервером. У протоколах HTTP/HTTPS кожна сесія складається

з запиту клієнта та відповіді сервера. Запит зазвичай включає метод запиту, URL-шлях та параметри, заголовки зі службовою інформацією (куки, токени тощо), тіло запиту, що може містити дані форм або JSON-об'єкти. Відповідь сервера має код статусу, який сигналізує про результат обробки, заголовки з інформацією про тип вмісту і політики кешування, тіло відповіді. Важливо враховувати специфіку протоколів, які використовуються у сучасних односторінкових або мобільних додатках, де активніше застосовуються API-запити та WebSocket-з'єднання. У таких випадках структура трафіку відрізняється від класичних веб-сторінок, і аналіз має фокусуватися на JSON-повідомленнях, структурі API-викликів і підтримці з'єднання.

Розуміння структури та динаміки трафіку дає можливість не лише виявити шкідливу активність, а й забезпечити адаптивний захист. Традиційні методи захисту досі залишаються одним із найпоширеніших та актуальних варіантів. Веб-брандмауери (WAF) аналізують запити в реальному часі, блокуючи відомі шаблони атак за допомогою сигнатурних баз. Брандмауери веб-додатків аналізують HTTP/HTTPS-запити й блокують потенційно шкідливі, базуючись на наборах сигнатур. Ці рішення ефективні проти стандартних загроз, таких як SQL-ін'єкції або XSS, проте мають обмеження щодо нових, невідомих векторів атак. Порівняно з традиційними сигнатурними методами, підходи з боку машинного навчання є більш гнучкими й здатними до самонавчання – це означає, що моделі можна оновлювати у відповідь на зміну поведінки атакувальників, що також зменшує навантаження на аналітиків безпеки. Сюди входить виявлення аномалій, побудоване на визначенні профілю нормальної поведінки веб-додатку, а також використання систем управління подіями безпеки (SIEM), які централізують логуювання, моніторинг і кореляцію подій у реальному часі.

Разом з перевагами сучасні підходи до виявлення атак часто не враховують динамічність та послідовність мережеских подій, що призводить до зниження точності моделей, більшість моделей машинного навчання виявляють проблеми з дисбалансом класів та наявністю шуму в даних, що впливає на їхню продуктивність, і ці питання вимагають вчасного вирішення.

Завданням даного дослідження є розробка моделі машинного навчання для прогнозування атак на веб-додатки на основі аналізу мережевого трафіку, яка буде враховувати структуру пакетів окремих видів трафіку, а також специфіку їх обробки за числовими та категоріальними характеристиками.

Аналіз останніх досліджень і публікацій

За останні роки застосування штучного інтелекту у сфері кібербезпеки стало предметом численних наукових досліджень, опублікованих у провідних виданнях, таких як IEEE, ACM, а також на популярній платформі arXiv. Найбільш цитовані роботи присвячені як теоретичним основам машинного навчання, так і практичним застосуванням для виявлення різних типів атак на мережеві та веб-додатки. Аналіз наукових публікацій за 2022-2025 роки дозволяє виділити три основні напрями, за якими розвиваються системи виявлення атак на основі використання штучного інтелекту: методи глибокого навчання (Deep Learning), класичні алгоритми машинного навчання (Machine Learning) та гібридні підходи.

Одним із ключових напрямків досліджень є використання глибокого навчання для аналізу мережевого трафіку. Зокрема, нейронні мережі, такі як рекурентні (RNN) [1, 2] та трансформери [3,4], показали високу ефективність у виявленні аномалій, DDoS-атак і навіть складних APT-кампаній. Наприклад, моделі на основі Long Short-Term Memory (LSTM) здатні аналізувати часові послідовності в трафіку, виявляючи незвичайні патерни, які можуть свідчити про атаку. Останні роботи також демонструють потенціал графових нейронних мереж (GNN) для аналізу взаємозв'язків між вузлами в мережі [5, 6], що особливо актуально для виявлення скоординованих атак.

Дослідники в працях [7,8] демонструють високу ефективність згорткових нейронних мереж (CNN) та мереж довгої короткострокової пам'яті (LSTM) для класифікації трафіку датасету CICIDS2017. Наприклад, у роботі [7] досягнуто точності 99.1% завдяки здатності нейромереж автоматично виділяти приховані ознаки.

Головним недоліком описаних вище методів є висока обчислювальна складність ("black-box problem") та потреба у потужному апаратному забезпеченні (GPU), що ускладнює їх інтеграцію в системи реального часу (Real-time WAF) з обмеженими ресурсами.

Окремий кластер досліджень присвячено аналізу текстових логів. Зокрема, Al-Omari та ін. (2024) у роботі «Anomaly Detection in HTTP Logs» [9] пропонують фреймворк для виявлення аномалій безпосередньо в журналах HTTP-серверів. Хоча цей метод ефективний для виявлення атак прикладного рівня (L7), він ігнорує мережеві параметри нижніх рівнів (TCP/IP flags, packet flow duration), що робить систему вразливою до розподілених атак на відмову в обслуговуванні (DDoS), які не завжди коректно відображаються в логах веб-сервера.

Потужний напрям – це класичні та ансамблеві методи машинного навчання, до яких належить і дане дослідження. У роботах [10-11] проводиться порівняльний аналіз алгоритмів SVM, k-NN та Random Forest. Перевага SVM, k-NN це простота реалізації, зрозуміла математична модель, недоліками є: низька швидкість на великих вибірках (k-NN), чутливість до шуму. Авторі дослідження [11] зазначають, що Random Forest демонструє найкращий баланс між швидкістю навчання та точністю класифікації на незбалансованих даних.

Попри велику кількість робіт, більшість досліджень [7, 8] розглядають датасет CICIDS2017 як основний для навчання та тестування моделей. При цьому завдання ставляться дуже узагальнено, намагаючись

класифікувати всі 14 типів атак одночасно, що часто призводить до зниження точності виявлення саме веб-загроз (Web Attacks) на фоні масових DDoS-атак. Тому існує потреба у створенні спеціалізованої моделі, оптимізованої саме для захисту веб-додатків, яка поєднувала б глибину аналізу пакетного трафіку (на відміну від аналізу логів [9]) з високою швидкістю ансамблевих методів (на відміну від важких DL-моделей). Саме цю задачу покликана вирішити дана робота, пропонуючи оптимізовану модель Random Forest для специфічного профілю загроз веб-застосунків.

Формулювання мети дослідження

Метою роботи є підвищення рівня захищеності веб-додатків шляхом розробки моделі машинного навчання на основі алгоритму Random Forest, здатної з високою точністю класифікувати мережевий трафік та виявляти аномалії, характерні для сучасних кібератак.

Для досягнення мети потрібно дослідити існуючі методи аналізу мережевого трафіку та алгоритми машинного навчання, що застосовуються в задачах безпеки веб-додатків; провести вибір та обґрунтування оптимальних типів моделей машинного навчання з урахуванням особливостей мережевого трафіку та задачі прогнозування атак; розробити та реалізувати модель машинного навчання для прогнозування атак на основі зібраних даних мережевого трафіку; провести тестування та оцінку якості розробленої моделі за допомогою відповідних метрик (Precision, Recall, F1-score).

Викладення основного матеріалу дослідження

Найбільш поширеним напрямком у застосуванні машинного навчання до кібербезпеки є контрольоване навчання (Supervised Learning), яке потребує попередньо розмічених даних (наприклад, «атака» чи «нормальний трафік»). Тут використовуються моделі:

- Random Forest – дозволяє будувати ансамбль дерев рішень і добре працює з ознаками трафіку;
- Support Vector Machine – ефективна для задач класифікації з чіткою границею між класами, особливо при високій розмірності простору ознак;
- XGBoost – градієнтний бустинг, який демонструє високу точність при виявленні атак завдяки оптимізації втрат і регуляризації;
- MLP (багатошарові перцептрони) – використовуються для більш складного представлення нелінійних залежностей у трафіку;
- Convolutional Neural Networks (CNN) – хоча більше відомі в комп'ютерному зорі, можуть бути адаптовані до задач безпеки, наприклад, при аналізі трафіку у вигляді часових рядів або образів.

Контрольоване навчання найбільш ефективне при наявності якісних і репрезентативних наборів даних, де класи атак добре розрізняються. Водночас основним викликом залишається отримання та оновлення таких даних, особливо для нових типів атак.

Для виявлення невідомих або атак нульового дня, коли немає маркованих прикладів, застосовують неконтрольоване навчання (Unsupervised Learning). Ці моделі не потребують розмітки та фокусуються на виявленні аномальної поведінки, відмінної від «нормального» трафіку:

- Isolation Forest – ізолює аномалії шляхом випадкового поділу ознак; добре працює з великими наборами даних;
- DBSCAN – кластеризує запити за відстанню та щільністю; незвичні запити автоматично класифікуються як шум (аномалії).
- Autoencoders – нейронні мережі, що навчаються реконструювати «нормальні» зразки; великі відхилення в реконструкції сигналізують про потенційну атаку.

Ці методи ідеально підходять для середовищ з постійно змінюваним трафіком, де оновлення сигнатур не встигає за новими загрозами.

Створення ефективної системи виявлення кібератак вимагає ретельної побудови експериментального пайплайну, який охоплює всі етапи – від завантаження даних до отримання прогнозів. Цей процес починається з підготовки даних, де особливу увагу приділяємо очищенню, нормалізації та балансуванню набору. Кожен етап має вирішальне значення для кінцевого результату, оскільки навіть найдосконаліші алгоритми не зможуть показати хороші результати на неякісних даних. Важливою частиною пайплайну є створення модульної структури, яка дозволяє легко змінювати окремі компоненти для експериментів з різними алгоритмами та підходами.

Для дослідження з побудови моделі було обрано виключно алгоритм Random Forest, що обумовлено низкою критичних переваг цього методу для задач аналізу мережевого трафіку [7, 8]. Random Forest є ансамблевим алгоритмом на основі дерев рішень, який особливо добре проявляє себе при роботі з даними, що мають численні ознаки та складні залежності між ними. Ця характеристика ідеально відповідає природі мережевого трафіку, де кожне з'єднання може описуватися десятками параметрів.

Головною перевагою Random Forest є його здатність ефективно працювати з даними, що мають високий рівень шуму та часткові відсутні значення – типові проблеми реальних наборів даних про мережевий трафік. Алгоритм

демонструє виняткову стійкість до перенавчання завдяки механізму випадкового вибору підмножин ознак для кожного дерева в ансамблі. Ця властивість є особливо важливою для нашої задачі, оскільки набір CICIDS2017 містить значну кількість корельованих ознак.

Для проведення експерименту було обрано набір даних CICIDS2017 (Canadian Institute for Cybersecurity Intrusion Detection System). Цей датасет є одним з найбільш репрезентативних для задач кібербезпеки, оскільки він містить сучасні сценарії атак (DoS, DDoS, Brute Force, Web Attacks, Infiltration) та повні пейлоади пакетів, на відміну від застарілих наборів на кшталт KDD99.

У рамках дослідження було виконано наступні етапи попередньої обробки даних (Data Preprocessing):

1) очищення даних (Data Cleaning): видалення дублікатів та записів із пропущеними значеннями (NaN), а також обробка значень *Infinity*, які виникають при розрахунку статистик потоку (наприклад, ділення на нуль при розрахунку швидкості пакетів);

2) кодування даних (Encoding): категоріальні дані (наприклад, текстові мітки класів атак) було перетворено у числовий формат;

3) нормалізація (Scaling): оскільки ознаки трафіку мають різні діапазони значень (наприклад, тривалість з'єднання та кількість байтів), було застосовано масштабування для покращення збіжності алгоритму;

4) балансування класів: використано підходи бібліотеки *Imbalanced-learn* для усунення дисбалансу між класом нормального трафіку (Benign), якого значно більше, та класами атак.

Після завершення всіх етапів очищення датасет було збережено у новий файл CICIDS2017_cleaned.csv. Фінальна кількість записів після усіх операцій очищення склала 2830628 рядків. Порівняно з початковим обсягом 2830744 рядки даних це становить 99,996% від оригінального набору, що вказує на відносно невелику частку аномальних записів. Отриманий очищений датасет містить узгоджені та коректні дані, готові до подальшого аналізу та використання в моделях машинного навчання.

Наступним етапом був експлораторний аналіз даних. Він є критично важливим етапом у процесі аналізу даних, який дозволяє досліднику ознайомитися з основними характеристиками датасету, виявити закономірності, аномалії та взаємозв'язки між змінними. У контексті дослідження мережевого трафіку, аналіз допомагає зрозуміти структуру даних, розподіл різних типів трафіку, а також виявити потенційні проблеми, які можуть вплинути на подальше моделювання.

Для кращого розуміння було створено не тільки текстовий варіант, а й візуальний. Візуалізація розподілу типів трафіку допомагає виявити дисбаланс класів, що може вплинути на вибір метрик оцінки моделі та стратегії навчання. Для визначення найбільш інформативних ознак використано ANOVA F-test, який оцінює вплив кожної ознаки на цільову змінну. Було відібрано топ-20 ознак із найвищими значеннями F-критерію (рис. 1).

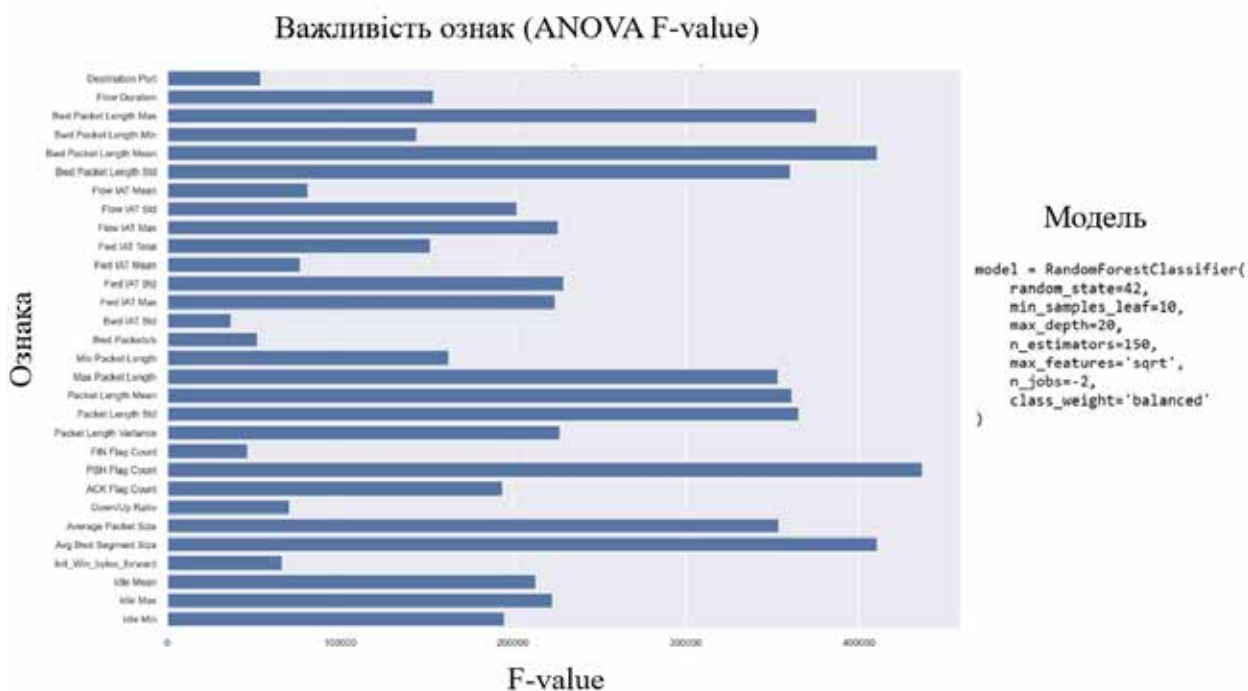


Рис. 1. Інформативні ознаки трафіку та модель класифікації

Практична реалізація моделі виконувалася мовою програмування Python 3.13 з використанням бібліотек Scikit-learn для побудови моделі та Pandas для маніпуляцій з даними. Після навчання моделі на тренувальній вибірці (70% даних) було проведено валідацію на тестовій вибірці (30%). Результати оцінки ефективності класифікатора наведено на рис. 2. Як видно з отриманих даних, модель досягла високої загальної точності у 98%.

Для детальнішого аналізу помилок було побудовано матрицю невідповідностей (Confusion Matrix). Модель Random Forest продемонструвала високу ефективність у класифікації мережевого трафіку на три категорії: 0 – атаки, 1 – звичайний трафік (BENIGN) та 2 – розвідувальні дії (Recon). Матриця невідповідностей показала, що модель правильно ідентифікувала 117091 випадків атак, 669953 випадків звичайного трафіку та 48006 випадків розвідувальних дій. При цьому спостерігається невелика кількість помилок класифікації: 1924 випадки звичайного трафіку були помилково класифіковані як атаки, а 11433 випадки атак були помилково віднесені до звичайного трафіку. Для класу Recon кількість помилок була мінімальною – лише 509 випадків.

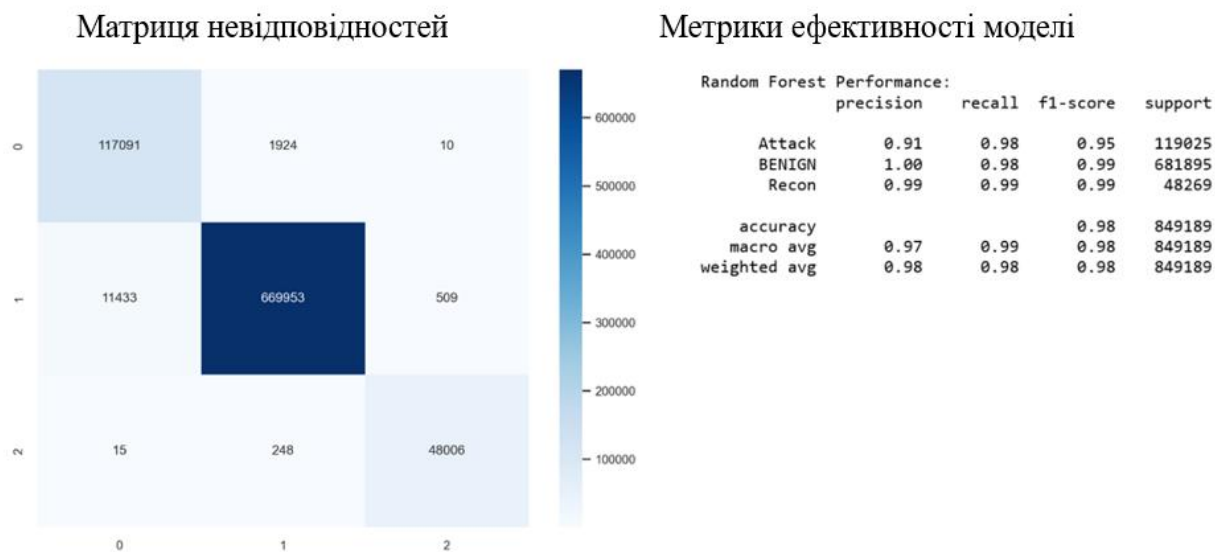


Рис. 2. Матриця невідповідностей та метрики ефективності моделі Random Forest

Аналіз матриці показує, що:

- 1) модель майже безпомилково класифікує нормальний трафік (Benign). Це критично важливо для уникнення блокування легітимних користувачів (мінімізація False Positives);
- 2) висока точність спостерігається для атак типу Recon (Port Scan) та DoS, які мають виражені статистичні аномалії (частота пакетів, тривалість сесії);
- 3) певні труднощі виникають при розрізненні підтипів веб-атак (наприклад, XSS та SQL Injection), що пояснюється схожістю їхніх ознак на рівні мережевих пакетів у зашифрованому трафіку (HTTPS).

Висновки

Завдяки здатності моделей машинного навчання аналізувати великі обсяги трафіку, виявляти приховані залежності та нетипові шаблони, такі підходи можуть розпізнати атаки, які залишаються непоміченими для класичних сигнатурних систем. Наприклад, під час автоматизованих атак з ботнетів, трафік з кожного окремого вузла може виглядати звичайним, однак моделі кластеризації або виявлення аномалій можуть виявити підозрілу синхронність або повторюваність дій. У випадку підвищення привілеїв ML-моделі можуть помітити нетипові шаблони доступу до адміністративних функцій з облікових записів звичайних користувачів.

У роботі було розроблено та досліджено модель виявлення аномалій у мережевому трафіку на базі алгоритму Random Forest. Використання набору даних CICIDS2017 дозволило наблизити умови експерименту до реальних. Отримані результати (точність 98%) підтверджують гіпотезу про те, що машинне навчання є ефективним інструментом для посилення безпеки веб-додатків. Модель демонструє високу здатність до узагальнення та може бути використана як другий ешелон захисту в комплексі з традиційними WAF. Подальші дослідження будуть спрямовані на оптимізацію гіперпараметрів моделі для зменшення часу обробки пакетів та інтеграцію підходів глибокого навчання для кращого розпізнавання складних веб-атак.

Список використаної літератури

1. Adefemi, K. O., Mutanga, M. B., & Alimi, O. A. (2025). A Hybrid CNN–GRU Deep Learning Model for IoT Network Intrusion Detection. *Journal of Sensor and Actuator Networks*, 14(5), 96. <https://doi.org/10.3390/jsan14050096>.
2. Alahmadi, A. A., Aljabri, M., Alhaidari, F., Alharthi, D. J., Rayani, G. E., Marghalani, L. A., Alotaibi, O. B., & Bajandouh, S. A. (2023). DDoS Attack Detection in IoT-Based Networks Using Machine Learning Models: A Survey and Research Directions. *Electronics*, 12(14), 3103. <https://doi.org/10.3390/electronics12143103>.
3. Long, Z., Yan, H., Shen, G. et al. A Transformer-based network intrusion detection approach for cloud security. *J Cloud Comp* 13, 5 (2024). <https://doi.org/10.1186/s13677-023-00574-9>.
4. Wu, Zihan & Zhang, Hong & Wang, Penghai & Sun, Zhibo. (2022). RTIDS: A Robust Transformer-Based Approach for Intrusion Detection System. *IEEE Access*. 10. 1-1. 10.1109/ACCESS.2022.3182333.
5. Bilot, T., Madhoun, N.E., Agha, K.A., & Zouaoui, A. (2023). Graph Neural Networks for Intrusion Detection: A Survey. *IEEE Access*, 11, 49114-49139.
6. Sun, Z., Teixeira, A. M. H., & Toor, S. (2024). GNN-IDS: Graph neural network based intrusion detection system. In *Proceedings of the 19th International Conference on Availability, Reliability and Security (ARES 2024)* (Article 14, 12 pages). <https://doi.org/10.1145/3664476.3664515>.
7. Kartiwi, M. CNN-LSTM: Hybrid Deep Neural Network for Network Intrusion Detection System. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2022.3206425>.
8. Psychogyios K, Papadakis A, Bourou S, Nikolaou N, Maniatis A, Zahariadis T. (2024). Deep Learning for Intrusion Detection Systems (IDSs) in Time Series Data. *Future Internet*. 16(3):73. <https://doi.org/10.3390/fi16030073>.
9. Ahmad, Waqas & Amjad, Muhammad. (2024). Anomaly Detection in HTTP Logs: Leveraging Machine Learning for Uncovering Anomalous Traffic Patterns with SIEM Integration. 622-629. 10.1109/IBCAST61650.2024.10877256.
10. Rahman, M. M., Al Shakil, S. & Mustakim, M. R. (2025) A survey on intrusion detection system in iot networks. *Cyber Secur. Appl.* 3, 100082. <https://doi.org/10.1016/j.csa.2024.100082>.
11. Pradhan, Reshamlal. (2022). Decision Tree Based Classifications on CICIDS 2017 Dataset for the Identification of DDoS, Botnet, and Web Attack. *NeuroQuantology*. 20. 4468-4475. 10.48047/NQ.2022.20.12.NQ77771.

References

1. Adefemi, K. O., Mutanga, M. B., & Alimi, O. A. (2025). A Hybrid CNN–GRU Deep Learning Model for IoT Network Intrusion Detection. *Journal of Sensor and Actuator Networks*, 14(5), 96. <https://doi.org/10.3390/jsan14050096>.
2. Alahmadi, A. A., Aljabri, M., Alhaidari, F., Alharthi, D. J., Rayani, G. E., Marghalani, L. A., Alotaibi, O. B., & Bajandouh, S. A. (2023). DDoS Attack Detection in IoT-Based Networks Using Machine Learning Models: A Survey and Research Directions. *Electronics*, 12(14), 3103. <https://doi.org/10.3390/electronics12143103>.
3. Long, Z., Yan, H., Shen, G. et al. A Transformer-based network intrusion detection approach for cloud security. *J Cloud Comp* 13, 5 (2024). <https://doi.org/10.1186/s13677-023-00574-9>.
4. Wu, Zihan & Zhang, Hong & Wang, Penghai & Sun, Zhibo. (2022). RTIDS: A Robust Transformer-Based Approach for Intrusion Detection System. *IEEE Access*. 10. 1-1. 10.1109/ACCESS.2022.3182333.
5. Bilot, T., Madhoun, N.E., Agha, K.A., & Zouaoui, A. (2023). Graph Neural Networks for Intrusion Detection: A Survey. *IEEE Access*, 11, 49114-49139.
6. Sun, Z., Teixeira, A. M. H., & Toor, S. (2024). GNN-IDS: Graph neural network based intrusion detection system. In *Proceedings of the 19th International Conference on Availability, Reliability and Security (ARES 2024)* (Article 14, 12 pages). <https://doi.org/10.1145/3664476.3664515>.
7. Kartiwi, M. CNN-LSTM: Hybrid Deep Neural Network for Network Intrusion Detection System. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2022.3206425>.
8. Psychogyios K, Papadakis A, Bourou S, Nikolaou N, Maniatis A, Zahariadis T. (2024). Deep Learning for Intrusion Detection Systems (IDSs) in Time Series Data. *Future Internet*. 16(3):73. <https://doi.org/10.3390/fi16030073>.
9. Ahmad, Waqas & Amjad, Muhammad. (2024). Anomaly Detection in HTTP Logs: Leveraging Machine Learning for Uncovering Anomalous Traffic Patterns with SIEM Integration. 622-629. 10.1109/IBCAST61650.2024.10877256.
10. Rahman, M. M., Al Shakil, S. & Mustakim, M. R. (2025) A survey on intrusion detection system in iot networks. *Cyber Secur. Appl.* 3, 100082. <https://doi.org/10.1016/j.csa.2024.100082>.
11. Pradhan, Reshamlal. (2022). Decision Tree Based Classifications on CICIDS 2017 Dataset for the Identification of DDoS, Botnet, and Web Attack. *NeuroQuantology*. 20. 4468-4475. 10.48047/NQ.2022.20.12.NQ77771.

Дата першого надходження рукопису до видання: 22.11.2025
Дата прийнятого до друку рукопису після рецензування: 17.12.2025
Дата публікації: 31.12.2025

І. С. ЧЕПУРНА

асистентка кафедри електронних обчислювальних машин

Харківський національний університет радіоелектроніки

ORCID: 0009-0008-2442-6221

МОДЕЛЬ ЗБОРУ МЕТРИК ОБЧИСЛЮВАЛЬНИХ РЕСУРСІВ БЛОКЧЕЙН-ОРІЄНТОВАНОЇ ІНФОРМАЦІЙНОЇ СИСТЕМИ У ХМАРНОМУ СЕРЕДОВИЩІ

У статті розглянуто задачу формалізації процесу збору метрик обчислювальних ресурсів у блокчейн-орієнтованій інформаційній системі, компоненти якої розгорнуто у хмарному середовищі. Показано, що наявні підходи до моніторингу хмарних та мікросервісних систем зосереджені на інструментах збору телеметрії та моніторингу, однак не забезпечують узгодженого опису багаторівневої архітектури та гетерогенних потоків метрик. Метою роботи є розроблення узагальненої математичної моделі, яка відображає багаторівневу структуру системи та забезпечує уніфіковане подання метрик обчислювальних ресурсів для подальшого аналізу й оптимізації. У межах дослідження побудовано формальну модель ресурсів одного хмарного середовища, введено множини рівнів та вузлів блокчейн-орієнтованої інформаційної системи, а також визначено вектори метрик для окремих вузлів. На цій основі запропоновано тензорне подання процесу збору метрик, яке одночасно враховує рівень архітектури, конкретний вузол і набір базових показників. Такий підхід дозволяє розглядати телеметрію системи як єдиний структурований об'єкт, спростити вибір підмножин даних та побудову агрегованих індикаторів стану. Адекватність моделі перевірено за структурними та інформаційними критеріями, а її працездатність продемонстровано на імітаційному прикладі трирівневої архітектури. Порівняння з традиційним аналізом розрізнених часових рядів підтверджує зменшення трудомісткості локалізації «вузьких місць» та підготовки рішень щодо масштабування і перерозподілу ресурсів. Запропонована модель може бути інтегрована з поширеними стеком моніторингу і слугувати основою для подальших досліджень, пов'язаних із прогнозуванням навантаження, адаптивним резервуванням та побудовою політик керування конфігурацією в мультихмарних та гетерогенних середовищах.

Ключові слова: інформаційна система, блокчейн, хмара, модель, тензор, моніторинг.

I. S. CHEPURNIA

Assistant at the Department of Electronic Computers

Kharkiv National University of Radio Electronics

ORCID: 0009-0008-2442-6221

MODEL OF COMPUTATIONAL RESOURCE METRIC COLLECTION FOR A BLOCKCHAIN-ORIENTED INFORMATION SYSTEM IN A CLOUD ENVIRONMENT

The paper addresses the problem of formalizing the process of collecting metrics of computational resources in a blockchain-oriented information system deployed in a cloud environment. It is shown that existing approaches to monitoring cloud and microservice systems focus on telemetry collection and SLA monitoring tools, but do not provide a consistent description of multi-level architectures and heterogeneous metric streams. The aim of the study is to develop a generalized mathematical model that represents the multi-level structure of the system and provides a unified representation of resource metrics for further analysis and optimisation. Within the proposed approach, a formal model of resources of a single cloud environment is constructed, sets of levels and nodes of the blockchain-oriented information system are introduced, and metric vectors for individual nodes are defined. On this basis, a tensor representation of the metric collection process is proposed, which simultaneously takes into account the architectural level, a particular node and the chosen set of basic indicators. This representation allows system telemetry to be treated as a single structured object, simplifies the selection of data subsets and the construction of aggregated state indicators. The adequacy of the model is verified using structural and informational criteria, and its applicability is demonstrated on a simulation example of a three-level architecture. A comparison with the traditional analysis of separate time series confirms a reduction in the effort required to localise bottlenecks and to prepare decisions on scaling and resource reallocation. The proposed model can be integrated with common monitoring stacks and used as a foundation for further research on workload prediction, adaptive resource reservation and configuration management policies in multi-cloud and heterogeneous environments.

Key words: information system, blockchain, cloud, model, tensor, monitoring.

Постановка проблеми

Сучасні блокчейн-орієнтовані інформаційні системи, що функціонують у хмарному середовищі, набули широкого поширення у фінансовій сфері, промисловому Інтернеті речей, логістиці, державному секторі та сервісах електронного урядування. Їх поширеність зумовлена поєднанням гнучкості й масштабованості хмарної інфраструктури з властивостями блокчейн-технологій – децентралізованим керуванням, незмінністю журналів транзакцій, прозорим аудитом операцій та підвищеним рівнем довіри між учасниками інформаційного обміну [1–3]. На практиці такі системи реалізуються як багатокомпонентні рішення, у яких обчислювальні, мережні та ресурси зберігання даних розподілені між хмарними дата-центрами, периферійними вузлами та спеціалізованими блокчейн-нодами.

Інформаційні системи такого класу характеризуються високою динамікою навантаження, неоднорідністю обчислювальних ресурсів і багаторівневою структурою, яка включає фізичну інфраструктуру, віртуалізовані ресурси, мікросервіси, смарт-контракти та прикладні сервіси, орієнтованих на кінцевого користувача. Забезпечення заданих показників продуктивності вимагає безперервного моніторингу стану обчислювальних ресурсів на всіх рівнях цієї структури [4–5]. При цьому обсяги телеметрії (метрики завантаження процесора, використання оперативної пам'яті, інтенсивності дискових операцій, параметри мережної взаємодії, характеристики обробки транзакцій тощо) у великих хмарних середовищах можуть сягати значних величин, що ускладнює їх обробку, зберігання та інтеграцію для подальшого аналізу в реальному масштабі часу.

Додатковою проблемою є різноманітність метрик, які збираються різними підсистемами моніторингу. Ці дані часто відрізняються форматом подання, часовою роздільною здатністю, контекстом інтерпретації та належать до різних рівнів архітектури інформаційної системи у хмарному середовищі. За відсутності узгодженої структури даних і єдиної моделі представлення метрик стає складно формувати цілісну картину стану системи, аналізувати вплив змін конфігурації на використання ресурсів і досліджувати траєкторію функціонування системи за умов змінного навантаження [4–7]. Це, у свою чергу, ускладнює синтез ефективних політик масштабування, адаптивного резервування обчислювальних ресурсів і своєчасного виявлення аномалій у роботі блокчейн-орієнтованих сервісів.

Останні дослідження демонструють потенціал використання блокчейн-технологій як базового механізму зберігання та верифікації транзакцій, зокрема як інструменту для прозорого обліку використання ресурсів, реєстрації подій порушення роботи інформаційної системи та підтримки децентралізованих механізмів керування ресурсами у хмарному середовищі [6–8]. Проте більшість наявних підходів фокусуються на окремих аспектах (розподіл задач, керування ресурсами, моніторинг тощо) і не пропонують узагальненої математичної моделі, яка б відображала багаторівневу структуру блокчейн-орієнтованої інформаційної системи у хмарному середовищі та забезпечувала уніфіковане подання множини метрик обчислювальних ресурсів для подальшого формального аналізу й оптимізації.

Отже, актуальною є задача побудови узагальненої моделі збору метрик обчислювальних ресурсів у блокчейн-орієнтованій інформаційній системі у хмарному середовищі, яка враховує її багаторівневу архітектуру, дозволяє структуровано представити метрики для різних компонентів інфраструктури та створює формальну основу для дослідження ефективності використання ресурсів, аналізу відмовостійкості й синтезу політик масштабування в умовах змінного та потенційно пікового навантаження.

Аналіз останніх досліджень і публікацій

У сучасних дослідженнях, присвячених моніторингу обчислювальних ресурсів у хмарних та мікросервісних середовищах, основна увага приділяється забезпеченню безперервного контролю стану компонентів інфраструктури та підтримці заданих показників якості сервісу. Запропоновані підходи орієнтовані на збір метрик завантаження процесора, використання оперативної пам'яті, дискових підсистем, мережних інтерфейсів і характеристик обробки запитів для подальшої оптимізації розподілу ресурсів та запобігання деградації продуктивності [9–12]. Зокрема, у дослідженні [9] приведено систему проактивного керування ресурсами для хмар мікросервісів, яка поєднує моніторинг контейнерів і кластерів з прогнозуванням навантаження з метою своєчасного масштабування. У [10] запропоновано стек моніторингу для edge- та хмарних інфраструктур, що забезпечує збір та аналіз великих обсягів телеметрії в режимі, наближеному до реального часу. Дослідження [11, 12] демонструють, що використання стандартних стеків у Kubernetes-кластерах дозволяє скоротити час виявлення інцидентів.

Окремий напрям становлять роботи, де блокчейн використовується як інфраструктурний механізм для децентралізованого керування ресурсами в хмарних середовищах. У [13] наведено огляд підходів до розподілу ресурсів у хмарних обчисленнях із використанням блокчейн-технологій. Розглядаються моделі, у яких інформація про стан ресурсів і рішення щодо їхнього виділення фіксуються у розподіленому реєстрі, що підвищує прозорість і довіру між учасниками. У дослідженні [8] пропонується блокчейн-орієнтована архітектура системи керування хмарними ресурсами, у якій транзакції щодо змін конфігурації інфраструктури, виділення ресурсів та подій

порушення політик системи реєструються у блокчейн-журналі, що забезпечує відтворюваність рішень та аудит історії використання ресурсів. Разом із тим, у зазначених роботах основна увага зосереджена на архітектурних рішеннях і протоколах узгодження, тоді як структурне подання множини метрик обчислювальних ресурсів для багаторівневих блокчейн-орієнтованих інформаційних систем формалізується лише частково.

Значна кількість досліджень присвячена проблемам неоднорідності й попередньої обробки даних моніторингу. У дослідженні [14] виконано систематизацію типів моніторингових даних для хмарних систем та показано, що для задач виявлення аномалій і аналізу поведінки систем необхідні спеціалізовані процедури очищення, нормалізації та синхронізації часових рядів. У науковій праці [15] акцентовано увагу на інтеграції гетерогенних даних із різних джерел, підкреслюючи, що різні формати, частота вимірювань та семантика атрибутів істотно ускладнюють побудову узгоджених моделей для аналітики. У [16] представлено порівняльне дослідження інструментів очищення та контролю якості даних на великих реальних наборах, де показано, що якість подальшого аналізу суттєво залежить від обраних процедур підготовки первинних даних.

Проведений аналіз [8–16] дозволяє зробити висновок, що, попри значний прогрес у розвитку інструментів моніторингу, блокчейн-орієнтованих архітектур для керування ресурсами та методів обробки гетерогенних даних, недостатньо дослідженими залишаються питання формального опису процесу збору метрик обчислювальних ресурсів у багаторівневих блокчейн-орієнтованих інформаційних системах у хмарному середовищі. У більшості робіт метрики розглядаються як множина ізольованих часових рядів або логів для окремих компонентів, без уніфікованого математичного подання, яке б явно відображало багаторівневу структуру системи та забезпечувало цілісне структурування метрик для подальшого аналізу, прогнозування та оптимізації використання обчислювальних ресурсів.

Формулювання мети дослідження

Мета дослідження полягає у розробленні узагальненої математичної моделі збору метрик обчислювальних ресурсів у блокчейн-орієнтованій інформаційній системі у хмарному середовищі, яка відображає її багаторівневу архітектуру та забезпечує уніфіковане структуроване подання гетерогенних метрик для подальшого аналізу, прогнозування та оптимізації використання обчислювальних ресурсів.

Викладення основного матеріалу дослідження

Вихідні припущення та модель ресурсів одного хмарного середовища

У даному дослідженні розглядається блокчейн-орієнтована інформаційна система, компоненти якої розгорнуті в одному хмарному середовищі. Нехай в межах цього середовища функціонує множина вузлів

$$V = \{v_1, v_2, \dots, v_N\}, \quad (1)$$

де N – кількість вузлів (віртуальних машин, контейнерів, сервісів тощо), що належать одному хмарному середовищу.

Нехай:

- R^c – сукупність обчислювальних ресурсів, доступних у хмарному середовищі у межах обраної моделі надання послуг;
- S^c – агреговані сервісні витрати (ресурси, що витрачаються на гіпервізори, системні служби, мережну інфраструктуру тощо).

Тоді умову стабільного та безперервного функціонування всієї сукупності вузлів (1) можна записати у вигляді балансового співвідношення:

$$\sum_{i=1}^N \bar{r}_i + S^c \leq R^c, \quad (2)$$

де $\bar{r}_i$ – вектор ресурсів, необхідних для стабільного функціонування вузла v_i в межах хмарного середовища. Співвідношення (2) задає узагальнену умову припустимості конфігурації інфраструктури, коли сумарні витрати ресурсів на функціонування всіх вузлів та сервісної інфраструктури не повинні перевищувати наявний обсяг ресурсів хмарного середовища.

У межах моделі виділення обчислювальних ресурсів розглядається множина типів ресурсів:

$$\Xi = \{\xi_1, \dots, \xi_M\}. \quad (3)$$

У (3) ξ_i можуть бути, наприклад, ξ_1 – параметри процесорного часу, ξ_2 – параметри оперативної пам'яті, ξ_3 – параметри дискового простору ξ_4 – параметри мережного інтерфейсу тощо. Для кожного вузла v_i вектор $\bar{r}_i$ можна подати у вигляді:

$$\bar{r}_i = (r_{i,\xi_1}, r_{i,\xi_2}, \dots, r_{i,\xi_M}), \quad (4)$$

де r_{i,ξ_m} – обсяг ресурсу типу ξ_m , виділений вузлу v_i при номінальному або середньому навантаженні відповідно до обраної моделі надання послуг.

Багаторівнева структура блокчейн-орієнтованої інформаційної системи

Блокчейн-орієнтована інформаційна система розглядається як багаторівнева структура, де на кожному рівні функціонують вузли певного типу (інфраструктурні вузли, вузли віртуалізації, контейнерні платформи, мікросервіси, блокчейн-ноди, прикладні сервіси тощо).

Нехай множина рівнів задається наступним чином:

$$L = \{1, 3, \dots, L_{\max}\}, \quad (5)$$

де $L_{\max}$ – кількість рівнів багаторівневої архітектури системи.

Для кожного рівня $\ell \in L$ вводиться множина вузлів, згідно з (1):

$$V_{\ell} = \{v_{\ell,1}, v_{\ell,2}, \dots, v_{\ell,N_{\ell}}\}, \quad (6)$$

де N_{ℓ} – кількість вузлів на рівні ℓ .

Тоді ресурсний вектор вузла $v_{\ell,j}$ можна записати з урахуванням (4) у вигляді:

$$\bar{r}_{\ell,j} = (r_{\ell,j,\xi_1}, r_{\ell,j,\xi_2}, \dots, r_{\ell,j,\xi_M}). \quad (7)$$

Баланс ресурсів на рівні хмарного середовища з урахуванням багаторівневої структури (7) має наступний вигляд:

$$\sum_{\ell=1}^{L_{\max}} \sum_{j=1}^{N_{\ell}} \bar{r}_{\ell,j} + S^c \leq R^c. \quad (8)$$

Формула (8) показує сумарне споживання ресурсів усіма вузлами на всіх рівнях архітектури.

Модель метрик обчислювальних ресурсів для вузла

Для аналізу стану інформаційної системи фактичне використання ресурсів на вузлах описується виділеними квотами $\bar{r}_{\ell,j}$ та множиною метрик, що характеризують набором функціональних дій кожного ресурсу у часі.

Нехай для вузла $v_{\ell,j}$ задано множину базових метрик, які вимірюються для кожного ресурсу типу ξ_m :

- завантаження процесора в середньому за інтервал;
- використання оперативної пам'яті;
- інтенсивність дискових операцій;
- пропускна здатність мережі;
- затримка обробки транзакцій тощо.

Тоді:

$$\bar{\mu}_{\ell,j}(t) = (\mu_{\ell,j,1}(t), \mu_{\ell,j,2}(t), \dots, \mu_{\ell,j,P}(t)) \quad (9)$$

є вектором метрик вузла $v_{\ell,j}$ в момент часу t .

У (9) P – кількість обраних базових метрик. При цьому кожен компонент $\mu_{\ell,j,p}(t)$ є функцією фактичного використання відповідних ресурсів:

$$\bar{\mu}_{\ell,j}(t) = F_{\ell,j}(\bar{r}_{\ell,j}(t)), \quad (10)$$

де $F_{\ell,j}(\cdot)$ – функція процесу збору та обчислення метрик для вузла $v_{\ell,j}$.

Тензорне подання процесу збору метрик у багаторівневій системі

Для тензорного подання процесу збору метрик у багаторівневій системі спочатку необхідно згрупувати метрики всіх вузлів на кожному рівні в матрицю:

$$M_{\ell}(t) = \begin{bmatrix} \bar{\mu}_{\ell,1}(t) \\ \bar{\mu}_{\ell,2}(t) \\ \vdots \\ \bar{\mu}_{\ell,N_{\ell}}(t) \end{bmatrix}, \quad \ell \in L. \quad (11)$$

Матриця $M_{\ell}(t)$ (11) має розмірність $N_{\ell} \times P$ і містить значення всіх базових метрик для вузлів рівня ℓ у момент часу t .

Для всієї блокчейн-орієнтованої інформаційної системи, яка функціонує в одному хмарному середовищі, стан метрик обчислювальних ресурсів у момент часу t можна подати у вигляді тензора:

$$\boxed{M(t)} = \{M_{\ell}(t)\}_{\ell=1}^{L_{\max}}. \quad (12)$$

Формула (12) відображає значення метрик для всіх вузлів на кожному рівні багаторівневої архітектури. У загальному випадку $\boxed{M(t)}$ можна розглядати як тривимірний масив розмірності $L_{\max} \times N_{\max} \times P$, в якому $N_{\max} = \max_{\ell} N_{\ell}$ – максимальна кількість вузлів на рівні, а відсутні вузли на окремих рівнях можуть позначатися нульовими або спеціальними значеннями.

Таке тензорне подання процесу збору метрик обчислювальних ресурсів (12) забезпечує:

- уніфіковану структуру даних для гетерогенних компонентів блокчейн-орієнтованої інформаційної системи у хмарному середовищі;
- можливість формального аналізу стану системи за рівнями, вузлами та типами метрик;
- основу для подальшого застосування методів агрегації, виявлення аномалій, прогнозування навантаження та синтезу політик масштабування й адаптивного резервування ресурсів.

Перевірка адекватності запропонованої моделі

Адекватність запропонованої моделі збору метрик обчислювальних ресурсів у блокчейн-орієнтованій інформаційній системі у хмарному середовищі перевіряється за двома групами критеріїв: структурними та інформаційними.

До структурних критеріїв належить відповідність багаторівневої схеми моделі реальній архітектурі системи: множина рівнів L повинна відображати виділені в інфраструктурі шари (фізичні сервери, віртуальні машини, контейнерні платформи, блокчейн-ноди, прикладні сервіси), а множини вузлів V_ℓ – конкретні екземпляри компонентів на кожному рівні. У рамках перевірки здійснюється відображення імітованої моделі конфігурації тестового стенду на множини L та V_ℓ , що дозволяє підтвердити, що кожному елементу інфраструктури відповідає єдиний вузол моделі, а агреговане співвідношення (8) коректно описує баланс використання ресурсів у хмарному середовищі.

Інформаційні критерії пов'язані з повнотою та узгодженістю метрик, що збираються. Вектор метрик $\vec{\mu}_{\ell,j}(t)$ для вузла $v_{\ell,j}$ розширюється до мінімально необхідного набору показників, які реально підтримуються системами моніторингу (наприклад, Prometheus-орієнтований стек, експортери метрик хмарної платформи, телеметрія блокчейн-нод). Перевірка адекватності полягає у демонстрації того, що всі фактично доступні метрики можуть бути однозначно віднесені до компонентів вектора $\vec{\mu}_{\ell,j}(t)$, а тензорне подання $\boxed{M(t)}$ дозволяє без втрати інформації відновити початкові часові ряди для окремих вузлів і рівнів.

Додатково перевіряється коректність інтерпретації моделі в типових експлуатаційних сценаріях: зростання навантаження на прикладному рівні відображається у $\boxed{M(t)}$ як узгоджена зміна відповідних метрик на рівнях мікросервісів, блокчейн-нод і інфраструктурних вузлів. Відмова вузла призводить до локальної зміни структури тензора (набуття нульового значення або маркування відповідних елементів), не порушуючи цілісності моделі в цілому. Така поведінка свідчить про адекватність запропонованої моделі з погляду відображення динаміки реальної інформаційної системи.

Експериментальна оцінка та порівняння з існуючими підходами

Експериментальна оцінка запропонованої моделі виконувалася в рамках імітаційного експерименту, що відтворює роботу блокчейн-орієнтованої інформаційної системи з багаторівневою архітектурою у хмарному середовищі. Розглядалося три рівні:

- рівень інфраструктури (L_1), до якого належать віртуальні машини;
- рівень блокчейн-нод (L_2), де здійснюється обробка й фіксація транзакцій;
- рівень мікросервісів (L_3), що реалізують прикладну логіку.

Для кожного рівня моделювалась множина вузлів, а протягом дискретного часу $t = 0, 1, \dots, 119$ генерувалися часові ряди завантаження процесора, що відображають перехід від номінального режиму до режиму підвищеного навантаження. На основі цих даних формувався тензор $\boxed{M(t)}$, який узгоджено описує стан обчислювальних ресурсів за рівнями, вузлами та часом.

На рис. 1 наведено динаміку завантаження процесорів вузлів на кожному з трьох рівнів. У початковому інтервалі часу спостерігається помірне коливання завантаження в межах 30-40 %, що відповідає номінальному режиму роботи інформаційної системи. В околі моменту $t \approx 40$ фіксується помітне зростання завантаження на всіх рівнях, яке відображає зміну режиму функціонування (збільшення інтенсивності запитів і транзакцій). Надалі криві демонструють хвилюподібну динаміку із збереженням підвищеного середнього рівня завантаження, причому для рівнів L_2 та L_3 характерні дещо вищі пікові значення, що відповідає концентрації навантаження на сервісних компонентах.

Рис. 2 ілюструє середнє завантаження процесорів для кожного рівня багаторівневої архітектури, обчислене на основі тензорного подання $\boxed{M(t)}$. Видно, що після переходу до режиму зростаючого навантаження середні значення для L_2 та L_3 стають вищими, ніж для L_1 , що відображає більшу чутливість цих рівнів до зміни інтенсивності транзакцій. Такий характер кривих відповідає очікуваній поведінці системи: першим реагує прикладний та блокчейн-рівень, тоді як інфраструктурний рівень демонструє більш згладжену динаміку.

Для демонстрації здатності моделі локалізувати «вузькі місця» побудовано карту розподілу завантаження процесора за рівнями та вузлами у піковий момент часу (рис. 3). По вертикальній осі відкладено рівні архітектури, по горизонтальній – вузли на кожному рівні. За інтенсивністю кольору можна виділити вузли з підвищеним завантаженням (зокрема, один із вузлів на рівні L_2 та кілька вузлів на рівні L_3) і вузли з відносно низьким

навантаженням, які потенційно можуть бути використані для перерозподілу обчислювальних ресурсів. Таким чином, тензорне подання $\mathbf{M}(t)$ дозволяє в єдиному форматі представити метрики для всієї системи та підтримує побудову узагальнених візуалізацій, що поєднують інформацію про рівень архітектури, окремі вузли та значення метрик.

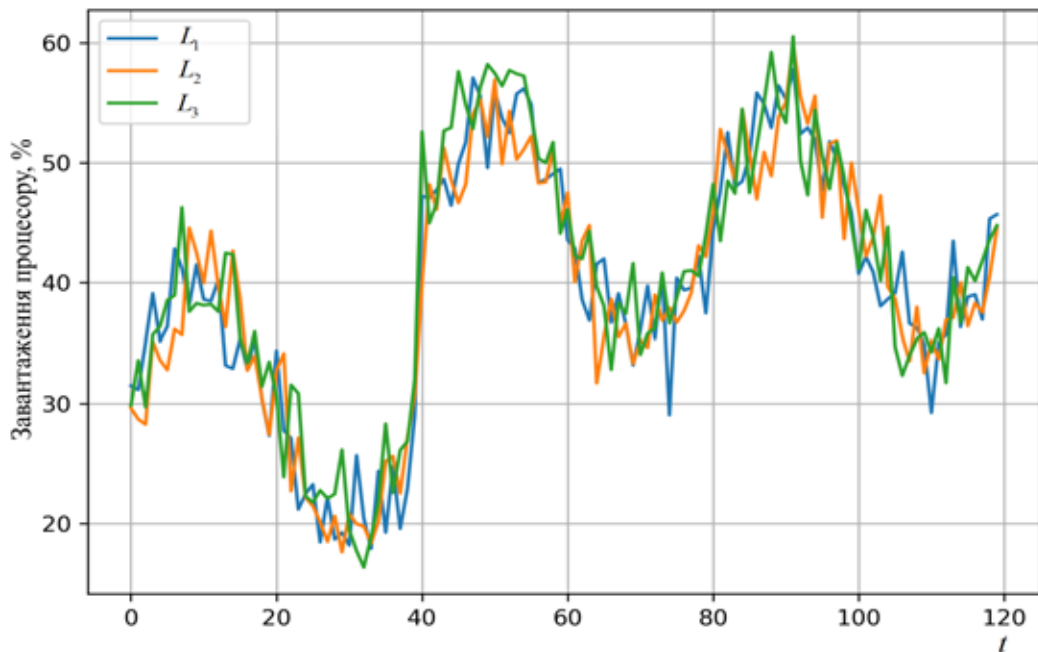


Рис. 1. Динаміка завантаження процесорів вузлів на кожному з трьох рівнів

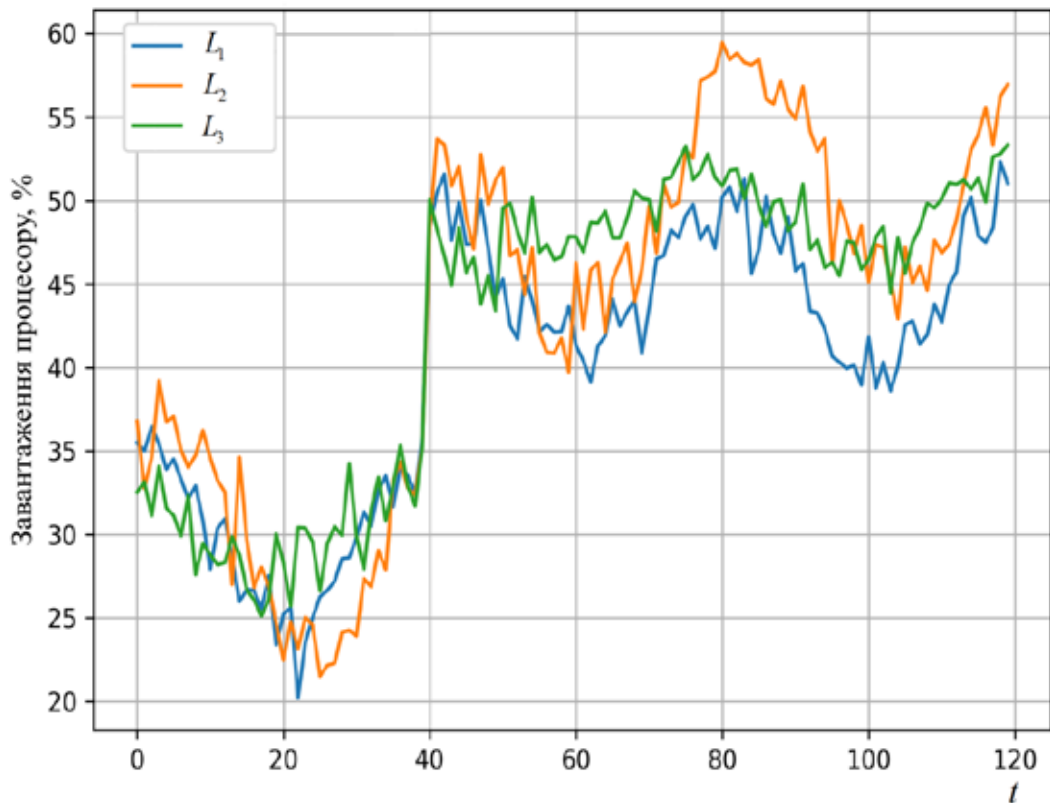


Рис. 2. Динаміка завантаження процесорів для кожного рівня багаторівневої архітектури

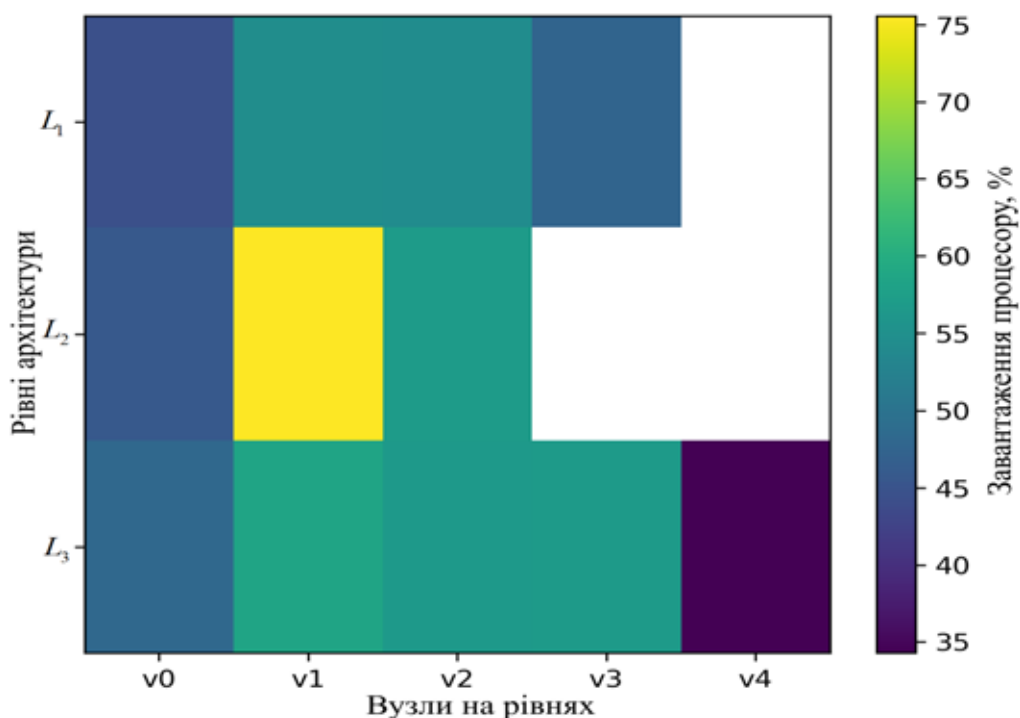


Рис. 3. Карту розподілу завантаження процесора за рівнями та вузлами у піковий момент часу

Порівняно з традиційним підходом, у якому аналіз моніторингових даних виконується на основі набору розрізнених структур агрегації даних для окремих компонентів, запропонована модель забезпечує формальний опис структури метрик для всієї інформаційної системи як єдиного об'єкта. Це спрощує вибір підмножин даних (за рівнем, вузлом, типом ресурсу), зменшує кількість кроків, необхідних для локалізації «вузьких місць», та створює передумови для подальшого застосування методів виявлення аномалій, прогнозування навантаження й синтезу політик масштабування та адаптивного резервування обчислювальних ресурсів у блокчейн-орієнтованих інформаційних системах у хмарному середовищі.

Висновки

У роботі розглянуто задачу формалізації процесу збору метрик обчислювальних ресурсів у блокчейн-орієнтованій інформаційній системі у хмарному середовищі. На основі аналізу сучасних підходів до моніторингу хмарних, мікросервісних та блокчейн-орієнтованих систем показано, що існуючі рішення здебільшого зосереджуються на інструментах збору телеметрії, SLA-моніторингу та механізмах розподілу ресурсів, але не пропонують узагальненої математичної моделі, яка б уніфіковано відображала багаторівневу структуру системи та множину гетерогенних метрик обчислювальних ресурсів.

Напрямок моделювання блокчейн-орієнтованих інформаційних систем у хмарному середовищі набув подальшого розвитку за рахунок запропонованих у цьому дослідженні засобів формалізації процесу збору метрик у вигляді тензорного подання $\mathbf{M}(t)$, яке одночасно враховує рівень архітектури, конкретний вузол та множину базових метрик. Це дозволило поєднати у єдиній моделі як ресурсні обмеження хмарного середовища, так і динаміку фактичного використання ресурсів на різних рівнях архітектури. Запропонований підхід забезпечує структуроване представлення гетерогенних даних моніторингу та створює формальну основу для їх подальшого аналізу.

Перевірка адекватності моделі за структурними та інформаційними критеріями показала, що множини рівнів, вузлів і метрик можуть бути узгоджено зіставлені з реальною конфігурацією блокчейн-орієнтованої системи у хмарному середовищі, а тензорне подання $\mathbf{M}(t)$ не призводить до втрати інформації щодо стану окремих компонентів. Імітаційний експеримент із трирівневою архітектурою продемонстрував, що модель придатна для локалізації «вузьких місць», побудови агрегованих показників стану за рівнями та аналізу реакції системи на зміну режиму навантаження. Порівняння з традиційним підходом, засвідчило перевагу запропонованого підходу з точки зору структурованості даних та зменшення трудомісткості аналітичних процедур.

Отримані результати мають як теоретичне, так і практичне значення. З теоретичної точки зору модель задає узагальнену схему опису метрик обчислювальних ресурсів для блокчейн-орієнтованих інформаційних систем у хмарному середовищі та може використовуватися як базис для побудови методів формального аналізу, оцінювання відмовостійкості та ефективності використання ресурсів. З практичної точки зору тензорне подання $\mathbf{M}(t)$

може бути інтегроване з існуючим стеком моніторингу (Prometheus, Grafana, системи логування), спрощуючи побудову багаторівневих візуалізацій, підтримку процедур локалізації проблемних вузлів і обґрунтування рішень щодо масштабування та перерозподілу ресурсів.

Подальші дослідження доцільно спрямувати на розширення запропонованої у цьому дослідженні моделі для мультимарних сценаріїв побудови обчислювального середовища та сценаріїв з динамічним міграційним перерозподілом сервісів між провайдерами (вендорами хмарних послуг), а також на інтеграцію тензорного подання метрик із методами прогнозування навантаження, адаптивного резервування ресурсів [17] та побудови політик управління конфігурацією блокчейн-орієнтованих інформаційних систем в гетерогенних середовищах [18-19].

Список використаної літератури

1. Dorsala M. R., Sastry V. N., Chapram S. Blockchain-based solutions for cloud computing: A survey. *Journal of Network and Computer Applications*. 2021. Vol. 196. 103246. DOI: <https://doi.org/10.1016/j.jnca.2021.103246>.
2. Ahmed W. Blockchain Integration in Modern Cloud Computing: A Comprehensive Survey of Security and Efficiency. *Premier Journal of Data Science*. 2025. DOI: <https://doi.org/10.70389/pjds.100003>.
3. A Systematic Review of Blockchain, AI, and Cloud Integration for Secure Digital Ecosystems / J. Singh et al. *International Journal of Networked and Distributed Computing*. 2025. Vol. 13, no. 2. DOI: <https://doi.org/10.1007/s44227-025-00072-1>.
4. Kanthed, S. Monitoring of Cloud Computing Environments: Concepts, Solutions, Trends, and Future Directions. *International Journal on Science and Technology*. 2024. Vol. 15, no. 1. DOI: <https://doi.org/10.71097/ijst.v15.i1.2836>.
5. Bliedy D., Khafagy M. H., Badry R. M. Resource Utilization Prediction Model for Cloud Datacentre: Survey. *International Journal of Advanced Computer Science and Applications*. 2025. Vol. 16, no. 3. DOI: <https://doi.org/10.14569/ijacsa.2025.0160380>.
6. Baranwal G., Kumar D., Vidyarthi D. P. Blockchain based resource allocation in cloud and distributed edge computing: A survey. *Computer Communications*. 2023. DOI: <https://doi.org/10.1016/j.comcom.2023.07.023>.
7. Гнатюк В. О., Зандер К. Ю. Методика формування структури центру збирання та оброблення даних під час моніторингу стану об'єктів критичної інфраструктури. *Інфокомунікаційні та комп'ютерні технології*. 2025. Т. 1, № 09. С. 9–17. DOI: <https://doi.org/10.36994/2788-5518-2025-01-09-01>.
8. Towards a Decentralized Blockchain-Based Resource Monitoring Solution For Distributed Environments / R. B. Dos Passos et al. *Journal of Internet Services and Applications*. 2024. Vol. 15, no. 1. P. 1–13. DOI: <https://doi.org/10.5753/jisa.2024.3813>.
9. Kravchenko P., Skriabin B., Dubinina O. Blockchain And Decentralized Systems. *Blockchain and decentralized systems*, 2019. 446 p.
10. Real-time Monitoring and Analysis of Edge and Cloud Resources / I. Korontanis et al. *HPDC '23: The 32nd International Symposium on High-Performance Parallel and Distributed Computing*, Orlando FL USA. New York, NY, USA, 2023. DOI: <https://doi.org/10.1145/3589010.3594892>.
11. Designing a Real-Time Monitoring System for the AWS Cloud: An Adaptive Dashboard-Based Approach with Prometheus and Grafana / A.W. Bello et al. *CITA 2025 – Emerging Technologies and Sustainable Agriculture, 26-28 June 2025, Cotonou, Benin*. URL: <https://ceur-ws.org/Vol-4036/Paper9.pdf>.
12. Monitoring tools for DevOps and microservices: A systematic grey literature review / L. Giamattei et al. *Journal of Systems and Software*. 2024. Vol. 208. P. 111906. DOI: <https://doi.org/10.1016/j.jss.2023.111906>.
13. Baranwal G., Kumar D., Vidyarthi D. P. Blockchain based resource allocation in cloud and distributed edge computing: A survey. *Computer Communications*. 2023. DOI: <https://doi.org/10.1016/j.comcom.2023.07.023>.
14. Monitoring data for Anomaly Detection in Cloud-Based Systems: A Systematic Mapping Study / A. Hrusto et al. *ACM Transactions on Software Engineering and Methodology*. 2025. DOI: <https://doi.org/10.1145/3744556>.
15. Putrama I. M., Martinek P. Heterogeneous data integration: Challenges and opportunities. *Data in Brief*. 2024. 110853. DOI: <https://doi.org/10.1016/j.dib.2024.110853> (date of access: 23.11.2025).
16. Performance and Scalability of Data Cleaning and Preprocessing Tools: A Benchmark on Large Real-World Datasets / P. Martins et al. *Data*. 2025. Vol. 10, no. 5. P. 68. DOI: <https://doi.org/10.3390/data10050068>.
17. Kuchuk N., Tkachov V. Self-healing Systems Modelling. *Advances in Self-healing Systems Monitoring and Data Processing*. Cham, 2022. P. 57–111. DOI: https://doi.org/10.1007/978-3-030-96546-4_2.
18. Фролов Д.Є. Ways of achieving fault tolerance of heterogeneous information systems under conditions of external influence. Проблеми інформатизації : тези доповідей дванадцятої міжнар. наук.-техн. конф., 21–22 листопада 2024 р. Том 2: секція 4. Баку–Харків–Бельсько-Бяла, 2024. С. 79.
19. Tokar L. O., Tsyliuryk V. Y., Solodilov V. V. Study of data replication process using Raft replication algorithm to maintain consistency in server cluster. *Radiotekhnika*. 2024. No. 217. P. 117–127. DOI: <https://doi.org/10.30837/rt.2024.2.217.10>.

References

1. Dorsala, M. R., Sastry, V. N., & Chapram, S. (2021). Blockchain-based solutions for cloud computing: A survey. *Journal of Network and Computer Applications*, 196, 103246. <https://doi.org/10.1016/j.jnca.2021.103246>
2. Ahmed, W. (2025). Blockchain integration in modern cloud computing: A comprehensive survey of security and efficiency. *Premier Journal of Data Science*. <https://doi.org/10.70389/pjds.100003>
3. Singh, J., et al. (2025). A systematic review of blockchain, AI, and cloud integration for secure digital ecosystems. *International Journal of Networked and Distributed Computing*, 13(2). <https://doi.org/10.1007/s44227-025-00072-1>
4. Kanthed, S. (2024). Monitoring of cloud computing environments: Concepts, solutions, trends, and future directions. *International Journal on Science and Technology*, 15(1). <https://doi.org/10.71097/ijst.v15.i1.2836>
5. Bliedy, D., Khafagy, M. H., & Badry, R. M. (2025). Resource utilization prediction model for cloud datacentre: Survey. *International Journal of Advanced Computer Science and Applications*, 16(3). <https://doi.org/10.14569/ijacsa.2025.0160380>
6. Baranwal, G., Kumar, D., & Vidyarthi, D. P. (2023). Blockchain based resource allocation in cloud and distributed edge computing: A survey. *Computer Communications*. <https://doi.org/10.1016/j.comcom.2023.07.023>
7. Hnatiuk, V. O., & Zander, K. Yu. (2025). Metodyka formuvannia struktury tsentru zbyrannia ta obroblennia danykh pid chas monitorynhu stanu ob'ektiv krytychnoi infrastruktury. *Infokomunikatsiini ta kompiuterni tekhnolohii*, 1(09), 9–17. <https://doi.org/10.36994/2788-5518-2025-01-09-01>
8. Dos Passos, R. B., et al. (2024). Towards a decentralized blockchain-based resource monitoring solution for distributed environments. *Journal of Internet Services and Applications*, 15(1), 1–13. <https://doi.org/10.5753/jisa.2024.3813>
9. Kravchenko, P., Skriabin, B., & Dubinina, O. (2019). Blockchain and decentralized systems. [Monograph, 446 pp.].
10. Korontanis, I., et al. (2023). Real-time monitoring and analysis of edge and cloud resources. In *Proceedings of the 32nd International Symposium on High-Performance Parallel and Distributed Computing (HPDC '23)*. ACM. <https://doi.org/10.1145/3589010.3594892>
11. Bello, A. W., et al. (2025). Designing a real-time monitoring system for the AWS cloud: An adaptive dashboard-based approach with Prometheus and Grafana. In *CITA 2025 – Emerging Technologies and Sustainable Agriculture* (26–28 June 2025, Cotonou, Benin). Retrieved from <https://ceur-ws.org/Vol-4036/Paper9.pdf>
12. Giamattei, L., et al. (2024). Monitoring tools for DevOps and microservices: A systematic grey literature review. *Journal of Systems and Software*, 208, 111906. <https://doi.org/10.1016/j.jss.2023.111906>
13. Baranwal, G., Kumar, D., & Vidyarthi, D. P. (2023). Blockchain based resource allocation in cloud and distributed edge computing: A survey. *Computer Communications*. <https://doi.org/10.1016/j.comcom.2023.07.023>
14. Hrusto, A., et al. (2025). Monitoring data for anomaly detection in cloud-based systems: A systematic mapping study. *ACM Transactions on Software Engineering and Methodology*. <https://doi.org/10.1145/3744556>
15. Putrama, I. M., & Martinek, P. (2024). Heterogeneous data integration: Challenges and opportunities. *Data in Brief*, 110853. <https://doi.org/10.1016/j.dib.2024.110853>
16. Martins, P., et al. (2025). Performance and scalability of data cleaning and preprocessing tools: A benchmark on large real-world datasets. *Data*, 10(5), 68. <https://doi.org/10.3390/data10050068>
17. Kuchuk, N., & Tkachov, V. (2022). Self-healing systems modelling. In *Advances in self-healing systems monitoring and data processing* (pp. 57–111). Springer. https://doi.org/10.1007/978-3-030-96546-4_2
18. Frolov, D. Ye. (2024). Ways of achieving fault tolerance of heterogeneous information systems under conditions of external influence. In *Problemy informatyzatsii: tezy dopovidei dvanadtsiatoi mizhnarodnoi naukovo-tekhnichnoi konferentsii, tom 2, sektsiia 4* (p. 79). Baku–Kharkiv–Bielsko-Biala.
19. Tokar, L. O., Tsyliuryk, V. Y., & Solodilov, V. V. (2024). Study of data replication process using Raft replication algorithm to maintain consistency in server cluster. *Radiotekhnika*, 217, 117–127. <https://doi.org/10.30837/rt.2024.2.217.10>

Дата першого надходження рукопису до видання: 27.11.2025
Дата прийнятого до друку рукопису після рецензування: 15.12.2025
Дата публікації: 31.12.2025

В. М. ЧИЖИК

кандидат технічних наук,
доцент кафедри транспортних систем і логістики
Харківський національний автомобільно-дорожній університет
ORCID: 0000-0003-0518-5246

М. В. ХАРЧЕНКО

кандидат економічних наук,
доцент кафедри менеджменту
Класичний приватний університет
ORCID: 0000-0001-7435-0642

А. О. ТВЕРДОХЛІБ

аспірант кафедри комп'ютерної інженерії
Державний університет інформаційно-комунікаційних технологій
ORCID: 0000-0002-6591-2866

ВИКОРИСТАННЯ БЛОКЧЕЙН-ТЕХНОЛОГІЙ ДЛЯ ПРОЗОРОСТІ ЛАНЦЮГІВ ПОСТАЧАННЯ В ТРАНСПОРТНІЙ ЛОГІСТИЦІ

Стаття присвячена дослідженню теоретичних основ і практичних аспектів упровадження блокчейн-технологій у сфері транспортної логістики з метою забезпечення прозорості, відстежуваності та безпеки даних у ланцюгах постачання. У роботі аналізується сучасний стан цифровізації логістичних процесів, зокрема застосування блокчейну, Інтернету речей (далі – IoT), смартконтрактів та аналітичних інструментів для підвищення ефективності управління товарними потоками. Особлива увага приділяється перевагам використання децентралізованої системи збереження даних, що дозволяє забезпечити невідворотність записів, захист від несанкціонованих змін та автоматизацію фінансових і договірних процесів у ланцюгах постачання. У статті розглядаються основні напрями застосування блокчейн-технологій, зокрема підвищення прозорості транзакцій, створення цифрових паспортів товарів для контролю походження та автентичності, відстеження переміщення вантажів у реальному часі, автоматизацію платежів та управління цифровими документами. Підкреслено важливість інтеграції блокчейн-рішень із сучасними логістичними платформами та технологіями IoT, що забезпечує повний контроль за параметрами продуктів, особливо у чутливих сферах, таких як харчова й фармацевтична продукція. Виокремлено основні перешкоди впровадження блокчейну в логістику, до яких належать інертність учасників ланцюга до нових технологій, необхідність висококваліфікованих фахівців, а також вибір платформ і програмного забезпечення для реалізації технології. У роботі аналізуються перспективи розвитку блокчейн-технологій у транспортно-логістичній сфері, зокрема їхній потенціал для оптимізації маршрутів доставки, прогнозування попиту та підвищення економічної ефективності підприємств. Стаття показує, що використання блокчейну сприяє не лише підвищенню операційної продуктивності та зниженню витрат, але й формуванню довіри між усіма учасниками ланцюга постачань, зміцненню репутації компаній і підвищенню лояльності кінцевих споживачів. Зроблено висновок про стратегічне значення блокчейн-технологій як інструменту цифрової трансформації логістики та управління ланцюгами постачання, який забезпечує прозорість, безпеку й стабільність у глобальному економічному середовищі.

Ключові слова: блокчейн, транспортна логістика, ланцюги постачання, прозорість, відстежуваність, безпека даних, смартконтракти, цифрові технології, Інтернет речей, оптимізація логістичних процесів.

V. M. CHYZHYK

PhD in Engineering Science (Transport Systems),
Associate Professor at the Department of Transport Systems and Logistics
Kharkiv National Automobile and Highway University
ORCID: 0000-0003-0518-5246

M. V. KHARCHENKO

Candidate of Economic Sciences,
Associate Professor at the Department of Management
Classic Private University
ORCID: 0000-0001-7435-0642

A. O. TVERDOKHLIB

Postgraduate Student at the Department of Computer Engineering
State University of Information and Communication Technologies
ORCID: 0000-0002-6591-2866

USING BLOCKCHAIN TECHNOLOGIES TO ENSURE TRANSPARENCY IN SUPPLY CHAINS WITHIN TRANSPORTATION LOGISTICS

The article is devoted to the study of theoretical foundations and practical aspects of implementing blockchain technologies in transportation logistics to ensure transparency, traceability, and data security in supply chains. The work analyzes the current state of digitization in logistics processes, including the use of blockchain, the Internet of Things, smart contracts, and analytical tools to enhance the efficiency of goods flow management. Special attention is given to the advantages of using a decentralized data storage system, which ensures the immutability of records, protection against unauthorized changes, and automation of financial and contractual processes within supply chains. The article examines key areas of blockchain application, including improving transaction transparency, creating digital passports for products to verify origin and authenticity, real-time tracking of cargo movements, automating payments, and managing digital documents. The study emphasizes the importance of integrating blockchain solutions with modern logistics platforms and Internet of Things technologies, providing full control over product parameters, particularly in sensitive sectors such as food and pharmaceuticals. The article also highlights major barriers to blockchain adoption in logistics, including the reluctance of supply chain participants to adopt new technologies, the need for highly skilled specialists, and the selection of appropriate platforms and software for implementation. The prospects for blockchain development in the transportation and logistics sector are analyzed, particularly its potential to optimize delivery routes, forecast demand, and improve the economic efficiency of enterprises. The article demonstrates that the use of blockchain not only enhances operational productivity and reduces costs but also builds trust among all supply chain participants, strengthens company reputations, and increases end-customer loyalty. In conclusion, the authors emphasize the strategic importance of blockchain technologies as a tool for digital transformation in logistics and supply chain management, ensuring transparency, security, and stability in the global economic environment.

Key words: blockchain, transportation logistics, supply chains, transparency, traceability, data security, smart contracts, digital technologies, Internet of Things, logistics process optimization.

Постановка проблеми

Сучасна транспортна логістика функціонує в умовах зростання складності глобальних ланцюгів постачання, що охоплюють численних учасників, різні етапи та географічні регіони. Це спричиняє виникнення ризиків, пов'язаних із затримками постачань, фальсифікацією документів, втручанням у дані, шахрайством і зниженням довіри між партнерами. Традиційні інформаційні системи управління логістичними процесами не забезпечують належного рівня прозорості, оскільки базуються на централізованих структурах даних і не гарантують повної автентичності інформації. В умовах цифрової трансформації все більшої уваги набувають блокчейн-технології, що забезпечують децентралізоване зберігання даних, незмінність записів і можливість простежування кожної операції. Вони створюють потенціал для підвищення ефективності контролю, зменшення транзакційних витрат і запобігання шахрайству в транспортній логістиці. Однак практичне впровадження блокчейну стикається з низкою проблем, зокрема відсутністю єдиних стандартів, високою вартістю реалізації, технічною складністю інтеграції з наявними логістичними системами та недостатньою готовністю учасників ринку до змін. Унаслідок цього постає наукова проблема, що полягає в необхідності глибокого аналізу потенціалу, обмежень та економічної доцільності застосування блокчейн-технологій для забезпечення прозорості ланцюгів постачання в транспортній логістиці, а також розроблення ефективних підходів до їх адаптації в практичну діяльність.

Аналіз останніх досліджень і публікацій

Значний внесок у дослідження застосування блокчейн-технологій у транспортній логістиці та управлінні ланцюгами постачання зроблено як вітчизняними, так і зарубіжними авторами. Зокрема, Ю. Рябчун, Д. Компанець, В. Хроленко та Н. Попович аналізують застосування блокчейн-технологій для забезпечення цілісності даних у логістичних системах, підкреслюючи важливість прозорості транзакцій та відстежуваності вантажів у реальному часі [1]. Перспективи впровадження блокчейну в транспортну логістику та управління ланцюгами постачань розглядає Т. Колодізева, акцентуючи на автоматизації транзакцій, створенні цифрових паспортів товарів та підвищенні прозорості ланцюга постачання [2]. Сутність цифрової трансформації в логістичній сфері

досліджують Є. Марінов та Є. Лісений, наголошуючи на зміні традиційних підходів до управління ланцюгами постачання та розподілу товарів. Автори підкреслюють посилення значення цифрових технологій у забезпеченні точності, ефективності та швидкості логістичних процесів, що дозволяє підприємствам адаптуватися до умов високої конкуренції й динамічного ринку [3]. Це джерело доповнює наукову дискусію щодо стратегічного використання інноваційних інструментів у логістиці, створюючи практичне підґрунтя для подальшої інтеграції таких технологій, як штучний інтелект, автоматизація та аналітика даних. На застосуванні блокчейну в складській логістиці зосереджуються В. Химич та Т. Білоконь, показуючи його вплив на управління запасами, контроль переміщення товарів та документообіг [4]. Вплив інноваційних технологій на ефективність та конкурентоспроможність підприємств аналізує Л. Якимишин, зазначаючи, що інтеграція блокчейн з аналітичними інструментами та штучним інтелектом дозволяє оптимізувати витрати й підвищити продуктивність логістичних процесів [5]. Розвиток інноваційних технологій в аграрній логістиці розглядає О. Школьний, підкреслюючи значення прозорості ланцюга постачання для підвищення довіри споживачів [6]. Економічний потенціал і виклики впровадження інноваційних технологій у транспортній логістиці досліджує Є. Марінов, наголошуючи на необхідності стратегічного планування та підготовки персоналу для ефективної інтеграції блокчейн-технологій [7]. Сучасні технології управління транспортною логістикою, зокрема IoT та цифрові платформи, вивчає О. Головіна [8]. Вплив Економіки 4.0 на глобальні логістичні операції транснаціональних корпорацій оцінюють В. Македон та О. Михайленко, підкреслюючи значення цифрових технологій для оптимізації ланцюгів постачання [9]. Цифровізацію логістики, нові технології управління та оптимізації процесів розглядають Н. Канцедал, О. Лега та Є. Морозов [10].

Попри значну увагу науковців до проблематики дослідження, низка аспектів залишається нерозв'язаною. Зокрема, це проблеми впровадження блокчейн-технологій у транспортній логістиці, пов'язані з відсутністю єдиних стандартів обміну даними, невизначеністю регуляторного середовища, високими витратами та недостатністю кваліфікованих кадрів. Обмежено вивченими є питання інтеграції блокчейну з корпоративними системами та формування економічно ефективних моделей впровадження.

Внесок цієї статті в розв'язання зазначених проблем полягає в узагальненні наукових підходів, визначенні пріоритетних напрямів оптимізації логістичних процесів та обґрунтуванні практичних рекомендацій щодо ефективного впровадження технологій у цю сферу.

Формулювання мети дослідження

Мета дослідження полягає в науковому обґрунтуванні та аналізі можливостей використання блокчейн-технологій для підвищення прозорості, надійності й ефективності ланцюгів постачання в транспортній логістиці, а також у визначенні шляхів їх практичної інтеграції в сучасну логістичну інфраструктуру.

Завдання дослідження:

- 1) проаналізувати теоретичні основи функціонування блокчейн-технологій і визначити їхні потенційні переваги для транспортно-логістичних процесів;
- 2) дослідити сучасний стан застосування блокчейну в галузі логістики та виявити основні перешкоди, що заважають його впровадженню;
- 3) обґрунтувати напрями та механізми використання блокчейн-технологій для забезпечення прозорості, відстежуваності й безпеки даних у ланцюгах постачання.

Викладення основного матеріалу дослідження

У сучасних цифрових логістичних системах збереження цілісності даних має вирішальне значення, адже саме достовірність інформації визначає рівень ефективності взаємодії між усіма учасниками ланцюга постачання. Водночас традиційні логістичні платформи часто стикаються з труднощами, пов'язаними з відсутністю довіри, недостатньою прозорістю та обмеженим контролем за обміном даними. Через це учасники логістичних процесів не завжди впевнені в точності отриманої інформації, що ускладнює координацію дій, знижує рівень співпраці та негативно впливає на загальну ефективність роботи логістичної системи [1, с. 170].

Цифрова трансформація економіки охопила всі її сектори, зокрема й логістику та управління ланцюгами постачання. У цьому контексті особливий науковий і практичний інтерес викликають сучасні цифрові технології, серед яких важливе місце посідає блокчейн. Використання цієї технології в транспортній логістиці набуває особливої значущості, адже саме тут перетинаються інтереси великої кількості учасників – власників і відправників чи отримувачів вантажів, логістичних посередників, перевізників, орендодавців і орендарів транспортних засобів, а також державних інституцій. Транзакції, що здійснюються під час транспортування вантажів у межах ланцюга постачання, потребують дотримання таких принципів, як прозорість, ефективність, зменшення витрат, конфіденційність, довіра, автентичність, кооперація, інтеграція, оптимізація та достовірність даних. Саме блокчейн-технології здатні забезпечити реалізацію цих принципів у системах транспортної логістики та управління ланцюгами постачання [2, с. 185].

Цифрова трансформація у сфері логістики є визначальним чинником, який трансформує усталені моделі управління ланцюгами постачання та процесами розподілу продукції. В умовах стрімкого технологічного прогресу та

підвищених вимог до ефективності, точності й оперативності виконання замовлень підприємства змушені адаптувати свою діяльність до нових умов функціонування [3, с. 3].

Завдяки децентралізованій архітектурі блокчейну інформація про товар захищена від хакерських атак та будь-яких несанкціонованих змін. Підrobка товару стає неможливою, оскільки кожна одиниця проходить перевірку якості, пакування та інших параметрів. Крім того, система спрощує проведення аудиту на складах підприємства завдяки збереженню повної історії транзакцій та забезпеченню дотримання нормативних вимог [4, с. 398].

Блокчейн ґрунтується на трьох основних принципах: децентралізації, криптографічному захисті та консенсусі. Децентралізація усуває єдиний центр управління, підвищуючи стійкість системи. Криптографічний захист забезпечує незмінність даних, а консенсусні механізми (Proof-of-Work, Proof-of-Stake тощо) гарантують узгодженість транзакцій між учасниками.

У логістиці технологія блокчейн зміцнює довіру між учасниками, забезпечуючи прозорий доступ до єдиної інформаційної бази в реальному часі, що мінімізує затримки та ймовірність помилок. Смартконтракти автоматизують виконання угод без посередників, оптимізуючи процеси й знижуючи витрати. Порівняльну характеристику типів блокчейн-систем у транспортній логістиці наведено в таблиці 1.

Таблиця 1

**Порівняльна характеристика основних типів блокчейн-систем
і їхнього застосування в транспортній логістиці**

Тип блокчейну	Основні характеристики	Потенційне застосування в логістиці	Переваги
Публічний	Відкритий доступ для всіх учасників, використання Proof-of-Work або Proof-of-Stake	Відстеження вантажів у глобальних ланцюгах постачання	Висока прозорість, стійкість до зловмисників
Приватний	Обмежений доступ, контроль учасників, можливість адаптації під потреби компанії	Управління внутрішньою логістикою, складування, корпоративні перевезення	Висока швидкість транзакцій, контроль над даними
Консорціум	Керований групою організацій, гібридні механізми консенсусу	Координація між партнерами в багатокомпанійних ланцюгах	Баланс між прозорістю та контролем, ефективна співпраця

Джерело: власна розробка авторів

Упровадження блокчейн-технологій у логістичні процеси також впливає на економічні показники підприємств. Застосування децентралізованих і автоматизованих систем дозволяє скоротити витрати на аудит, перевірку достовірності документів та взаємодію з посередниками. Крім того, блокчейн сприяє зниженню ризиків фінансових втрат унаслідок шахрайства або помилок у документації, що безпосередньо впливає на рентабельність транспортно-логістичних операцій.

Іншим аспектом є підвищення екологічної ефективності ланцюгів постачання. Використання блокчейн-технологій у поєднанні з аналітичними системами дозволяє оптимізувати маршрути транспортування, планувати завантаження транспортних засобів і прогнозувати потребу в ресурсах, що зменшує споживання пального та викиди парникових газів. Такий підхід сприяє не лише економічній, але й соціальній відповідальності компаній у сфері транспортної логістики (табл. 2).

Таблиця 2

Потенційні переваги використання блокчейн-технологій у транспортно-логістичних процесах

Категорія переваг	Опис	Практичний ефект
Прозорість	Усі учасники отримують доступ до незмінної інформації про вантажі	Підвищення довіри між партнерами, зменшення кількості конфліктів
Безпека даних	Захист від несанкціонованих змін та кібератак	Мінімізація ризиків шахрайства, збереження цілісності інформації
Автоматизація процесів	Виконання смартконтрактів, автоматичне оновлення статусів	Скорочення часу на оформлення документів і контроль виконання угод
Оптимізація ресурсів	Аналіз великих даних, прогнозування потреб та маршрутів	Зниження витрат на транспортування та складування
Відстежуваність	Реєстрація всіх етапів руху товару	Можливість швидкого реагування на порушення або затримки

Джерело: власна розробка авторів

Розглядаючи потенційні обмеження застосування блокчейну в транспортно-логістичних системах, варто виокремити технічні, організаційні та економічні чинники, які можуть стримувати ефективне впровадження цієї технології. Технічні обмеження охоплюють проблеми масштабованості й продуктивності при обробленні великої кількості транзакцій, високі вимоги до обчислювальних ресурсів та енергоспоживання. Організаційні

перешкоди пов'язані з необхідністю інтеграції блокчейну з наявними ERP-системами, стандартизації форматів даних і навчання персоналу. Економічні чинники стосуються високих початкових витрат на впровадження технології, а також потреби в постійному обслуговуванні та підтримуванні мережі.

Попри зазначені обмеження, використання блокчейну в транспортно-логістичних процесах відкриває значний потенціал для підвищення їхньої прозорості, безпеки та ефективності. Технологія забезпечує надійне відстеження вантажів, ідентифікацію товарів і здійснення фінансових операцій, створюючи незмінні записи про переміщення продукції та формуючи цифрові паспорти для перевірки їхньої автентичності. Смартконтракти автоматизують оплату на основі виконання умов доставлення, а цифрові документи зберігаються без ризику несанкціонованих змін. Крім того, блокчейн підтримує відновлення стану системи через резервні копії даних і журнали транзакцій, що зміцнює довіру між усіма учасниками ланцюга постачання [5, с. 90].

Управління інформаційними потоками та оптимізація капіталу, пов'язаного з товарними запасами, є основними завданнями стратегічного управління ланцюгами постачань. Накопичення значних обсягів зерна ускладнює його експорт через інфраструктурні, логістичні та регуляторні перешкоди, тоді як для задоволення внутрішнього попиту виникає потреба в імпорті продукції тваринництва, яку можна виробляти з наявного зерна. Загострення проблем ресурсного дефіциту, глобального потепління та деградації довкілля підкреслює необхідність впровадження нових технологічних рішень. Технологія блокчейн дає змогу виробникам налагоджувати прозорі взаємозв'язки з постачальниками, зміцнювати конкурентні позиції та підвищувати якість послуг, надаючи споживачам достовірну інформацію про продукцію та її походження. Це особливо важливо для ланцюгів постачання свіжих овочів і фруктів, де спостерігаються значні втрати під час транспортування та зберігання. Інтеграція блокчейну з технологіями IoT забезпечує контроль параметрів виробництва та збереження якості молочних продуктів, а в ланцюгах постачання м'яса птиці сприяє підтриманню репутації бренду та зміцненню довіри споживачів [6, с. 363].

Для ефективного управління транспортними послугами необхідна розвинена логістична інфраструктура. Наявність кваліфікованих водіїв та диспетчерів, а також належний технічний стан транспортних засобів і їхня постійна готовність до роботи дозволяють уникнути проблем у наданні транспортно-логістичних послуг. Орієнтація на підвищення якості обслуговування та задоволення потреб клієнтів є важливою для ефективності транспортної системи. Використання сучасних технологій у поєднанні з професійним персоналом сприяє зростанню економічного потенціалу та позитивно впливає на розвиток економіки країни [7, с. 7]. Оскільки транспортна логістика має пріоритетне значення для забезпечення сталого економічного розвитку як на національному, так і на міжнародному рівні, одним із головних напрямів впровадження інноваційних технологій у цій сфері є автоматизація та оптимізація логістичних процесів, а також інтеграція систем контролю й моніторингу транспортних засобів. Використання спеціалізованого програмного забезпечення та систем управління дозволяє прискорити й спростити процеси матеріального планування та координації, а також оперативно й ефективно реагувати на нові виклики [8, с. 35].

Транспортні та логістичні компанії мають вирішальне значення у впровадженні інновацій, особливо у сфері цифровізації та інтелектуалізації глобальних ланцюгів постачання. Разом із розвитком концепції «Економіки 4.0» це зумовлює суттєві зміни в логістичних процесах. Завдяки модернізації логістичних систем транснаціональні компанії змогли зміцнити свої провідні позиції на ринку [9, с. 276].

На сучасному етапі блокчейн використовується в логістиці у різних сферах. Зокрема, великі компанії та транснаціональні корпорації застосовують його для оптимізації управління запасами, забезпечення прозорості міжнародних перевезень та контролю якості продукції. Яскравими прикладами є Maersk та IBM, які реалізували платформу TradeLens для відстеження вантажів у реальному часі. Компанія Walmart використовує блокчейн для контролю ланцюга постачання харчових продуктів, що дозволяє зменшити ризики псування та покращити безпеку споживачів. DHL інтегрує блокчейн із системами IoT для моніторингу умов перевезення чутливих вантажів, зокрема фармацевтичних препаратів, контролюючи температуру, вологість та місце перебування.

Використання цієї технології в транспортній логістиці дає змогу відстежувати маршрут вантажу від виробника до кінцевого споживача, забезпечуючи повну історію переміщень та мінімізуючи ризик підробок. Попри це, впровадження блокчейну стикається з такими труднощами, як висока вартість інтеграції, відсутність єдиних стандартів і регуляторних норм, а також недостатній рівень цифрової грамотності персоналу. Багато компаній, зокрема представники середнього бізнесу, не відмовляються від традиційних методів управління через занепокоєння щодо стабільності, надійності та безпеки впровадження нових цифрових систем. У таблиці 3 наведено основні напрями застосування блокчейну в логістиці та основні перешкоди, що обмежують ефективність його використання.

Таким чином, блокчейн у логістиці демонструє значний потенціал для підвищення ефективності та безпеки ланцюгів постачання. Проте для його широкого впровадження необхідне подолання економічних, технічних та організаційних перешкод, а також створення єдиних стандартів і підвищення цифрової компетентності персоналу. Успішна інтеграція технології дозволить компаніям оптимізувати процеси, зміцнити довіру учасників ринку та забезпечити конкурентні переваги в глобальних ланцюгах постачання.

Таблиця 3

**Основні напрями застосування блокчейну в логістиці та основні перешкоди,
що обмежують ефективність його використання**

Напрями застосування блокчейну	Переваги	Основні перешкоди для впровадження
Відстеження вантажів та логістичних операцій	Прозорість, зменшення ризиків шахрайства, контроль маршруту	Високі витрати на інтеграцію, складність стандартизації
Ідентифікація та контроль якості товарів	Надійність даних, підтвердження автентичності продукції	Недостатня цифрова грамотність персоналу, обмежена інтеграція
Автоматизація фінансових транзакцій через смартконтракти	Швидкість і безпека платежів, зниження адміністративних витрат	Побоювання щодо стабільності системи, правові обмеження
Моніторинг критичних параметрів (IoT)	Оптимізація умов транспортування, забезпечення якості	Висока вартість обладнання, складність технічної інтеграції

Джерело: власна розробка авторів

Цифровізація логістики є важливим чинником підвищення ефективності та зниження витрат у сучасному бізнес-середовищі. Використання сучасних технологій, таких як системи управління ланцюгами постачання (SCM), автоматизовані складські системи (WMS), IoT, штучний інтелект (далі – ШІ), блокчейн та інші, сприяє поліпшенню прогнозування попиту, оптимізації маршрутів і контролю товарних потоків. Це дозволяє підвищити точність операцій, скоротити витрати на транспортування та зберігання, а також забезпечити прозорість і надійність логістичних процесів, що особливо актуально в умовах глобалізації ринків [10, с. 49].

Варто також зазначити, що впровадження інформаційних технологій має вплив не лише на операційну діяльність окремих компаній, а й на загальний економічний розвиток. Інтеграція цифрових рішень сприяє підвищенню продуктивності галузі, створенню нових робочих місць та зміцненню національної конкурентоспроможності. У глобальному масштабі це дозволяє підприємствам ефективно адаптуватися до умов цифрової економіки, зменшуючи розрив між розвиненими країнами та країнами, що розвиваються [11, с. 6].

Найсуттєвішою перешкодою для впровадження цієї технології є консерватизм виробників щодо нововведень, ефективність яких не має безпосереднього матеріального підтвердження. Використання блокчейн-технології потребує глибоких знань сучасних програмних середовищ і принципів їхнього практичного застосування. З огляду на це, важливим завданням є не лише вибір оптимальної платформи для реалізації технології, а й підготовка фахівців, здатних ефективно працювати з відповідними інструментами [12, с. 173].

Важливо забезпечити інтеграцію нових технологій із наявними системами управління та моніторингу, що дозволить зберегти безперервність операцій і мінімізувати проблеми під час переходу на нові платформи. Ефективне впровадження таких рішень потребує не лише технічної сумісності, а й належного рівня підготовки персоналу. Для результативного застосування ШІ необхідне регулярне навчання та підвищення кваліфікації працівників, що передбачає організацію курсів, семінарів і тренінгів, спрямованих на опанування сучасних технологій і методів роботи з ШІ [13, с. 772].

Оскільки технології ШІ швидко розвиваються, вони відкривають нові можливості для різних галузей, зокрема морських перевезень. Одним із перспективних напрямів їхнього використання є цифрові двійники (digital twins) – віртуальні моделі фізичних об'єктів, які дають змогу відтворювати й аналізувати їхню поведінку в реальному часі. Застосування цифрових двійників суден і маршрутів забезпечує проведення симуляцій та оптимізацію перевезень з урахуванням різних сценаріїв, що сприяє зниженню ризиків і підвищенню ефективності морської логістики [14, с. 37].

Митна логістика як практична діяльність охоплює комплекс функцій, що складаються з логістичних операцій, спрямованих на досягнення визначених цілей. До них належить забезпечення ефективного та швидкого переміщення товарів через митний кордон, а також зниження витрат на митні процедури при дотриманні вимог і норм. Усі ці функції спрямовані на оптимізацію роботи всіх учасників митного процесу під час переміщення вантажів та підтримку ефективної зовнішньоторговельної діяльності [15, с. 179].

Смартконтракти дозволяють автоматично виконувати фінансові або логістичні операції після настання певних умов, що підвищує швидкість обробки даних та знижує ймовірність людської помилки. Інтеграція з IoT забезпечує збір даних про стан вантажу, умови транспортування, температуру та вологість, що дозволяє підтримувати якість продуктів у реальному часі. Платформи на базі блокчейну дозволяють об'єднувати всі ці дані в єдину систему, де кожен учасник має доступ до актуальної та достовірної інформації.

Отже, використання блокчейн-технологій у ланцюгах постачання сприяє підвищенню ефективності, надійності та прозорості логістичних процесів. Упровадження таких рішень дозволяє компаніям знизити витрати, підвищити довіру між учасниками ланцюга й забезпечити відповідність вимогам безпеки та якості. Застосування інтегрованих механізмів, таких як смартконтракти, цифрові паспорти товарів та датчики IoT, робить блокчейн ефективним інструментом управління сучасними складними логістичними системами. У перспективі розвиток стандартів та платформ на базі блокчейну сприятиме масштабуванню цих технологій та їх більш широкому застосуванню в глобальних ланцюгах постачання.

Таблиця 4

Основні функції технології блокчейн та відповідні механізми їх реалізації в ланцюгах постачання

Функції	Механізми реалізації	Основні технології	Очікуваний ефект	Приклади застосування
Прозорість	Регістрація всіх транзакцій у блоках, надання доступу всім учасникам ланцюга	Блокчейн-платформи, смартконтракти	Зменшення ризиків шахрайства, підвищення довіри між учасниками ланцюга	Контроль переміщення вантажів у транспортних компаніях
Відстежуваність	Створення цифрових паспортів товарів, інтеграція з IoT та сенсорами	IoT, цифрові двійники, QR-коди	Контроль якості та походження товарів, швидке виявлення проблем	Моніторинг свіжих продуктів та фармацевтичних препаратів
Безпека даних	Децентралізоване зберігання, криптографічне шифрування, використання смартконтрактів	Криптографія, блокчейн, смартконтракти	Захист від несанкціонованих змін, автоматизація виконання угод	Автоматизовані платежі та договори між постачальниками та перевізниками
Оптимізація логістичних процесів	Автоматичне відстеження маршрутів, прогнозування попиту, контроль запасів	ШІ, машинне навчання, аналітика даних	Зниження витрат на зберігання та транспортування, підвищення ефективності	Планування маршрутів доставки та управління складськими запасами
Підвищення рівня довіри споживачів	Забезпечення відкритого доступу до інформації про продукт	Блокчейн, цифрові паспорти товарів	Покращення репутації бренду, підвищення лояльності клієнтів	Перевірка автентичності органічних та преміальних товарів

Джерело: власна розробка авторів

Висновки

Використання блокчейн-технологій у ланцюгах постачання забезпечує новий рівень прозорості, відстежуваності та безпеки. Децентралізоване зберігання даних і незмінність записів знижують ризик шахрайства та підвищують рівень довіри між учасниками. Інтеграція з технологіями IoT дає змогу контролювати рух і стан товарів у реальному часі, підвищуючи точність та оперативність логістики.

Смартконтракти автоматизують виконання фінансових і договірних зобов'язань, скорочуючи час транзакцій і адміністративні витрати. Це особливо важливо для складних ланцюгів постачання з великою кількістю учасників.

Блокчейн також зміцнює довіру споживачів завдяки відкритому доступу до інформації про походження та якість продукції, оптимізує маршрути доставки й зменшує втрати. Такий підхід підвищує економічну ефективність підприємств і їхню конкурентоспроможність, перетворюючи технологію на стратегічний інструмент розвитку логістики.

Подальші дослідження у сфері використання блокчейн-технологій для прозорості ланцюгів постачання в транспортній логістиці можуть зосередитися на розробленні стандартів безпеки й методик оцінювання економічної ефективності впровадження блокчейн-рішень у різних сегментах логістики.

Список використаної літератури

- Рябчун Ю., Компанець Д., Хроленко В., Попович Н. Застосування блокчейн-технологій в забезпеченні цілісності даних у логістичних системах. *Управління розвитком складних систем*. 2025. № 62. С. 170–176. <https://doi.org/10.32347/2412-9933.2025.62.170-176>
- Колодізева Т. О. Перспективи впровадження блокчейн-технологій в транспортну логістику та управління ланцюгами поставок. *Бізнес-Інформ*. 2023. № 6. С. 184–190. DOI: <https://doi.org/10.32983/2222-4459-2023-6-184-190>
- Марінов Є., Лісений Є. Цифрова трансформація в логістиці. *Економіка та суспільство*. 2024. № 66. DOI: <https://doi.org/10.32782/2524-0072/2024-66-53>
- Химич В. В., Білоконь Т. М. Використання блокчейн в складській логістиці. *Маркетинг і логістика в системі менеджменту: виклики цифрової глобалізації: тези доп. XV Міжнар. наук.-практ. конф., пам'яті проф. Є. Крикавського* (м. Львів, 17–18 жовтня 2024 р.). Львів, 2024. С. 398–399. URL: <https://science.lpnu.ua/sites/default/files/attachments/2024/oct/36215/abstractsoftheconferenceml20241111.pdf> (дата звернення: 15.10.2025).
- Якимишин Л. Я. Інновації у логістиці: вплив технологій на ефективність та конкурентоспроможність підприємства. *Маркетингові та логістичні технології: інновації для забезпечення ефективності бізнес-процесів*: монографія. Тернопіль, 2024. С. 88–98. URL: <https://elartu.tntu.edu.ua/handle/lib/46773> (дата звернення: 15.10.2025).
- Школьнік О. О. Розвиток інноваційних технологій в аграрній логістиці. *Менеджмент XXI століття: проблеми і перспективи: матеріали XIV Всеукр. наук.-практ. конф.* (м. Умань, 17 жовтня 2024 р.). Умань, 2024. С. 362–364. URL: <https://lib.udau.edu.ua:8443/server/api/core/bitstreams/1bac7e4b-6168-4480-8d5f-7185cf7c17a1/content?page=362> (дата звернення: 15.10.2025).

7. Марінов Є. А. Інноваційні технології у транспортній логістиці: економічний потенціал і виклики впровадження. *Академічні візії*. 2024. № 30. URL: <https://academy-vision.org/index.php/av/article/view/1367> (дата звернення: 15.10.2025).
8. Головіна О. Сучасні технології в управлінні транспортною логістикою. *International Science Journal of Management, Economics & Finance*. 2023. № 2 (3). С. 35–42. DOI: 10.46299/j.isjmef.202302023.04
9. Македон В., Михайленко О. Вплив економіки 4.0 на глобальні логістичні операції транснаціональних корпорацій. *Економічний простір*. 2024. № 189. С. 276–282. DOI: <https://doi.org/10.32782/2224-6282/189-49>
10. Канцедал Н., Лега О., Морозов Є. Цифровізація логістики: нові технології для покращення управління та оптимізації. *Економічний простір*. 2025. № 199. С. 45–51. DOI: <https://doi.org/10.30838/EP.199.45-51>
11. Цимбалістова О., Черніхова О. Роль і значення інформаційних технологій у впровадженні логістичної методики управління. *Економіка та суспільство*. 2024. № 70. DOI: <https://doi.org/10.32782/2524-0072/2024-70-117>
12. Потапова Н. Технології блокчейн в інформаційній логістиці. *Прикладні аспекти сучасних міждисциплінарних досліджень: Матеріали I Міжнародної науково-практичної конференції* (м. Вінниця, 18 листопада 2022 р.). Вінниця, 2022. С. 172–174. URL: <https://jpasmd.donnu.edu.ua/article/view/12954> (дата звернення: 15.10.2025).
13. Коростін О. Аналіз викликів та можливостей застосування штучного інтелекту в управлінні морськими вантажними потоками. *Наука і техніка сьогоденні. Серія «Техніка»*. 2024. № 7(35). С. 762–775. DOI: [https://doi.org/10.52058/2786-6025-2024-7\(35\)-762-775](https://doi.org/10.52058/2786-6025-2024-7(35)-762-775)
14. Коростін О. Оптимізація маршрутів морських перевезень за допомогою штучного інтелекту: аналіз можливостей та викликів. *Комп'ютерно-інтегровані технології: освіта, наука, виробництво*. 2024. № 56. С. 31–38. DOI: <https://doi.org/10.36910/6775-2524-0560-2024-56-03>
15. Коробкова О. М., Коробкова З. С. Митна логістика при організації міжнародних перевезень вантажів. *Вчені записки ТНУ імені В.І. Вернадського. Серія: Технічні науки*. 2024. Т 35 (74). № 3. С. 177–182. DOI: <https://doi.org/10.32782/2663-5941/2024.3.2/27>

References

1. Riabchun, Yu., Kompanets, D., Khrolenko, V., & Popovych, N. (2025) Zastosuvannia blokchein-tekhnologii v zabezpechenni tsilnistosti danykh u lohistychnykh systemakh [Application of blockchain technologies to ensure data integrity in logistics systems]. *Upravlinnia rozvytkom skladnykh system*, no. 62, pp. 170–176. <https://doi.org/10.32347/2412-9933.2025.62.170-176>
2. Kolodzieieva, T. O. (2023) Perspektyvy vprovadzhennia blokchein-tekhnologii v transportnu lohistyku ta upravlinnia lantsiuhamy postavok [Prospects for implementing blockchain technology in transport logistics and supply chain management]. *Biznes-Inform*, no. 6, pp. 184–190. <https://doi.org/10.32983/2222-4459-2023-6-184-190>
3. Marinov, Ye., & Lisenyi, Ye. (2024) Tsyfrova transformatsiia v lohistytsi [Digital transformation in logistics]. *Ekonomika ta suspilstvo*, no. 66. <https://doi.org/10.32782/2524-0072/2024-66-53>
4. Khymych, V. V., & Bilokon, T. M. (2024) Vykorystannia blokchein v skladskii lohistytsi [Using blockchain in warehouse logistics]. *Marketynh i lohistyka v systemi menedzhmentu: vyklyky tsyfrovoi hlobalizatsii: tezy dop. XV Mizhnar. nauk.-prakt. konf., pamiati prof. Ye. Krykavskoho* (m. Lviv, 17–18 zhovtnia 2024 r.) (pp. 398–399). Lviv. Retrieved from <https://science.lpnu.ua/sites/default/files/attachments/2024/oct/36215/abstractsoftheconferenceml20241111.pdf>
5. Yakymyshyn, L. Ya. (2024) Innovatsii u lohistytsi: vplyv tekhnologii na efektyvnist ta konkurentospromozhnist pidpriemstva [Innovations in logistics: The impact of technologies on enterprise efficiency and competitiveness]. In *Marketynhovi ta lohistychni tekhnologii: innovatsii dlia zabezpechennia efektyvnosti biznes-protsesiv: monohrafiia* (pp. 88–98). Ternopil. Retrieved from <https://elartu.tntu.edu.ua/handle/lib/46773>
6. Shkolnyi, O. O. (2024) Rozvytok innovatsiinykh tekhnologii v aharnii lohistytsi [Development of innovative technologies in agricultural logistics]. *Menedzhment XXI stolittia: problemy i perspektyvy : materialy XIV Vseukr. nauk.-prakt. konf.* (m. Uman, 17 zhovtnia 2024 r.) (pp. 362–364). Uman. Retrieved from <https://lib.udau.edu.ua:8443/server/api/core/bitstreams/1bac7e4b-6168-4480-8d5f-7185cf7c17a1/content#page=362>
7. Marinov, Ye. A. (2024) Innovatsiini tekhnologii u transportnii lohistytsi: ekonomichniy potentsial i vyklyky vprovadzhennia [Innovative technologies in transport logistics: Economic potential and implementation challenges]. *Akademichni vizii*, no. 30. Retrieved from <https://academy-vision.org/index.php/av/article/view/1367>
8. Holovina, O. (2023) Suchasni tekhnologii v upravlinni transportnoi lohistykoii [Modern technologies in transport logistics management]. *International Science Journal of Management, Economics & Finance*, vol. 2, no. 3, pp. 35–42. <https://doi.org/10.46299/j.isjmef.202302023.04>
9. Makedon, V., & Mykhailenko, O. (2024) Vplyv ekonomiky 4.0 na hlobalni lohistychni operatsii transnatsionalnykh korporatsii [The impact of Industry 4.0 on global logistics operations of transnational corporations]. *Ekonomichnyi prostir*, no. 189, pp. 276–282. <https://doi.org/10.32782/2224-6282/189-49>

10. Kantsedal, N., Leha, O., & Morozov, Ye. (2025) Tsyfrovizatsiia lohistyky: novi tekhnolohii dlia pokrashchennia upravlinnia ta optymizatsii [Logistics digitalization: New technologies for management and optimization]. *Ekonomichnyi prostir*, no. 199, pp. 45–51. <https://doi.org/10.30838/EP.199.45-51>
11. Tsymbalistova, O., & Chernikhova, O. (2024) Rol i znachennia informatsiinykh tekhnolohii u vprovadzhenni lohistychnoi metodyky upravlinnia [The role and importance of information technologies in implementing logistic management methodology]. *Ekonomika ta suspilstvo*, no. 70. <https://doi.org/10.32782/2524-0072/2024-70-117>
12. Potapova, N. (2022) Tekhnolohii blokchein v informatsiinii lohistytsi [Blockchain technologies in information logistics]. *Prykladni aspekty suchasnykh mizhdystsyplinarnykh doslidzhen: Materialy I Mizhnarodnoi naukovo-praktychnoi konferentsii* (m. Vinnytsia, 18 lystopada 2022 r) (pp. 172–174). Vinnytsia. Retrieved from <https://jpasmd.donnu.edu.ua/article/view/12954>
13. Korostin, O. (2024) Analiz vyklykiv ta mozhlyvosti zastosuvannia shtuchnoho intelektu v upravlinni morskymy vantazhnymy potokamy [Analysis of challenges and opportunities of AI in managing maritime cargo flows]. *Nauka i tekhnika sohodni. Seriia «Tekhnika»*, vol. 7, no. 35, pp. 762–775. [https://doi.org/10.52058/2786-6025-2024-7\(35\)-762-775](https://doi.org/10.52058/2786-6025-2024-7(35)-762-775)
14. Korostin, O. (2024) Optymizatsiia marshrutiv morskyykh perevezen za dopomohoiu shtuchnoho intelektu: analiz mozhlyvosti ta vyklykiv [Optimization of maritime transportation routes using artificial intelligence: Analysis of opportunities and challenges]. *Kompiuterno-intehrovani tekhnolohii: osvita, nauka, vyrobnytstvo*, no. 56, pp. 31–38. <https://doi.org/10.36910/6775-2524-0560-2024-56-03>
15. Korobkova, O. M., & Korobkova, Z. S. (2024) Mytna lohistyka pry orhanizatsii mizhnarodnykh perevezen vantazhiv [The role of customs logistics in international transportation]. *Vcheni zapysky TNU imeni V. I. Vernadskoho. Seriia: Tekhnichni nauky*. vol. 35 (74), no. 3, pp. 177–182. <https://doi.org/10.32782/2663-5941/2024.3.2/27>

Дата першого надходження рукопису до видання: 15.11.2025
Дата прийнятого до друку рукопису після рецензування: 12.12.2025
Дата публікації: 31.12.2025

Я. В. ЧУМАЧЕНКО

кандидат технічних наук, доцент,

доцент кафедри комп'ютерних наук

Львівський національний лісотехнічний університет

ORCID: 0000-0002-0447-4874

РОЗРОБКА КОМП'ЮТЕРНИХ АЛГОРИТМІВ В ПРОЄКТАХ ДОСЛІДЖЕННЯ ХВИЛЬОВИХ ПРОЦЕСІВ У ДИСКРЕТНИХ СЕРЕДОВИЩАХ

У наданій статті була практично реалізована задача розвинути принципи граничного переходу від дискретних моделей до континуальних в задачах на динаміку пружних конструкцій; дослідити більш ретельно дискретні моделі та існуючі методи вирішення задач по даній тематиці; удосконалити вже існуючі рішення шляхом побудови таких дискретних моделей, які б були максимально наближені до відомих континуальних та мали б деякі суттєві переваги при розв'язках задач. Розробити для подальшого застосування програмний модуль для проєкту, що дозволяє побудувати такі дискретні моделі системи, що досліджується, які б забезпечували стабільність системи під дією прикладеної сили, незалежно від заданих початкових умов, а також відсутність похибки обчислень на етапі вирішення задачі.

Для досягнення мети було розглянуто усі характерні залежності сил, діючих на систему, зокрема сили, що залежать від часу, положення системи та швидкості. Відображена архітектура програмного продукту *DiscretModels*, що поєднує у собі підсистеми графічної візуалізації коливальних систем, що вивчається, розрахункову підсистему, яка реалізує запропонований у роботі підхід для побудови дискретних моделей системи під динамічним навантаженням, та підсистему задання та редагування основних параметрів системи. Побудовані: підсистема *Form* – основна підсистема проєкту, що є контейнером для графічної та розрахункової підсистем. Головною функцією даної складової є зв'язок користувача з усіма можливостями, що надає розроблений проєкт. Підсистема *Param* – підсистема проєкту, що забезпечує введення та редагування основних параметрів системи, відповідно до обраної задачі. Підсистема *Graf* – розроблена динамічна бібліотека, що входить до складу програмного проєкту, з метою формування нового типу графічного компоненту на основі *System.Windows.Forms.Control*. Під терміном „система, що вивчається” в даному контексті розуміється одномірний осцилятор та стержень, представлений у вигляді ланцюжка мас із дискретними пружинами, під динамічним навантаженням.

При виконанні даного дослідження на основі вже існуючих рішень було запропоновано новий підхід побудови дискретних моделей системи, що вивчається. Розроблено програмний продукт, який успішно реалізує запропонований підхід та порівнює отримані результати із результатами існуючих рішень для підтвердження ефективності проєкту. Будь-який програмний продукт має обмеженості у його використанні та обсягу розв'язуваних задач. У цьому методі доведено, що деякі обмеженості можна вирішити застосувавши до реалізації додаткові умови перевірки та критерії виявлення помилок.

Зроблено висновки про переваги та недоліки кожного з методів континуальних та дискретних моделей. Показано, що найбільш перспективним є метод дискретизації із постійним кроком зміни сил, що діють у системі. Але використання такого варіанту потребує відповідного програмного алгоритму для визначення моментів стрибкоподібної зміни сил, що обумовлює перспективи подальшого розвитку проєкту.

Ключові слова: хвильові процеси, дискретні моделі, континуальні моделі, механіка суцільного середовища, діюча сила, динамічне навантаження, дискретизація сили.

YA. V. CHUMACHENKO

Candidate of Technical Sciences, Associate Professor,

Associate Professor at the Department of Computer Science

Lviv National Forestry University

ORCID: 0000-0002-0447-4874

DEVELOPMENT OF COMPUTATIONAL ALGORITHMS FOR THE STUDY OF WAVE PROCESSES IN DISCRETE MEDIA

In the provided article, the task was practically implemented to develop the principles of limiting transition from discrete models to continual models in problems concerning the dynamics of elastic structures; to study discrete models and existing methods for solving problems on this topic more thoroughly; to improve already existing solutions by

constructing such discrete models that would be maximally close to known continual models and would have some significant advantages in problem solutions. To develop a software module for further application in the project, which allows for the construction of such discrete models of the system under study that would ensure system stability under the action of the applied force, regardless of the set initial conditions, as well as the absence of calculation errors at the stage of problem solution.

To achieve the goal, all characteristic dependencies of forces acting on the system were considered, including forces that depend on time, system position, and velocity. The architecture of the DiscretModels software product is presented, which combines subsystems for the graphical visualization of the oscillations of the system under study, a calculation subsystem that implements the proposed approach for constructing discrete models of the system under dynamic loading, and a subsystem for setting and editing the main system parameters. The following were built: the Form subsystem – the main subsystem of the project, which is a container for the graphical and calculation subsystems. The main function of this component is the user's connection with all the possibilities provided by the developed project. The Param subsystem – a project subsystem that ensures the input and editing of the main system parameters, according to the selected problem. The Graf subsystem – a developed dynamic library that is part of the software project, intended for forming a new type of graphical component based on System.Windows.Forms.Control. The term "system under study" in this context refers to a one-dimensional oscillator and a rod, represented as a chain of masses with discrete springs, under dynamic loading.

During the execution of this research, a new approach to constructing discrete models of the system under study was proposed, based on already existing solutions. A software product was developed that successfully implements the proposed approach and compares the obtained results with the results of existing solutions to confirm the project's effectiveness. Any software product has limitations in its use and the scope of solvable problems. In this method, it is proven that some limitations can be overcome by applying additional checking conditions and error detection criteria to the implementation.

Conclusions were drawn regarding the advantages and disadvantages of each of the methods of continual and discrete models. It is shown that the most promising is the discretization method with a constant step change of forces acting in the system. However, the use of this option requires a corresponding software algorithm to determine the moments of abrupt force changes, which defines the prospects for the project's further development.

Key words: wave processes, discrete models, continual models, mechanics of the continuous environment, operating force, dynamic loading, force digitization.

Постановка проблеми

У сучасній механіці суцільного середовища використовуються два підходи. Найбільш розповсюджений континуальний, з використанням ДУ. Також широко розповсюджені дискретні моделі, які більш зручно використовувати для чисельних розрахунків на сучасних комп'ютерах. При цьому передбачається, що моделі двох видів тісно зв'язані між собою і дають якісно однакову інформацію про процеси, що вивчаються. Але розповсюдженою є точка зору, що дискретні моделі є зіпсованим варіантом континуальних моделей.

У класичній механіці суцільного середовища фізично нескінченно малий об'єм розглядається як матеріальна точка, що не має розмірів. Відповідна континуальна модель вельми успішно описує рух суцільного середовища, а також розподіл напруги і деформацій в досить гладких областях при досить плавних навантаженнях.

У тих же випадках, коли характерний розмір збурення має порядок характерного розміру мікроструктури, класичні континуальні моделі стають або не зовсім адекватними, або їх використання приводить до значних математичних труднощів. Так, облік мікроструктурних ефектів важливий при моделюванні ультрадисперсних і нанокристалічних матеріалів, при описі високочастотних вібрацій в складних кристалічних решітках органічних речовин, у вершинах тріщин і на фронті хвилі руйнування, при структурних перетвореннях і виникненні дефектів. У подібних ситуаціях в тій або іншій мірі доводиться враховувати дискретний характер структури речовини.

Але виявляється, що не завжди виконується граничний перехід між результатами, отриманими на ґрунті дискретних та континуальних моделей. Вирішення цієї проблеми можливе двома способами. Перший – вдосконалення відповідних континуальних моделей. Такий підхід, здатний відображати специфіку внутрішньої структури речовини, запропонований, наприклад, у роботі Андріанова І.В. Другий – передбачає вдосконалення дискретної моделі, як запропоновано у роботі Шамровського О.Д [1].

При другому підході відпадає необхідність вживання процедури дискретизації, неминучої при застосуванні стандартних чисельних методів до вирішення задач в рамках континуальних моделей. У вказаному напрямі отримані певні результати, проте, в цілому, розвиток дискретних моделей механіки суцільних середовищ істотно відстає від розвитку континуальних моделей.

Аналіз останніх досліджень і публікацій

Дослідження останніх років у механіці суцільного середовища зосереджені на побудові узагальнених континуумів та мультимасштабному моделюванні для подолання обмежень класичних теорій. Наприклад у роботі Meng Z., Zhang S. A. [2] представлено уніфіковану схему для моделювання як суцільних, так і розривних матеріалів (наприклад, з тріщинами) за допомогою неявного поєднання методу дискретних елементів (DEM) та методу скінченних елементів (FEM). Монографія Neff, P. [3] присвячена теоретичному дослідженню та порівнянню

дискретних і континуальних моделей для аналізу складних метаматеріалів та їхніх незвичайних механічних властивостей. Зокрема, активізувався інтерес до мікрополярних (Коссера) та нелокальних моделей, які дозволяють враховувати обертальні ступені свободи та взаємодію на відстані, що критично важливо для наноматеріалів. Проблема граничного переходу між дискретними ґратковими моделями (наприклад, моделями частинок) та континуальними (МГД, Коссера, Коші) залишається ключовою. Дослідження показують, що гомогенізація (усереднення) дискретних моделей з обертальними ступенями свободи часто призводить до континууму типу Коссера, а не простішого типу Коші, як це добре показано в роботі Wang R., Fan F., Sun J. [4]. Це підкреслює необхідність уточнення принципів континуалізації. У сфері чисельних методів розвиваються гібридні та монолітні схеми, що поєднують метод скінченних елементів (МСЕ) та метод дискретних елементів (МДЕ), відомі як FDEM (Finite Discrete Element Method). Такі методи необхідні для моделювання переходу матеріалу з континуального стану (до руйнування) у дискретний (після фрагментації). Вони демонструють, що дискретні підходи можуть бути інтегровані в єдиний варіаційний фреймворк, що свідчить про їхню рівноправність з континуальними як це було впроваджено в роботі Liu Q., Li Q., Shi D. [6]. Як показали публікації Bedair O. K., El-Sayed O. E. [8] також активно розробляються покращені континуальні моделі для дискретних середовищ, які враховують нелокальну взаємодію та вирішують проблему нескінченності амплітуди коливань. Це підтверджує актуальність роботи з побудови стійких дискретних моделей, які завдяки вдосконаленому оцифруванню сили можуть досягати високої точності, характерної для узагальнених континуальних теорій.

Формулювання мети дослідження

Дослідити принципи граничного переходу від дискретних моделей до континуальних в задачах динаміки пружних конструкцій. Удосконалити вже існуючі рішення шляхом побудови таких дискретних моделей, які б були максимально наближені до відомих континуальних та мали б деякі переваги при рішеннях задач.

Розробити програмний продукт, який дозволяє будувати такі дискретні моделі системи, що вивчається, які б забезпечували стабільність системи під дією прикладеної сили, незалежно від заданих початкових умов, а також забезпечує відсутність росту похибки обчислень на етапі вирішення задачі.

Викладення основного матеріалу дослідження

Дослідивши предметну область серед усіх представлених методів вирішення подібної проблеми, запропоновано видозмінювати дискретну модель, використовуючи дискретизацію пружної сили, що веде до ліквідації ефекту Гібсу на фронті хвилі. Завдяки цьому було отримано на основі дискретної моделі результати, аналогічні результатам континуальної моделі.

Розроблена система являє собою програмний продукт *DiscretModels*, що поєднує у собі підсистеми графічної візуалізації коливань системи, що вивчається, розрахункову підсистему, яка реалізує запропонований у роботі підхід для побудови дискретних моделей системи під динамічним навантаженням, та підсистему задання та редагування основних параметрів системи. Підсистема *Form* представляє собою основний інтерфейс програми, що дозволяє користувачу задавати основні параметри системи, розраховувати та відображати результати рішення відповідних задач. Підсистема *Param* – підсистема проекту, що забезпечує введення та редагування основних параметрів системи, відповідно до обраної задачі. Підсистема *Graf* – розроблена динамічна бібліотека, що входить до складу програмного проекту, з метою формування нового типу графічного компоненту на основі *System.Windows.Forms.Control*.

До можливостей бібліотеки *Graf* відноситься: автоматичне масштабування вхідного набору точок для розміщення їх на екрані; розподіл набору точок на об'єкти з можливістю персоналізувати вивід кожного об'єкту на екран; відображення результатів як у вигляді ізометрії, так і у Декартовій площині; анімаційне відображення контенту.

Специфікація даних описує наступні сутності: сутність «Сила від швидкості» представлена класом *"DiscretSpeed"*, який призначено для виконання дискретизації сили, що залежить від швидкості, що змінюється за лінійним законом, сутність «Сила за часом» представлена класом *"DiscretFunc"*, який призначено для дискретизації сили за часом, що змінюється за нелінійним законом $F = A \sin(kt)$, сутність «Одномірний осцилятор» представлена класом *"DiscretOsc"*, який призначено для дослідження поведінки системи одномірного осцилятора під дією тільки пружної сили, тобто вільних коливань системи, сутність «Ланцюжок мас під змінною силою» представлена класом *"DiscretSteel"*, який призначено для дослідження розповсюдження хвилі уздовж стержня під дією постійної та змінної сили.

Розглянемо звичайний одномірний осцилятор, що складається з вантажа маси m , з'єднаного пружним зв'язком жорсткості C із нерухомою опорою. Якщо матеріальна точка рухається під дією змінної сили, необхідно дискретизувати цю силу, тобто представити її у вигляді кусочно-постійного навантаження (Рис. 1). Вирішимо таку задачу. Нехай у момент часу t_i вантаж знаходиться в положенні x_i і має швидкість v_i . Треба знайти найближчий момент часу t_{i+1} , при якому або відбувається стрибкоподібна зміна сили, або відбувається зупинка вантажу.

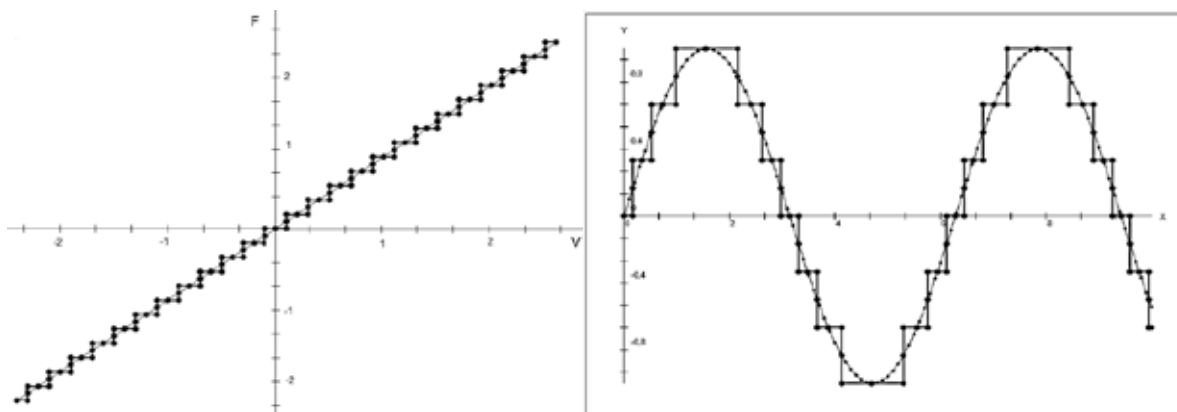


Рис. 1. Залежність пружної сили від деформації

Записавши рівняння руху вантажу:

$$v = v_i + a_i(t - t_i), \quad a_i = -\frac{P}{m} \text{round}\left(\frac{x_i}{\Delta}\right).$$

$$x = x_i + v_i(t - t_i) + \frac{1}{2}a_i(t - t_i)^2. \quad (1)$$

визначимо найближчу до x_i точку, в якій відбувається стрибкоподібна зміна сили, знаходячи середину відповідного відрізка. Якщо вантаж рухається з позитивною швидкістю v , то шуканою найближчою крапкою є права точка знайденого відрізка, а якщо з негативною – то ліва. Якщо в положенні x_i швидкість дорівнює нулю, то для визначення потрібної точки використовується знак прискорення. Позначимо її через x_i^+ і запишемо:

$$x_i^+ = x_i + v_i(t^+ - t_i) + \frac{1}{2}a_i(t^+ - t_i)^2. \quad (2)$$

Вирішуючи це рівняння, ми знаходимо той момент часу t^+ , в який вантаж досягає, з положення x_i , положення x_i^+ . Знак плюс або мінус перед коренем вибирається так, щоб знайдене значення t^+ було більше, ніж t_i . Якщо таких значень виявляється два, то вибирається менше з них. Звернемо увагу на те, що попадання вантажу в положення x_i^+ супроводжується переходом вантажу з одного відрізка в сусідній – справа або зліва залежно від напрямку руху. При цьому, природно, стрибкоподібно змінюється сила, що діє на вантаж, і відповідне прискорення.

Підставляючи знайдене значення часу в (1) знаходимо швидкість v_i^+ , що досягається вантажем в положенні x_i^+ . Негативне значення дискримінанта означає, що вантаж не може досягти кордону відрізка, тобто зупиняється по дорозі до цього кордону. Тоді момент зупинки знаходиться як:

$$t^+ = t_i - \frac{v_i}{a_i}. \quad (3)$$

Координату зупинки знаходимо, підставляючи знайдене значення часу t^+ в (1). Повторюючи описані дії, починаючи з початкових умов, і задаючи нові початкові умови на кожному кроці, як результати кінця попереднього кроку, можна послідовно визначати ключові точки.

Таким чином, дискретна модель має самостійну цінність, причому в рамках даної моделі результати виходять без накопичення похибки. Результат дії даного алгоритму показано на рисунку 2.

Із результатів видно, що навіть при досить великому кроці дискретизації сили, дискретна модель є абсолютно стабільною, на відміну від моделей, що використовують диференціальні рівняння та чисельні методи. Також це дуже економить ресурси комп'ютера.

Продовженням даної задачі є побудова дискретної моделі одномірного осцилятора за умов, при яких окрім пружної сили до нього прикладені ще дві змінні сили різної природи (тобто вимушені коливання): сила, залежна від швидкості або сила тертя (змінюється за лінійним законом) та сила залежна від часу (змінюється за періодичним законом). Дані залежності було дискретизовано за допомогою нового запропонованого методу, при якому використовуються стрибки сили постійного значення. Моменти часу стрибків сили вираховуються за відомими формулами (Рис. 3).

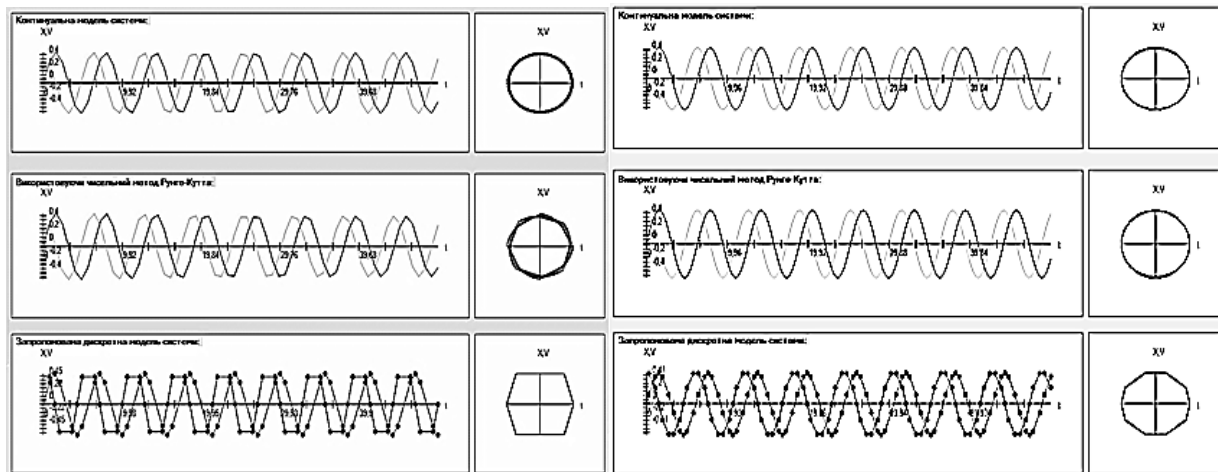


Рис. 2. Різні моделі поведінки системи під дією пружної сили: крок дискретизації 1 (зліва) та 0,1 (справа).

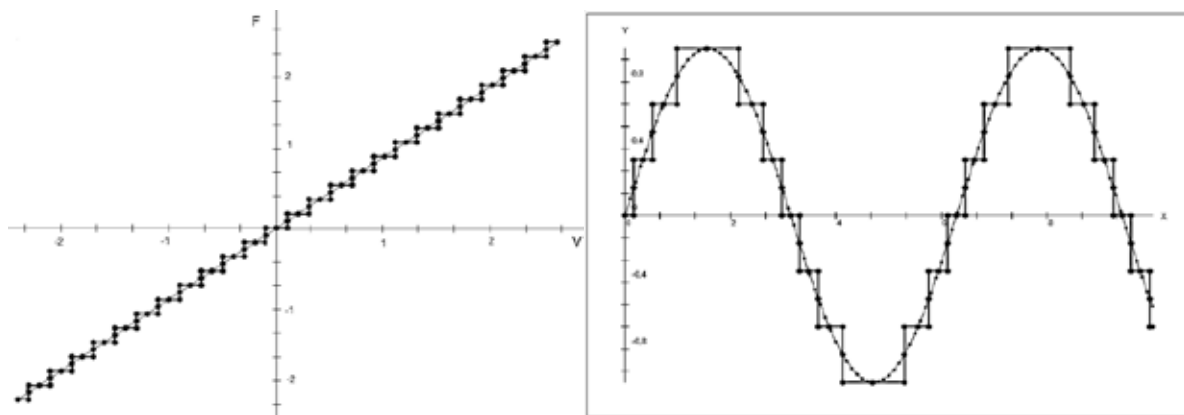


Рис. 3. Дискретизація сили тертя(зліва) та сили за часом(справа)

Поведінка системи під дією результуючої сили показано на рисунках 4-5:

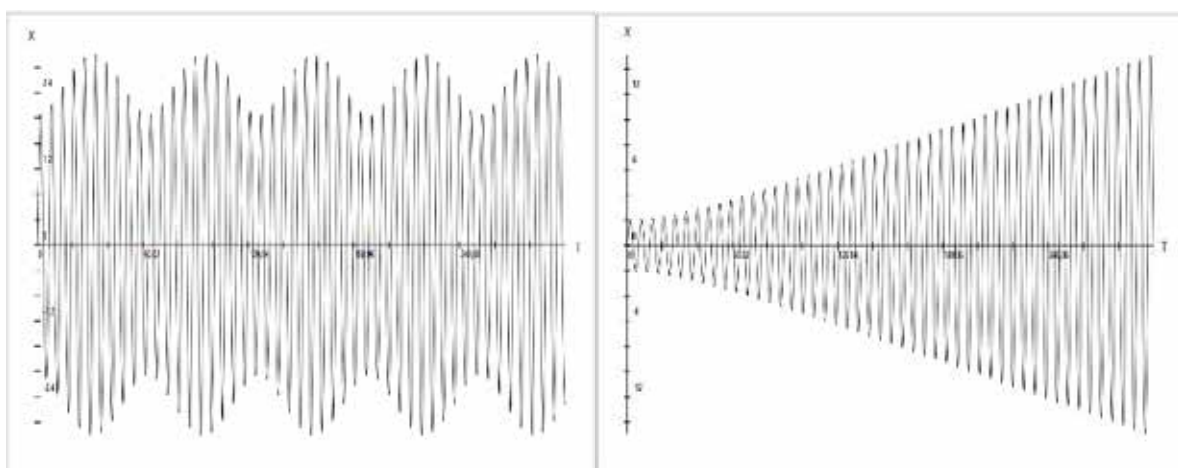


Рис. 4. Графік вимушених коливань на основі *континуальної моделі*:
результуюча сила випадок $k \neq \omega$ (зліва) і $k = \omega$ (справа)

При порівнянні отриманих результатів досліджень, бачимо що результати на основі дискретної моделі практично аналогічні результатам на основі континуальної моделі.

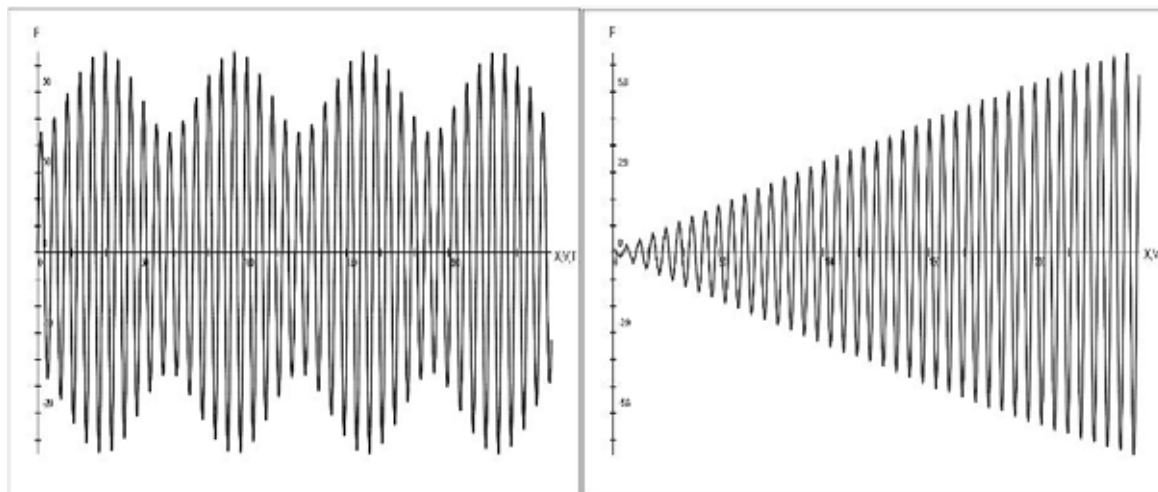


Рис. 5. Графік вимушених коливань на основі дискретної моделі: результуюча сила випадок $k \neq \omega$ (зліва) і $k = \omega$ (справа)

Висновки

Побудовано новий спосіб створення дискретних моделей динамічних систем, який не потребує використання диференціальних рівнянь. Вивчено умови, за яких використання дискретної моделі не приводить до зростання похибки обчислень. Розроблено дискретні моделі для конкретних прикладів рішення динамічних задач, що містять у собі не тільки опис системи, а й одночасно й алгоритм розв'язку задач за допомогою таких моделей. Отриманий спосіб побудови дискретних моделей дозволяє використовувати ці моделі разом із відповідними континуальними із забезпеченням граничного переходу для результатів, отриманих двома підходами. Розроблено програмний продукт, що дозволяє будувати дискретні та континуальні моделі окремих задач та порівнювати результати отримані на основі цих моделей.

Список використаної літератури

1. Шамровський О. Д. До динаміки пружних конструкцій: застосування дискретних моделей на основі граничного переходу / О. Д. Шамровський // Вісник Запорізького національного технічного університету. 2010. Вип. 2(55). С. 132–138.
2. Meng Z., Zhang S. A unified scheme for continuous and discontinuous materials using implicit Discrete Element Method and Finite Element Method. *Computer Methods in Applied Mechanics and Engineering*. 2025. Vol. 433. Art. 117365. DOI: 10.1016/j.cma.2024.117365.
3. Neff, P. *Discrete and Continuum Models for Complex Metamaterials*. Cambridge University Press, 2024. 345 p.
4. Wang R., Fan F., Sun J. Homogenization of discrete fine-scale mechanical models with rotational degrees of freedom into a Cosserat or a Cauchy continuum. *Journal of the Mechanics and Physics of Solids*. 2025. Vol. 182. Art. 105470. DOI: 10.1016/j.jmps.2024.105470.
5. Zeng F., Zhang P. A micro-polar continuum model for the dynamic analysis of nanocrystalline materials. *International Journal of Solids and Structures*. 2023. Vol. 280. Art. 112440. DOI: 10.1016/j.ijsolstr.2023.112440.
6. Liu Q., Li Q., Shi D. A unified implicit framework for the combined Finite/Discrete Element Method (FDEM). *International Journal for Numerical Methods in Engineering*. 2023. Vol. 134, Iss. 10. e02927. DOI: 10.1002/nme.6927.
7. Коваленко О. М. Комп'ютерне моделювання як один із сучасних методів геокосмічних досліджень / О. М. Коваленко // Вісник ХНУРЕ. 2023. № 1. С. 57–65.
8. Bedair O. K., El-Sayed O. E. Comparison of discrete and continuum models for wave propagation in periodic structures. *Applied Mathematical Modelling*. 2024. Vol. 129. Art. 106399. DOI: 10.1016/j.apm.2024.106399.

References

1. Shamrovsky, O. D. (2010). To the dynamics of elastic structures: Application of discrete models based on limiting transition. *Visnyk Zaporizhzhia National Technical University*, *(2)*55, 132–138.
2. Meng, Z., & Zhang, S. (2025). A unified scheme for continuous and discontinuous materials using implicit Discrete Element Method and Finite Element Method. *Computer Methods in Applied Mechanics and Engineering*, 433, 117365. <https://doi.org/10.1016/j.cma.2024.117365>
3. Neff, P. (2024). *Discrete and Continuum Models for Complex Metamaterials*. Cambridge University Press.

4. Wang, R., Fan, F., & Sun, J. (2025). Homogenization of discrete fine-scale mechanical models with rotational degrees of freedom into a Cosserat or a Cauchy continuum. *Journal of the Mechanics and Physics of Solids*, 182, 105470. <https://doi.org/10.1016/j.jmps.2024.105470>
5. Zeng, F., & Zhang, P. (2023). A micro-polar continuum model for the dynamic analysis of nanocrystalline materials. *International Journal of Solids and Structures*, 280, 112440. <https://doi.org/10.1016/j.ijsolstr.2023.112440>
6. Liu, Q., Li, Q., & Shi, D. (2023). A unified implicit framework for the combined Finite/Discrete Element Method (FDEM). *International Journal for Numerical Methods in Engineering*, 134(10), e02927. <https://doi.org/10.1002/nme.6927>
7. Kovalenko, O. M. (2023). Computer modeling as one of the modern methods of geospace research. *Visnyk KhNURE*, (1), 57–65.
8. Bedair, O. K., & El-Sayed, O. E. (2024). Comparison of discrete and continuum models for wave propagation in periodic structures. *Applied Mathematical Modelling*, 129, 106399. <https://doi.org/10.1016/j.apm.2024.106399>

Дата першого надходження рукопису до видання: 17.11.2025

Дата прийнятого до друку рукопису після рецензування: 15.12.2025

Дата публікації: 31.12.2025

В. П. ШАПОВАЛ

аспірант кафедри інформаційних технологій проектування
Черкаський державний технологічний університет
ORCID: 0009-0000-5334-4841

Я. В. ТАРАСЕНКО

доктор технічних наук, доцент,
провідний науковий співробітник
Державний науково-дослідний інститут випробувань
і сертифікації озброєння та військової техніки
ORCID: 0000-0002-5902-8628

МЕТОД ІНТЕЛЕКТУАЛЬНОГО ВИМІРЮВАННЯ ПСИХОЛІНГВІСТИЧНИХ ПОКАЗНИКІВ ДЛЯ ДИНАМІЧНИХ СИСТЕМ БЕЗПЕРЕРВНОГО МОНІТОРИНГУ ПСИХОЛОГІЧНОГО СТАНУ

У статті вирішено актуальну науково-практичну задачу психолінгвістичного уточнення, визначених за відеорядом психологічних функціональних станів, яке проводиться за рахунок інтелектуального вимірювання психологічних показників для мультимодальних динамічних систем безперервного моніторингу психологічного стану співробітників. Розроблено та описано у формалізованому вигляді двофазний метод інтелектуального вимірювання психолінгвістичних показників. Фаза первинного психолінгвістичного аналізу забезпечує формування індивідуального психологічного профілю. Фаза глибинного психолінгвістичного дослідження дозволяє на основі регресійного аналізу уточнити результати відеоконтролю первинних ознак психологічного стану при кореляції з визначеним після першої фази психологічним профілем. Вагові коефіцієнти вимірювань у мовній (текстовій) та невербальній (відео) модальностях узгоджуються за рахунок використання адаптивної ф'южн-моделі інтеграції шляхом персоналізації психолінгвістичних показників. Наведено структурно-логічну схему методу для наочної демонстрації передачі параметрів за етапами його функціонування. Було проведено попередню валідацію методу за рахунок симуляційного експерименту. За результатами експерименту запропонований метод демонструє покращене інтегральне значення точності вимірювання психолінгвістичних показників на 4-7% у порівнянні з найбільш ефективними сучасними аналогічними підходами. Отримані експериментальні значення доводять ефективність запропонованого адаптивного вагового узгодження вимірювань у текстовій та відео модальностях з урахуванням персоналізації психолінгвістичних показників. Запропонований метод потребує подальшої валідації на емпіричних даних. Подальші дослідження пропонується проводити у напрямку інтеграції додаткових модальностей та вдосконалення адаптивного оновлення вагових коефіцієнтів. Практичне значення полягає у використанні методу в мультимодальних системах безперервного інтелектуального моніторингу психологічного стану співробітників державних та комерційних структур.

Ключові слова: інтелектуальне вимірювання, психофізіологічний стан, ф'южн-модель, мультимодальні системи, системи безперервного моніторингу, автоматизація психолінгвістики.

V. P. SHAPOVAL

Postgraduate Student at the Department of Information Technology Design
Cherkasy State Technological University
ORCID: 0009-0000-5334-4841

YA. V. TARASENKO

Doctor of Engineering Science, Associate Professor,
Leading Researcher
State Scientific Research Institute of Armament
and Military Equipment Testing and Certification
ORCID: 0000-0002-5902-8628

METHOD FOR INTELLIGENT MEASUREMENT OF PSYCHOLINGUISTIC INDICATORS FOR DYNAMIC SYSTEMS OF CONTINUOUS PSYCHOLOGICAL STATE MONITORING

This article addresses a relevant scientific and practical problem of psycholinguistic refinement of psychological functional states identified from video data. This refinement is achieved through intelligent measurement of psychological indicators in multimodal dynamic systems for continuous monitoring of employees' psychological states. A two-phase method for intelligent measurement of psycholinguistic indicators is developed and formalized. The first phase, primary psycholinguistic analysis, ensures the formation of an individual psychological profile. The second phase, deep psycholinguistic analysis, allows the refinement of primary psychological state features obtained from video monitoring using regression analysis, correlating them with the psychological profile determined after the first phase. Measurement weights in the linguistic (text) and non-verbal (video) modalities are harmonized via an adaptive fusion integration model, enabling personalization of psycholinguistic indicators. A structural-logical scheme of the method is provided to visually demonstrate the transfer of parameters across its functional stages. Preliminary validation of the method was conducted through a simulation experiment. The results indicate that the proposed method improves the integral accuracy of psycholinguistic indicator measurement by 4–7% compared to the most effective contemporary approaches. Experimental findings confirm the effectiveness of adaptive weight harmonization across text and video modalities with consideration of psycholinguistic personalization. The proposed method requires further validation on empirical data. Future research is suggested in the direction of integrating additional modalities and enhancing adaptive updating of weight coefficients. The practical significance lies in its application for multimodal systems of continuous intelligent monitoring of employees' psychological states in both governmental and commercial organizations.

Key words: intelligent measurement, psychophysiological state, fusion model, multimodal systems, continuous monitoring systems, psycholinguistics automation.

Постановка проблеми

Інциденти, пов'язані з внутрішніми загрозами інформаційній безпеці державних і комерційних підприємств, установ та організацій потребують все більшої уваги з кожним роком. У сучасних умовах в Україні та світі спостерігається значне зростання кількості та вагомості випадків зловмисної чи випадкової діяльності персоналу, яка нанесла шкоду репутації установи чи супроводжувалась фінансовими збитками. Відповідно до звіту дослідницької організації Future Market Insights [1] прогнозоване зростання ринку засобів захисту з 2025 до 2035 року оцінюється в 24433,7 мільйонів доларів із середньорічним зростанням в 17,7%. Прогнозовані значення витрат демонструють актуальність досліджень в напрямку розробки нових інформаційних технологій захисту і моніторингу для запобігання чи виявлення інсайдерських атак.

У роботі [2] проаналізовані причини поведінки персоналу, яку можна розцінювати як загрозу вчинення інсайдерської атаки. Встановлено, що особистісні якості, психологічні стани та ситуативні фактори впливають на реалізації загрози інсайдерських атак. Виникає потреба виконання психологічного моніторингу для виявлення психологічних девіацій в процесі виконання функціональних обов'язків персоналу чи наявності особистісних якостей, які потенційно можуть спричинити реалізацію подібних дій.

Періодичний психологічний моніторинг не може надати точних та своєчасних результатів. Виникає потреба вирішення задачі безперервного психологічного моніторингу із застосуванням інформаційних технологій інтелектуального аналізу психологічного стану співробітників. Важливим при цьому є забезпечення саме адаптивного безперервного моніторингу з урахуванням інтелектуального вимірювання особистісних психологічних показників.

Аналіз останніх досліджень і публікацій

Сучасні методи класичного психологічного тестування проходять етап цифровізації. У роботі [3] розглядаються питання застосування інформаційних систем для нейропсихологічного тестування. Найбільш сучасні технології, включно їх машинним навчанням також використовуються у таких системах з метою забезпечення адаптивності. У роботі [4] використовується підхід автоматичної генерації елементів тестування на основі великих мовних моделей (Large Language Model – LLM). Такі системи безпосередньо орієнтовані на психологічний аналіз відповідно до існуючих шкал. Отримані дані не можуть бути точними та характеризуються періодичністю тестування. Основним обмеженням застосування подібних методів вимірювання психологічних показників є частота моніторингу, яка може бути в одних випадках недостатньою, а в інших – надмірною. Застосування найбільш сучасних технологій при використанні класичних психологічних тестів не забезпечує реалізацію необхідних для задач психологічного моніторингу властивостей, які прямо чи опосередковано впливають на характеристику точності. До таких властивостей належить достовірність результатів, глибина вимірювання параметрів, мультимодальність і динамічність.

Сучасні наукові дослідження виділяють психолінгвістичне тестування як один з інструментів забезпечення перелічених вище властивостей завдяки можливості мультимодального аналізу поведінкових, когнітивних та

мовних ознак. Такий підхід дає змогу виявляти психологічні зміни у процесі мовленнєвої діяльності та характеризуватися динамічністю вимірювання психологічних характеристик.

Психолінгвістичний аналіз ґрунтується на різних підходах, де текст виступає основним джерелом інформації. Письмові тексти, створені на довільну тематику використовуються у роботі [5]. Непрямі підходи, на основі LLM застосовуються у роботі [6] для аналізу емоційних змін у часі за психолінгвістичною характеристикою. Методи штучного інтелекту застосовуються для визначення психоемоційних відхилень на основі тексту в роботі [7]. Точність таких підходів для безперервного моніторингу є недостатньою, оскільки вимірювання враховує лише один канал збору інформації, яким виступає текст. Існують роботи, присвячені мультимодальним підходам [8], які комбінують аналіз тексту та акустичної інформації. Акустичний сигнал у такому випадку є окремою модальністю, яка доповнює текст, але вимірюються мовні показники, як паралінгвістичні маркери, які належать до одного класу – психолінгвістичних показників. Якщо розглядати інші канали отримання інформації про психологічний стан особи як відеоряд, використаний в роботі [9] для визначення первинних ознак психологічного стану, то виникає проблема мультимодальної інтеграції таких каналів як відео та текстові дані.

Спостерігається недостатня кількість робіт щодо мультимодального калібрування та інтелектуального узгодження каналів отримання інформації про психологічний стан. Згідно отриманих у роботі [10] результатів, не існує єдиної універсальної моделі. Ефективність інтелектуальної складової залежить від задачі та способу обробки інформації. Виникає потреба забезпечення точності вимірювання психологічних показників з мультимодальною інтеграцією відео та текстового каналів отримання інформації для задач динамічного безперервного моніторингу психологічного стану співробітників. Існує необхідність психолінгвістичного уточнення, визначення за відеорядом психологічних функціональних станів.

Формулювання мети дослідження

Метою статті є підвищення точності визначення індивідуальних особистісних якостей на основі інтелектуального вимірювання психолінгвістичних показників у задачах динамічного безперервного моніторингу психологічного стану.

Викладення основного матеріалу дослідження

Метод інтелектуального вимірювання психолінгвістичних показників є двофазним. Перша фаза передбачає проведення первинного психолінгвістичного аналізу. Аналіз виконується з метою формування індивідуального психологічного профілю особи при формуванні вхідного контролю. В якості вхідних параметрів передається написаний співробітником на довільну тему текст у вигляді:

$$T = \{t_1, t_2, \dots, t_N\}, \quad (1)$$

де t_N – речення N тексту T .

Фаза первинного психолінгвістичного аналізу складається з двох етапів: вимірювання індивідуальних психологічних показників (передбачає вимірювання психологічних коефіцієнтів та визначення вагових значень цих коефіцієнтів) і формування індивідуального психологічного профілю.

Для вимірювання лінгвістичних психологічних коефіцієнтів P^l Текст аналізується комплексно на трьох рівнях: синтаксичному (S_i), лексико-семантичному (L_i) та дискурсному (D_i). На синтаксичному рівні аналізується когнітивна складність мовлення. Досліджується складність та довжина мовних конструкцій, частота вживання складнопідрядних та складносурядних речень, імперативних конструкцій. На лексико-семантичному рівні оцінюється емоційне забарвлення та різноманіття вживання мовних конструкцій. На дискурсному рівні досліджується когерентність та логічна єдність тексту.

На першому етапі першої фази здійснюється кількісне вимірювання індивідуальних психолінгвістичних показників. Вимірювання кожного лінгвістичного психологічного коефіцієнта P_i^l виконується за принципом інтеграції результатів тривірневого аналізу тексту:

$$P_i^l = f_i(T; \Theta_i) = w_i^S S_i + w_i^L L_i + w_i^D D_i, \quad (2)$$

де $\Theta_i = \{w_i^S, w_i^L, w_i^D\}$ – множина вагових коефіцієнтів, які визначають вплив синтаксичного, лексико-семантичного та дискурсного рівнів на i -й психологічний коефіцієнт.

Визначення вказаних вагових коефіцієнтів виконується на основі мінімізації середньоквадратичної похибки між еталонними психологічними оцінками психологічних коефіцієнтів P_i^{et} та індивідуальними значеннями лінгвістичних психологічних коефіцієнтів:

$$\min_{\Theta_i} J_i = \frac{1}{N} \sum_{k=1}^N (P_{i,k}^{et} - (w_i^S S_{i,k} + w_i^L L_{i,k} + w_i^D D_{i,k}))^2, \quad (3)$$

при виконанні умов нормування: $w_i^S + w_i^L + w_i^D = 1$, $w_i^S, w_i^L, w_i^D \geq 1$.

Значення вагових коефіцієнтів формуються на основі регресійного аналізу для кількісної оцінки впливу кожного рівня психолінгвістичного аналізу в загальний психологічний коефіцієнт.

Наступним кроком етапу вимірювання індивідуальних психологічних показників є визначення вагових значень для лінгвістичних психологічних коефіцієнтів. Вагові коефіцієнти визначаються за аналогією з ваговими коефіцієнтами при відеомоніторингу [9] на основі факторного аналізу:

$$g_i^l = \frac{q_i n_i}{\sum_{j=1}^m q_j n_j}, \quad (4)$$

де q_i – частка бінарних показників, рівних 1 для i -го фактора; n_i – кількість ознак у групі i -го фактора; m – загальна кількість факторів.

На другому етапі першої фази виконується формування індивідуального психологічного профілю з урахуванням психологічних коефіцієнтів та їх вагових значень, отриманих на першому етапі:

$$H = \{(P_1^l, g_1^l), (P_2^l, g_2^l), \dots, (P_m^l, g_m^l)\}. \quad (5)$$

Сформований індивідуальний психологічний профіль необхідний для формування психологічного відбитку співробітника з метою подальшого динамічного відслідковування девіацій. Дані використовуються як вихідні дані методу та передаються на подальшу обробку до другої фази.

Друга фаза методу передбачає проведення уточнюючого глибинного психолінгвістичного аналізу. Розподіляється на два етапи: проведення повторного психолінгвістичного аналізу з акцентом на змінені психологічні коефіцієнти, виявлені в процесі відеомоніторингу первинних ознак психологічного стану та узгодження вагових коефіцієнтів.

На першому етапі другої фази проводиться повторний аналіз тексту з урахуванням даних відеомоніторингу. На вхід передаються психологічні коефіцієнти P_n , визначені в процесі відеомоніторингу, їх вагові коефіцієнти g_n та новий текст T^{new} для аналізу. Визначення лінгвістичних психологічних коефіцієнтів виконується наступним чином:

$$P_i^{new} = f_i(T^{new}; \Theta_i; \tilde{P}_i) = w_i^S S_i + w_i^L L_i + w_i^D D_i + \gamma_i \tilde{P}_i, \quad (6)$$

де γ_i – ваговий коефіцієнт впливу уточнених психологічних станів $\tilde{P}_i$ на глибинний психолінгвістичний аналіз.

Уточнення психологічних коефіцієнтів виконується за принципом регресійного узгодження з урахуванням лінгвістичних психологічних коефіцієнтів та коефіцієнтів функціональних станів. Формально уточнення описується рівнянням:

$$\tilde{P}_i = \beta_{0i} + \beta_{1i} P_i^l + \beta_{2i} P_i + \varepsilon_i, \quad (7)$$

де P_i – коефіцієнт функціонального психологічного стану; $\beta_{0i}, \beta_{1i}, \beta_{2i}$ – параметри регресійної моделі; ε_i – похибка моделі.

Таке регресійне узгодження забезпечує персоналізацію шляхом кореляції функціональних станів з результатами первинного психолінгвістичного аналізу, що дозволяє адаптивно враховувати індивідуальні психологічні особливості особи при первинній фіксації психологічних девіацій.

Динамічна корекція вагових коефіцієнтів передбачає використання адаптивної ф'южн-моделі, яка забезпечує властивість інтелектуалізації. Така властивість проявляється у здатності автоматичної зміни вагових коефіцієнтів. Формально властивість описується формулою:

$$g_i'(t) = \alpha_i(t) g_i^l + (1 - \alpha_i(t)) g_i, \quad (8)$$

де $g_i'(t)$ – уточнений ваговий коефіцієнт; g_i – ваговий коефіцієнт функціонального стану, визначений на етапі відеоконтролю; $\alpha_i(t)$ – адаптивний параметр узгодження, який змінюється в часі на інтервалі $[0, 1]$ в залежності від достовірності інформації отриманої з різних модальностей.

У результаті інтегральний психологічний коефіцієнт визначається як середнє зважене, що об'єднує текстову та відео модальності. Ваги залежать від поточної достовірності інформації, отриманої з різних модальностей. Психологічні показники формують за формулою:

$$P_i(t) = g_i'(t) P_i^{new} + (1 - g_i'(t)) \tilde{P}_i. \quad (9)$$

Принцип інтелектуальної адаптації у методі при інтеграції текстової та відео модальностей описується формулою:

$$\Theta_i'(t) = \lambda \Theta_i + (1 - \lambda) \Theta_i^V, \quad (10)$$

де $\Theta'_i(t)$ – оновлені лінгвістичні вагові коефіцієнти; Θ'_i – невербальні вагові коефіцієнти; λ – коефіцієнт узгодження, що приймає значення на інтервалі $[0,1]$ в залежності від достовірності інформації у двох модальностях.

Коефіцієнт узгодження $\alpha_i(t) = \lambda$ визначається на основі достовірності каналів за формулою:

$$\lambda_i(t) = \frac{1 - \frac{\sigma^2(P_i^l)}{\max(\sigma^2(P_i^l))}}{1 - \frac{\sigma^2(P_i^l)}{\max(\sigma^2(P_i^l))} + \text{corr}(P_i, \hat{P}_i)}, \quad (11)$$

де $\sigma^2(P_i^l)$ – дисперсія лінгвістичного психологічного коефіцієнта; $\hat{P}_i$ – прогнозоване значення показника на основі відеомоніторингу.

Коефіцієнт узгодження у ф'южн-моделі забезпечує динамічне балансування вагових значень між модальностями відповідно до їх достовірності.

Структурно-логічна схема запропонованого двофазного методу з урахуванням функціональних зв'язків між етапами кожної фази та передачі параметрів між ними представлена на рисунку 1.

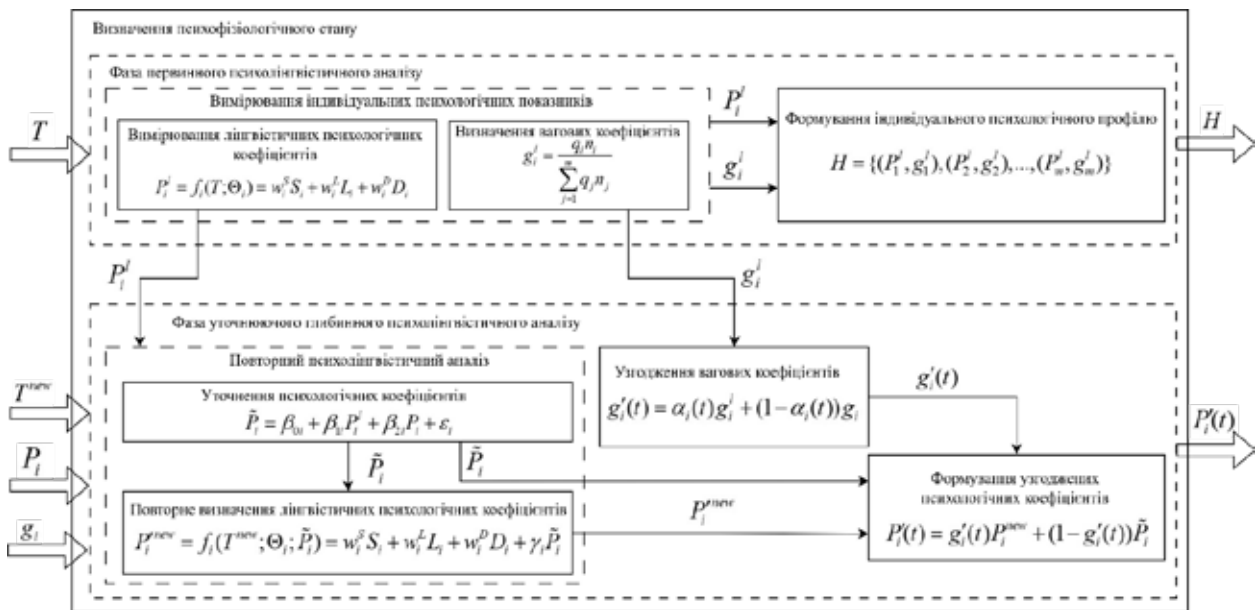


Рис. 1. Структурно-логічна схема методу інтелектуального вимірювання психолінгвістичних показників

З метою оцінювання точності визначення індивідуальних особистісних якостей було проведено симуляційний експеримент у середовищі Python 3.11 із використанням бібліотек Numpy, Pandas, scikit-learn. Моделювалась робота розробленого методу та трьох найбільш релевантних сучасних аналогічних підходів: з урахуванням невизначеності (UA) [11], орієнтований на увагу / знання (DAKG) [12] та мультимодальний з комбінуванням LLM і виразів обличчя (LLMF) [13].

Симуляції показали, що ключову роль відіграє кількість змодельованих суб'єктів (персоналу). Якщо суб'єкт один, то кращий результат точності демонстрував підхід з урахуванням невизначеності. Для задач безперервного моніторингу психологічного стану персоналу важлива наявність більшої кількості суб'єктів. У ході симуляційного експерименту було змодельовано наступні умови: 213 унікальних суб'єктів, активними з яких були одночасно від 120 до 166 осіб. Завдяки цьому було виконано імітацію динамічної заміни кадрового складу. Було згенеровано 9000 спостережень з додаванням шуму та випадкових психологічних девіацій. Вимірювання виконувались у мовній та невербальній модальності. Було враховано різну ступінь достовірності каналів вимірювання психологічних факторів. Для кожного спостереження виконувались обчислення таких похибок: середньоквадратична (RMSE), середня абсолютна (MAE) та коефіцієнти детермінації (R^2). Також обчислювалось відсоткове значення хибнопозитивних (FP) та хибнонегативних (FN) визначень зміни індивідуальних психологічних характеристик. Візуальне представлення у вигляді порівняльної діаграми зображено на рисунку 2.

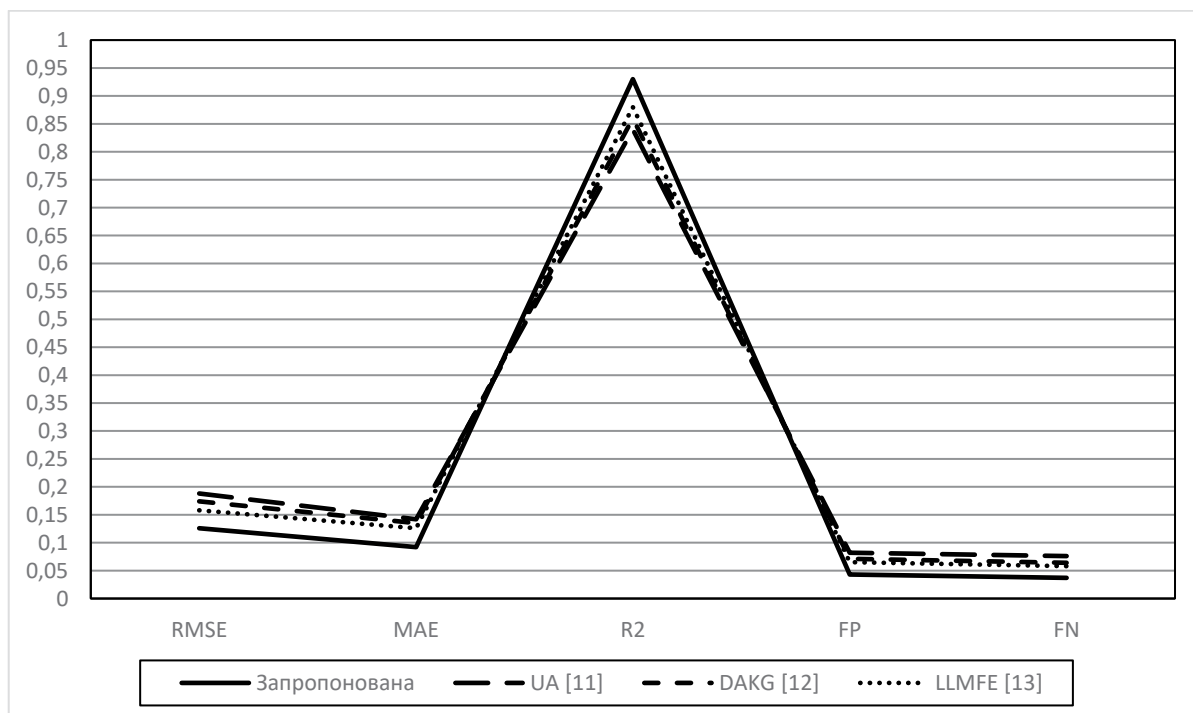


Рис. 2. Порівняльна діаграма показників точності розробленого методу та аналогів

На основі отриманих значень було обчислено інтегральний показник точності за формулою:

$$I_{acc} = 1 - \frac{\omega_1 \cdot RMSE + \omega_2 \cdot MAE + \omega_3 \cdot (FP + FN) / 2}{3} + \omega_4 \cdot R^2, \quad (12)$$

де $\omega_1 = 0,3$; $\omega_2 = 0,3$; $\omega_3 = 0,2$; $\omega_4 = 0,2$.

Інтегральний показник точності для розробленого методу складає 0,935. Для UA – 0,862; DAKG – 0,877; LLMFE – 0,894.

Результати симуляційного експерименту довели підвищення точності для розробленого методу на 4-7% у порівнянні з найкращими сучасними підходами. Низькі значення FP та FN дозволяють використовувати даний метод у системах безперервного моніторингу та зумовлює високий рівень довіри до результатів вимірювань.

Висновки

Таким чином, у роботі було вирішено актуальну науково-практичну задачу психолінгвістичного уточнення, визначених за відеорядом психологічних функціональних станів у безперервному психологічному моніторингу.

Вперше запропоновано двофазний метод інтелектуального вимірювання психолінгвістичних показників на основі вагового узгодження вимірювань у мовній та невербальній модальностях за рахунок використання адаптивної ф'южн-моделі інтеграції шляхом персоналізації психолінгвістичних показників, що забезпечує підвищення точності визначення психологічних коефіцієнтів при динамічному уточненні їх вагових значень відповідно до поточної достовірності каналів спостереження.

Результати попередньої валідації розробленого методу шляхом виконання симуляційного експерименту, який моделює умови динамічного безперервного моніторингу психологічного стану персоналу доводять покращення точності за інтегральним показником оцінювання на 4-7% у порівнянні з найбільш ефективними сучасними аналогічними підходами. Отримані дані дозволили підтвердити ефективність адаптивного вагового узгодження вимірювань у відео та текстовій модальностях з урахуванням персоналізації психолінгвістичних показників.

Практичне значення запропонованого методу полягає у можливості його використання при побудові мультимодальних інформаційних систем безперервного інтелектуального моніторингу психологічного стану співробітників державних та комерційних структур. Існує можливість інтеграції в уже існуючі інформаційні системи моніторингу персоналу.

Запропонований метод потребує подальшої валідації на емпіричних даних. Перспективою подальших досліджень є додавання інших модальностей та вдосконалення адаптивного оновлення ваг на їх основі, а також зменшення похибки дистанційних інтелектуальних вимірювань психологічних коефіцієнтів.

Список використаної літератури

1. Future Market Insights. Insider threat protection market analysis size and share forecast outlook 2025-2035 / S. Saha. Future Market Insights, 2025. 520 p. URL: <https://www.futuremarketinsights.com/reports/insider-threat-protection-market> (дата звернення: 01.11.2025).
2. Saddica M., Ruohonen Ju. SoK: the psychology of insider threats. *EAI Endorsed Transactions on Security and Safety*. 2025. Vol. 9, № 1. URL: <https://doi.org/10.4108/eetss.v9i1.9298>
3. Moving toward the digitalization of neuropsychological tests: an exploratory study on usability and operator perception / M.G. Maggio et al. *Digital Health*. 2025. Vol. 11. URL: <https://10.1177/20552076251334449> (дата звернення: 03.11.2025).
4. Lee P., Son M., Jia Z. AI-powered automatic item generation for psychological tests: a conceptual framework for an LLM-based multiagent AIG system. *Journal of Business and Psychology*. 2025. URL: <https://10.1007/s10869-025-10067-y> (дата звернення: 03.11.2025).
5. Psychometric evaluation of large language model embeddings for personality trait prediction / J. Maharjan et al. *Journal of Medical Internet Research*. 2025. Vol. 27. URL: <https://doi.org/10.2196/75347> (дата звернення: 03.11.2025).
6. Towards dynamic theory of mind: evaluating LLM adaptation to temporal evolution of human states / Y. Xiao et al. *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Vienna, Austria, Stroudsburg, PA, USA, 2025. P. 24036–24057. DOI: <https://doi.org/10.18653/v1/2025.acl-long.1171>
7. Sert B., Ulker S.V. A review of LWIC and machine learning approaches on mental health diagnosis. *Social Review of Technology and Change*. 2023. Vol. 1, № 2. P. 71–92.
8. When LLMs meets acoustic landmarks: an efficient approach to integrate speech into large language models for depression detection / X. Zhang et al. *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, Miami, Florida, USA. Stroudsburg, PA, USA, 2024. P. 146–158. DOI: <https://doi.org/10.18653/v1/2024.emnlp-main.8>
9. Шаповал В.П., Тарасенко Я.В. Метод інтелектуального відеоконтролю первинних ознак психологічного стану. *Вчені записки ТНУ імені В.І. Вернадського. Серія: технічні науки*. 2025. Том 36 (75), № 2. С. 222–227. DOI: <https://doi.org/10.32782/2663-5941/2025.2.2/30>
10. Machine learning for multimodal mental health detection: a systematic review of passive sensing approaches / L.S. Khoo et al. *Sensors*. 2024. Vol. 24, № 2. P. 348. DOI: <https://doi.org/10.3390/s24020348>
11. Uncertainty-aware multi-modal random network prediction / H. Wang et al. *Lecture notes in computer science*. Cham, 2022. P. 200–217. DOI: https://doi.org/10.1007/978-3-031-19836-6_12
12. Knowledge-guided dynamic modality attention fusion framework for multimodal sentiment analysis / X. Feng et al. *Findings of the association for computational linguistics: EMNLP 2024*, Miami, Florida, USA. Stroudsburg, PA, USA, 2024. P. 14755–14766. DOI: <https://doi.org/10.18653/v1/2024.findings-emnlp.865>
13. Harnessing multimodal approaches for depression detection using large language models and facial expressions / M. Sadeghi et al. *NPJ Mental Health Research*. Vol. 3, № 1. DOI: <https://doi.org/10.1038/s44184-024-00112-8>

References

1. Saha, S. (2025). *Insider threat protection market analysis size and share forecast outlook 2025-2035*. Future Market Insights. 520 p. <https://www.futuremarketinsights.com/reports/insider-threat-protection-market>
2. Saddica, M., & Ruohonen, Ju. (2025). SoK: the psychology of insider threats. *EAI Endorsed Transactions on Security and Safety*, 9(1). <https://doi.org/10.4108/eetss.v9i1.9298>
3. Maggio, M.G., Giambo, F.M., Barbera, M., De Pasquale, P., Bruno, F., Calderone, A., Rizzo, A., & Calabro, R.S. (2025). Moving toward the digitalization of neuropsychological tests: an exploratory study on usability and operator perception. *Digital Health*, 11. <https://10.1177/20552076251334449>
4. Lee, P., Son, M., & Jia, Z. (2025). AI-powered automatic item generation for psychological tests: a conceptual framework for an LLM-based multiagent AIG system. *Journal of Business and Psychology*. <https://10.1007/s10869-025-10067-y>
5. Maharjan, J., Jin, R., Zhu, J., Kenne, D. (2025). Psychometric evaluation of large language model embeddings for personality trait prediction. *Journal of Medical Internet Research*, 27. <https://doi.org/10.2196/75347>
6. Xiao, Y., Wang, J., Xu, Q., Song, C., Xu, C., Cheng, Y., Li, W., & Liu, P. (2025). Towards dynamic theory of mind: evaluating LLM adaptation to temporal evolution of human states. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 24036–24057). <https://doi.org/10.18653/v1/2025.acl-long.1171>
7. Sert, B., & Ulker, S.V. (2023). A review of LWIC and machine learning approaches on mental health diagnosis. *Social Review of Technology and Change*, 1(2), 71–92.
8. Zhang, X., Liu, H., Xu, K., Zhang, Q., Liu, D., Ahmed, B., & Epps, J. (2024). When LLMs meets acoustic landmarks: an efficient approach to integrate speech into large language models for depression detection. In *Proceedings of the*

2024 Conference on Empirical Methods in Natural Language Processing (pp. 146–158). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2024.emnlp-main.8>

9. Shapoval, V.P., & Tarasenko, Y.V. (2025). Metod intelektualnoho videokontroliu pervynnykh oznak psykhologichnoho stanu [Method of intelligent video monitoring of primary signs of psychological state]. *Vcheni zpyssky TNU imeni V.I. Vernadskoho. Serija: tekhnichni nauky – Scientific notes of Taurida National V.I. Vernadsky University. Series: Technical Sciences*, 36(75), 2, 222–348. <https://doi.org/10.32782/2663-5941/2025.2.2/30>

10. Khoo, L.S., Lim, M.K., Chong, C.Y., & McNaney, R. (2024). Machine learning for multimodal mental health detection: a systematic review of passive sensing approaches. *Sensors*, 24(2), 348. <https://doi.org/10.3390/s24020348>

11. Wang, H., Zhang, J., Chen, Y., Ma, C., Avery, J., Hull, L., & Carneiro, G. (2022). Uncertainty-aware multi-modal random network prediction. In *Lecture notes in computer science* (pp. 200–217). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-19836-6_12

12. Feng, X., Lin, Y., He, L., Li, Y., Chang, L., & Zhou, Y. (2024). Knowledge-guided dynamic modality attention fusion framework for multimodal sentiment analysis. In *Findings of the association for computational linguistics: EMNLP 2024* (pp. 14755–14766). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2024.findings-emnlp.865>

13. Sadeghi, M., Richer, R., Egger, B., Schindler-Gmelch, L., Rupp, L.H., Rahimi, F., Berking, M., & Eskofier, B.M. (2024). Harnessing multimodal approaches for depression detection using large language models and facial expressions. *NPJ Mental Health Research*, 3(1). <https://doi.org/10.1038/s44184-024-00112-8>

Дата першого надходження рукопису до видання: 11.11.2025
Дата прийнятого до друку рукопису після рецензування: 08.12.2025
Дата публікації: 31.12.2025

О. М. ШУШУРА

доктор технічних наук, професор,
професор кафедри цифрових технологій в енергетиці
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
ORCID: 0000-0003-3200-720X

Б. О. СУХЕНКО

бакалавр кафедри цифрових технологій в енергетиці
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
ORCID: 0009-0007-2508-4350

С. О. СОЛОМАХА

кандидат економічних наук,
доцент кафедри інформаційних систем та технологій
Державний університет інформаційно-комунікаційних технологій
ORCID: 0009-0002-8358-8657

Л. А. АСЄЄВА

доктор філософії,
доцент кафедри комп'ютерної інженерії
Державний університет інформаційно-комунікаційних технологій
ORCID: 0000-0001-5954-4211

ІНТЕГРАЦІЯ ПРОГРАМНИХ РЕЄСТРАТОРІВ РОЗРАХУНКОВИХ ОПЕРАЦІЙ З ERP-СИСТЕМАМИ НА ПРИКЛАДІ ODOO

У статті розглянуто архітектуру, моделі процесів та програмне забезпечення для інтеграції українських програмних реєстраторів розрахункових операцій (ПРРО) з ERP-системами на прикладі модуля Point of Sale (Точка продажу) системи Odoо з провайдерами «Вчасно.Каса» та «Чекбокс». Особливу увагу приділено відмінностям у проведенні фіскалізації через різні типи API провайдерів, проілюстрованим на моделях процесів провайдера «Чекбокс».

Необхідність інтеграції ПРРО з ERP-системами зумовлена стрімким розвитком автоматизації фіскалізації в Україні та потребою малого й середнього бізнесу в гнучких рішеннях. Така інтеграція дає змогу підприємцям одночасно використовувати потужні системи управління ресурсами та взаємовідносинами з клієнтами (CRM/ERP) і виконувати в них усі фіскальні та нефіскальні операції точки продажу з обов'язковою реєстрацією операцій у Державній податковій службі України (ДПС), а також працювати з периферійним обладнанням.

У роботі формалізовано вимоги до інтеграції ПРРО з ERP-системами, розроблено архітектуру підсистеми інтеграції та моделі процесів фіскалізації розрахункових операцій із застосуванням ПРРО. Запропонована модульна архітектура, побудована за схемою «хаб + адаптери», дозволяє розробникам-інтеграторам швидко й зручно підключати нових провайдерів ПРРО до підсистеми точки продажу, забезпечуючи при цьому стійкість, масштабованість та легкість супроводу системи.

На основі розробленої архітектури створено програмне забезпечення, що складається з центрального хаба та адаптерів для двох провайдерів. Реалізовано підтримку як хмарного (віддаленого), так і локального API, що забезпечує роботу в офлайн-режимі в межах, дозволених законодавством України. Система підтримує повний цикл фіскальних і нефіскальних операцій, стандартних для точки продажу. Тестування на обох типах API підтвердило високу швидкодію та надійність розробленого рішення навіть за великої кількості одночасних торгових операцій.

Ключові слова: ERP, інтеграція ПРРО, Чекбокс, Вчасно.Каса, фіскалізація, API, інформаційна система, Odoо, Python, JavaScript.

O. M. SHUSHURA

Doctor of Technical Sciences, Professor,
Professor at the Department of Digital Technologies in Energy
National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute”
ORCID: 0000-0003-3200-720X

B. O. SUKHENKO

Bachelor at the Department of Digital Technologies in Energy
National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute”
ORCID: 0009-0007-2508-4350

S. A. SOLOMAKHA

Candidate of Economic Sciences,
Associate Professor at the Department of Information Systems and Technologies
State University of Information and Communication Technologies
ORCID: 0009-0002-8358-8657

L. A. ASIEIEVA

PhD, Associate Professor at the Department of Computer Engineering
State University of Information and Communication Technologies
ORCID: 0000-0001-5954-4211

INTEGRATION OF SOFTWARE FISCAL REGISTRARS WITH ERP SYSTEMS ON THE EXAMPLE OF ODOO

The article examines the architecture, process models, and software for integrating Ukrainian software fiscal registrars (PRRO) with ERP systems, using the Point of Sale module of the Odoo system as an example, together with the PRRO providers «Vchasno.Kasa» and «Checkbox». Particular attention is paid to the differences in fiscalization processes when using different types of provider APIs, illustrated through the process models of the «Checkbox» provider.

The need for PRRO integration with ERP systems is driven by the rapid development of fiscalization automation in Ukraine and the demand from small and medium-sized businesses for flexible solutions. Such integration enables entrepreneurs to simultaneously use powerful resource and customer relationship management systems (ERP/CRM) while performing all fiscal and non-fiscal point-of-sale operations directly within them, with mandatory registration of transactions in the State Tax Service of Ukraine (STS), as well as supporting the operation of peripheral equipment.

The study formalizes the requirements for PRRO integration with ERP systems, proposes an integration subsystem architecture, and develops models of fiscalization processes for settlement transactions using PRRO. The proposed modular architecture, built on the «hub + adapters» pattern, allows integration developers to quickly and conveniently connect new PRRO providers to the point-of-sale subsystem while ensuring system resilience, scalability, and ease of maintenance.

Based on the developed architecture, software has been created consisting of a central hub and adapters for the two providers. Support has been implemented for both cloud-based (remote) and local APIs, enabling offline operation within the limits permitted by Ukrainian legislation. The system supports the full cycle of fiscal and non-fiscal operations typical for a point of sale. Testing on both API types confirmed the high performance and reliability of the developed solution, even under a heavy load of simultaneous retail transactions.

Key words: ERP, SFR integration, Checkbox, Vchasno.Kasa, fiscalization, API, information system, Odoo, Python, JavaScript.

Постановка проблеми

З 1 січня 2022 року набув чинності Закон України «Про застосування реєстраторів розрахункових операцій у сфері торгівлі, громадського харчування та послуг» в редакції № 1073-IX [1], який зробив використання програмних реєстраторів розрахункових операцій (ПРРО) обов’язковим для більшості суб’єктів господарювання. Станом на квітень 2025 року щоденно фіскалізується понад 28 мільйонів чеків, що на 0,7 млн більше, ніж у березні попереднього року [2]. Цей показник продовжує зростати завдяки розширенню переліку платників та переходу бізнесу на цифрові рішення.

Варто відмітити, що жодна ERP/CRM-система не має зараз функціональних можливостей для роботи з провайдерами ПРРО – для цього слугують програми самих провайдерів. Багатьом підприємцям доводиться одночасно використовувати ПРРО, а потім окремо вести звітність в ERP-системі. Вирішення задачі інтеграції ПРРО в такі системи дозволить підвищити ефективність малого та середнього бізнесу.

Кількість провайдерів ПРРО також активно збільшується та розвиваються їх функції. Це зумовлює важливість такої архітектури інтеграції, що дозволила б легко і швидко інтегрувати нових провайдерів до доступних з повноцінним збереженням всіх необхідних фіскальних та нефіскальних функціональних можливостей точки продажу. Тому розробка новітньої модульної архітектури «хаб + адаптери» дасть змогу розробникам-інтеграторам швидко і зручно підключати до системи нових провайдерів ПРРО. Варто відмітити, що дану архітектуру можна використовувати в багатьох задачах інтеграції зовнішніх підсистем в ERP/CRM-системи.

Популярна open-source ERP-система Odoo активно використовується українським малим і середнім бізнесом для обліку та автоматизації точок продажу. Стандартний модуль точки продажу Point of Sale (POS), не має вбудованої інтеграції з ПРРО, що створює значну прогалину в автоматизації процесів та відповідності законодавству.

Аналіз останніх досліджень і публікацій

Задача інтеграції програмних реєстраторів розрахункових операцій (ПРРО) з ERP-системами, зокрема ODOO, зумовлена швидким розвитком автоматизації фіскалізації в Україні та потребою в гнучких рішеннях для малого та середнього бізнесу. Тим не менш, системи ПРРО є доволі новим засобом, тому досліджень з цієї тематики небагато.

У статті [3] аналізується впровадження ПРРО як інструменту детінізації торгівлі в Україні. Досліджуються причини переходу підприємців на ПРРО, їхні переваги, проводиться дослідження проблемних аспектів впровадження реєстраторів розрахункових операцій фізичними-особами підприємцями, а також порівняння провайдерів ПРРО. Системи ПРРО хоч і є зручним засобом для малого та середнього бізнесу, але не дозволяють повноцінно виконувати управління ресурсами підприємства і вести повноцінну звітність. Цю проблему має вирішити інтеграція провайдерів ПРРО з ERP-системами.

У роботі [4] з фіскальної децентралізації та цифровізації податкових систем аналізуються концепції фіскального регулювання в умовах цифрової економіки України, де підкреслюється роль програмних комплексів для оптимізації бюджетних ресурсів та їхнє використання. Автори пропонують моделі, що дозволяють підвищити ефективність фіскалізації через API-інтеграцію. Ця робота демонструє, що актуальність ПРРО є зростаючою і значно оптимізує ресурси. Тим не менш, варто зазначити, що залишається невирішеним питання інтеграції ПРРО з ERP-платформами, хоча це є важливим аспектом сьогодення: самі по собі ПРРО не надають систему управління ресурсами і тому розробники вимушені власноруч інтегрувати їх в свої ERP-системи.

У дослідженні [5] розглядається інтеграція бізнес-додатків з блокчейн-технологіями на прикладі Odoo та Hyperledger Fabric, з акцентом на API-моделі для верифікації даних, що актуально для фіскалізації в ERP-системах. Автори пропонують proof-of-concept для оптимізації зовнішніх інтеграцій. Але запропонована оптимізація не дозволить зробити модульну систему максимально масштабованою. Це є дуже важливим в той час, коли кількість провайдерів ПРРО збільшується, а підприємства хочуть мати свободу вибору в плані вибору свого провайдера.

Автори статті [6] розглядають імплементацію Odoo 9 Community Edition з BI-інструментами для виробничого підприємства, моделюючи бізнес-процеси (в тому числі продаж і облік). Особлива увага в статті приділена модульній архітектурі для інтеграції з локальними системами. Однак при цьому не приділяється достатньої уваги тому, що у випадку інтеграції багатьох подібних підсистем архітектура має бути по типу «хаб + адаптери», а саме один модуль може (і має) бути центральним, виконувати спільні функції, а інші – наслідувати і розширювати його. Тому задача подальшого розвитку архітектури та моделей інтеграції українських програмних реєстраторів розрахункових операцій (ПРРО) з ERP-системами є актуальною.

Формулювання мети дослідження

Мета роботи полягає у розробці архітектури, моделей та програмного забезпечення підсистеми інтеграції програмних реєстраторів розрахункових операцій з ERP-системами на прикладі інформаційної системи Odoo для забезпечення повної відповідності вимогам законодавства України та автоматизації фіскалізації розрахункових операцій у малих і середніх підприємствах.

Для досягнення поставленої мети необхідно:

- формалізувати вимоги до реалізації інтеграції ПРРО з ERP-системами;
- сформувати архітектуру підсистеми інтеграції та моделі процесів фіскалізації розрахункових операцій;
- розробити програмне забезпечення модулів Odoo для взаємодії з провайдерами програмних реєстраторів розрахункових операцій «Вчасно.Каса» і «Чекбокс»;
- провести тестування розробленого програмного забезпечення.

Викладення основного матеріалу дослідження

Функціональні вимоги до підсистеми інтеграції ПРРО з ERP-системами формалізовано у вигляді діаграми прецедентів UML на рисунку 1.

Як видно з рисунку 1, виділено дві основні ролі користувачів: касир та адміністратор торгової точки. Касир виконує фіскальні операції (продаж, повернення, внесення/видача готівки, X/Z-звіти), а адміністратор налаштовує роботу ПРРО.

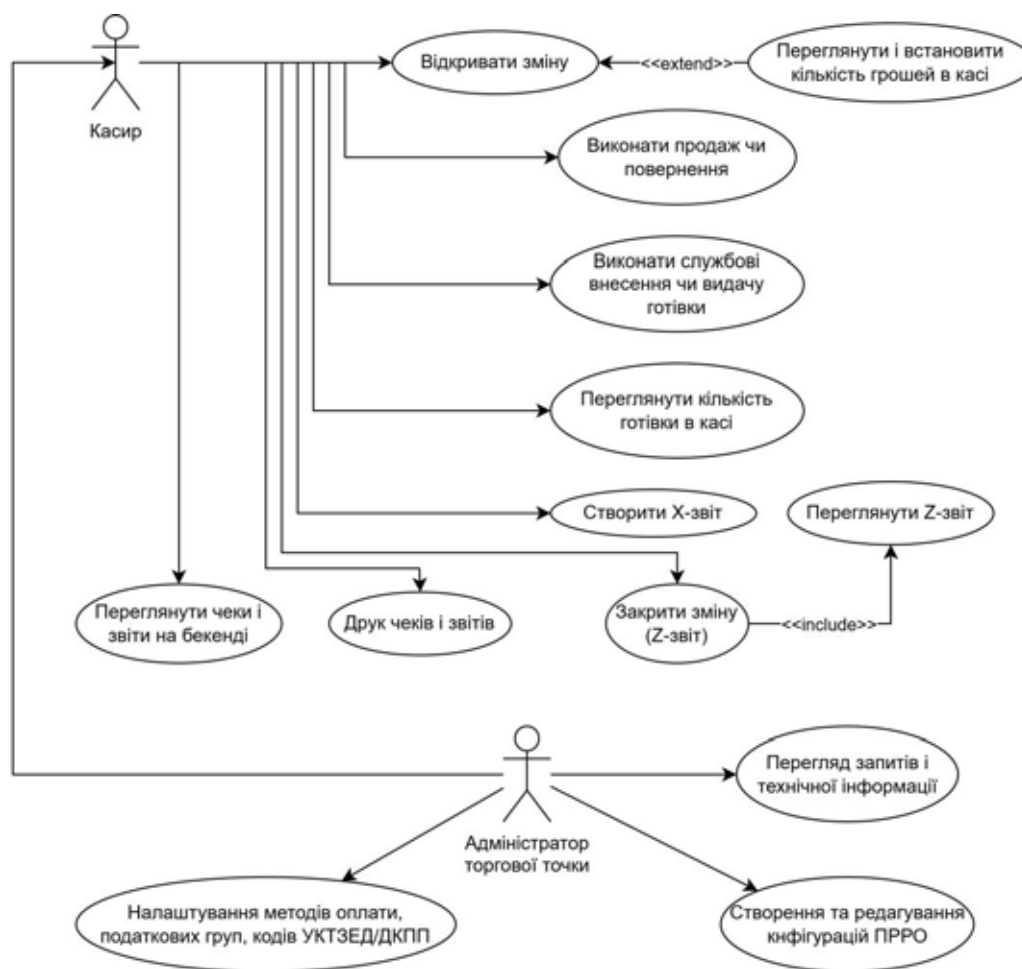


Рис. 1. Діаграма прецедентів розробленої системи

На основі наведених функціональних вимог розроблено архітектуру системи. Вона включає в себе два типи модулів: центральний модуль «ПРРО» та модулі провайдерів.

Центральний модуль «ПРРО» містить в собі всі спільні та необхідні для роботи налаштування, конфігурації провайдерів ПРРО і реалізацію функцій, подібних до контролю часу.

Модулі провайдерів відповідають за, власне, всі фіскальні і нефіскальні операції точки продажу. Весь набір операцій, доступних в підсистемі інтеграції, можна побачити на рисунку 1.

Така архітектура забезпечує необхідну легкість і зручність як для користувача, який може обирати будь-якого провайдера, так і для інтегратора, який використовує вже існуючий центральний модуль і реалізує лише зв'язок з API нового провайдера ПРРО. На основі зазначеного підходу розроблено представлену на рисунку 2 архітектуру підсистеми інтеграції ПРРО провайдерів з провайдерами «Вчасно.Каса» та «Чекбокс» з інформаційною системою Odoo.

Як видно на рисунку 2, модуль-хаб `us_prro` є центральним, він залежить напряму від модуля «Точка продажу» Odoo і розширює його. Модулі провайдерів наслідують центральний та розширюють його і працюють з API своїх провайдерів. Варто відмітити, що використання подібної архітектури можливе в майже будь-яких ERP/CRM-системах, які містять в собі модуль Точки Продажу.

З врахуванням запропонованих архітектурних рішень необхідно розробити моделі проведення фіскальних операцій. Через різні можливості використання торговими точками підприємств мережі Інтернет необхідно розглянути варіанти роботи з локальними API ПРРО (коли доступ до Інтернет нестабільний) та хмарними API. Слід зазначити, що відмінності в реалізації цих процесів різними провайдерами ПРРО незначні. Тому розглянемо моделі фіскалізації продажу на прикладі хмарного та локального API провайдера «Чекбокс». Ці моделі у вигляді діаграм послідовності UML наведено на рисунках 3 та 4 відповідно.

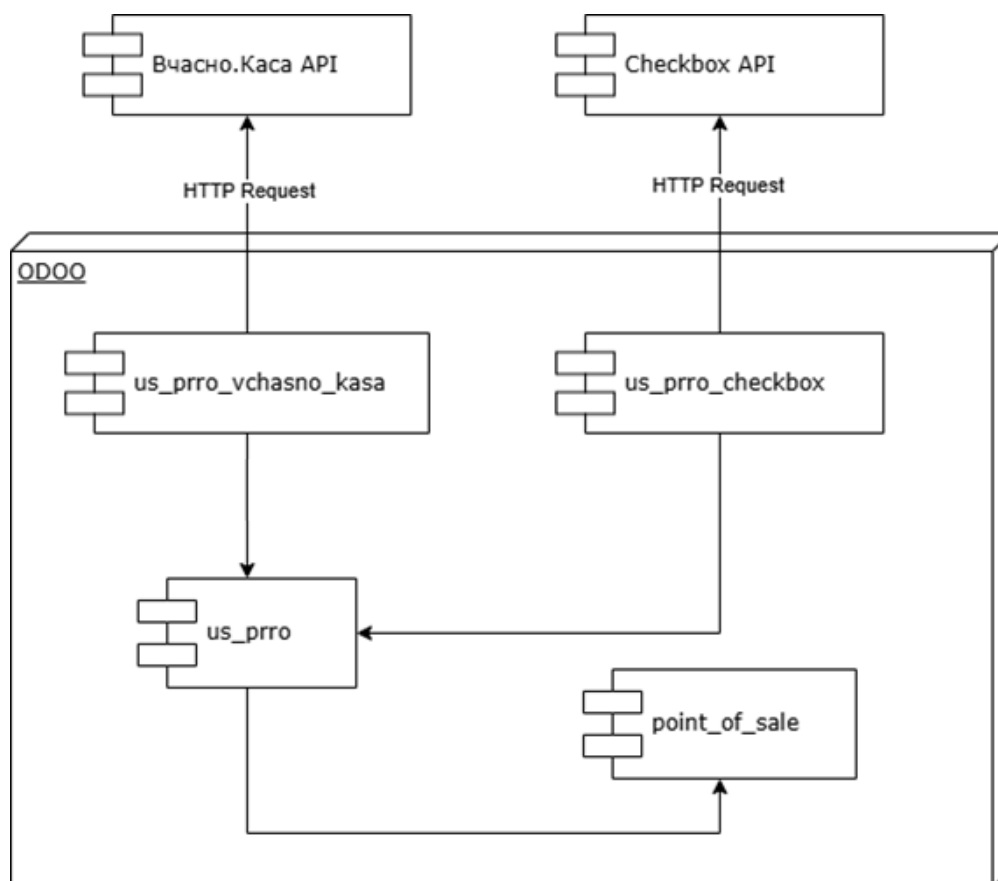


Рис. 2. Архітектура модульної системи інтеграції для Odoo

З моделі, представленої на рисунку 3, бачимо, що у випадку роботи з хмарним API спочатку виконується збір і форматування даних під запит, після чого дані надсилаються на сервер Odoo, а сервер робить POST-запит на фіскалізацію та отримує статус CREATED для чеку. Після цього потрібно дочекатись фіскалізації даних – сервер з невеликим таймаутом робить запити з id чеку, доки не отримає статус DONE. Отримавши цей статус, сервер виконує запит на отримання текстової відформатованої версії чеку і передає всі дані назад на користувацьку частину. Дані зберігаються в поточному замовленні, доки воно не фіналізується стандартною логікою Odoo, а потім надсилаються знову на сервер для збереження.

Як видно на рисунку 4, модель процесу фіскалізації через API провайдера, що знаходиться локально на комп'ютері торгової точки, значно відрізняється. Після форматування чеку виконується запит на генерацію UUID чеку на сервер, який у випадку прямого API генерувався автоматично. Тепер система надсилає на «Каса Менеджер» (локальне API) запит на фіскалізацію і відразу отримує відповідь щодо його успішності, і далі вже відбувається аналогічна прямому API логіка збереження. Підсистема «Каса Менеджер» робить всі відповідні запити асинхронно. Якщо інтернет з'єднання нема, то вона видає фіскальний номер чеку з пулу зарезервованих номерів і надсилає дані на сервер вже тоді, коли з'єднання буде.

Зазначені моделі покладено в основу програмного забезпечення, розробленого на мові програмування Python з використанням OWL-фреймворка Odoo для реалізації реактивного інтерфейсу за допомогою JavaScript [7]. Програмне забезпечення реалізує автоматичне заокруглення при виконанні готівкових операцій, обробку всіх кодів оплат відповідно до діючої нормативної бази [8], підтримує обробку акцизних марок на товар. Зручний користувацький інтерфейс модулів забезпечує швидку роботу, гнучкість у роботі з ПРРО, дозволяє відслідковувати та обробляти всі необхідні дані.

Тестування проводилось на точці продажу з середнім навантаженням приблизно 50 чеків на годину. Усі операції, як в режимі онлайн, так і в офлайн, були виконані успішно, помилок фіскалізації не зафіксовано.

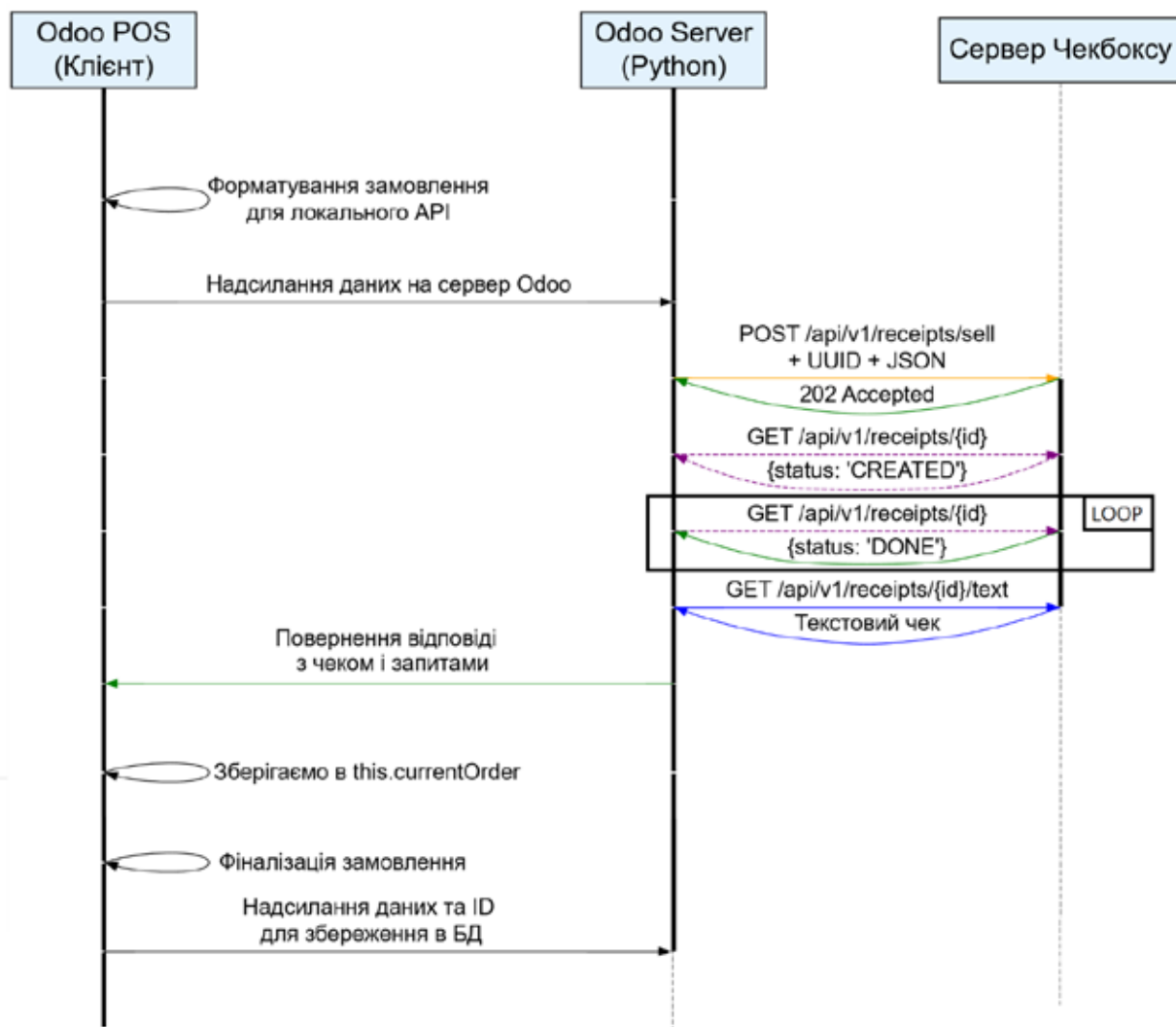


Рис. 3. Модель процесу фіскалізації продажу через пряме (хмарне) API

Висновки

Таким чином, розроблено архітектуру, моделі та програмне забезпечення підсистеми інтеграції програмних реєстраторів розрахункових операцій з ERP-системами на прикладі інформаційної системи Odoo з повною відповідністю вимогам чинного законодавства України.

Формалізовано вимоги до інтеграції ПРРО з ERP-системами та запропоновано модульну архітектуру «хаб + адаптери», яка забезпечує швидке підключення нових провайдерів і незалежність спільної логіки від конкретної реалізації провайдера.

Сформовано моделі процесів фіскалізації для випадків хмарного та локального API з урахуванням офлайн-режиму, чергування операцій та контролю лімітів, установлених нормативною базою України.

Розроблено програмне забезпечення трьох Odoo-модулів (центрального хаба та адаптери для провайдерів «Вчасно.Каса» й «Чекбокс»), що підтримують повний набір фіскальних і нефіскальних операцій точки продажу підприємства. Проведене тестування показало повну готовність цієї підсистеми Odoo до практичного використання.

Розроблене рішення забезпечує надійну інтеграцію ПРРО з ERP-системами в умовах нестабільного інтернет-з'єднання, що має особливу актуальність для малого та середнього бізнесу України.

Подальші дослідження в цій області доцільно спрямувати на створення універсального абстрактного шару інтеграції ПРРО, незалежного від конкретної ERP-платформи, та розширення підтримки нових провайдерів ПРРО.

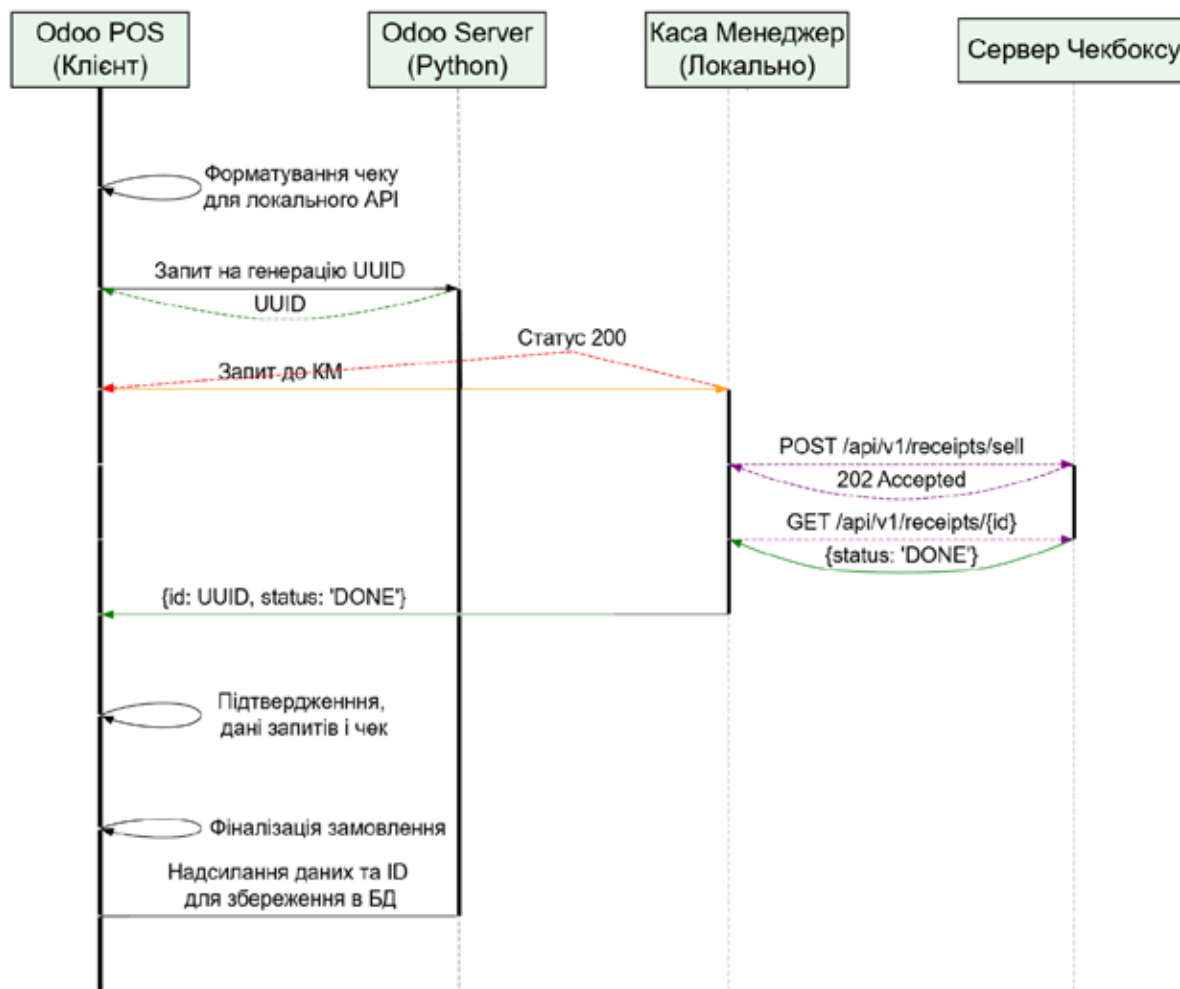


Рис. 4. Модель процесу фіскалізації продажу через локальне API

Список використаної літератури

1. Про застосування реєстраторів розрахункових операцій у сфері торгівлі, громадського харчування та послуг : Закон України від 06.07.1995 № 265/95-ВР. URL: <https://zakon.rada.gov.ua/laws/show/265/95-вр#Text> (дата звернення: 29.10.2025).
2. Як змінювався ринок ПРРО після 1 березня. Вчасно. URL: <https://vchasno.ua/yak-zminyvsya-rynok-prro-pislya-1-bereznia> (дата звернення: 14.11.2025).
3. Belova, I., Farion, V., & Yaroshchuk, O. (2023). Features of the implementation of settlement transactions registrars in ukraine: accounting and financial aspects. *Ekonomichnyy analiz*, 2(2), 209–226. <https://doi.org/10.35774/econa2023.02.209>.
4. Lutsenko, A., Vykluk, M., Skoryk, M., & Hromova, T. (2020). Fiscal Regulation Concept Formation Of The Ukraine's Economy Development. *Studies of Applied Economics*, 38(4), 1–10. <https://doi.org/10.25115/eea.v38i4.3997>.
5. Belhi, A., Gasmi, H., Bouras, A., Aouni, B., & Khalil, I. (2021). Integration of Business Applications with the Blockchain: Odoo and Hyperledger Fabric Open Source Proof of Concept. *IFAC-PapersOnLine*, 54(1), 817–824. <https://doi.org/10.1016/j.ifacol.2021.08.185>.
6. Wu, J.-Y., & Chen, S.-L. (2020). Odoo ERP with Business Intelligence Tool for a Small-Medium Enterprise: a scenario case study. In *Proceedings of the 2020 11th International Conference on E-Education, E-Business, E-Management, and E-Learning* (pp. 1–6). New York: ACM. <https://doi.org/10.1145/3377571.3377607>.
7. OWL Components. Odoo Documentation 18.0. URL: https://www.odoo.com/documentation/18.0/developer/reference/frontend/owl_components.html (дата звернення: 10.11.2025).
8. API Наказ № 601 – зміни в чеку. Checkbox Wiki. URL: https://wiki.checkbox.ua/uk/api/nakaz_601#зміни-в-чеку-згідно-наказу-по601 (дата звернення: 05.11.2025).

References

1. Verkhovna Rada of Ukraine. (1995, July 6). Pro zastosuvannia reiestratoriv rozrakhunkovykh operatsii u sferi torhivli, hromadskoho kharchuvannia ta posluh [On the use of registrars of settlement operations in the field of trade, public catering and services] (Law No. 265/95-VR). <https://zakon.rada.gov.ua/laws/show/265/95-вр#Text>.
2. Vchasno. (n.d.). Yak zminyvsya rynek PRRO pislia 1 bereznia [How the PRRO market changed after March 1]. Retrieved November 14, 2025, from <https://vchasno.ua/yak-zminyvsya-rynok-prro-pislya-1-bereznia>.
3. Belova, I., Farion, V., & Yaroshchuk, O. (2023). Features of the implementation of settlement transactions registrars in Ukraine: Accounting and financial aspects. *Ekonomichnyy Analiz*, 2(2), 209–226. <https://doi.org/10.35774/econa2023.02.209>.
4. Lutsenko, A., Vykluk, M., Skoryk, M., & Hromova, T. (2020). Fiscal regulation concept formation of the Ukraine's economy development. *Studies of Applied Economics*, 38(4), Article 3997. <https://doi.org/10.25115/eea.v38i4.3997>.
5. Belhi, A., Gasmi, H., Bouras, A., Aouni, B., & Khalil, I. (2021). Integration of business applications with the blockchain: Odoo and Hyperledger Fabric open source proof of concept. *IFAC-PapersOnLine*, 54(1), 817–824. <https://doi.org/10.1016/j.ifacol.2021.08.185>.
6. Wu, J.-Y., & Chen, S.-L. (2020). Odoo ERP with business intelligence tool for a small-medium enterprise: A scenario case study. In *Proceedings of the 2020 11th International Conference on E-Education, E-Business, E-Management, and E-Learning (IC4E 2020)* (pp. 1–6). Association for Computing Machinery. <https://doi.org/10.1145/3377571.3377607>.
7. Odoo. (n.d.). OWL components. In Odoo documentation 18.0. Retrieved November 10, 2025, from https://www.odoo.com/documentation/18.0/developer/reference/frontend/owl_components.html.
8. Checkbox. (n.d.). API Nakaz № 601 – zminy v cheku [API Order No. 601 – Changes in the receipt]. In *Checkbox Wiki*. Retrieved November 5, 2025, from https://wiki.checkbox.ua/uk/api/nakaz_601#зміни-в-чеку-згідно-наказу-по601.

Дата першого надходження рукопису до видання: 19.11.2025
Дата прийнятого до друку рукопису після рецензування: 12.12.2025
Дата публікації: 31.12.2025

СФЕРА ОБСЛУГОВУВАННЯ

УДК 338.48:004.738.5

DOI <https://doi.org/10.35546/kntu2078-4481.2025.4.3.37>

О. О. ДОЛИНСЬКА

доктор філософії, доцент
Хмельницька гуманітарно-педагогічна академія
ORCID: 0000-0002-7977-5700

І. М. ШОРОБУРА

доктор педагогічних наук, професор
Хмельницька гуманітарно-педагогічна академія
ORCID: 0000-0003-3728-7968

Т. Г. ГІЛЬБЕРГ

кандидат географічних наук, доцент
Хмельницька гуманітарно-педагогічна академія
ORCID: 0000-0001-8746-3515

QR-КОДИ ЯК ІНСТРУМЕНТ ЦИФРОВОГО МАРКЕТИНГУ ТА ІНФОРМАЦІЙНО-НАВІГАЦІЙНОЇ ПІДТРИМКИ В СУЧАСНОМУ ТУРИЗМІ

У статті досліджено роль QR-кодів як ефективного інструменту цифрового маркетингу та інформаційно-навігаційної підтримки в сучасному туризмі. Визначено основні напрями використання QR-технологій у туристичній діяльності, зокрема в екскурсійному обслуговуванні, маркетинговому просуванні дестинацій, інформаційно-навігаційному забезпеченні та гастрономічному туризмі. Показано, що QR-коди виконують комплекс інформаційних, комунікаційних, рекламних та аналітичних функцій, забезпечуючи інтерактивну взаємодію між туристом і туристичним брендом. Особливу увагу приділено безконтактним технологіям, які набули поширення під час пандемії COVID-19 і стали ключовим елементом цифрової інфраструктури туризму. Зазначено, що впровадження QR-кодів сприяє підвищенню ефективності туристичних послуг, розвитку «розумного туризму» та екологічності завдяки зменшенню використання друкованих матеріалів. Проаналізовано міжнародний досвід використання QR-кодів у туристичних практиках, де такі технології активно інтегруються в системи навігації, музейні експозиції та рекламні платформи. Зазначено, що в умовах цифрової трансформації суспільства QR-коди стають важливим елементом розвитку «розумного туризму», орієнтованого на комфорт, мобільність, безпеку та екологічність туристичних подорожей. Вони забезпечують туристам можливість миттєво отримувати актуальні дані про туристичні об'єкти, маршрути, події, бронювання послуг, а також доступ до інтерактивних карт, аудіо- й відеогідів. Використання QR-технологій сприяє підвищенню ефективності маркетингових кампаній, лояльності клієнтів, оптимізації витрат на друковану продукцію та зростанню конкурентоспроможності туристичних підприємств. Наголошено, що для українського туристичного ринку QR-коди стають важливим інструментом цифрової трансформації, що поєднує маркетинг, сервіс і навігацію в єдиний інтерактивний простір. Їх застосування забезпечує туристам доступ до мультимедійного контенту, сприяє підвищенню рівня комфорту, безпеки та залученості користувачів. Використання QR-кодів у туризмі розглядається як стратегічний напрям інноваційного розвитку галузі, орієнтований на потреби сучасного споживача та тенденції глобальної цифровізації.

Ключові слова: QR-коди, цифровий маркетинг, туризм, цифровізація, безконтактні технології.

О. О. DOLYNSKA

PhD, Associate Professor
Khmelnyskyi Humanitarian-Pedagogical Academy
ORCID: 0000-0002-7977-5700

I. M. SHOROBURA

Doctor of Pedagogical Sciences, Professor
Khmelnyskyi Humanitarian-Pedagogical Academy
ORCID: 0000-0003-3728-7968

Т. Н. HILBERH

Candidate of Geographical Sciences, Associate Professor

Khmelnytskyi Humanitarian-Pedagogical Academy

ORCID: 0000-0001-8746-3515

QR-CODES AS A TOOL OF DIGITAL MARKETING AND INFORMATION-NAVIGATIONAL SUPPORT IN MODERN TOURISM

The article examines the role of QR-codes as an effective tool of digital marketing and information-navigation support in contemporary tourism. The main directions of QR technology application in tourism activities are identified, including excursion services, destination marketing, information and navigation systems, and gastronomic tourism. It is demonstrated that QR-codes perform a complex set of informational, communicative, advertising, and analytical functions, ensuring interactive interaction between tourists and tourism brands. Special attention is paid to contactless technologies that became widespread during the COVID-19 pandemic and have since become a key component of the digital tourism infrastructure. It is noted that the implementation of QR-codes contributes to improving the efficiency of tourism services, fostering the development of "smart tourism," and promoting environmental sustainability by reducing the use of printed materials. The study analyzes international practices of QR-code implementation in tourism across the EU, the USA, and Asia, where such technologies are actively integrated into navigation systems, museum exhibitions, and advertising platforms. It is emphasized that under the conditions of digital transformation, QR-codes are becoming an essential element in the development of "smart tourism," oriented toward comfort, mobility, safety, and environmental responsibility. QR-codes provide tourists with instant access to up-to-date information about attractions, routes, events, and service bookings, as well as to interactive maps, audio, and video guides. The use of QR technologies enhances the effectiveness of marketing campaigns, customer loyalty, cost optimization in printed materials, and the competitiveness of tourism enterprises. It is underlined that for the Ukrainian tourism market, QR-codes represent an important instrument of digital transformation that integrates marketing, service, and navigation into a unified interactive environment. Their application provides tourists with access to multimedia content and contributes to increased comfort, safety, and user engagement. The use of QR-codes in tourism is considered a strategic direction for the innovative development of the industry, aligned with the needs of modern consumers and global digitalization trends.

Key words: QR-codes, digital marketing, tourism, digitalization, contactless technologies.

Постановка проблеми

Сучасний етап розвитку туристичної індустрії характеризується активною цифровою трансформацією, що змінює традиційні підходи до організації, просування та споживання туристичних послуг. Одним із ключових інструментів цієї трансформації є QR-коди, які забезпечують миттєвий доступ до інформації, сприяють інтерактивності комунікації між туристом і інформацією про дестинацію, а також формують нову модель поведінки споживача в цифровому середовищі. В умовах глобальної конкуренції та зростання вимог до якості туристичного продукту впровадження QR-технологій стає важливим чинником підвищення ефективності маркетингових стратегій, інформаційно-навігаційного забезпечення та безпеки туристів. Попри широке розповсюдження QR-кодів у світовій практиці, в Україні їх потенціал у туризмі використовується недостатньо системно. Відсутні цілісні методичні підходи до інтеграції QR-технологій у діяльність туристичних підприємств і дестинацій, не сформовано єдину концепцію їх застосування в маркетингових, екскурсійних та навігаційних цілях. Більшість ініціатив мають локальний характер і не забезпечують комплексного ефекту цифровізації туристичної інфраструктури. Крім того, залишається нерозкритим питання оцінювання ефективності QR-кодів як інструменту цифрового маркетингу – зокрема, їхнього впливу на поведінку споживачів, рівень залученості туристів та формування позитивного іміджу дестинацій. Недостатньо вивченими є аспекти інтеграції QR-кодів із іншими елементами смарт-туризму, такими як доповнена реальність, геолокаційні сервіси, мобільні додатки та системи аналітики даних.

Актуальність дослідження зумовлена необхідністю теоретичного обґрунтування та практичного впровадження QR-кодів як ефективного інструменту цифрового маркетингу й інформаційно-навігаційної підтримки в туризмі. Вирішення цієї проблеми сприятиме підвищенню конкурентоспроможності туристичних підприємств, удосконаленню сервісу, розширенню можливостей комунікації з туристами та розвитку інноваційної туристичної інфраструктури в Україні відповідно до світових тенденцій.

Аналіз останніх досліджень і публікацій

Наукові праці, присвячені впровадженню маркетингових інструментів у діяльність туристичних підприємств та застосуванню цифрового маркетингу задля підвищення ефективності функціонування туристичного бізнесу, представлено у працях М. Барни [1], М. Мальської [6], І. Мельник [1] та інших учених. Дослідники наголошують на необхідності модернізації традиційних підходів до просування туристичних продуктів відповідно до вимог цифрової доби.

Водночас питання цифровізації туристичної сфери та застосування сучасних технологічних рішень знаходять відображення не лише у працях вітчизняних науковців, але й у міжнародних аналітичних звітах. У доповідях

Всесвітньої туристичної організації (UNWTO) [7] та матеріалах Державного агентства розвитку туризму України (ДАРТ) [2] акцентується увага на важливості використання цифрових інновацій як ключового чинника сталого розвитку туристичної галузі, підвищення якості послуг, ефективності управління туристичними DESTINATIONAMI та забезпечення їхньої конкурентоспроможності на глобальному ринку.

Формулювання мети дослідження

Метою дослідження є теоретичне обґрунтування та аналіз можливостей використання QR-кодів як ефективного інструменту цифрового маркетингу й інформаційно-навігаційної підтримки у сучасному туризмі, а також виявлення напрямів їх практичного застосування для підвищення якості туристичних послуг, конкурентоспроможності DESTINATIONAMI та рівня цифровізації туристичної інфраструктури.

Викладення основного матеріалу дослідження

У сучасних умовах цифрової трансформації туристичної галузі QR-коди стали ефективним інструментом цифрового маркетингу, який поєднує зручність, мобільність та інтерактивність. Вони забезпечують швидкий доступ до інформації про туристичні продукти, послуги чи події, сприяють зростанню лояльності споживачів і підвищенню конкурентоспроможності туристичних підприємств. QR-коди в туризмі – це сучасний інструмент цифрової комунікації, який активно використовується для підвищення якості туристичних послуг, зручності подорожей та інтерактивної взаємодії з туристами [5].

QR-коди (Quick Response Codes) – це двовимірні штрих-коди, що містять зашифровану інформацію, яку можна швидко зчитати за допомогою камери смартфона або спеціального сканера.

QR-коди можуть містити текст, посилання на вебсайти, контакти, геолокаційні дані, зображення, аудіо чи відеофайли. Їхня головна перевага – швидкість доступу до інформації без необхідності ручного введення даних.

З маркетингової точки зору, QR-коди виконують кілька ключових функцій:

1. Інформаційна – дозволяють туристам миттєво отримати детальні дані про локацію, маршрут, подію або акційну пропозицію.

2. Комунікативна – сприяють налагодженню прямої взаємодії між туристом і туристичним брендом, зокрема через посилання на офіційні сайти, соціальні мережі чи відеопрезентації.

3. Рекламна – QR-коди розміщуються на буклетах, білбордах, сувенірній продукції, що підвищує охоплення цільової аудиторії без значних витрат.

4. Аналітична – за допомогою систем аналітики можна відстежувати кількість сканувань, географію користувачів та ефективність рекламних кампаній.

Основні напрями використання QR-кодів у туризмі:

Екскурсійна діяльність. QR-коди встановлюються біля історичних пам'яток, музеїв, архітектурних об'єктів. Турист, відсканувавши код, отримує інформацію про об'єкт, аудіогід, відеоматеріали чи віртуальний тур. Це створює інтерактивну форму самостійного пізнання і розширює можливості екскурсій без участі гідів. У сучасних умовах цифрової трансформації туристичної сфери QR-коди стають одним із найефективніших інструментів для підвищення якості та інтерактивності екскурсійної діяльності. Вони поєднують у собі інформаційні, комунікаційні та маркетингові функції, дозволяючи екскурсантам отримувати додаткові знання у зручній, динамічній та технологічно доступній формі.

Інформаційна функція QR-кодів полягає у забезпеченні миттєвого доступу до текстових, аудіо- чи відеоматеріалів, що розкривають зміст екскурсійного об'єкта. За допомогою смартфона відвідувач може відсканувати код, розміщений біля історичної пам'ятки, музею або експозиції, і дізнатися більше про її історію, архітектуру, культурну цінність. Це значно розширює інформаційні можливості екскурсовода та надає туристу свободу у виборі темпу сприйняття матеріалу.

Інноваційний аспект використання QR-кодів полягає в інтеграції технологій доповненої реальності, 3D-моделей, аудіогідів та інтерактивних карт. У результаті екскурсія набуває мультимедійного формату, що особливо привабливо для молодшої аудиторії, яка цінує інтерактивність і самостійність у пізнавальному процесі. Маркетингова цінність QR-кодів полягає у можливості просування туристичних продуктів і брендів DESTINATIONAMI. Зокрема, за допомогою QR-кодів можна направити користувача на офіційний сайт туристичної організації, сторінки у соціальних мережах або онлайн-магазини сувенірів. Це створює додаткові канали комунікації між туристом і туристичним підприємством, сприяє підвищенню лояльності та зацікавленості відвідувачів [3].

Переваги для екскурсоводів та організаторів турів також очевидні. QR-коди дозволяють оптимізувати структуру екскурсії, скоротити час на пояснення базових фактів, зосереджуючись на емоційному та аналітичному аспектах подачі інформації. Крім того, вони є економічно вигідними: створення та розміщення кодів потребує мінімальних витрат у порівнянні з друкованими матеріалами.

З позицій сталого розвитку туризму, використання QR-кодів сприяє екологічності екскурсійного процесу – зменшується потреба у паперових буклетах, мапах та проспектах. Це відповідає сучасним екологічним тенденціям та формує позитивний імідж DESTINATIONAMI як «зеленої» та інноваційної.

QR-коди виступають важливим елементом сучасної екскурсії, який поєднує технологічність, зручність і пізнавальну насиченість. Їх використання не лише підвищує ефективність екскурсійного обслуговування, а й формує нову культуру сприйняття туристичного продукту, орієнтовану на інтерактивність, самостійність і цифрову грамотність туристів.

Інформаційно-навігаційна підтримка. У сучасному туристичному середовищі інформаційно-навігаційна підтримка є невід'ємною складовою комфортного та безпечного пересування туристів [4]. Одним із найефективніших інструментів у цьому напрямі стали QR-коди, які забезпечують швидкий доступ до актуальної інформації, значно спрощують орієнтацію в просторі та підвищують якість туристичного обслуговування. QR-коди розміщуються на туристичних картах, інформаційних стендах, транспорті, зупинках громадського транспорту, у готелях, ресторанах, музеях, біля пам'яток архітектури та природних об'єктів. Сканування коду за допомогою смартфона дає змогу туристу миттєво отримати доступ до офіційних сайтів туристичних DESTINACIЙ, інтерактивних карт, сервісів бронювання квитків та проживання, розкладів подій, екскурсійних маршрутів або довідкової інформації про об'єкт.

Такі технології сприяють цифровізації туристичної інфраструктури, адже дозволяють зменшити потребу у друкованих матеріалах, зробити інформацію динамічною, мультимедійною та багатомовною. Наприклад, за допомогою QR-коду турист може отримати аудіогід кількома мовами, переглянути 3D-модель історичного об'єкта, або перейти до віртуального туру. Це не лише полегшує навігацію, а й робить процес пізнання більш інтерактивним та захопливим.

Важливим аспектом є й те, що QR-коди сприяють розвитку розумних туристичних міст, де поєднуються цифрові технології, аналітика даних і мобільні сервіси для створення зручного та персоналізованого туристичного досвіду. Такі інструменти дозволяють туристам самостійно планувати маршрути, отримувати інформацію про рівень завантаженості локацій, користуватися навчальними або ігровими елементами гейміфікації, що підвищує їхню залученість.

QR-коди як елемент інформаційно-навігаційної підтримки відіграють важливу роль у модернізації туристичної інфраструктури, формують позитивний імідж DESTINACIЙ, підвищують доступність інформації та сприяють створенню комфортного туристичного середовища. Їх використання є проявом інноваційного підходу до розвитку туризму, орієнтованого на потреби сучасного споживача та тенденції цифрової трансформації.

Маркетинг і просування туристичних DESTINACIЙ. У сучасних умовах цифрової трансформації маркетинг туристичних DESTINACIЙ набуває нових форм і підходів. Одним із інноваційних інструментів, що активно використовується для просування туристичних продуктів і формування позитивного іміджу територій, є QR-коди. Їх застосування дає змогу поєднати традиційні та цифрові канали комунікації, забезпечуючи туристам швидкий доступ до актуальної інформації.

Туристичні організації, органи місцевого самоврядування та оператори туристичного бізнесу розміщують QR-коди на рекламних носіях.

Використання QR-кодів у маркетингових стратегіях сприяє підвищенню ефективності просування туристичних DESTINACIЙ, адже дозволяє: забезпечити доступ до мультимедійного контенту; оперативно оновлювати інформацію без потреби перевидання друкованих матеріалів; збирати аналітичні дані про кількість сканувань, географію користувачів, рівень зацікавленості; інтегрувати маркетингові кампанії в соціальні медіа та системи лояльності; персоналізувати комунікацію з туристом.

Таким чином, QR-коди виступають інструментом інтерактивного маркетингу, що поєднує інформаційну, рекламну та навігаційну функції. Їх застосування не лише підвищує конкурентоспроможність туристичних DESTINACIЙ, а й сприяє розвитку «розумного туризму», орієнтованого на використання цифрових технологій для створення комфортного, інформативного та безпечного туристичного середовища.

Гастрономічний туризм. У ресторанах QR-коди ведуть до меню, інформації про місцеві продукти, історію страв чи навіть відеорецептів. Це підсилює автентичність і сприяє формуванню культурного досвіду туриста.

Музейні та тематичні маршрути. Використання QR-кодів у рамках тематичних екскурсій дає змогу створювати «розумні маршрути» – турист може рухатися за інтерактивними підказками, відкриваючи нові локації та контент на своєму пристрої.

Пандемія COVID-19 стала потужним каталізатором впровадження безконтактних технологій у різних сферах суспільного життя, зокрема й у туризмі. Одним із найпоширеніших інструментів, що забезпечують безпечну взаємодію між туристами та представниками туристичної індустрії, стали QR-коди. Вони дозволяють здійснювати широкий спектр операцій без необхідності фізичного контакту, що значно підвищує рівень гігієнічної безпеки та довіри з боку споживачів. QR-коди активно використовуються для електронної ідентифікації туристів, реєстрації у готелях, проходження онлайн-чекінів, отримання віртуальних квитків до музеїв, виставок, концертів, тематичних парків чи інших туристичних об'єктів. Такий підхід мінімізує ризики передачі інфекцій через паперові документи або фізичні точки продажу. В умовах зростання вимог до безпеки подібні технології стали невід'ємним елементом цифрової інфраструктури туризму.

Особливого поширення набули безконтактні платежі, що реалізуються за допомогою QR-кодів, мобільних додатків або електронних гарантії. Це не лише зручно, але й сприяє прозорості фінансових операцій, зниженню витрат на обробку готівки та підвищенню ефективності роботи підприємств туристичної сфери.

Важливою перевагою безконтактних технологій є також підвищення рівня кібербезпеки. Сучасні системи QR-аутентифікації мають багаторівневий захист, що дозволяє уникнути підробок, несанкціонованого доступу або втрати персональних даних. При цьому користувачі отримують можливість самостійно контролювати процес передачі інформації, що формує додаткове відчуття захищеності.

Крім безпеки, використання безконтактних технологій сприяє покращенню туристичного досвіду. Завдяки QR-кодам туристи можуть отримувати інтерактивну інформацію про об'єкти, завантажувати аудіогід, переглядати віртуальні експозиції чи бронювати квитки в один дотик. Це створює зручне, сучасне, екологічно безпечне середовище для подорожей і стимулює розвиток цифрового туризму як складової сталого розвитку індустрії.

Безконтактні технології, зокрема використання QR-кодів, стали не просто тимчасовою відповіддю на виклики пандемії, а сталим трендом цифрової трансформації туризму, що поєднує комфорт, безпеку та інноваційність у взаємодії між туристом і дестинацією.

Сучасний розвиток цифрових технологій створює передумови для активного впровадження QR-кодів у різні сегменти туристичної діяльності. QR-коди виступають ефективним інструментом цифровізації туристичної інфраструктури, забезпечуючи оперативність і зручність доступу до інформації, знижуючи витрати на друк поліграфічних матеріалів, підвищуючи інтерактивність та залученість туристів, а також дозволяючи здійснювати аналітику поведінки відвідувачів на основі статистики сканувань.

Інтеграція QR-технологій у туристичну сферу є складовою концепції «розумного туризму», що сприяє підвищенню доступності інформації, зручності комунікації та безпеки подорожей. У країнах Європейського Союзу QR-коди стають елементом цифрової туристичної ідентифікації територій. Так, у Парижі (Франція) QR-коди розміщуються біля визначних пам'яток та на туристичних картах, спрямовуючи користувача на офіційні ресурси з віртуальними турами та історичними довідками. У Римі (Італія) QR-коди забезпечують мультимедійне супроводження туристів, надаючи відео, описи та маршрути історичних об'єктів після сканування коду. У Празі (Чехія) QR-коди інтегровані в систему навігації старого міста, дозволяючи завантажувати карти, аудіогід та розклад культурних подій. У Іспанії використовують QR-коди на сувенірах, квитках та рекламних матеріалах для переходу на офіційні туристичні портали міста.

Азійські країни демонструють високий рівень впровадження QR-технологій у туристичній сфері. У Японії QR-коди застосовуються в музеях, храмах, на вокзалах і зупинках громадського транспорту, забезпечуючи миттєвий доступ до багатомовної інформації про об'єкт, його історію та культурне значення.

США (Нью-Йорк) вони застосовуються в музеях (Metropolitan Museum of Art, MoMA) для перегляду додаткових матеріалів, онлайн-каталогів та відеогідів. У національних парках QR-коди розміщуються на туристичних стежках та в інформаційних центрах, надаючи карти, правила безпеки, рекомендації щодо маршрутів і відомості про природні об'єкти.

Таким чином, використання QR-кодів у туризмі відображає глобальну тенденцію цифровізації туристичного простору. Вони сприяють формуванню інтерактивного туристичного досвіду, підвищенню доступності інформації, зниженню витрат на друк матеріалів і підвищенню привабливості туристичних дестинацій. Українська практика демонструє поступову інтеграцію QR-кодів у міські туристичні системи, що відповідає світовим трендам розвитку інноваційних технологій у сфері туризму.

Висновки

Сучасний розвиток туристичної індустрії відбувається в умовах глибокої цифрової трансформації, що зумовлює появу нових підходів до організації, просування та споживання туристичних послуг. У цьому контексті QR-коди виступають одним із найефективніших інструментів цифрового маркетингу, інформаційної підтримки та навігації, здатних забезпечити інтерактивну взаємодію між туристом і дестинацією, підвищити рівень задоволеності споживачів і конкурентоспроможність туристичних підприємств.

Застосування QR-кодів у туризмі сприяє підвищенню ефективності маркетингових стратегій та просуванню туристичних дестинацій; удосконаленню екскурсійного обслуговування через інтеграцію мультимедійного контенту, аудіогідів та доповненої реальності; розвитку системи інформаційно-навігаційної підтримки, що забезпечує комфорт і безпеку пересування туристів; популяризації безконтактних технологій, які відповідають вимогам постпандемічного періоду та принципам сталого розвитку; цифровізації туристичної інфраструктури та створенню передумов для розвитку «розумного туризму». Зарубіжний досвід (Франції, Італії, Чехії, Іспанії, Японії, США) доводить ефективність використання QR-кодів для підвищення інформативності, доступності та привабливості туристичних дестинацій. Водночас в Україні застосування цих технологій поки що має епізодичний характер і потребує системного підходу, зокрема розроблення єдиної концепції інтеграції QR-технологій у сферу туризму, методики оцінювання їх ефективності та адаптації міжнародного досвіду до національних умов. QR-коди слід розглядати не лише як технічний інструмент комунікації, а як складову інноваційної екосистеми туризму, що

поєднує маркетинг, навігацію, інформаційну безпеку й екологічність. Їх широке впровадження здатне суттєво підвищити якість туристичних послуг, посилити конкурентні переваги українських дестинацій і сприяти формуванню позитивного іміджу України як сучасної, цифрово-орієнтованої туристичної держави.

Перспективними напрямками подальших досліджень є: розроблення методичних засад впровадження QR-кодів у маркетингову діяльність туристичних підприємств; аналіз впливу QR-технологій на поведінку туристів і рівень їхньої залученості; вивчення можливостей інтеграції QR-кодів із технологіями доповненої реальності, геолокаційними сервісами та аналітичними системами смарт-туризму.

Список використаної літератури

1. Барна М. Ю., Мельник І. М. Стратегія цифровізації готельно-ресторанного бізнесу. Економіка та суспільство. 2025. № 71. <https://doi.org/10.32782/2524-0072/2025-71-66>
2. Державне агентство розвитку туризму. URL: <https://www.tourism.gov.ua/> (дата звернення: 18.10.2025).
3. Долинська О. О. Сучасні тенденції організації туристичних подорожей у контексті цифрової трансформації та сталого розвитку. Економіка та суспільство. 2025. № 77. DOI: <https://doi.org/10.32782/2524-0072/2025-77-15>
4. Долинська О. О., Шоробура І. М., Гільберг Т. Г. Роль реклами у просуванні туристичного продукту. Індустрія туризму і гостинності в Центральній та Східній Європі. Львів. 2025. № 13. С. 11-17.
5. Долинська О. О., Шоробура І. М., Біницька О. П. Інновації в туризмі. Наукові записки Тернопільського національного педагогічного університету імені Володимира Гнатюка. Сер. Географія. Тернопіль. 2023. № 2. (55). С. 58-64.
6. Мальська М. П., Мандюк Н. Л. Основи маркетингу у туризмі. Київ: «Центр учбової літератури», 2016. 336 с.
7. UNWTO. Офіційний сайт. URL: <https://www.unwto.org>. (дата звернення: 18.10.2025).

References

1. Barna, M. Yu., & Melnyk, I. M. (2025). Stratehiia tsyfrovizatsii hotelno-restorannoho biznesu [Strategy of digitalization of the hotel and restaurant business]. Ekonomika ta suspilstvo [Economy and Society], (71). <https://doi.org/10.32782/2524-0072/2025-71-66> [in Ukrainian].
2. Derzhavne ahentstvo rozvytku turyzmu [State Agency for Tourism Development]. (2025). Retrieved October 18, 2025, from <https://www.tourism.gov.ua/> [in Ukrainian].
3. Dolynska, O. O. (2025). Suchasni tendentsii orhanizatsii turystychnykh podorozhei u konteksti tsyfrovoy transformatsii ta staloho rozvytku [Modern trends in organizing tourist trips in the context of digital transformation and sustainable development]. Ekonomika ta suspilstvo [Economy and Society], (77). <https://doi.org/10.32782/2524-0072/2025-77-15> [in Ukrainian].
4. Dolynska, O. O., Shorobura, I. M., & Hilberh, T. H. (2025). Rol reklamy u prosuvanni turystychnoho produktu [The role of advertising in promoting a tourist product]. Industriia turyzmu i hostynnosti v Tsentralnii ta Skhidnii Yevropi [Tourism and Hospitality Industry in Central and Eastern Europe], (13), 11–17. [in Ukrainian].
5. Dolynska, O. O., Shorobura, I. M., & Binytska, O. P. (2023). Innovatsii v turyzmi [Innovations in tourism]. Naukovi zapysky Ternopilskoho natsionalnoho pedahohichnoho universytetu imeni Volodymyra Hnatiuka. Seriya: Neohrafiia [Scientific Notes of Ternopil Volodymyr Hnatiuk National Pedagogical University. Geography Series], (2[55]), 58–64. [in Ukrainian].
6. Malska, M. P., & Mandiuk, N. L. (2016). Osnovy marketynhu u turyzmi [Fundamentals of marketing in tourism]. Kyiv: Tsentr uchbovoi literatury. [in Ukrainian].
7. UNWTO. (2025). Ofitsiyni sait [Official website]. Retrieved October 18, 2025, from <https://www.unwto.org> [in English].

Дата першого надходження рукопису до видання: 23.11.2025
Дата прийнятого до друку рукопису після рецензування: 12.12.2025
Дата публікації: 31.12.2025

В. М. ПОДОЛЯК

кандидат технічних наук, доцент,
доцент кафедри туризму та готельно-ресторанної справи
Луцький національний технічний університет
ORCID: 0000-0002-8521-1258

ПРОЄКТУВАННЯ ІННОВАЦІЙНОГО КОМПЛЕКСУ УСТАТКУВАННЯ МОБІЛЬНОЇ КАВ'ЯРНІ З УРАХУВАННЯМ РЕСУРСНОЇ АВТОНОМНОСТІ

У статті висвітлено комплексний підхід до проєктування інноваційної мобільної кав'ярні, ключовою особливістю якої є повна ресурсна автономність завдяки використанню відновлюваних джерел енергії. Актуальність дослідження обумовлена стабільним зростанням ринку кав'ярень в Україні на тлі підвищення попиту на екологічно чисті послуги, а також необхідністю вирішення гострої проблеми залежності мобільних точок продажу від стаціонарних електромереж або використання некомфортних гучних паливних генераторів.

Метою роботи є розробка та детальне техніко-економічне обґрунтування бізнес-моделі автономного закладу, що може функціонувати у рекреаційних зонах без шкоди для довкілля. У ході дослідження автором проведено аналіз ринкових тенденцій та виконано стратегічний SWOT-аналіз, який дозволив виявити сильні сторони проєкту (автономність, мобільність, екологічність) та нівелювати загрози, які виникають в процесі роботи. На основі цього сформовано концепцію закладу, що включає ергономічне зонування обмеженого простору трейлера, «Lean»-підхід до організації робочого місця бариста та маркетингову стратегію з унікальною торговою пропозицією.

Центральне місце в роботі займає технічне обґрунтування системи автономного енергозабезпечення. Запропоновано та розраховано гібридну модель енергозабезпечення, що поєднує сонячні панелі, вітрогенератори та системи збереження енергії. Розрахунки енергобалансу доводять, що система здатна генерувати близько 42,0 кВт·год/день при піковому споживанні 13,9 кВт·год/день, створюючи значний резерв потужності для безперебійної роботи закладу.

Економічна частина дослідження містить калькуляцію бюджету проєкту (початкові інвестиції – 1,2 млн грн) та прогноз фінансових результатів. Розрахунковий термін окупності у 6–7 місяців підтверджує високу інвестиційну привабливість та рентабельність розробленої моделі в сучасних економічних умовах.

Ключові слова: мобільна кав'ярня, проєктування, відновлювані джерела енергії, устаткування кав'ярень, техніко-економічне обґрунтування, бізнес-модель, розрахунок окупності, маркетингова стратегія.

V. M. PODOLIAK

Candidate of Technical Sciences,
Associate Professor at the Department of Tourism and Hotel
and Restaurant Business
Lutsk National Technical University
ORCID: 0000-0002-8521-1258

DESIGNING AN INNOVATIVE EQUIPMENT COMPLEX FOR A MOBILE COFFEE SHOP CONSIDERING RESOURCE AUTONOMY

The article presents a comprehensive approach to designing an innovative mobile coffee shop, the key feature of which is full resource autonomy achieved through the use of renewable energy sources. The relevance of the study is driven by the steady growth of the coffee shop market in Ukraine amidst increasing demand for eco-friendly services, as well as the need to solve the acute problem of mobile points of sale depending on stationary power grids or using noisy fuel generators that cause discomfort.

The aim of the work is the development and detailed techno-economic justification of a business model for an autonomous establishment capable of operating in recreational areas without harming the environment. The study involves an analysis of market trends and a strategic SWOT analysis, which allowed identifying the project's strengths (autonomy, mobility, eco-friendliness) and mitigating threats that arise during operation. Based on this, a concept for the establishment was formed, including ergonomic zoning of the trailer's limited space, a "Lean" approach to organizing the barista's workplace, and a marketing strategy with a unique selling proposition.

The central focus of the work is the technical justification of the autonomous energy supply system. A hybrid energy supply model combining solar panels, wind generators, and energy storage systems is proposed and calculated. Energy

balance calculations demonstrate that the system is capable of generating approximately 42.0 kWh/day with a peak consumption of 13.9 kWh/day, creating a significant power reserve for the establishment's uninterrupted operation.

The economic part of the research contains a calculation of the project budget (initial investments – 1.2 million UAH) and a forecast of financial results. The calculated payback period of 6–7 months confirms the high investment attractiveness and profitability of the developed model in current economic conditions.

Key words: mobile coffee shop, design, renewable energy sources, coffee shop equipment, feasibility study, business model, payback calculation, marketing strategy.

Постановка проблеми

На даному етапі розвитку економіки сфера харчування, зокрема ресторани й кафе, демонструє стабільне зростання. Основною рушійною силою цього зростання є попит на каву та похідні від неї напої. Сучасні тенденції в підході до проектування закладів такого типу, включають не лише різноманітність дизайну закладів та унікальний досвід, але й поступове впровадження екологічних форматів при підборі устаткування та мобільність в обслуговуванні. Ринок кав'ярень у містах та зонах відпочинку, стабільно збільшується. Спостерігається значний попит на каву «з собою» та на заклади з особливим дизайном.

На даний час традиційні мобільні кав'ярні стикаються з наступними проблемами: логістичною залежністю від під'єднання до стаціонарних електромереж та екологічною – через використання гучних бензинових чи дизельних генераторів. Це створює перепони для розміщення таких закладів в екологічно чистих зонах (парки, рекреації) та негативно впливає на імідж. Постає проблема розробки такої бізнес-моделі мобільного закладу, яка б поєднувала високу мобільність, якість продукту та повну енергетичну й екологічну автономність.

Аналіз останніх досліджень і публікацій

Теоретичною основою дослідження стали праці вітчизняних та закордонних фахівців у сферах маркетингу, сталого розвитку та ресторанного бізнесу. Зокрема, питання загальних маркетингових стратегій для підприємств малого бізнесу висвітлено у працях Н. Савич та А. Михайлова [1]. Аспекти впровадження екологічних ініціатив у бізнес-процеси та теоретичні засади еко-маркетингу розглядаються І. Петровою та С. Гончаруком [2], що є важливим для обґрунтування екологічної складової в процесі проектування. Сучасні тенденції кавової культури в Україні, що підтверджують актуальність обраного напрямку, аналізує Ю. Ковальчук [3]. Окрему увагу в літературі приділено сучасним інструментам просування, таким як SMM [4], інфлюенс-маркетинг [7] та програми лояльності [8], які є невід'ємною частиною не лише проектного частини, але й маркетингового плану проєкту. Міжнародні асоціації та бізнес-видання [6] підтверджують глобальний тренд на сталий розвиток.

Однак, незважаючи на значну кількість публікацій, присвячених окремим аспектам (маркетингу, екології, бізнес-плануванню), виявлено суттєву наукову прогалину. Відсутні комплексні дослідження, що пропонують *цілісну техніко-економічну модель* мобільного закладу на відновлюваних джерелах енергії. Більшість існуючих праць не поєднують технічні розрахунки енергетичного балансу з економічними показниками окупності та маркетинговою стратегією. Таким чином, проблема розробки науково обґрунтованої, життєздатної бізнес-моделі автономної мобільної кав'ярні залишається недостатньо вирішеною та потребує детального обґрунтування.

Формування мети дослідження

Метою дослідження є розробка та комплексне техніко-економічне обґрунтування бізнес-моделі автономної мобільної кав'ярні, що функціонує на базі відновлюваних джерел енергії.

Викладення основного матеріалу дослідження

Досягнення поставленої мети вимагає послідовного вирішення кількох завдань: розробки концептуальної моделі закладу, проведення стратегічного аналізу, технічного обґрунтування енергосистеми та економічного розрахунку ефективності проєкту.

1. Концептуальна модель та брендинг.

Основою проєкту є створення екологічно чистого бізнесу – мобільної кав'ярні, що працює виключно на відновлюваних джерелах енергії (сонячні панелі, вітрогенератори та системи збереження енергії). Це рішення є не лише відповіддю на зростаючий споживчий попит на екологічно свідомі підходи в роботі мобільних закладів, але й вирішує ключову операційну проблему традиційних закладів такого типу – залежність від стаціонарних електромереж та гучних неекологічних генеруючих системах (бензинових/дизельних генераторів). Таким чином, ресурсна автономність об'єкту стає головною конкурентною перевагою та унікальною торговою пропозицією (УТП) проєкту, дозволяючи розмішувати заклад у локаціях без доступу до електромереж (парки, фестивалі, рекреаційні зони) з мінімальним впливом на довкілля.

Концепція проєкту знаходить своє пряме матеріальне втілення в дизайні інтер'єру, розробленому з урахуванням ергономічності, екологічності та технологічності (рис. 1). Це створює цілісний клієнтський досвід, де фізичне оточення (т.зв. «servicescape») підсилює та доводить бренд-повідомлення про інноваційність та сталий розвиток.



Рис. 1. Інтер'єр автономної мобільної кав'ярні

В умовах обмеженого простору мобільного трейлера, пропонується чітке просторове зонування, оптимізоване за принципами "Lean" (ощадливого виробництва):

- *Робоча зона бариста*: Оптимальне розміщення професійного кавового обладнання, холодильних установок та зон зберігання. Проектування зони зосереджене на мінімізації рухів бариста та максимальній швидкості обслуговування, що є критичним показником ефективності для формату «to-go».
- *Зона обслуговування*: Компактна барна стійка з інтегрованою POS-системою. Вона слугує не лише для фінансових операцій, але і як центр збору даних (CRM) та управління запасами в реальному часі.
- *Зона комфорту*: Передбачено мобільні відкидні стійки для споживання напоїв на місці. Це рішення підкреслює пріоритет моделі "grab-and-go", водночас надаючи базову опцію для короткочасного перебування клієнтів.

В оформленні домінують натуральні та екологічно чисті матеріали (оброблене дерево, метал). Проектна пропозиція полягає у використанні тактильно приємних поверхонь, що створює сенсорний досвід, який вигідно відрізняє заклад від стандартизованих пластикових фудтраків. Колірна гама (зелений, коричневий, бежевий) виконує функцію невербальної комунікації: зелений акцентує увагу на екологічності; коричневий та бежевий асоціюються з кавою, теплом та затишком. Освітлення забезпечується виключно енергоощадними LED-лампами з світловою гамою максимально наближеною до природного спектру.

Брендинг проекту підкреслює його інноваційність. Центральним елементом є логотип (рис. 2), що слугує візуальним абстрактом усієї бізнес-моделі – кавовий фудтрак із сонячними панелями, вітрогенератором та системами накопичення енергії. Це дозволяє миттєво ідентифікувати УТП закладу навіть при побіжному погляді, що є важливим для мобільної точки торгівлі.

2. Стратегічний аналіз та маркетингове позиціонування.

Стратегічне планування проекту вимагає глибокого аналізу його внутрішнього потенціалу та зовнішніх ринкових умов. Для цього було проведено SWOT-аналіз, що став базисом для формування маркетингової стратегії та позиціонування.

Проведення SWOT-аналізу дозволяє здійснити систематизація сильних (S), слабких (W) сторін, можливостей (O) та загроз (T) проекту (представлена в табл. 1) та не лише констатувати фактори, але й виявити їхній синергетичний ефект та сформулювати ключові стратегічні напрями.

Простий перелік факторів SWOT є недостатнім; ключовим є їх перехресний аналіз для формування проактивних стратегій.

Стратегії S-O (Сила-Можливість). Це найбільш пріоритетний напрям «агресивного розвитку», що використовує переваги проекту для завоювання ринку.

- (S1, S2, S3 + O2, O3). Використання повної автономності та екологічності для укладання ексклюзивних контрактів з організаторами фестивалів, еко-заходів та парковими адміністраціями. Позиціонування як єдиного «зеленого» кейтерингу, що є перевагою для організаторів, які дбають про свій імідж.

- (S4, S5 + O1). Створення привабливої франшизи. Інноваційний імідж (S4) та доведені низькі операційні витрати (S5) є потужним аргументом для потенційних франчайзі.



Рис. 2. Концепція логотипу автономної мобільної кав'ярні

Таблиця 1

SWOT-аналіз проєкту мобільної автономної кав'ярні

Сильні сторони (S)	Слабкі сторони (W)	Можливості (O)	Загрози (T)
S1. Повна ресурсна автономність	W1. Висока початкова вартість обладнання	O1. Франчайзинг та масштабування проєкту	T1. Висока конкуренція з традиційними кав'ярнями
S2. Екологічність та енергоефективність (УТП)	W2. Обмежена пропускна здатність (об'єм виробництва)	O2. Співпраця з великими фестивалями, івентами, ярмарками	T2. Пряма залежність генерації енергії від погодних умов
S3. Висока мобільність та гнучкість у виборі локації	W3. Потенційна складність в обслуговуванні інноваційного обладнання	O3. Зростання ринкового попиту на екологічні ініціативи та "зелений" бізнес	T3. Юридичні та бюрократичні ризики (отримання дозволів на торгівлю)
S4. Сучасний дизайн та інноваційний імідж	W4. Обмежена потужність енергосистеми в пікові години	O4. Розширення асортименту (холодні напої, десерти)	T4. Зростання цін на якісну кавову сировину
S5. Мінімальні операційні витрати (без оренди та комунальних платежів)			

Стратегії W-O (Слабкість-Можливість). Використання ринкових можливостей для подолання внутрішніх недоліків.

– ($W1 + O3$). Компенсація високої початкової вартості (W1) через маркетингове використання тренду на екологічність (O3). Висока ціна входу в проєкт обґрунтовується преміальним «зеленим» позиціонуванням, що дозволяє встановлювати вищу кінцеву ціну на продукт, перетворюючи витрати на інвестицію в бренд.

Стратегії S-T (Сила-Загроза). Використання сильних сторін для нейтралізації зовнішніх загроз.

– ($S3 + T1$). Нейтралізація високої конкуренції (T1) за рахунок унікальної мобільності (S3). Проєкт не конкурує «лоб в лоб» зі стаціонарними закладами, а створює власні «блакитні океани» попиту в локаціях, фізично недоступних для конкурентів (набережні, віддалені куточки парків, тимчасові заходи).

– ($S2, S4 + T4$). Зниження загрози від зростання цін на сировину (T4) шляхом сильного бренду (S2, S4). Клієнти платять не стільки за каву, скільки за належність до «еко-спільноти» та інноваційний досвід, що підвищує їхню лояльність та знижує цінову еластичність попиту.

Стратегії W-T (Слабкість-Загроза). Це захисні стратегії, спрямовані на виживання.

– ($W4 + T2$). Усвідомлення комбінованої загрози (погана погода ($T2$) знижує генерацію енергії, водночас пікові навантаження ($W4$) можуть вимкнути систему) диктує необхідність технічного рішення – встановлення акумуляторних батарей зі значним запасом ємності.

Виходячи з аналізу, маркетингова стратегія базується на диференціації та нішевому позиціонуванні.

Позиціонування: «Для екологічно свідомих мешканців міста та відвідувачів заходів, наша автономна кав'ярня є єдиним мобільним закладом, що пропонує якісну каву та унікальний досвід без будь-якої шкоди для довкілля, завдяки 100% роботі на відновлюваних джерелах енергії».

Слоган: «Кава з майбутнього».

При проектуванні закладу мобільного формату було проаналізована цільова аудиторія потенційних користувачів:

1. *Первинний сегмент («Еко-новатори»)*. Молодь 18-30 років, студенти, IT-фахівці, представники креативних індустрій. Високо цінують інновації, екологічність та «історію» бренду. Активні в Instagram, TikTok. Вони є ядром лояльної аудиторії та «амбасадорами» бренду.

2. *Вторинний сегмент («Гості подій»)*. Відвідувачі фестивалів, концертів, міських свят. Їхня основна потреба – зручність, швидкість та отримання нових вражень. Мобільність кав'ярні, її можливість оперативно змінювати локацію спрямована саме на цей сегмент.

3. *Третинний сегмент («Свідомі споживачі»)*. Люди старшого віку (30-45+), часто з дітьми (відвідувачі парків), які свідомо обирають «зелені» та якісні продукти. Для них важлива відсутність шуму від генератора та чиста локація.

Маркетингова стратегія базується на позиціонуванні кав'ярні як символу якості, інновацій та сталого розвитку тому крім екологічної складової ми проектуємо також *маркетингову стратегію (Marketing Mix 4P)*.

1. *Product (Продукт)* Не лише кава, а комплексний «продукт-враження». Окрім якісних кавових напоїв (з можливим залученням локальних обсмажувальників), продукт включає: екологічну упаковку (біорозкладні стакани, дерев'яні мішалки), унікальний дизайн («servicescape») та сам факт купівлі «кави від сонця та вітру».

2. *Price (Ціна)*. Стратегія преміального ціноутворення. Ціна має бути вищою за середньоринкову на 10-15%. Це обґрунтовано:

а) високими початковими інвестиціями;

б) унікальною ціннісною пропозицією (екологічність, інноваційність); в) відсутністю прямих конкурентів у обраних локаціях.

3. *Place (Місце/Дистрибуція)*. Ключова перевага. Пропонується «кочова» (nomadic) модель дистрибуції:

– *Будні*. «Якірні» локації з високим трафіком (біля ВНЗ, офісних центрів, у центральних парках).

– *Вихідні*. Робота виключно на подіях (фестивалі, ярмарки) або у великих рекреаційних зонах.

– *Digital-mісце*. Обов'язкова реєстрація та щоденне оновлення статусу в Google Maps, Foursquare, та сторіз в Instagram з геолокацією, щоб клієнти могли знайти кав'ярню.

4. *Promotion (Промування)*.

– *Digital-маркетинг*: Instagram та TikTok як основні канали. Візуальний контент, що демонструє технологічність (робота сонячних панелей), естетику та процес приготування кави. Таргетована реклама на локацію.

– *PR / Storytelling*: Замість прямої реклами – розповідь історії ("чому ми це робимо"). Публікації в місцевих ЗМІ та у еко-блогерів про «першу в місті автономну кав'ярню».

– *Програма лояльності*. Цифрова бонусна система (наприклад, «Кожна 10-та кава безкоштовно»), що стимулює повторні візити та формує базу постійних клієнтів.

3. *Технічне обґрунтування енергетичної автономності*.

Це ключовий елемент, що доводить технічну спроможність проекту. Стратегічно обґрунтованим є вибір на користь комбінованої (гібридної) системи, що поєднує сонячні панелі, вітрогенератори та систему збереження енергії. Використання виключно сонячних панелей робить бізнес вразливим у зимовий період, похмурі дні та вечірні години. Використання виключно вітрогенераторів є ненадійним через мінливість вітрових умов.

Запропонована гібридна модель нівелює недоліки обох систем, забезпечуючи значно вищу стабільність генерації: сонячні панелі ефективні вдень, тоді як вітрогенератори можуть працювати вночі та у вітряну похмуру погоду, а система збереження енергії збільшує тривалість роботи кав'ярні. Це забезпечує безперервність бізнес-процесів та пряму нівелює ключову загрозу ($T2$), виявлену у SWOT-аналізі. Розрахунок добового енергобалансу базується на піковому споживанні обладнання та порівнянні його з потенційною генерацією (табл. 2).

Згідно розрахунків, загальне пікове споживання обладнання кав'ярні становить 13.9 кВт·год/день, а запропонована комбінована система генерації (5 сонячних панелей та 2 вітрогенератори) здатна виробляти в середньому 42.0 кВт·год/день. Запропонований трикратний запас потужності є не просто надлишком, а критично важливим стратегічним активом. По-перше, він дозволяє ефективно заряджати акумуляторні батареї (запланована ємність – 19.2 кВт·год) для гарантованої роботи в періоди повної відсутності сонця та вітру (наприклад, декілька похмурих

безвітряних днів). По-друге, цей резерв забезпечує можливість майбутнього розширення асортименту (можливість О4) шляхом підключення додаткового обладнання (наприклад, гриль, потужний блендер) без необхідності модернізації всієї енергосистеми. Це нівелює ключову технічну загрозу та забезпечує гнучкість бізнес-моделі.

Таблиця 2

Споживання електроенергії	Потужність, кВт	Час роботи, год/день	Споживання, кВт·год/день
Кавомашина	1.5	6	9.0
Кавомолка	0.5	2	1.0
Холодильник	0.2	12	2.4
Освітлення (LED)	0.05	10	0.5
Вентиляція та ін.	-	-	1.0
Всього споживання			13.9
Генерація електроенергії	Кількість, шт.	Потужність, кВт	Генерація, кВт·год/день
Сонячні панелі	5	3.0 (сумарно)	18.0 (при 6 год. сонця)
Вітрогенератори	2	2.0 (сумарно)	24.0 (при 12 год. вітру)
Всього генерація (комбінована)			~ 42.0

4. Економічне обґрунтування та аналіз окупності.

Фінальним етапом є визначення економічної доцільності проєкту. Розрахунок бюджету та окупності представлено в табл. 3.

Таблиця 3

Бюджет та розрахунок окупності проєкту

Показник	Сума, грн.	Примітка
Початкові інвестиції	1 200 000	
Мобільний трейлер (кемпер)	340 000	
Система енергозабезпечення	386 000	(Панелі, вітрогенератори, акумулятори, інвертор)
Кавове обладнання (POS-система)	300 000	
Логістика та оформлення	65 000	
Маркетинг (старт) та резерв	109 000	
Щомісячні операційні витрати	63 762	
Дозвіл на виїзду торгівлю	15 000	
Закупівля (продукти, пакування)	30 000	
Податки ФОП (вкл. ЄСВ)	2 862	(станом на 02.2025)
Маркетинг (операційний)	8 000	
Амортизація та ін.	7 900	
Прогнозовані доходи	280 000	(Середній чек 100 грн. x 100 клієнтів/день)
Фінансові результати		
Місячна виручка	280 000	
Місячний чистий прибуток	216 238	(280 000 – 63 762)
Термін окупності	~ 6-7 міс.	(1 200 000 / 216 238 ≈ 5.55 міс. + резерв)

Аналіз структури інвестицій показує, що «Система енергозабезпечення» (386 000 грн) становить 32,2% від загальних вкладень, що підтверджує висновок SWOT-аналізу про високу початкову вартість як ключову слабку сторону. Однак, розрахунок операційних показників доводить високу потенційну рентабельність. При досягненні консервативного прогнозу у 100 клієнтів на день із середнім чеком 100 грн, щомісячний чистий прибуток становитиме 216 238 грн.

Розрахунковий термін окупності ($1\,200\,000 / 216\,238 \approx 5.55$ місяців) є надзвичайно конкурентним для даного проєкту у сфері громадського харчування. Враховуючи необхідний час на досягнення планових показників, реалістичний горизонт повернення інвестицій становить 6-7 місяців. Це доводить, що високі початкові капітало-вкладення в автономність є економічно виправданими та компенсуються високою маржинальністю та гнучкістю бізнес-моделі.

Висновки

Проведене дослідження дозволило розробити та комплексно обґрунтувати бізнес-модель автономної мобільної кав'ярні. Доведено технічну спроможність проєкту завдяки використанню комбінованої системи енергозабезпечення (сонячні панелі, вітрогенератори та системи збереження енергії), яка здатна покрити енергоспоживання обладнання зі значним резервом потужності. Економічні розрахунки підтверджують інвестиційну привабливість

проєкту: при загальній сумі інвестицій 1 200 000 грн. та очікуваному чистому прибутку 216 238 грн. на місяць, термін окупності становить близько 6-7 місяців.

Розроблена модель поєднує екологічність (відсутність викидів та шуму від генераторів), мобільність (незалежність від електромереж) та економічну ефективність (зниження експлуатаційних витрат). Перспективи подальших досліджень полягають у вивченні можливостей масштабування даного проєкту через франчайзинг та адаптації моделі для інших типів мобільних закладів харчування.

Список використаної літератури

1. Савич Н., Михайлов А. *Маркетинг для малого бізнесу*. Київ: Бізнес-Альянс, 2020.
2. Петрова І., Гончарук С. *Основи еко-маркетингу: теорія та практика*. Львів: Економіка, 2019.
3. Ковальчук Ю. *Кавова культура в Україні: тенденції та перспективи*. Харків: Наукова думка, 2021.
4. Левченко О., Степанова І. *Соціальні медіа та їх використання для бізнесу*. Київ: IT-Прес, 2018.
5. Офіційний сайт «Association of Coffee Producers». URL: <https://www.coffee-association.org> (дата звернення: 09.11.2025).
6. Як створити успішний бізнес на основі сталого розвитку. *Business & Sustainability*. 2020.
7. Літвин М. Інфлюенс-маркетинг: Тренди 2023 року. *Технології маркетингу*. 2023.
8. Яковенко С. Підвищення ефективності маркетингових кампаній через програми лояльності. *Маркетингові технології*. 2022.

References

1. Savych, N., & Mykhailov, A. (2020). *Marketing for Small Business*. Kyiv: Biznes-Alians.
2. Petrova, I., & Honcharuk, S. (2019). *Fundamentals of Eco-Marketing: Theory and Practice*. Lviv: Ekonomika.
3. Kovalchuk, Yu. (2021). *Coffee Culture in Ukraine: Trends and Prospects*. Kharkiv: Naukova dumka.
4. Levchenko, O., & Stepanova, I. (2018). *Social Media and Its Use for Business*. Kyiv: IT-Pres.
5. Official website of the "Association of Coffee Producers". URL: <https://www.coffee-association.org> (accessed: 09.11.2025).
6. How to create a successful business based on sustainable development. (2020). *Business & Sustainability*.
7. Lytvyn, M. (2023). Influencer Marketing: Trends of 2023. *Marketing Technologies*.
8. Yakovenko, S. (2022). Increasing the effectiveness of marketing campaigns through loyalty programs. *Marketing Technologies*.

Дата першого надходження рукопису до видання: 25.11.2025
Дата прийнятого до друку рукопису після рецензування: 15.12.2025
Дата публікації: 31.12.2025

СОЦІАЛЬНІ ТА ПОВЕДІНКОВІ НАУКИ

УДК 330.341.1:334.012.64(477)

DOI <https://doi.org/10.35546/kntu2078-4481.2025.4.3.39>

Т. Ю. ОСАДЧА

викладач-стажист кафедри фінансів, обліку та оподаткування,
аспірант кафедри економіки, підприємництва та економічної безпеки
Херсонський національний технічний університет
ORCID: 0009-0009-4772-443X

**МАКРОЕКОНОМІЧНІ ДЕТЕРМІНАНТИ РОЗВИТКУ ПІДПРИЄМНИЦТВА
В УКРАЇНІ: АНАЛІЗ ПОТОЧНОГО СТАНУ ТА ПРОГНОЗНІ ОРІЄНТИРИ**

Військові дії на території України спричиняють значні коливання макроекономічних показників, таких як сукупний державний борг та його відношення до ВВП, індекс інфляції та девальвації, кількість безробітних та показник мінімальної заробітної плати. Ці показники безпосередньо впливають на розвиток та масштабування підприємств в Україні. Метою статті є визначення та аналіз впливу основних макроекономічних детермінант на розвиток підприємництва в Україні, а також обґрунтування напрямів удосконалення державної політики щодо формування сприятливих умов для його функціонування.

У статті досліджено динаміку сукупного державного боргу та ВВП України, а також відношення сукупного державного боргу до ВВП. Головними наслідками зростання сукупного державного боргу є зростання обсягів внутрішніх і зовнішніх запозичень для фінансування оборонних потреб та соціальних видатків, а також девальваційні процеси, які вплинули на номінальну оцінку боргу у національній валюті. Проведено аналіз показників індексу інфляції та індексу девальвації гривні відносно долару США. Наведено динаміку кількості безробітних, яка має тенденцію до зменшення. Це свідчить про перехід частини безробітних до неформальної зайнятості та адаптацією ринку праці до нових умов діяльності. Здійснено порівняльний аналіз динаміки мінімальної заробітної плати в Україні та деяких країнах Європи. Такий аналіз свідчить про низький рівень соціальних стандартів в Україні, порівняно з країнами Європейського Союзу. Проведено опис прогнозу економічного та соціального розвитку України на 2026-2028 роки, який розробило Міністерство економіки, довкілля та сільського господарства. З метою підтримки та стимулювання розвитку підприємництва в Україні, запропоновано такі інструменти державної підтримки, як грантові програми, пільгове кредитування, підвищення освіченості населення та необхідність прийняття ефективних управлінських рішень.

Ключові слова: підприємництво, макроекономічні показники, сукупний державний борг, ВВП, індекс інфляції, індекс девальвації, кількість безробітних, мінімальна заробітна плата, державна підтримка.

T. YU. OSADCHA

Trainee Lecturer at the Department of Finance, Accounting and Taxation,
Postgraduate Student at the Department of Economics, Entrepreneurship
and Economic Security
Kherson National Technical University
ORCID: 0009-0009-4772-443X

**MACROECONOMIC DETERMINANTS OF ENTREPRENEURSHIP DEVELOPMENT IN UKRAINE:
ANALYSIS OF THE CURRENT SITUATION AND FORECAST GUIDELINES**

Military actions on the territory of Ukraine cause significant fluctuations in macroeconomic indicators, such as total public debt and its ratio to GDP, inflation and devaluation indices, unemployment rates, and minimum wage indicators. These indicators directly affect the development and scaling of enterprises in Ukraine. The purpose of this article is to identify and analyze the impact of key macroeconomic determinants on the development of entrepreneurship in Ukraine, as well to justify directions for improving state policy on creating favorable conditions for its functioning.

The article examines the dynamics of Ukraine's total public debt and GDP, as well as the ratio of total public debt to GDP. The main consequences of the growth of aggregate public debt are an increase in domestic and external borrowing to finance defense needs and social expenditures, as well as devaluation processes that have affected the nominal valuation of debt in national currency. An analysis of the inflation index and the hryvnia devaluation index relative to the US dollar is provided. The dynamics of the number of unemployed, which tends to decrease, are presented. This

indicates the transition of some of the unemployed to informal employment and the adaptation of the labor market to new conditions of activity. A comparative analysis of the dynamics of the minimum wage in Ukraine and some European countries was carried out. This analysis indicates a low level of social standards in Ukraine compared to European Union countries. A description of the forecast for Ukraine's economic and social development for 2026-2028, developed by the Ministry of Economy, Environment, and Agriculture, is provided. In order to support and stimulate the development of entrepreneurship in Ukraine, the following instruments of state support are proposed: grant programs, preferential lending, raising the level of education of the population, and the need to make effective management decisions.

Key words: entrepreneurship, macroeconomic indicators, total public debt, GDP, inflation index, devaluation index, number of unemployed, minimum wage, state support.

Постановка проблеми

У сучасних змінних економічних реаліях, коли світова спільнота зіштовхується з новими викликами, кризами та трансформаційними можливостями, значення підприємств як ключового елементу формування стійкої, конкурентоспроможної та інноваційно орієнтованої економіки зростає як ніколи раніше. В умовах постійної турбулентності, зумовленої як глобальними процесами (геополітичними конфліктами, енергетичними та фінансовими кризами, зміною ланцюгів постачання), так і локальними факторами, вітчизняні підприємства потребують міцного інституційного та фінансового підґрунтя для забезпечення своєї життєздатності та стратегічного розвитку.

Аналіз останніх досліджень і публікацій

Вплив ключових макроекономічних показників на розвиток підприємництва в Україні посідає важливе місце у працях вітчизняних науковців, таких як О. Гарват [1], Л. Сідельникова [3], Д. Крилов [7] та інші. Разом із тим, зазначений науковий напрям зберігає свою актуальність і тепер, оскільки функціонування сучасного підприємницького середовища відбувається в умовах воєнного стану, що супроводжується посиленням економічної турбулентності та підвищенням рівня ризиків і невизначеності.

Формулювання мети дослідження

Метою статті є визначення та аналіз впливу основних макроекономічних детермінант на розвиток підприємництва в Україні, а також обґрунтування напрямів удосконалення державної політики щодо формування сприятливих умов для його функціонування.

Викладення основного матеріалу дослідження

Сьогодні Україна переживає найтяжче випробування за всю історію своєї незалежності – повномасштабну війну, яка стала потужним дестабілізуючим чинником для економічного середовища. Зростання державних витрат на військові потреби, невизначеність у зовнішньоекономічній діяльності, обмеження у міжнародній торгівлі та інвестиціях спричиняють поступове послаблення кожної з ланок економічної системи. У цих умовах перед вітчизняним підприємницьким сектором постає нагальна потреба в адаптації до нових викликів, у формуванні гнучких моделей управління та залученні інноваційних підходів до організації бізнес-процесів.

У своєму дослідженні, О. Гарват зазначає, що «рівень підприємницької активності має прямий вплив на ключові макроекономічні показники країни» [1, с. 70]. Зокрема, у дослідженні йдеться про «зростання внутрішнього валового продукту (далі – ВВП), створення нових робочих місць, сприяння реалізації заходів соціального захисту населення, формування платоспроможного попиту, а також збільшення надходжень до Державного бюджету України» [1, с. 70].

Макроекономічна стабільність виступає однією з ключових передумов ефективного функціонування підприємницького сектору. Її порушення виявляється, зокрема, у постійному нарощуванні обсягів державного боргу, наявності хронічного бюджетного дефіциту та погіршенні кредитного рейтингу держави. Сукупна дія цих чинників обмежує приплив інвестиційних ресурсів в економіку та знижує рівень довіри економічних суб'єктів до здійснення підприємницької діяльності.

На рисунку 1 розглянемо динаміку сукупного державного боргу та ВВП України, а також відношення сукупного державного боргу до ВВП протягом 2020-2024 років.

З рисунку 1 можна побачити, що протягом 2020-2024 років сукупний державний борг та показник ВВП мали стійку тенденцію до зростання. Зокрема, співвідношення сукупного державного боргу до ВВП зросло з 48,9% у 2021 році до 91,2% у 2024 році, що свідчить про суттєве збільшення боргового навантаження на економіку. Максимальне значення показника відношення сукупного державного боргу до ВВП у 2024 році перевищує критичний поріг, рекомендований міжнародними фінансовими інституціями (60% ВВП), що сигналізує про зниження фіскальної стійкості та підвищення ризиків боргової залежності [3, с. 302]. Однак, на період дії воєнного стану вищезазначена норма призупинена.

Надмірний рівень державного боргу зумовлює необхідність спрямування значної частини бюджетних ресурсів на обслуговування боргових зобов'язань, що істотно обмежує фіскальні можливості уряду щодо фінансування програм стимулювання підприємницької діяльності, розвитку інфраструктури та впровадження інноваційних проєктів.

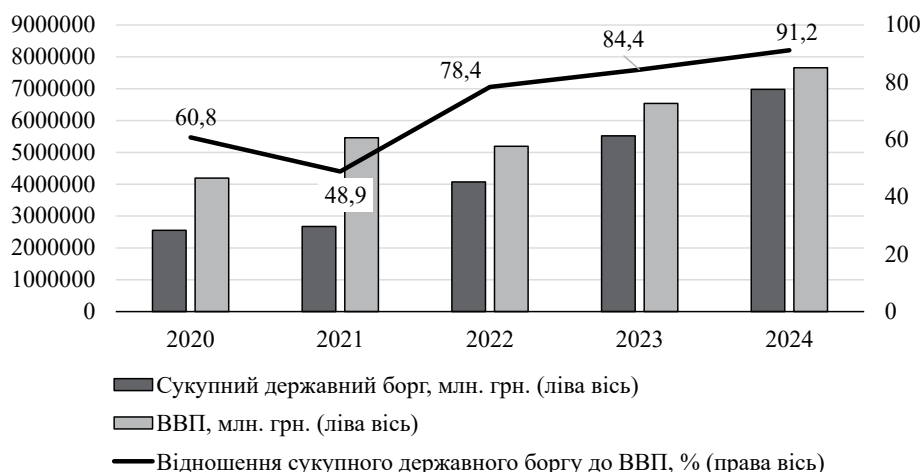


Рис. 1. Динаміка сукупного державного боргу та ВВП України, а також відношення сукупного державного боргу до ВВП у період з 2020 по 2024 роки

Джерело: Розроблено та систематизовано авторами за даними [2]

Серед чинників збільшення рівня державного боргу виокремимо такі: скорочення темпів економічного зростання унаслідок повномасштабної війни; зростання обсягів внутрішніх і зовнішніх запозичень для фінансування оборонних потреб та соціальних видатків; девальваційні процеси, що вплинули на номінальну оцінку боргу у національній валюті (рис. 2).

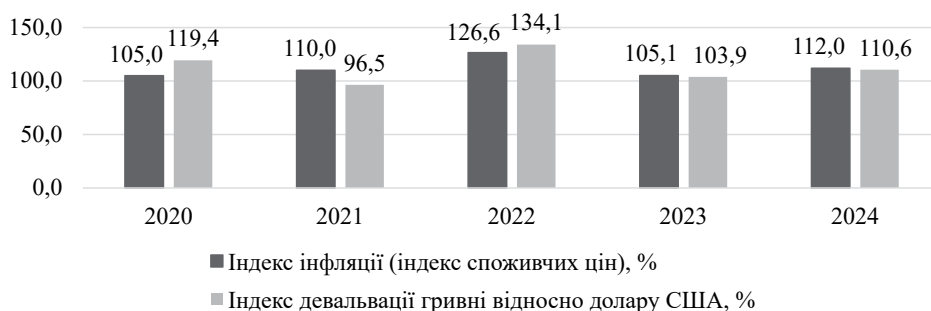


Рис. 2. Індекс інфляції (індекс споживчих цін) та індекс девальвації гривні відносно долару США у 2020-2024 роках

Джерело: Розроблено та систематизовано авторами за даними [4]

Аналіз показників індексу інфляції (індексу споживчих цін) та індексу девальвації гривні відносно долару США у 2020-2024 роках свідчить про наявність істотних коливань макроекономічної стабільності в Україні. У 2020-2021 роках спостерігалось відносне утримання інфляційних процесів на помірному рівні (105,0% та 110,0% відповідно). Протягом 2022 року, під впливом воєнних дій, ситуація суттєво погіршилася: індекс споживчих цін зріс до 126,6%, а індекс девальвації гривні досяг 134,1%. У 2023 році показники поступово стабілізувалися – інфляція знизилася до 105,1%, а девальвація гривні сповільнилася до 103,9%, що свідчить про ефективність заходів монетарної політики Національного банку України, спрямованих на утримання курсової стабільності та пом'якшення інфляційного тиску. Показники 2024 року демонструють збереження відносної макроекономічної рівноваги.

Війна суттєво змінила структуру та динаміку національного ринку праці, спричинивши глибокі дисбаланси між попитом і пропозицією робочої сили. Хоч економіка поступово адаптується до нових реалій воєнного часу, спостерігається значне навантаження на систему зайнятості, що зумовлює збереження високого рівня безробіття. Якщо у 2020 році їх чисельність становила 459,2 тис. осіб, то вже у 2024 році цей показник знизився до 94,2 тис. осіб, тобто майже у п'ять разів [5]. Таке різке скорочення безробітних пояснюється адаптацією ринку праці до нових економічних умов, активізацією державних програм зайнятості (зокрема «єРобота»), а також зменшенням кількості осіб працездатного віку внаслідок міграційних процесів і мобілізації.

Важливим елементом державного регулювання доходів є мінімальна заробітна плата, яка впливає на мотивацію до праці, купівельну спроможність і рівень бідності. В умовах повномасштабної війни Україна зіткнулася

з масовою міграцією працездатного населення до безпечніших регіонів України та за кордон, що зумовило структурні зміни на ринку праці та потребу в адаптації державної політики зайнятості.

Якщо порівнювати розмір мінімальної заробітної плати в Україні та в деяких країнах Європи, то даний показник в Україні відстає від європейських країн. У 2024 році в Україні мінімальна заробітна плата становить лише 184 євро, що понад у десять разів менше, ніж у таких країнах, як Німеччина (2000 євро) чи Велика Британія (2013 євро). Навіть у порівнянні рівня мінімальної заробітної плати в Україні з країнами Центральної Європи, такими як Польща (998 євро) та Чехія (755 євро), розрив залишається значним [6]. Можна зробити висновок, що в Україні зберігається низький рівень соціальних стандартів, порівняно з країнами Європейського Союзу, що свідчить про необхідність підвищення мінімальної заробітної плати, аби забезпечити гідний рівень життя та стимулювати трудову активність населення.

Д. Крилов наголошує, що ключовим завданням післявоєнного етапу відновлення національної економіки має стати не лише стабілізація ринку праці, а й комплексне усунення соціально-економічних диспропорцій, що обмежують його ефективне функціонування [7].

Міністерство економіки, довілля та сільського господарства розробило «Прогноз економічного і соціального розвитку України на 2026-2028 роки» [8].

Даний прогноз економічного та соціального розвитку України на 2026-2028 роки базується на двох альтернативних сценаріях, що прямо залежать від часових меж покращення безпекової ситуації в країні, як ключового чинника економічної стабільності [8].

Перший сценарій є базовим та оптимістичним, який передбачає стабілізацію безпекового середовища вже з 2026 року. У цьому випадку очікується активізація інвестиційної діяльності, зростання прямих іноземних інвестицій, прискорена модернізація інфраструктури, поліпшення бюджетної політики, зниження дефіциту та зростання соціальних видатків. Це створить позитивні мультиплікативні ефекти: підвищення доходів, зменшення безробіття та зміцнення довіри бізнесу й населення.

Другий сценарій є стриманим та консервативним, який передбачає покращення безпекової ситуації лише з 2027 року. Він характеризується повільнішими темпами економічного зростання, вищою інфляцією, тиском на державні фінанси через необхідність підтримувати оборонні видатки, а також обмеженими можливостями для розвитку інфраструктури та інновацій.

Можна підсумувати, що у будь-якому з вищезазначених сценаріїв подальший економічний розвиток України буде залежати від швидкості стабілізації безпеки в державі, масштабів інвестиційного відновлення та показників підприємницької активності. Варто зазначити, що покращення макроекономічного середовища сприятиме підвищенню довіри фінансових інститутів і розширенню можливостей для інвестування. Водночас важливим завданням держави залишається проведення структурних реформ, спрямованих на підвищення ефективності управління, підтримку підприємництва та інтеграцію у європейський економічний простір.

Висновки

Проведене дослідження показує, що незважаючи на високий рівень ризиків та вразливість підприємницького сектору до деструктивних впливів воєнних дій, підприємницька активність в Україні не лише зберігається, а й демонструє ознаки поступової адаптації та відновлення. З метою активізації бізнес-ініціатив уряд реалізує широкий спектр державних програм, проєктів і стратегічних заходів, спрямованих на створення ефективної системи підтримки підприємництва. Йдеться не лише про загальнодержавні програми, а й про спеціалізовані ініціативи, орієнтовані на розвиток малого та середнього бізнесу – як ключових драйверів економічного зростання [9].

Серед основних інструментів державної підтримки варто виокремити грантові програми, пільгове кредитування, податкові преференції, а також розгалужену мережу консультаційно-освітніх сервісів, що сприяють підвищенню підприємницької компетентності населення. В умовах триваючих воєнних викликів та необхідності відбудови національної економіки такі заходи набувають критичного значення, оскільки саме підприємницький сектор здатний забезпечити відновлення виробничого потенціалу, створення нових робочих місць, залучення інвестицій і формування конкурентних переваг України на міжнародному ринку. Суттєву увагу держава приділяє цифровізації управлінських процесів, розвитку електронних сервісів для бізнесу, зниженню адміністративного навантаження на підприємців, упровадженню інноваційних технологій і підтримці стартап-екосистем.

Отже, державна політика підтримки підприємництва виступає не лише засобом економічної стабілізації, а й стратегічним чинником посткризового відновлення та довгострокового розвитку України. Вона сприяє формуванню інноваційної моделі економічного зростання, підвищенню інвестиційної привабливості держави та інтеграції українського бізнесу у світові економічні процеси.

Список використаної літератури

1. Гарват О. Аналіз підприємницької активності в Україні. *Вісник Хмельницького національного університету. Економічні науки*. 2022. № 6. Том 2. DOI: [https://doi.org/10.31891/2307-5740-2022-312-6\(2\)-13](https://doi.org/10.31891/2307-5740-2022-312-6(2)-13) (дата звернення: 10.11.2025).

2. Динаміка сукупного державного боргу і ВВП України з 2020 по 2024 роки. Сайт Міністерства фінансів України. URL: <https://index.minfin.com.ua/ua/finance/debtgov/> (дата звернення: 10.11.2025).
3. Сідельникова Л. П. Сучасні виклики та новітні асиметрії фіскального простору держави. *Актуальні проблеми розвитку економіки регіону*. 2024. Вип. 20. Т. 1. С. 293-302. DOI: <https://doi.org/10.15330/apred.1.20.293-305> (дата звернення: 10.11.2025).
4. Індекс інфляції та девальвація гривні відносно долару США з 2020 по 2024 роки. Сайт Міністерства фінансів України. URL: <https://index.minfin.com.ua/ua/economy/index/devaluation/> (дата звернення: 10.11.2025).
5. Кількість зареєстрованих безробітних в Україні з 2020 по 2024 роки. Сайт Міністерства фінансів України. URL: <https://index.minfin.com.ua/ua/labour/unemploy/register/2014/> (дата звернення: 10.11.2025).
6. Зарплати і ціни в Україні та Європі: порівнюємо, як все змінилося за час війни. *Главком*. URL: <https://glavcom.ua/economics/finances/zarplati-i-tsini-v-ukrajini-ta-jevropi-porivnjujemo-jak-vse-zminilosja-za-chas-vijni-1043669.html> (дата звернення: 10.11.2025).
7. Крилов Д.В. Сучасні аспекти розвитку ринку праці в Україні. *Проблеми сучасних трансформацій. Серія: економіка та управління*. 2023. № 9. <https://doi.org/10.54929/2786-5738-2023-9-03-04> (дата звернення: 10.11.2025).
8. Прогноз економічного і соціального розвитку України на 2026-2028 роки. Сайт Міністерства економіки, довкілля та сільського господарства України. URL: <https://me.gov.ua/Documents/Detail/04033e6d-ae91-47c0-b5c2-438c23bb7220?lang=uk-UA&title=PrognozEkonomichnogoISotsialnogoRozvitkuUkrainiNa2026-2028-Roki> (дата звернення: 10.11.2025).
9. Боровік Л. В., Осадча Т. Ю. Перспективи розвитку підприємницької діяльності в Україні в контексті державної підтримки. *Збірник наукових праць Національного університету кораблебудування імені Адмірала Макарова*. 2024. № 3(496). DOI: 10.15589/znp2024.3(496).17 (дата звернення: 10.11.2025).

References

1. Harvat O. (2022). Analiz pidpryyemnytskoyi aktyvnosti v Ukraini [Analysis of entrepreneurship in Ukraine]. *Herald of Khmelnytskyi National University. Economic Science*, vol. 6, 2. DOI: [https://doi.org/10.31891/2307-5740-2022-312-6\(2\)-13](https://doi.org/10.31891/2307-5740-2022-312-6(2)-13) (accessed November 10, 2025).
2. Dynamika suкупного derzhavnoho borhu I VVP Ukrainy z 202 po 2024 roky [Dynamics of Ukraine's total public debt and from 2020 to 2024]. *Website of the Ministry of Finance of Ukraine*. URL: <https://index.minfin.com.ua/ua/finance/debtgov/> (accessed November 10, 2025).
3. Sidelnykova L. (2024). Suchasni vyklyky ta novitni asymetriyi fiskalnoho prostoru derzhavy [Modern challenges and newest asymmetries of the fiscal space of the state]. *The actual problems of regional economy development*, vol. 20, 1, pp. 293-302. DOI: <https://doi.org/10.15330/apred.1.20.293-305> (accessed November 10, 2025).
4. Indeks inflyatsiyi ta devalvatsiya hryvni vidnosno dolaru SSHA z 2020 po 2024 roky [Inflation index and devaluation of the hryvnia against the US dollar from 2020 to 2024] *Website of the Ministry of Finance of Ukraine*. URL: <https://index.minfin.com.ua/ua/economy/index/devaluation/> (accessed November 10, 2025).
5. Kilkist zareyestrovanykh bezrobitnykh v Ukraini z 2020 po 2024 roky [Number of registered unemployed in Ukraine from 2020 to 2025]. *Website of the Ministry of Finance of Ukraine*. URL: <https://index.minfin.com.ua/ua/labour/unemploy/register/2014/> (accessed November 10, 2025).
6. Zarplaty i tsiny v Ukraini ta Yevropi: porivnyuyemo, yak vse zminylosya za chas viyny [Salaries and prices in Ukraine and Europe: comparing how things have changed during the war]. *Glavcom*. URL: <https://glavcom.ua/economics/finances/zarplati-i-tsini-v-ukrajini-ta-jevropi-porivnjujemo-jak-vse-zminilosja-za-chas-vijni-1043669.html> (accessed November 10, 2025).
7. Krylov D. (2023). Suchasni aspekty rozvytku rynku pratsi v Ukraini [Modern aspects of labor market development in Ukraine]. *Problems of Modern Transformations. Series: Economics and Management*, vol. 9. <https://doi.org/10.54929/2786-5738-2023-9-03-04> (accessed November 10, 2025).
8. Prohnoz ekonomichnoho I sotsialnoho rozvytku Ukrainy na 2026-2028 roky [Forecast of economic and social development of Ukraine for 2026-2028]. *Website of the Ministry of Economy, Environment and Agriculture of Ukraine*. URL: <https://me.gov.ua/Documents/Detail/04033e6d-ae91-47c0-b5c2-438c23bb7220?lang=uk-UA&title=PrognozEkonomichnogoISotsialnogoRozvitkuUkrainiNa2026-2028-Roki> (accessed November 10, 2025).
9. Borovik L. V., Osadcha T. Y. (2024). Perspektyvy rozvytku pidpryyemnytskoyi diyalnosti v Ukraini v konteksti derzhavnoyi pidtrymky [Prospects for the development of entrepreneurship in Ukraine in the context of state support]. *Collection of Scientific Papers of Admiral Makarov National University of Shipbuilding*, vol. № 3(496). DOI: 10.15589/znp2024.3(496).17 (accessed November 10, 2025).

Дата першого надходження рукопису до видання: 29.11.2025
Дата прийнятого до друку рукопису після рецензування: 15.12.2025
Дата публікації: 31.12.2025

С. І. ПАСТУШЕНКО

доктор технічних наук, професор,
професор кафедри прикладної механіки та технічного сервісу
Відокремлений підрозділ
Національного університету біоресурсів і природокористування України
«Бережанський агротехнічний інститут»
ORCID: 0000-0002-1996-3558

Г. Д. БЕРЕГОВА

доктор філософських наук, професор,
професор кафедри економіки, підприємництва та економічної безпеки
Херсонський національний технічний університет
ORCID: 0000-0002-2418-7178

УНІВЕРСИТЕТ ЯК СИНТЕЗ ДОБІРНИХ ПРАКТИК ВИЩОЇ ПРОФЕСІЙНОЇ ТА ГУМАНІТАРНОЇ ОСВІТИ

У роботі розкрито сформований на основі поєднання практик вищої професійної та гуманітарної освіти які діють в Україні та закордоном, алгоритм створення стратегії реформації та розвитку ЗВО. Актуальність такої моделі та її змістовного наповнення було перевірено у період з часу її розробки у 2018 році та до теперішнього часу, часу дії правового режиму воєнного стану. Окреслено тенденції, аспекти, концепції, проблеми пов'язані зі сферою розвитку високоосвіченої особистості, а саме: набуття індивідуальних та професійних якостей; творення середовища академічної свободи та вільного прояву й розвитку суб'єктів освітньої взаємодії; наділення правами та інструментами щодо досягнення покликання; надання найкращих шансів для міждисциплінарних проривів; реалізація репутації ЗВО, як надійного роботодавця; реальне виконання принципу «навчання протягом життя»; розвиток міцної власної етики, як основи етично стійкої фахової практики; плекання середовища ЗВО, що єднає освіту, науку та розвиток особистості, які утворюють його єдину цілісність тощо. Також запропоновано шляхи формування якостей співробітників, лідерів ЗВО, які має нести «тягар» всіх окреслених аспектів його діяльності, вибудовуючи свою роботу на основі лідерського служіння. Повідомлено короткий зміст головних положень стратегії розвитку ЗВО, які пропонувалися автором у 2018 році. Особлива увага матеріалу статті відводиться важливим аспектам професійної та гуманітарної діяльності ЗВО на яких варто зосередитися у повоєнній Україні. До їх числа віднесено: освітнє служіння суспільству крізь призму моральності; автономія ЗВО; інноваційна трансформація освіти, науки та трудових відносин; гідність в академічній кар'єрі; культура національної ідентичності; місія ЗВО у повоєнній Україні. У висновках зазначаються позитивні наслідки реформації сучасного ЗВО на основі практик професійної та гуманітарної освіти, що уможливило особистісне та фахове зростання кожної зі сторін, залучених у цей процес. Підкреслено, що сучасний ЗВО має бути плодом злиття великого бачення з деталями, що уособлюють окремі аспекти його діяльності.

Ключові слова: особистість, заклад вищої освіти, служіння, зростання, зміни, аспекти діяльності, цінності.

S. I. PASTUSHENKO

Doctor of Technical Sciences, Professor,
Professor at the Department of Applied Mechanics and Technical Services
Separated Subdivision
of National University of Life and Environmental Sciences of Ukraine
“Berezhany Agrotechnical Institute”
ORCID: 0000-0002-1996-3558

H. D. BEREHOVA

Doctor of Philosophical Sciences, Professor,
Professor at the Department of Economics, Entrepreneurship
and Economic Security
Kherson National Technical University
ORCID: 0000-0002-2418-7178

UNIVERSITY AS A SYNTHESIS OF SELECTED PRACTICES OF HIGHER PROFESSIONAL AND HUMANITARIAN EDUCATION

The paper reveals an algorithm for developing a strategy for the reformation and development of higher education institutions, based on the synthesis of practices of higher professional and humanitarian education applied in Ukraine and abroad. The relevance of this model and its substantive content has been verified during the period from its development in 2018 to the present, including under the legal regime of martial law. The study outlines trends, aspects, concepts, and issues related to the development of a highly educated individual, namely: the acquisition of personal and professional qualities; creation of an environment of academic freedom and open self-expression and development of educational interaction participants; empowerment with rights and tools for achieving one's calling; providing the best opportunities for interdisciplinary breakthroughs; establishing the university's reputation as a reliable employer; real implementation of the principle of "lifelong learning"; cultivation of strong personal ethics as the basis for ethically resilient professional practice; and fostering a university environment that unites education, science, and personal development into a coherent whole. The paper also proposes ways to cultivate the qualities of university staff and leaders who must bear the "burden" of all these aspects of activity, building their work on the principles of servant leadership. The article presents a brief summary of the main provisions of the university development strategy proposed by the author in 2018. Particular attention is given to key aspects of professional and humanitarian activities of higher education institutions that should be prioritized in post-war Ukraine, including: educational service to society through the lens of morality; university autonomy; innovative transformation of education, science, and labor relations; dignity in academic careers; the culture of national identity; and the mission of the university in post-war Ukraine. The conclusions highlight the positive outcomes of reforming modern universities based on the practices of professional and humanitarian education, which enable both personal and professional growth for all participants in the process. It is emphasized that a modern university should be the result of merging a grand vision with the detailed aspects that embody its diverse activities.

Key words: *personality, higher education institution, service, growth, transformation, activity aspects, values.*

Постановка проблеми

23 лютого 2022 року, за день до початку найбільшого збройного конфлікту у Європі після Другої світової війни, коли відбулося широкомасштабне вторгнення сусідньої держави на територію незалежної країни, розпорядженням Кабінету Міністрів України схвалено Стратегію розвитку вищої освіти в Україні [1]. Документ визначив напрями зміцнення та основні характеристики які мають бути сформовані системою до 2032 року. Ним ідентифіковано наявні ключові проблеми вищої освіти держави, які разом із зовнішніми та внутрішніми викликами негативно впливають на можливості її подальшого розвитку. Єдина ключова проблема яка не врахована Стратегією – це війна.

Місією вищої освіти згідно з [1], є інтелектуальний, культурний і професійний розвиток особистості, формування ЗВО людського капіталу, створення ними інтелектуальних та інноваційних продуктів, просування етичних принципів, згуртування суспільства для утвердження України як рівноправного члена європейської спільноти.

Основний суб'єкт, котрого сотні разів згадує цей документ – це заклад вищої освіти (ЗВО). Адже саме від цієї «бойової одиниці» системи залежать її досягнення в цілому. Саме в ЗВО формуються особливі елементи, які він пропонує в презентованих освітніх програмах (ОП), професійні та загальні компетентності майбутнього фахівця.

Формулювання мети дослідження

Матеріал роботи було побудовано на основі огляду ряду проєктів та ініціатив, які з'явилися останніми роками у сфері вищої освіти України та закордоння з метою узагальнення результатів, уточнення концепцій і відкритих питань задля визначення головних аспектів діяльності, як сукупності професійних та гуманітарних практик спроможних разом сформувати алгоритм незворотних змін, наближаючи розв'язання проблем повоєнної реформації українських ЗВО.

Іншими словами, робота має на меті допомогти краще зрозуміти вузькі місця й потреби інституцій освітньої галузі та презентує потенціал, який мають й зуміють набути ті з них, які обирають шлях реформації на основі традиційних цінностей й висунутих ідей, перетворюючи ЗВО у простір цілісного розвитку й зростання кожної зі сторін, залучених у цей процес.

Викладення основного матеріалу дослідження

Безпосередньо в освітньому середовищі ЗВО здобувач повстає як високоосвічена особистість, набуваючи професійні та індивідуальні якості в ході взаємодії з викладачами, одногрупниками та іншими суб'єктами освітнього процесу. Саме в ньому він отримує «наповнення», що відповідає Стандарту вищої освіти для обраної ним ОП, виконуючи встановлений обсяг навантаження у кредитах ЕКТС, необхідних для здобуття ступеня. Специфіка ОП, як правило визначається місією ЗВО, його стратегічними цілями, цінностями, принципами діяльності, властивостями сильних сторін та інтересів навчально-наукових структур ЗВО. Нерідко специфіка ОП є наслідком лімітованих або перспективних дій, які виникають із того, що прийнято закладом освіти, позначено місцевими та регіональними особливостями, або до чого призводить уніфікація ОП з іншими ЗВО. Водночас уніфікація обмежує адаптивність, інакше кажучи, здатність реагувати на індивідуальність здобувачів освіти, на зміну потреб

роботодавців регіону, на зниження інноваційної переваги ОП, як потенціалу покращення освітнього процесу і результатів навчання.

Нерідко найліпші ідеї та нові потреби в компетенціях виникають за межами чинних дисциплінарних кордонів, і в такому разі завданням ЗВО має бути підтримка таких освітніх ініціатив та надання здобувачам найкращих можливостей для міждисциплінарних проривів, оскільки вони дають знання, формують навички та впевненість у собі, допомагаючи розкрити потенціал та обрати шлях, який відповідає їх інтересам та здібностям. Величезні обсяги післявоєнного оновлення України, наповнення економіки держави фахівцями із якостями гнучкої адаптації до потреб суспільства, демографічний спад – є окремими з тих проблем, які покликані вирішити міждисциплінарні ОП [2].

На думку В. Рябенка [3, с. 10-11], яка співзвучна з думкою автора цієї статті, у ЗВО повинен бути вибудований системний підхід щодо подолання проблематики розвитку особистості котра проходить крізь соціокультурне середовище ЗВО, який дозволить формувати її не лише у вузьких рамках навчального процесу, а й у суспільному і глобальному контекстах. Основою такої діяльності є академічна свобода і щирі можливості для вільного прояву й розвитку не лише здобувача, а й викладача, як основного суб'єкта взаємодії. Через соціальне оточення та взаємовідносини у такому освітньому середовищі проявляється, які люди насправді відійшли від ручного управління цим процесом з позицій авторитарної педагогіки, а не просто проголошують про таку перспективу.

Одним з першорядних цілей ЗВО є підготовка майбутніх фахівців до швидкозмінного ринку праці. Однак визначення того, які знання, уміння і навички має набути здобувач та як це робити, створює значні труднощі навіть для країн, котрі мають високі позиції в рейтингах, що оцінюють якість вищої освіти та освітні можливості загалом. Підтвердженням такої думки є документ [4] робочої групи з питань працевлаштування Coimbra Group (CG), що об'єднує 42 провідних університети Європи з 22 країн-учасниць.

Підготовка здобувачів вищої освіти до вирішення завдань і протистояння викликам, що повстають у сучасному світі та пов'язані з соціальним порядком, здатністю існувати без зовнішнього втручання або контролю, збереження природи, і головне, способів відповіді на них, у моральній філософії актуалізують через аналіз поняття покликання. Для українців воно передусім полягає у вільному розвитку і духовній свободі, які приносять особисте задоволення, повноту і користь суспільству. Філософію покликання – «сродної праці», розробляв видатний український філософ, поет, музикант і педагог Григорій Сковорода. Визнаючи, що працювати за покликанням – мабуть, найбільше задоволення у житті [5], він як би звертає увагу прийдешніх поколінь українців, щоб побачили неохитне у феномені покликання, який за його переконанням має духовні корені, адже «Дух Божий живе в нас» [6, с. 418] і ґрунтується на етичних аксіомах в контексті проблем і викликів сучасності.

Наділити здобувача вищої освіти повноваженнями, інструментами й можливостями щодо досягнення свого покликання, є одним з основних завдань ЗВО, що забезпечує основу для появи новітніх сміливих ідей, проведенню неформальних експериментів та реалізації інноваційних практик, які будуть перевірені та відкориговані викладачами відповідно до чинних наукових стандартів. Найкращі дослідження вимагають часу для занурення в них, і ЗВО має прагнути створити можливості для своїх дослідників працювати над такими ідеями.

Особливим аспектом забезпечення якості вищої освіти будь-якого ЗВО – є людський капітал, що складається з викладачів, науковців, навчально-допоміжного, адміністративно-господарського персоналу та працівників функціональних структурних підрозділів. Кожна з цих груп працівників ЗВО виконує свої, лише їй притаманні функції у складі єдиного організму ЗВО. Вони разом, у злагодженій єдності, створюють освітнє середовище, що уможливило отримання здобувачем якісної вищої освіти. Слід відзначити, що тільки за умов постійного академічного, професійного, наукового та особистісного розвитку кожної групи працівників [7], ЗВО буде здатний презентувати здобувачам актуальний та дійсно якісний освітній продукт.

На додаток, розробка і реалізація репутації надійного роботодавця ЗВО, яка на думку автора цієї публікації має обов'язково містити у собі збереження, залучення і розвиток працівників високого морально-етичного і фахового рівня які володіють сучасними просвітницькими й дослідницькими компетенціями, забезпечення для них гідних умов праці та плекання комфортного клімату у колективі, що сприятиме успішній реалізації стратегії закладу, чинячи як магніт, що «притягує» і утримує талановиті кадри.

Ще одним аспектом діяльності ЗВО – є розвиток дієвої системи особистісного зростання кожної зі сторін освітнього процесу. За визначенням особистість – це людина, що має найвищий ступінь духовного розвитку, вкритій системі сформованих цінностей, світорозуміння, поважний людські якості, які дозволяють особистості, вносити більш цінний вклад в розвиток суспільства, ніж середньостатистична людина [8].

Сегменти які потребують сталого зростання в кожного з учасників освітнього процесу охоплюють духовний, професійний, творчий, науково-інноваційний, соціальний та інші компоненти їх життєдіяльності. Концепція, що сприяє виконанню такого завдання в багатьох з сегментів має назву «Lifelong learning», або принцип «навчання протягом життя» [9], який було закладено у тіло Статті 3 Закону України «Про вищу освіту», як один з принципів державної політики у сфері вищої освіти. Його реалізація допомагає оновленню знань, умінь, навичок, передбачає необхідність постійного професійного розвитку задля актуалізації компетентностей індивіда, підвищення

здатності їх застосування, покращення особистісних якостей та інтелектуальної активності. Щорічні результати зростання викладача виражаються в його рейтинговій оцінці, яка є прийнятою ЗВО системою оцінювання його розвитку/діяльності, що базується на певних критеріях. Вони відображають професійну компетентність, результативність роботи та внесок викладача у: навчальну, методичну, наукову та інноваційну, організаційну та виховну, фінансово-господарську, міжнародну та інші види діяльності [10]. Рейтингова оцінка може враховувати самоосвіту, підвищення кваліфікації, участь у професійних програмах та інші форми фахового зростання, що підтверджують його розвиток.

Особисте лідерство в університеті вимагає високого рівня професіоналізму, етики та моралі, що демонструється такими рисами, як доброчесність, справедливість та прозорість. Ці якості є вирішальними для подолання моральних дилем, побудови довіри та створення позитивного середовища в ЗВО. Університетам потрібні лідери з сильними особистими моральними якостями, які можуть виступати взірцями для наслідування, керувати процесом прийняття етичних рішень на кожному внутрішньому рівні закладу та просувати назовні ідеали університету. Дослідження даного аспекту діяльності [11] проводяться в контексті зростання кожної зі сторін освітнього процесу у практичних надбаннях, для чого постає закономірність включення в ОП які презентує ЗВО, повноцінної підготовки спрямованої на розвиток етичних лідерів. Одним з головних висновків таких досліджень є те, що більша нетерпимість до відсутності ділової етики призводить до кращого вирішення ситуації, де потрібно зробити вибір між двома розбіжностями. Він виділяє важливість розвитку сильної особистої етики, яка може перерости в етично стійку професійну практику. Проведення теоретичних досліджень з даного напрямку, сприяє розумінню взаємозв'язку між особистою етикою та деонтологічним лідерством, стверджуючи, що особиста етика є важливою передумовою етичного лідерства в організаціях [12].

Університетське середовище має бути тим живим організмом, в якому розквітає єдність трьох його компонентів: освіти, науки та розвитку особистості, які утворюють єдину цілісність ЗВО. Призначення такого середовища полягає у тому, щоб діючи як осередок принципів гуманізму та свободи, демократії й відкритості до комунікації плекати академічну спільноту, готувати високоосвічених відповідальних лідерів-фахівців для суспільства, які спроможні вирішувати будь-які завдання його сучасного буття. В Україні прикладом співзвучного підходу до формування подібного середовища є Каразінський університет (м. Харків). Його стратегічні цілі та наміри містять саме такі пріоритети розвитку ЗВО до 2030 року.

Розглянемо головні якості співробітників, лідерів ЗВО, які мають нести «тягар» всіх вище окреслених аспектів діяльності ЗВО, тих котрі прагнуть позитивних змін за допомогою втілення в університетське життя синтезу кращих практик вищої професійної та гуманітарної освіти, що дозволяє бачити об'єкт опрацювання у його цілісності та взаємозв'язку всіх частин.

Взірцевим документом, на нашу думку, що визначає ключові цінності та елементи які формують рамку ефективного адміністрування для ЗВО у Великій Британії є «Кодекс управління вищою освітою» [13], що ґрунтується на принципі «застосовувати або пояснювати». Інакше кажучи, ЗВО надається набір цінностей та елементів, з яких вони можуть вибирати, які частини Кодексу до них застосовуються. Однак, від них очікується обґрунтування мотивів їх вибору. В основу надання освіти у Великій Британії закладено набір основних цінностей, визначених Кодексом, які очікуються від лідерів і спільнот ЗВО: прозорість, доброчесність, свобода слова та академічна свобода, участь та справедливі результати для всіх, високоякісні наукові дослідження та викладання, зростання у соціальній та культурній сферах, колегіальність, співпраця. Ці цінності, разом із «Принципами громадського життя Нолана» (Nolan Principles of Public Life [14]), створюють етичну основу для особистої поведінки всіх рівнів адміністрації та рад ЗВО, і є серцевиною цього Кодексу. Акцентуємо, що принципи взяті з урядового вебсайту є основою етичних стандартів, які очікують від державних службовців. Цим кроком наголошується, що робота у вищій освіті будується на лідерському служінні спорідненому з державною службою.

Ще одним ціннісним елементом, навколо якого точаться дискусії у вищій освіті – є духовність. В дослідженні [15] проведений аналіз вказує на імперативність визнання духовної частини освіти, шляхом відмежування її від релігійності. Набуття цієї особливої цінності особистістю не має залежати від глобалізованого погляду що домінує, адже пізнання, яким користуються ЗВО діє на основі державного права, ставлячи науку по один бік, а духовні питання – по інший. Але пізнавальна духовність лежить по обидва боки.

Цікавим є порівняння наведених вище цінностей які очікуються від лідерів і спільнот ЗВО з тими, які визнають у бізнесі, адже вони визначають корпоративну культуру, впливають на прийняття рішень та формують взаємодію як всередині компанії, так і з зовнішнім середовищем, слугуючи орієнтиром для співробітників, допомагаючи забезпечити єдність у команді та сприяючи досягненню спільних цілей. Їх порівняння дозволяє зробити висновок, що об'єкти порівняння показують суттєву схожість та мають багато спільного. Авторки дослідження [16], розглядаючи проблематику взаємозв'язку цінностей людини й цінностей організації пропонують етапність впровадження цінностей у корпоративну культуру організації. Ключовими (першими) елементами такої програми вони визначають, по-перше, розробку бачення (стратегічної мети) і місії, як гарантів стійкого сталого становища організації, що фокусують наміри кожного працівника на досягнення загальної мети, по-друге, визначення

першорядних цінностей організації, які повинні бути «вживлені» в ідеологію та культуру управління, систему норм поведінки в ній. І тільки на базі цих основних і ряду наступних робочих етапів, розробляється «Стратегія перетворень», як програма трансформації й ключових показників стосовно трьох рівнів: усієї організації, її структурних підрозділів і кожного працівника.

Подібний алгоритм створення «Стратегії реформації та розвитку ЗВО» (Стратегія), який автор статті пропонував впровадити у діяльність ЗВО, як синтез кращих практик вищої професійної та гуманітарної освіти, був застосований ним у 2018 році, як кандидатом на заміщення посади ректора одного з університетів України.

В тілі Стратегії реформації й розвитку ЗВО було сформовано 2 блоки:

перший – стратегічна ціль, місія, наші цінності й принципи діяльності, який не підлягав змінам, адже був наріжним каменем програми кандидата;

другий – стратегічні завдання, генеруючи які, автор не старався увести в них все, на що мав бачення, а тільки завдання, близькі йому, як дороговказ змін, що виходять з основ Стратегії наведених у першому блоці.

Всі інші завдання для тексту документа пропонувалося сформулювати безпосередньо від небайдужих до самих себе членів колективу – студентів, викладачів і співробітників, оскільки укладачем пропонувалася програма, яка не нав'язує бачення кандидата на посаду, а закликає до того, щоб бачення будь-кого з працівників мало перспективу бути реалізованим для них же. За своєю структурою стратегічні завдання є деталізованим планом дій для досягнення довгострокових цілей, що випливають з місії та цінностей, та поділялися на дев'ять розділів: формування духовно-культурних і мовних цінностей; академічна доброчесність; розвиток кадрового потенціалу; студентське самоврядування; виховання студентської молоді; соціальний, навчально-освітній, дослідно-інноваційний, інформатизаційний, фінансово-економічний та адміністративний устрої. За задумом автора, кожен з розділів стратегічних завдань мав бути наповнений пропозиціями співробітників їхнього бачення цих завдань, або прийняттям ними того, що запропонували інші, яке стає безумовним, коли за нього проголосувала більшість.

Анотація змісту головних положень місії й стратегічної цілі.

1. Наріжним каменем всіх видів діяльності університету ставиться студент – суб'єкт і продукт нашої діяльності – головний наш роботодавець.

2. Впровадження підходу до освіти – кожне обличчя є індивідуальним.

3. Просування в регіоні, Україні та світі індивідуальної, ексклюзивної, виняткової траєкторії розвитку університету.

4. Підпорядкування університетської структури менеджменту та адміністрування – першим трьома пунктам. Всі служать головній справі – підготовці лідерів для аграрного середовища України та інших держав.

Ціннісними орієнтирами університетського середовища в документі є:

I. Створення середовища – живого співтовариства ЗВО, в якому вільно просуваються знання, проводяться інтелектуальні відкриття і будуються дружні відносини в житті.

II. Наділення здобувачів вищої освіти повноваженнями на досягнення мети, маючи час власного життя, при підготовці до місії свого життя.

III. Ми докладаємо особливих зусиль щодо утримання, зацікавлення, пошуку й залучення до роботи найкращих людей для виконання місії ЗВО.

IV. Ми заохочуємо рух за духовним, професійним, творчим, науково-інноваційним і соціально орієнтованим розвитком і відзначаємо кожного з партнерів освітнього процесу за єдиним критерієм – наслідками їх зростання.

V. Як проголошується в преамбулі Конституції України, ми також усвідомлюємо відповідальність перед Богом, на основі чого вимагаємо від себе персональне лідерство, почуття обов'язку, високих стандартів професійної та людської етики й моралі.

VI. Ми підтверджуємо принцип гуманізму, демократії й відкритості до комунікації, як інструменти, що дозволяють забезпечити реалізацію місії ЗВО.

VII. Ми стверджуємо, що якість, які повинні мати претенденти на роботу є: здібність, правдивість, духовність і нульова терпимість до корупції.

Принципи діяльності через які ЗВО реалізовує своє призначення:

a. формування індивідуальності та безперечної особливості поміж ЗВО;

b. активну життєву позицію учасників освітнього процесу в реалізації місії та цінностей ЗВО на основі розуміння, що вони є владою, здійснюючи її через колегіальні та виконавчі органи ЗВО, інституту, факультету, кафедри;

c. систему адміністрування, що ґрунтується на принципах гуманізму, демократії та відкритості до комунікації, яка гарантує право свободи думки й слова, совісті та переконань учасникам освітнього процесу, мотивацію участі у розробці рішень виконавчих органів, які показують їхню волю та інтереси;

d. запровадження прозорості, вільної, чесної та конкурентної процедури обрання уповноважених осіб, що представлятимуть в колегіальному органі інтереси здобувачів, викладачів та інших груп учасників освітнього процесу;

е. принцип рівності – всі законні рішення виконавчих і колегіальних органів, мають однакову силу і застосовуються однаково для всіх партнерів;

ф. недискримінаційний, неупереджений підхід щодо розподілу прав й повноважень між партнерами, а не плекання управлінської єдиноначальності;

г. перетворення зі статусу бюджетної установи у статус некомерційної установи, як крок економічної дерегуляції, що збільшує можливості розвитку університету та відповідальність всіх учасників освітнього процесу.

Авторське бачення Стратегії реформації та розвитку даного ЗВО, як «запитання-відповідь» було конкретизовано в інтерв'ю каналу НІС-ТВ [17].

Розглянемо ряд важливих, аспектів професійної та гуманітарної освітньої діяльності ЗВО на яких варто зосередитися у повоєнній Україні.

Освітнє служіння суспільству крізь призму моральності. Вища освіта покликана підготувати здобувача до життя у сучасній повоєнній Україні. Але в таких реаліях, щоб стати гідною людиною, самих лише професійних знань недостатньо – потрібно зростати духовно, знаходити моральні орієнтири та цінності, вчитися справедливості та добру, допомозі та емпатії, адже без цього зрозуміти душі людей навколо, пошкоджені пекельним вогнем війни, буде неможливо. Нинішні викладачі, зрощуючи молодь, мають ділитися не лише знаннями, а й мудрістю життя, власним досвідом того, чому навчають. Вони повинні сприяти у пошуку свого місця у світі, утверджувати в студентів морально-інтелектуальні дороговкази-орієнтири, які приведуть їх до Істини. Навчати добрих справ потрібно не лише словами, але й власним прикладом та вчинками. Успішне виконання освіти свого призначення – формування морально свідомого суспільства України через виховання морально свідомих його членів, залежить насамперед, від особистої моральної гідності й професіоналізму представників освітньої галузі [18].

Автономія ЗВО. Зовнішня автономія ЗВО – як процес децентралізації управління вищою освітою, стосовно якого триває полеміка в політичному та освітньому середовищах, стосовно таких сфер як організаційна та фінансова, не є прерогативою рівня ЗВО. Його рівнем є проблеми внутрішньої сфери – автономії структурних підрозділів університету, практики що унормовується самим закладом керуючись законами, статутом ЗВО та чинними нормативно-правовими актами. Така реформація ієрархічного ладу дозволяє одержати переваги вільного, відповідального та багатовекторного їх вдосконалення. Інститут, факультет, центр тощо набувають делегованих ЗВО повноважень незалежно вирішувати справи свого «внутрішнього світу», професійного й особистісного зростання в тілі закладу. Вони мають право налагоджування прямої комунікації із партнерами з освітньої, наукової й бізнесової сфер не тільки України, але й закордоння. Місією ЗВО для автономних структурних підрозділів стає координація, піклування і допомога у зростанні в кожній з галузей підготовки фахівців, що єднають групи споріднених спеціальностей, наприклад, в педагогіці, інженерії, економіці тощо. Це активізує науковців, викладачів, співробітників, здобувачів. Користуючись полем такої автономії, вони налагоджуватимуть порозуміння з бізнесом, іншими партнерами в Україні й за кордоном, як у науково-практичних, так і в освітніх проектах.

Інноваційна трансформація. Цей комплексний процес задля оновлення та покращення кожної ланки ЗВО на основі втілення нових ідей, технологій, продуктів, методів і моделей для досягнення якісного зростання, підвищення ефективності та у підсумку, переходу на новий рівень всієї організації. Інноваційній трансформації, на нашу думку, належить підпорядкувати будь-який з напрямів діяльності ЗВО, наприклад так, як наведено нижче.

Освітній процес. З огляду на актуалізацію висловленої вище ідеї щодо освітньої моделі – «кожне обличчя є індивідуальним» пріоритетним стає педагогічний підхід, де центральне місце посідає особистість здобувача, його самоцінність, суб'єктний досвід та індивідуальні потреби. Метою такого підходу є не тільки засвоєння знань, а й розвиток пізнавальних здібностей, плекання ключових компетентностей, самоосвіти та самореалізації студента, що досягається через співпрацю, створення ситуацій вибору та адаптацію методики навчання до його індивідуальних можливостей.

Наукова творчість. Цей напрям інноваційної трансформації потребує побудови, діяльності й розвитку реальної, а не уявної інноваційної структури ЗВО, наприклад, парку інновацій – місця концентрації талантів, компаній, фондів, ідей, подій, середовища для зростання. Або центру науково-освітньої діяльності, як-то інноваційний хаб, де відбувається співпраця між наукою, бізнесом та державою, або освітній хаб, де можна отримати додаткові знання. Особлива цінність таких структур для ЗВО – у розкритті та розвитку талантів, місця доступу до засобів і сервісів від провідних технологічних корпорацій, місця налагодження контактів з тими, хто пройшов шлях від ідеї до компанії. Ще один напрям – зрощення творення у співпраці середовища колективного інтелекту, який дозволяє вирішувати завдання ефективніше ніж окремі його члени: сукупними знаннями, шляхом координації, обміном інформацією та самоорганізацією груп вільних учасників. Проте «мудрість колективу» руйнується, коли його члени втрачають відмінність, інакше кажучи, перестають мислити самостійно та починають слідувати за групою.

Трудові відносини. Забезпечення балансу між інтересами адміністрації й працівників, інтелектуалізацією трудової діяльності та її креативністю, що включає гнучкість мислення, здатність бачити нестандартне у звичному, а також наполегливість і віру у власні сили для втілення задумів. Такий підхід інноваційних змін перетворює специфіку трудових відносин, активуючи у ЗВО процеси цифровізації з використанням штучного інтелекту,

створення цифрових екосистем, які зумовлюють змістові зрушення його діяльності. Вважаємо конче перспективним кроком впровадження в ЗВО за прикладом [19] інноваційної для України, адаптованої до умов окремого ЗВО, моделі трудових відносин, що будуються на служінні, культурі спільного блага, відмовившись від моделі побудованої на власному престижі та інстинкті самозбереження. Адже коли авторитарний керманіч думає, що знає, що правильно, він ніколи не зможе використовувати мудрість колективного інтелекту людей. Це зовсім інша культура та зовсім інший спосіб мислення.

Гідність в академічній кар'єрі. Зміни у світі й Україні зробили академічну кар'єру складнішою, у порівнянні з її нещодавнім престижем. Повага до дослідників з роками послабилася, особливо у державній системі ЗВО та НДІ, адже просування їх академічної кар'єри суттєво залежить від державних інвестицій. Наприклад, написання статей застосовується для підтримки рейтингу наукової продукції та, як наслідок, отримання фінансування від державних установ. В цьому контексті гідність на роботі – це концепція, пов'язана зі здатністю до автономії та свободи для утвердження соціальної ідентичності, переконань і цінностей. Вона передбачає доступ до гідних умов праці та надання перспектив зростання, щоб зробити діяльність значущою та наділяючи плідну працю сенсом і суспільною корисністю. Занепад статусу викладача сприяв низьким зарплатам, зростанню бюрократії та відштовхнув молодь від академічної кар'єри, і нині, під час війни, це лише погіршилося. Це зумовлює розробку та реалізацію нових практик приведення у відповідність оплати праці не лише можливостям держави, а й потенціалом, що створюється ЗВО. СENS академічної кар'єри ґрунтується на ціннісному підході, який прагне трансформувати індивідів і націю, а не лише виробляти знання. Успіх не слід вимірювати виключно чисельними показниками, такими як кількість публікацій, кількість порад або обсяг мережі впливу, що часто призводить до вигорання. Натомість фокус має бути зміщений у бік якості роботи, чесного розподілу завдань і надання часу для творчості – через політику ЗВО, що сприяє: збалансованому навантаженню; боротьбі з академічною перевтомою в рамках кар'єри; різноманітним кар'єрним шляхам, що включають інновації, виробничі консультування, культурне зростання, видавництво тощо.

Культура національної ідентичності. Її концепція посідає чільне місце у вихованні колективного почуття приналежності до національної держави. Вона є винятково важливою для сприяння соціальній інтеграції та єдності, особливо в разі культурної різноманітності, етнічної та духовної спадщини. У багатьох країнах така культура є провідною у вихованні патріотичних цінностей, національної єдності та розуміння громадянських обов'язків у здобувачів. Маастрихтський договір, який є одним з договорів які сформували конституційну основу ЄС, визнає важливість національної ідентичності його держав-членів. Виховання у цій сфері повинно відбуватися на принципах відповідальності, патріотичності, моральності, духовності, нетерпіння до нечестя/корупції, доброчесності, відсутності поневолення. Елементи такої діяльності визначено Законом України «Про основні засади державної політики у сфері утвердження української національної та громадянської ідентичності» (ЗУ), який почав діяти з січня 2023 року. Ним встановлено, що надважливим елементом державної політики у цій сфері є національно-патріотичне виховання, що поєднує громадянсько-патріотичне, духовно-моральне виховання та співпрацю із закордонними українцями. Вагомим відгуком ЗВО на завдання, встановлені ЗУ, буде не лише зрощування такої культури безпосередньо в університеті, а й участь у формуванні та діяльності Координаційних рад з питань утвердження ідентичності при місцевих органах виконавчої влади та самоврядування. Такі ради своїми пропозиціями та рекомендаціями вплинуть на ухвалення рішень щодо державної політики у визначеній ЗУ сфері. А саме: формування громадянської позиції на засадах поваги до прав людини, духовних цінностей українського народу, національної самобутності, духовної єдності українців, відстоювання своїх духовно-моральних і державницьких позицій тощо.

Місія ЗВО у повоєнній Україні. Осмислення шляхів і інструментів побудови вертикалі відповідальності ЗВО у зціленні ран війни, подоланні її руйнівного впливу, творенні нової та модернізованої України відбувається як на внутрішньому рівні – серед зацікавлених у співпраці з освітнім закладом окремих осіб, груп чи організацій, так і на зовнішньому рівні – в середовищі територіальних громад зони впливу ЗВО задля реабілітації травм війни. Основою такої діяльності з одного боку, є комунікація з людьми, координація на місцях, соціалізація та психологічна адаптація українців до повоєнного життя. З іншого – комплекс заходів, що передбачає участь у розбудові центрів реабілітації, у реалізації проєктів відновлення (фізичного, психосоціального, духовного), доступність та інклюзивність такої діяльності.

Увічнення пам'яті про жертвовність і самовідданість захисників України, які віддали життя боронячи наше право існувати, право на свободу і на особисту недоторканність належить до основних обов'язків ЗВО у сфері національно-патріотичного виховання молоді.

Ветерани часто мають цінні лідерські та управлінські навички, як люди які проходили надскладні ситуації, та брак вищої освіти може перешкоджати їх працевлаштуванню, наприклад, в ІТ-галузі, на державній службі тощо. Тому цей аспект діяльності ЗВО полягає у пріоритетності освіти ветеранів та їх працевлаштування, шляхом створення особливих освітніх траєкторій, та активної комунікації з роботодавцями про їх цінність як фахівців.

Війна суворо привернула увагу українського суспільства до ролі сім'ї та традиційних сімейних цінностей. Сім'я стала опорою у воєнні часи важких випробувань, а традиційні ролі чоловіка-захисника-годувальника та

жінки-волонтерки-берегині (дому, родини) виявилися панівними. Отже, для ЗВО, теми демографічної кризи, соціальних викликів та допомоги у відбудові зруйнованого житла сім'ям українців мають бути пріоритетними, оскільки вони є такими для стратегії національної та соціальної безпеки держави.

Наслідками військової агресії суттєво пошкоджено абіотичні компоненти екосистеми, фауна і флора української території, що по новому актуалізує проблему екологізації освіти й науки, щоб заохочувати ініціативи (відкриття нових і міждисциплінарних ОП, нові наукові й прикладні проекти тощо), спрямовані на зцілення довкілля від руйнівних наслідків бойових дій та зведення до мінімуму тривалої шкоди природі від ран війни.

Вища освіта повинна бути драйвером в оновленні повоєнної України, готуючи фахівців, утримуючи в собі талановитих особистостей, залучаючи видатних науковців закордоння до освітнього процесу, приваблюючи повертатися з-за кордону вимушених біженців та українців з тимчасово окупованих територій. Засобами реалізації такої політики є створення: атмосфери доброзичливості, безпеки, достатності, коли більше не має потреби шукати свою цінність ззовні; середовища професійної та гуманітарної ексклюзивності; стану ґрунтового відчуття непопорушної єдності післявоєнного українського суспільства на прикладі свого ЗВО.

Висновки

Парадигма реформації сучасного ЗВО на основі добірних практик професійної та гуманітарної освіти відкриває перспективи для особистісного та професійного зростання кожної зі сторін, залучених у цей процес. Такий спосіб дій допомагає перетворити ЗВО у простір цілісного розвитку особистості, вчитися справедливості та добру, допомозі та емпатії, знаходити моральні орієнтири та цінності, розвивати душевну чутливість та критичне мислення, зростати духовно, опановувати методи продуктивної комунікації, що є запорукою добрих наслідків власного зростання.

Інтеграція представлених у статті структурованих ідей професійної й гуманітарної діяльності ЗВО, із наведеними цільовими освітніми практиками, дозволяє змістовно змінити університет з «кузні кадрів» на джерело творення особистостей, одночасно, позитивно змінюючи самих «творців» живого середовища ЗВО, які матимуть свої унікальні кар'єрні шляхи, психологічні, духовні та соціальні характеристики, спілкуючись з іншими не лише мовою інтелекту, а й мовою впливів і мовою серця.

Проведений огляд доводить, що стратегічна ціль, місія, цінності й принципи діяльності розроблені та презентовані автором в «Стратегії реформації та розвитку університету» у 2018 році, набули більшого наповнення у змістовності стратегічних завдань, щодо їх втілення у життя в умовах сьогоденних українських реалій та реалій повоєнної України.

Сучасний ЗВО повинен бути плодом злиття великого бачення з деталями, що уособлюють окремі аспекти його діяльності в контексті цього бачення, як би розділивши загальну роботу на видимі окремі перемоги, делегувавши повноваження для цього на щабель автономних структурних підрозділів, задля підвищення рівня всієї команди. Адже призначення ЗВО – творити чисте, світле, морально й інтелектуально стійке середовище в душах і серцях людей, а потім одягати їх у фахові знання, навички й уміння.

Список використаної літератури

1. Про схвалення Стратегії розвитку вищої освіти в Україні на 2022-2032 роки : Розпорядження КМУ від 23.02.2022 р. № 286-р. URL: <https://zakon.rada.gov.ua/laws/show/286-2022-p#Text>
2. Моделі міждисциплінарних та міжгалузевих освітніх та освітньо-наукових програм: виклики, можливості та варіанти впровадження : збірник матеріалів II міжн. конф., м. Одеса, 5-6.07.2021 р. Одеса : ОНУ імені І. І. Мечникова, 2021. 180 с.
3. Особистість студента і соціокультурне середовище університету в суспільному контексті : матеріали Всеукр. науково-практичної конф., м. Київ, 02.06.2017 р. Вступне слово В.І. Рябченка : Київ, 2017. 153 с.
4. Preparing graduates and lifelong learners to an ever-changing job market Coimbra Group input to the European Commission's Call for Evidence for an EU Quality Jobs Roadmap. URL: <https://www.coimbra-group.eu/wp-content/uploads/CG-INPUT-Preparing-graduates-and-lifelong-learners-to-an-ever-changing-job-market-EU-Quality-Jobs-Roadmap-consultation-1.pdf>
5. Мулярчук, Є. (2023). Філософія Сковороди і покликання сучасної людини. Філософська думка, 2, 132–143. <https://doi.org/10.15407/fd2023.02.132>
6. Сковорода Григорій. Твори: У 2 т. Київ : АТ «Обереги», 1994. (Гарвард, бібл-ка давнього укр. письменства). Т. 1 / Передм. О. Мишанича. 528 с. URL: https://shron1.chtyvo.org.ua/Skovoroda/Tvory_u_2_tomakh_Tom_2.pdf
7. Рязанов М.Р. Кадровий аспект забезпечення якості вищої освіти. Економіка та суспільство, Вип. 42. 2022. DOI: <https://doi.org/10.32782/2524-0072/2022-42-13>
8. Людина і особистість. Агентство Marictr. URL: <https://studies.in.ua/soc-ekzam/3261-lyudina-osobistst.html>
9. Теоретичні основи і технологія професійного розвитку науково-педагогічних працівників університетів в умовах інтеграції вищої освіти і науки : монографія / О.Г. Ярошенко та ін. ; за заг. ред. О.Г. Ярошенко. Київ : Інститут вищої освіти НАПН України, 2019. 236 с. DOI: <https://doi.org/10.31874/978-617-7486-26-7-2019>

10. Положення про рейтингову систему оцінювання діяльності науково-педагогічних працівників та структурних підрозділів НУБіП України. СУ СМЯ НУБіП України 7.2 – 01. Київ. 2025. URL: https://nubip.edu.ua/sites/default/files/u187/polozhennya_pro_reytingovu_sistemu_2025.pdf
11. Sánchez, A. V., Castro, M. C., Galindo-Domínguez, H., & Fuentes, E. P. (2024). Personal Ethics and Ethical Leadership in The Face of The Resolution of Dilemmas and Problematic Situations in Chilean Graduates. *Revista De Gestão – RGSA*, 18(6), e07479. DOI: <https://doi.org/10.24857/rgsa.v18n6-121>
12. Рикачевський О.В. Морально-педагогічні засади формування деонтологічної культури. Інноваційна педагогіка. 2022. Т. 2, Вип. 48. С. 25-30. DOI: <https://doi.org/10.32843/2663-6085/2022/48.2.5>
13. Бородієнко, О. В. Принципи публічно-громадського управління у вищій освіті в контексті забезпечення ефективності освітньої діяльності. Наукові записки. Серія: Педагогічні науки, Вип. № 209, 2023. С. 22-29. DOI: <https://doi.org/10.36550/2415-7988-2022-1-209-22-29>
14. The Seven Principles of Public Life. British government. URL: <https://www.gov.uk/government/publications/the-7-principles-of-public-life>
15. Цибулько О.С. Духовність у освітньому просторі: проблеми впровадження та реалізації. Освітній дискурс: зб. наук. праць. 2020. Вип. 19(1). С. 95-104. DOI: [https://doi.org/10.33930/ed.2019.5007.19\(1\)-9](https://doi.org/10.33930/ed.2019.5007.19(1)-9)
16. Грішнова О.А. Взаємозв'язок цінностей людини й цінностей організації в корпоративній системі. Формування ринкової економіки: зб. наук. праць. Спец. вип.: у 3 т. Соціал.-трудові відносини: теорія і практика. Київ. : КНЕУ. Т. 1. 2010. 468 с.
17. Добрий вечір 05.09.2018 Пастушенко. Телерадіокомпанія НІС-ТВ. URL: <https://www.youtube.com/watch?v=dSrFjF6vWPE&feature=youtu.be>
18. Пастушенко С. Освітнє служіння суспільству крізь призму моральності. Європейське майбутнє: філос.-освіт. студ.: Зб. наук. праць ХНТУ /за ред. Г. Д. Берегової. Херсон. 2024. Ч.2. С. 286-294. URL: [https://kntu.net.ua/ukr/content/download/115890/650762/file/+Матеріали%20конференції_Ч2%20статті%20Херсон\).pdf](https://kntu.net.ua/ukr/content/download/115890/650762/file/+Матеріали%20конференції_Ч2%20статті%20Херсон).pdf)
19. Служіння – третя місія вищої школи. Український католицький ун-т. URL: <https://ucu.edu.ua/news/sluzhinnya-tretya-misiya-vyshhoi-shkoly>

References

1. KМУ (2022) Pro skhvalennia Stratehii rozvytku vyshchoi osvity v Ukraini na 2022-2032 roky : Rozporiadzhennia KМУ vid 23.02.2022r. № 286-r. URL: <https://zakon.rada.gov.ua/laws/show/286-2022-p#Text> (in Ukrainian).
2. ONU (2021) Modeli mizhdystsyplinarnykh ta mizhhaluzevykh osvity i osvitno-naukovykh proham: vyklyky, mozhlyvosti ta varianty vprovadzhennia : zbirnyk materialiv II mizhn. konf., m. Odesa, 5-6.07.2021 r. Odesa : ONU imeni I. I. Mechnykova. 180 p. (in Ukrainian).
3. Riabchenko V. I. (2017) Osobystist studenta i sotsiokulturne seredovyshche universytetu v suspilnomu konteksti : materialy Vseukr. naukovopraktychnoi konf., m. Kyiv, 02.06.2017 r. Vstupne slovo V.I. Riabchenko : Kyiv. 153 p. (in Ukrainian).
4. Preparing graduates and lifelong learners to an ever-changing job market Coimbra Group input to the European Commission's Call for Evidence for an EU Quality Jobs Roadmap. URL: <https://www.coimbra-group.eu/wp-content/uploads/CG-INPUT-Preparing-graduates-and-lifelong-learners-to-an-ever-changing-job-market-EU-Quality-Jobs-Roadmap-consultation-1.pdf>
5. Muliarchuk, Ye. (2023). Filosofiia Skovorody i poklykannia suchasnoi liudyny. Filosofska dumka, Vol. 2, pp. 132–143. <https://doi.org/10.15407/fd2023.02.132> (in Ukrainian).
6. Skovoroda Hryhori. (1994) Tvory: U 2 t. Kyiv : AT «Oberehy», (Harvard, bibl-ka davnoho ukr. pysmenstva). T. 1 / Peredm. O. Myshanycha. 528 p. URL: https://shron1.chtyvo.org.ua/Skovoroda/Tvory_u_2_tomakh_Tom_2.pdf (in Ukrainian).
7. Riazanov M.R. (2022) Kadrovyi aspekt zabezpechennia yakosti vyshchoi osvity. Ekonomika ta suspilstvo, Vol. 42. DOI: <https://doi.org/10.32782/2524-0072/2022-42-13> (in Ukrainian).
8. Liudyna i osobystist. Ahentstvo Mahistr. URL: <https://studies.in.ua/soc-ekzam/3261-lyudyna-osobystst.html> (in Ukrainian).
9. IVO NAPNU (2019) Teoretychni osnovy i tekhnolohiia profesiinoho rozvytku naukovopedagogichnykh pratsivnykiv universytetiv v umovakh intehratsii vyshchoi osvity i nauky : monohrafiia / O.H. Yaroshenko ta in. ; za zah. red. O.H. Yaroshenko. Kyiv : Instytut vyshchoi osvity NAPNUkrainy, 236 p. DOI: <https://doi.org/10.31874/978-617-7486-26-7-2019> (in Ukrainian).
10. NUBiPU (2025) Polozhennia pro reitynhovu systemu otsiniuvannia diialnosti naukovopedagogichnykh pratsivnykiv ta strukturnykh pidrozdiliv NUBiP Ukrainy. SU SMIA NUBiP Ukrainy 7.2 – 01. Kyiv. URL: https://nubip.edu.ua/sites/default/files/u187/polozhennya_pro_reytingovu_sistemu_2025.pdf (in Ukrainian).
11. Sánchez, A. V., Castro, M. C., Galindo-Domínguez, H., & Fuentes, E. P. (2024). Personal Ethics and Ethical Leadership in The Face of The Resolution of Dilemmas and Problematic Situations in Chilean Graduates. *Revista De Gestão – RGSA*, 18(6), e07479. DOI: <https://doi.org/10.24857/rgsa.v18n6-121>

12. Rykachevskyi O.V. (2022) Moralno-pedahohichni zasady formuvannia deontolohichnoi kultury. Innovatsiina pedahohika. T. 2, Vol. 48. pp. 25-30. DOI: <https://doi.org/10.32843/2663-6085/2022/48.2.5> (in Ukrainian).
13. Borodiienko, O. V. (2023) Pryntsypy publichno-hromadskoho upravlinnia u vyshchii osviti v konteksti zabezpechennia efektyvnosti osvitnoi diialnosti. Naukovi zapysky. Seriya: Pedahohichni nauky, Vol. № 209. pp. 22-29. DOI: <https://doi.org/10.36550/2415-7988-2022-1-209-22-29> (in Ukrainian).
14. The Seven Principles of Public Life. British government. URL: <https://www.gov.uk/government/publications/the-7-principles-of-public-life>
15. Tsybulko O.S. (2020) Dukhovnist u osvitnomu prostori: problemy vprovadzhennia ta realizatsii. Osvitnii dyskurs: zb. nauk. prats. Vol. 19(1). pp. 95-104. DOI: [https://doi.org/10.33930/ed.2019.5007.19\(1\)-9](https://doi.org/10.33930/ed.2019.5007.19(1)-9) (in Ukrainian).
16. Hrishnova O.A. (2010) Vzaiemozviazok tsinnosti liudyny y tsinnosti orhanizatsii v korporatyvni systemi. Formuvannia rynkovoi ekonomiky: zb. nauk. prats. Spets. vyp.: u 3 t. Sotsial.-trudovi vidnosyny: teoriia i praktyka. Kyiv. : KNEU. T. 1. 468 p. (in Ukrainian).
17. Dobryi vechir 05.09.2018 Pastushenko. Teleradiokompaniia NIS-TV. URL: <https://www.youtube.com/watch?v=dSrFjF6vWPE&feature=youtu.be>
18. Pastushenko S. (2024) Osvitnie sluzhinnia suspilstvu kriz pryzmu moralnosti. Yevropeiske maibutnie: filos.-osvit. stud.: Zb. nauk. prats KhNTU /za red. H. D. Berehovo. Kherson. Ch.2. S.286-294. URL: [https://kntu.net.ua/ukr/content/download/115890/650762/file/+Матеріали%20конференції_Ч2%20статті%20Херсон\).pdf](https://kntu.net.ua/ukr/content/download/115890/650762/file/+Матеріали%20конференції_Ч2%20статті%20Херсон).pdf) (in Ukrainian).
19. Sluzhinnia – tretia misiia vyshchoi shkoly. Ukrainskyi katolytskyi un-t. URL: <https://ucu.edu.ua/news/sluzhinnya-tretya-misiya-vyshchoyi-shkoly> (in Ukrainian).

Дата першого надходження рукопису до видання: 24.11.2025
Дата прийнятого до друку рукопису після рецензування: 19.12.2025
Дата публікації: 31.12.2025

Н. В. ШАНДОВА

доктор економічних наук, професор,
завідувач кафедри економіки, підприємництва та економічної безпеки
Херсонський національний технічний університет
ORCID: 0000-0002-6278-1143

Н. В. МЄШКОВА-КРАВЧЕНКО

старший викладач кафедри економіки, підприємництва
та економічної безпеки
Херсонський національний технічний університет
ORCID: 0000-0001-5296-8687

ОЦІНКА СОЦІАЛЬНО-ЕКОНОМІЧНОЇ БЕЗПЕКИ ПІДПРИЄМСТВА

У статті розглянуто важливість оцінки соціально-економічної безпеки підприємства як елементу її забезпечення для створення умов для ефективної діяльності і стабільного розвитку підприємства.

Розглянуто підходи до оцінювання соціально-економічної безпеки підприємства. Указано на зростання значимості соціального аспекту безпеки підприємства та її оцінки. Зазначено про доцільність застосування графоаналітичного методу для оцінки соціально-економічної безпеки підприємства, який дозволяє наочно представити аналітичні розрахунки, що є основою прийняття обґрунтованих управлінських рішень.

Запропоновано адаптовану графоаналітичну модель оцінки соціально-економічної безпеки підприємства, яка дозволяє оцінити безпеку за векторами: операційна безпека, соціальна безпека, техніко-технологічна безпека, фінансова безпека. Показники за кожною групою відібрані таким чином, щоб для їх розрахунку була можливість отримувати інформацію з відкритих джерел, зокрема, з фінансової звітності підприємства. Показники характеризують стан, динаміку, ефективність використання ресурсів підприємства, абсолютний розмір та динаміку результатів діяльності, фінансову стабільність підприємства. Виокремлено соціальну безпеку як ресурсну складову і як спрямування до досягнення мети – сталого розвитку підприємства. Використання даної моделі дозволяє оцінити соціально-економічну безпеку підприємства в динаміці та в порівнянні з іншими суб'єктами ринку. Показники сформовані так, що є можливість порівняти їх на галузевому, регіональному рівні на основі статистичної інформації Державної служби статистики України. Модель передбачає визначення інтегрального показника соціально-економічної безпеки підприємства.

Здійснено оцінку соціально-економічної безпеки за запропонованою моделлю молокопереробного підприємства в динаміці та як учасника групи компаній.

Ключові слова: соціально-економічна безпека, підприємство, оцінка, модель, графоаналітичний метод, операційна безпека, соціальна безпека, техніко-технологічна безпека, фінансова безпека.

N. V. SHANDOVA

Doctor of Economics, Professor,
Head of the Department of Economics, Entrepreneurship
and Economic Security
Kherson National Technical University
ORCID: 0000-0002-6278-1143

N. V. MIESHKOVA-KRAVCHENKO

Senior Lecturer at the Department of Economics, Entrepreneurship
and Economic Security
Kherson National Technical University
ORCID: 0000-0001-5296-8687

ASSESSMENT OF SOCIO-ECONOMIC SECURITY OF AN ENTERPRISE

The article considers the importance of assessing the socio-economic security of an enterprise as an element of its provision for creating conditions for effective activity and stable development of an enterprise.

Approaches to assessing the socio-economic security of an enterprise are considered. The growing importance of the social aspect of enterprise security and its assessment is indicated. The expediency of using a graph-analytical method

for assessing the socio-economic security of an enterprise is indicated, which allows a visual presentation of analytical calculations, which is the basis for making informed management decisions.

An adapted graph-analytical model for assessing the socio-economic security of an enterprise is proposed, which allows assessing security by vectors: operational security, social security, technical and technological security, financial security. Indicators for each group are selected in such a way that for their calculation it was possible to obtain information from open sources, in particular, from the financial statements of the enterprise. The indicators characterize the state, dynamics, efficiency of use of enterprise resources, absolute size and dynamics of activity results, financial stability of the enterprise. Social security is singled out as a resource component and as a direction to achieve the goal – sustainable development of the enterprise. The use of this model allows assessing the socio-economic security of the enterprise in dynamics and in comparison with other market entities. The indicators are formed in such a way that it is possible to compare them at the industry, regional level based on statistical information of the State Statistics Service of Ukraine. The model provides the determination of an integral indicator of the socio-economic security of the enterprise.

An assessment of socio-economic security was carried out according to the proposed model of a dairy processing enterprise in dynamics and as a member of a group of companies.

Key words: socio-economic security, enterprise, assessment, model, graph-analytical method, operational security, social security, technical and technological security, financial security.

Постановка проблеми

Безпекові питання є важливими повсякчас і стосовно різних об'єктів і суб'єктів. Підприємство як ключовий суб'єкт господарювання відіграє важливу роль у формуванні ланцюжка: національна безпека – безпека регіону – безпека територіальної громади – безпека підприємства – безпека людини. Саме акцент на безпеку людини посилює соціальну складову загалом, і стосовно підприємства зокрема. Тож економічна безпека підприємства, яка активно досліджувалася останніми роками, трансформується в соціально-економічну безпеку підприємства. Ефективне формування системи соціально-економічної безпеки підприємства та управління нею потребує розуміння сутності даної категорії та можливості оцінки. Актуальність розгляду окреслених питань зростає в кризових ситуаціях, особливо під час військових дій і в повоєнному відродженні.

Аналіз останніх досліджень і публікацій

Дослідження безпеки підприємства вченими та практиками пов'язано у більшій мірі з економічною безпекою як цілісною категорією, так і в розбивці за окремими складовими, які різняться в окремих дослідженнях [1]. Останнім часом все більше приділяється уваги соціальній безпеці як складовій економічної безпеки, або як окремій категорії. Частина дослідників, поєднуючи ці дві категорії, розглядає соціально-економічну безпеку підприємства, зокрема Коленда Н.В. [2], Копитко М.І., Маланчук А.М. [3].

Щодо оцінки соціально-економічної безпеки підприємства, то вона найчастіше здійснюється окремо за складовими: економічна безпека та соціальна безпека [4], при цьому більш описано саме оцінку економічної безпеки [4], або соціальну безпеку розглядають як складову економічної безпеки [6].

Графоаналітичний метод, який найчастіше використовуються для оцінки потенціалу підприємства, його конкурентоспроможності [8-11] був адаптований авторами [12] для оцінки економічної безпеки підприємства. Однак, використання даного підходу для оцінки саме соціально-економічної безпеки в наукових джерелах не представлено.

Формулювання мети дослідження

Мета статті полягає в розробці та апробації адаптованої графоаналітичної моделі для комплексної оцінки рівня соціально-економічної безпеки підприємства.

Викладення основного матеріалу дослідження

Оцінювання соціально-економічної безпеки підприємства потребує, перш за все, розуміння самої категорії. Найчастіше оперують категорією економічна безпека, рідше – соціально-економічна безпека.

Грунтовне дослідження щодо сутності «економічної безпеки підприємства» проведено Боса І. [1, с. 305], де виокремлено такі групи підходів: ресурсний підхід, управлінський підхід, концепція економічної безпеки з урахуванням гармонізації інтересів, концепція економічної безпеки з урахуванням цільових завдань.

За Босою І., «економічна безпека підприємства – це стан господарюючого об'єкта, що характеризується його здатністю найефективніше використовувати доступні ресурси для забезпечення стабільної економічної та господарської діяльності та запланованого розвитку» [1, с. 306].

Значимість соціальної безпеки зростає, але її реалізація нерозривно пов'язана з економічною безпекою, то ж доцільніше розглядати саме соціально-економічну безпеку підприємства.

За Колендою Н.В., «соціально-економічна безпека підприємства – це стан, який характеризується надійною захищеністю соціально-економічної системи від наявних та потенційних загроз із внутрішнього та зовнішнього середовища, який дає можливість для самоорганізації та саморозвитку, збереження та розвитку трудового потенціалу, гарантує стійкість та стабільність, відповідний соціальний стан та рівень життєзабезпечення в умовах відповідного рівня прибутковості та узгодженості й реалізації економічних інтересів» [2, с. 676].

Копитко М.І., Маланчук А.М. розглядають соціально-економічну безпеку підприємства як «стан, при якому формуються безпекові умови для забезпечення захисту соціальних та економічних складових від негативного впливу ризиків, загроз та небезпек та відсутності реакції на різні форми виклику» [3, с. 19].

Для оцінки соціально-економічної безпеки підприємства найчастіше застосовують методичні підходи, які застосовують до оцінки економічної безпеки підприємства. Найпоширеніші підходи до оцінювання економічної безпеки підприємства, а саме індикаторний, ресурсно-функціональний, підхід на основі теорії економічних ризиків, програмно-цільовий охарактеризовано Ж. Шило [5]. Субота М.В. аналізує такі підходи до оцінювання соціально-економічної безпеки сталого розвитку суб'єктів господарської діяльності: ресурсно-функціональний, індикаторний, доходно-ресурсний (інвестиційний), стратегічний, сценарний, комплексний (системний) підхід [6]. При цьому вказано на важливість саме комплексного оцінювання рівня економічної безпеки підприємства, методу якого і запропонована авторами.

Лагодієнко В.В., Суханова А.В. для оцінки економічної безпеки підприємства використовують індикатори виробництва, фінансові індикатори і соціальні індикатори [7].

Субота М.В. серед груп показників оцінювання соціально-економічної безпеки виділяє: фінансову, інноваційно-технологічну, інформаційно-управлінську, матеріально-технічну, інформаційну, екологічну, внутрішню соціальну, кадрову, економіко-ринкову [6].

Важливим при оцінці є вибір показників: з одного боку – намагання охопити якнайбільше аспектів діяльності підприємства, з іншого – можливість швидко оцінити стан соціально-економічної безпеки, при цьому потрібно мати можливість порівняти з іншими суб'єктами ринку та з середніми показниками галузі, регіону, країни. Тож доцільно, щоб дані індикатори або вихідна інформація до їх розрахунку була у відкритому доступі.

Зокрема, науковці запропонували для оцінки соціально-економічної безпеки підприємства на основі офіційних статистичних даних групи з чотирьох показників окремо для економічної безпеки (рентабельність власного капіталу, коефіцієнт автономії, коефіцієнт загальної ліквідності, рівень рентабельності всієї діяльності підприємств) та соціальної безпеки (питома вага витрат на персонал у витратах на виробництво, рентабельність персоналу, коефіцієнт повної зайнятості персоналу, витрати на оплату праці у розрахунку на 1 зайнятого працівника за рік) [4, с. 491-492].

Одним із методів оцінки та аналізу певних систем та об'єктів є графоаналітичний метод, що передбачає використання аналітичних розрахунків і графічного зображення. Перевагами методу є розгляд системи як цілого, з урахуванням взаємозв'язків між елементами, поєднання математичних та статистичних методів при розрахунку показників і наочність представлення даних у вигляді графів, що спрощує сприйняття.

В економічній сфері найчастіше графоаналітичний метод використовують для оцінки потенціалу підприємства. Запропонований Репіною І.М. [8] графоаналітичний метод «Квадрат потенціалу», який передбачав оцінку потенціалу підприємства на основі аналізу конкурентів за групами показників: «Виробництво і збут продукції», «Організаційна структура і менеджмент», «Маркетинг», «Фінанси», набув поширення. Надалі були запропоновані модифікації як стосовно вибору назв векторів і показників, так і оцінки в динаміці [9] та визначення інтегрального показника [10]. Строченко Н.І., Ковальова О.М. [11] створили модель ресурсного потенціалу на основі аналізу шести груп показників.

Ураховуючи, що у трактуванні як потенціалу підприємства, так і економічної безпеки підприємства одним із підходів є ресурсний, та і формування оптимального потенціалу підприємства є основою забезпечення високого рівня економічної безпеки, було запропоновано модель оцінки економічної безпеки підприємства як адаптований графоаналітичний метод «Квадрат потенціалу» [12]. Векторами у моделі були визначені: «Виробництво і реалізація», «Кадрова безпека», «Техніко-технологічна безпека», «Фінансова безпека».

Як вказувалося раніше, наразі доцільно розглядати саме соціально-економічну безпеку підприємства. Тож пропонується модель оцінки соціально-економічної безпеки, у якій видозмінені вектори квадрату, саме:

- 1 вектор – Операційна безпека;
- 2 вектор – Соціальна безпека;
- 3 вектор – Техніко-технологічна безпека;
- 4 вектор – Фінансова безпека.

Перелік показників у групах розширено, у перших трьох до 10, у четвертій – до 12 показників.

Виокремлення в якості векторів саме цих чотирьох складових безпеки із значної кількості, які розглядаються науковцями, пов'язано з наступним:

- зменшення до чотирьох передбачає можливість зосередитися на найбільш важливих аспектах соціально-економічної безпеки;
- стан захищеності залежить перш за все від результатів операційної (основної) діяльності;
- забезпечення результатів пов'язано з ресурсним забезпеченням і його захищеністю, що окреслює соціальна безпека і техніко-технологічна безпека;

– окремо виділена саме соціальна безпека, яка поєднує як ресурсний аспект, так і результат, цілеспрямування;
 – можливість забезпечення та покращення стану соціально-економічної безпеки напряму залежить від фінансової спроможності, від фінансової стійкості, що обумовлює розгляд фінансової безпеки.

Перелік показників сформовано так, щоб вихідну інформацію можна було отримати з відкритих джерел, а саме з річної фінансової звітності підприємств. При цьому частину окреслених показників можна розрахувати на основі статистичної інформації, що акумулюється у Державній службі статистики України [13], що дозволяє, за необхідності, оцінити стан конкретного підприємства в галузевому аспекті, на регіональному рівні, в рамках національної економіки.

Апробацію моделі здійснено за матеріалами ТДВ «Яготинський маслозавод» [14] в динаміці та як учасника АТ «Молочний альянс» [15], до якого входять підприємства з виробництва сирів, цільномолочної та кисломолочної продукції, підприємства зі збору та обробки молока та молочної продукції. Коефіцієнт чутливості визначено експертним шляхом.

Для прийняття ефективних рішень щодо забезпечення та підвищення рівня соціально-економічної безпеки підприємства важливо оцінити ситуацію в динаміці. На основі матеріалів річної фінансової звітності ТДВ «Яготинський маслозавод» за 2021-2024 роки [14] здійснено оцінку соціально-економічної безпеки підприємства за графоаналітичним методом.

Операційну безпеку оцінюємо на основі абсолютних результатів, динаміки змін та рентабельності (табл. 1).

Таблиця 1

Вихідна інформація для розрахунку вектору «Операційна безпека»

Показник	Коефіцієнт чутливості	2021		2022		2023		2024	
		значення	місце	значення	місце	значення	місце	значення	місце
Чистий дохід, тис. грн.	1,1	3103661	4	3736091	3	4857266	2	5635593	1
Темп зміни чистого доходу	1,05	0,980	4	1,204	2	1,300	1	1,160	3
Валовий прибуток, тис. грн.	1,1	319157	4	645244	3	877957	1	763220	2
Темп зміни валового прибутку	1,05	0,628	4	2,0217	1	1,361	2	0,869	3
Прибуток від операційної діяльності, тис. грн.	1,1	20788	4	291991	2	444037	1	237773	3
Темп зміни прибутку від операційної діяльності	1,1	0,093	4	14,046	1	1,521	2	0,535	3
Темп зміни собівартості реалізованої продукції	1,05	1,048	1	1,110	2	1,207	3	1,224	4
Рентабельність активів, %	1,15	1,33	4	16,21	2	21,18	1	10,59	3
Рентабельність продажу, %	1,15	10,28	4	17,27	2	18,08	1	13,54	3
Рентабельність операційної діяльності, %	1,15	0,67	4	8,36	2	10,02	1	4,32	3

Джерело: складено та розраховано авторами на основі [14]

Оцінювання соціальної безпеки підприємства повинно бути спрямоване на врахування соціально-економічних аспектів життя людини, працівника підприємства, що, звичайно, оцінити тільки показниками, що базуються на об'єктивній інформації, складно. Тож в подальшому, для більш детального аналізу стану соціальної безпеки підприємства потрібно отримувати інформацію також шляхом опитування. Однак, наразі дана методика передбачає використання інформації з відкритих джерел, фінансової звітності підприємства.

Лагодієнко В.В., Суханова А.В. до соціальних індикаторів відносять: рівень оплати праці відносно середнього показника по промисловості або в економіці в цілому, рівень заборгованості по зарплаті, втрати робочого часу, структуру кадрового потенціалу (вікова, кваліфікаційна) [7, с. 324].

Строченко Н.І., Ковальова О.М при формуванні моделі ресурсного потенціалу підприємства виділяють групу показників соціального розвитку підприємства, а саме: коефіцієнт зростання заробітної плати, коефіцієнт плинності кадрів, рівень освіти працівників підприємства, питома вага управлінського персоналу, коефіцієнт використання річного фонду робочого часу [11, с. 46].

Нами запропоновано наступний перелік показників для оцінки соціальної безпеки підприємства, визначених на основі фінансової звітності підприємства, які корелюються з показниками, які можна визначити на основі статистичної інформації (табл. 2).

Показники ресурсної групи «Техніко-технологічна безпека» охоплюють характеристики стану, руху, ефективності використання основних засобів, нематеріальних активів (як індикатора інноваційної складової) та рівня матеріальних витрат.

Показники, що формують блок фінансової безпеки є як абсолютними (чистий фінансовий результат) і в динаміці, так і відносними – показники фінансової стійкості, платоспроможності і ділової активності (табл. 4).

Таблиця 2

Вихідна інформація для розрахунку вектору «Соціальна безпека»

Показник	Коефіцієнт чутливості	2021		2022		2023		2024	
		значення	місце	значення	місце	значення	місце	значення	місце
Чисельність персоналу, осіб	1,1	901	3	900	4	921	2	928	1
Темп зміни чисельності персоналу	1,05	0,901	4	0,999	3	1,023	1	1,008	2
Чистий дохід на одного працюючого, тис. грн. /особу	1,2	3444,68	4	4151,21	3	5273,90	2	6072,84	1
Темп зміни чистого доходу на одного працюючого	1,1	0,999	4	1,205	2	1,270	1	1,151	3
Рентабельність персоналу, тис. грн/особу	1,05	-	4	172,93	2	277,05	1	102,98	3
Частка витрат на оплату праці з відрахуваннями на соціальні заходи у витратах, %	1,05	7,603	3	7,715	2	7,724	1	7,358	4
Витрати на оплату праці на одного працівника, тис. грн. на особу	1,1	187,10	4	217,69	3	271,22	2	321,82	1
Темп зміни витрат на оплату праці на одного працівника	1,2	1,049	4	1,163	3	1,246	1	1,187	2
Рентабельність витрат на оплату праці, %	1,05	0,123	4	1,490	2	1,778	1	0,796	3
Темп зміни рентабельності витрат на оплату праці	1,1	0,090	4	12,114	1	1,193	2	0,448	3

Джерело: складено та розраховано авторами на основі [14]

Таблиця 3

Вихідна інформація для розрахунку вектору «Техніко-технологічна безпека»

Показник	Коефіцієнт чутливості	2021		2022		2023		2024	
		значення	місце	значення	місце	значення	місце	значення	місце
Активи на кінець року	1,1	1574620	4	2029019	3	2164580	2	2327696	1
Темп зміни активів	1,05	1,142	2	1,289	1	1,067	4	1,075	3
Коефіцієнт приросту основних засобів	1,1	0,052	3	0,047	4	0,097	2	0,195	1
Коефіцієнт зносу основних засобів на кінець року	1,15	0,461	1	0,519	3	0,538	4	0,491	2
Частка основних засобів в активах на кінець року	1,05	0,265	1	0,193	3	0,192	4	0,244	2
Частка активних основних засобів на кінець року	1,1	0,756	1	0,740	2	0,718	3	0,687	4
Віддача основних засобів	1,2	4,118	4	4,712	3	5,678	1	5,590	2
Коефіцієнт приросту нематеріальних активів	1,1	-0,032	3	0,004	2	-0,216	4	0,013	1
Частка нематеріальних активів в активах	1,05	0,0018	1	0,0008	2	0,0006	3	0,0005	4
Частка матеріальних витрат у витратах	1,1	0,822	1	0,835	4	0,830	2	0,833	3

Джерело: складено та розраховано авторами на основі [14]

Довжину векторів моделі визначаємо за формулою:

$$B_{pj} = 100 - \left(\sum_{i=1}^n a_{ij} \times k_i - \sum_{i=1}^n k_i \right) \times \frac{100}{\sum_{i=1}^n k_i \times (m-1)},$$

де B_{pj} – р-й вектор для j-го періоду; a_{ij} – місце j-го періоду за i-м показником; k_i – коефіцієнт чутливості i-го показника; n – кількість показників у групі; m – кількість періодів.

Результати розрахунків відображено графічно на рис. 1.

Таблиця 4

Вихідна інформація для розрахунку вектору «Фінансова безпека»

Показник	Коефіцієнт чутливості	2021		2022		2023		2024	
		значення	місце	значення	місце	значення	місце	значення	місце
Чистий фінансовий результат, тис. грн.	1,05	-63319	4	155639	2	255160	1	95563	3
Зміна чистого фінансового результату, тис. грн.	1,05	-192866	4	218958	1	99521	2	-159597	3
<i>Фінансова стійкість</i>									
Коефіцієнт автономії	1,1	0,336	4	0,338	3	0,434	1	0,350	2
Коефіцієнт фінансової стійкості	1,1	0,493	2	0,463	3	0,545	1	0,443	4
Коефіцієнт маневреності власного капіталу	1,1	0,651	3	0,752	2	0,779	1	0,540	4
<i>Платоспроможність</i>									
Коефіцієнт поточної ліквідності (покриття)	1,1	1,432	3	1,472	2	1,743	1	1,340	4
Коефіцієнт швидкої ліквідності	1,1	1,224	3	1,321	2	1,548	1	1,145	4
Коефіцієнт абсолютної ліквідності	1,1	0,040	2	0,007	4	0,042	1	0,022	3
<i>Ділова активність</i>									
Коефіцієнт оборотності активів	1,05	1,987	4	2,074	3	2,317	2	2,509	1
Коефіцієнт оборотності запасів	1,05	18,704	4	22,566	3	27,131	1	25,293	2
Коефіцієнт оборотності дебіторської заборгованості	1,1	3,468	2	3,157	4	3,341	3	3,839	1
Коефіцієнт оборотності кредиторської заборгованості	1,1	6,825	4	6,876	3	8,424	2	9,265	1

Джерело: складено та розраховано авторами на основі [14]

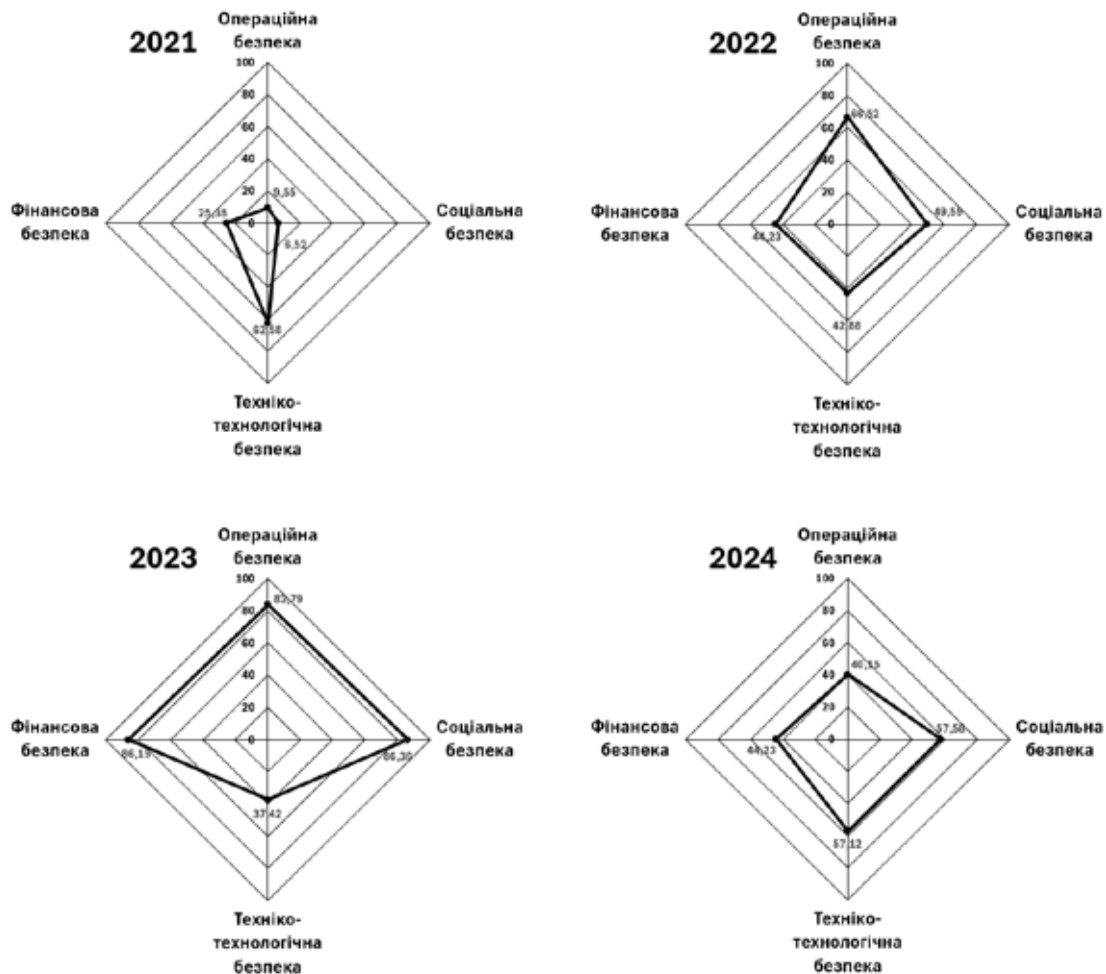


Рис. 1. Модель соціально-економічної безпеки ТДВ «Яготинський маслозавод» в динаміці

Джерело: побудовано на основі розрахунків авторів

Загальний рівень соціально-економічної безпеки підприємства за певний період (ЕБ) визначаємо за формулою:

$$EB = \frac{1}{2} \times \left(\left(\sum_{p=1}^{l-1} B_p \times B_{p+1} \right) + B_l \times B_1 \right)$$

де B_p – р-й вектор;
 l – кількість векторів.

Результати розрахунків відображено в таблиці 5.

Таблиця 5

Інтегральний показник соціально-економічної безпеки в динаміці

Рік	2021	2022	2023	2024
Інтегральний показник соціально-економічної безпеки	1150,48	5129,77	10454,97	4951,53

Джерело: розраховано авторами

Результати дослідження свідчать, що рівень соціально-економічної безпеки ТДВ «Яготинський маслозавод» зростає в період з 2021 року до 2023 року. Однак, у 2024 році відбулося різке зниження інтегрального показника соціально-економічної безпеки. Якщо розглянути графічну модель, то відслідковуємо, що у 2024 році підвищився рівень тільки техніко-технологічної безпеки, за всіма іншими складовими відбулося зниження рівня, при цьому за складовою «Соціальна безпека» в 1,5 рази, за складовими «Операційна безпека» і «Фінансова безпека» приблизно в 2 рази. Тож прийняття управлінських рішень потребує розгляду динаміки окремих показників в групах і, за необхідності, проведення додаткового дослідження.

ТДВ «Яготинський маслозавод» (підприємство 1) входить до складу групи компаній «Молочний альянс» поряд з ТДВ «Баштанський сирзавод» (підприємство 2), ТДВ «Золотоніський маслоробний комбінат» (підприємство 3), ТОВ «Пирятинський сирзавод» (підприємство 4). Тож для прийняття обґрунтованих управлінських рішень в рамках групи компаній доцільно за запропонованою методикою оцінити рівень соціально-економічної безпеки даних підприємств.

Довжину векторів моделі визначаємо за формулою:

$$B_{pj} = 100 - \left(\sum_{i=1}^n a_{ij} \times k_i - \sum_{i=1}^n k_i \right) \times \frac{100}{\sum_{i=1}^n k_i \times (m-1)},$$

де B_{pj} – р-й вектор для j-го підприємства;
 a_{ij} – місце j-го підприємства за i-м показником;
 k_i – коефіцієнт чутливості i-го показника;
 n – кількість показників у групі;
 m – кількість підприємств.

Результати розрахунків за 2024 рік відображено графічно на рис. 2.

Інтегральний показник соціально-економічної безпеки вказує, що найвищий рівень економічної безпеки серед підприємств групи компаній «Молочний альянс» у ТДВ «Яготинський маслозавод», найнижчий – у ТДВ «Баштанський сирзавод». Володіння інформацією про загальний рівень та окремі складові соціально-економічної безпеки підприємства дозволяє приймати обґрунтовані управлінські рішення. Важливо також те, що оцінку соціально-економічної безпеки підприємства можна здійснити у порівнянні з конкурентами на основі відкритої інформації, переважну більшість показників можна порівняти з показниками на рівні галузі, регіону, країни, розрахованими на основі статистичної інформації Державної служби статистики України.

Висновки

Забезпечення соціально-економічної безпеки підприємства є важливим для забезпечення ефективної діяльності і стабільного розвитку підприємства, особливо в умовах нестабільного середовища. Сформулювати відповідний рівень соціально-економічної безпеки підприємства можливо на основі розуміння її стану та напрямів покращення.

Науковці пропонують певний перелік підходів до оцінювання соціально-економічної безпеки підприємства. Застосування графоаналітичного методу дозволяє на основі аналітичної інформації наочно представити результати, що спрощує прийняття управлінського рішення.

Запропоновано адаптовану графоаналітичну модель для оцінки соціально-економічної безпеки підприємства, що включає оцінку за векторами: операційна безпека, соціальна безпека, техніко-технологічна безпека, фінансова безпека та інтегральну оцінку. Показники відібрані таким чином, щоб інформацію для їх розрахунку можна було отримати з відкритих джерел, зокрема з фінансової звітності підприємства. Використання моделі дозволяє оцінювати соціально-економічну безпеку підприємства в динаміці, в порівнянні з іншими суб'єктами ринку на

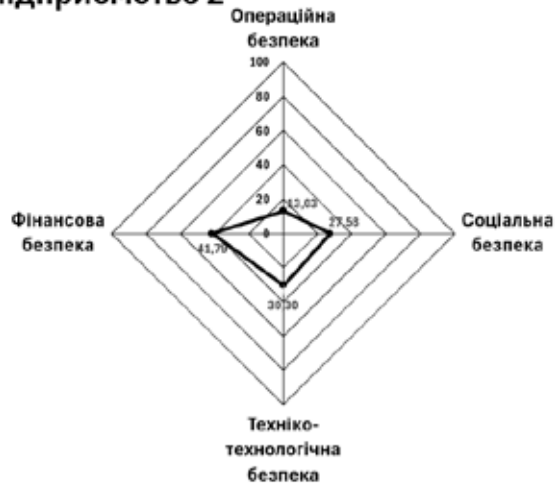
галузевому, регіональному рівні, і приймати обґрунтовані управлінські рішення щодо забезпечення сталого розвитку підприємства.

Подальші дослідження слід спрямувати на удосконалення інтегральної оцінки соціально-економічної безпеки підприємства.

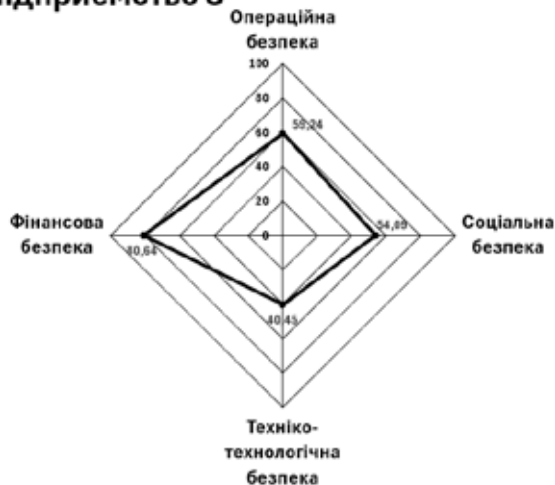
Підприємство 1



Підприємство 2



Підприємство 3



Підприємство 4

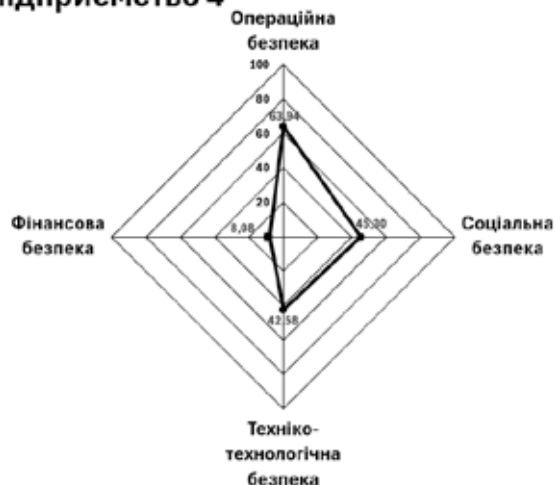


Рис. 2. Модель соціально-економічної безпеки підприємств групи компаній «Молочний альянс»

Джерело: побудовано на основі розрахунків авторів

Результати розрахунків загального рівня соціально-економічної безпеки підприємств відображено в таблиці 6.

Таблиця 6

Інтегральний показник соціально-економічної безпеки в динаміці

Підприємство	Підприємство 1	Підприємство 2	Підприємство 3	Підприємство 4
Інтегральний показник соціально-економічної безпеки	10721,78	1502,90	6715,62	2843,02

Джерело: розраховано авторами

Список використаної літератури

- Боса І. Етимологія поняття «економічної безпеки підприємства». *Вісник Хмельницького національного університету. Серія: Економічні науки*. 2022. № 6(2). Т2. С. 300-308. DOI: [https://doi.org/10.31891/2307-5740-2022-312-6\(2\)-50](https://doi.org/10.31891/2307-5740-2022-312-6(2)-50)
- Коленда Н.В. Поняття соціально-економічної безпеки підприємства. *Економіка і суспільство*. 2016. Випуск 7. С. 672-678.

3. Копитко М.І., Маланчук А.М. Дослідження сутності соціально-економічної безпеки підприємства в сучасних умовах розвитку. *Науковий вісник Львівського державного університету внутрішніх справ. Серія економічна*. 2023. № 1. С. 18-22. DOI: <https://doi.org/10.32782/2311-844X/2023-1-3>
4. Шубалий О.М., Бровченко Б.Ю. Оцінка соціально-економічної безпеки підприємств в Україні в умовах негативних екстерналій. *Економічні науки. Серія «Регіональна економіка»*. 2024 № 21 (83). Том 1. С. 489-495. DOI: [https://doi.org/10.36910/2707-6296-2024-21\(83\)-57](https://doi.org/10.36910/2707-6296-2024-21(83)-57)
5. Жанна Шило. Методика оцінювання рівня економічної безпеки підприємства. *Міжнародний науковий журнал з менеджменту, економіки та фінансів*. 2022. № 4. Том 1. С. 17-25. DOI: <https://doi.org/10.46299/j.isjmef.20220104.03>
6. Субота М. В. Методологія оцінювання соціально-економічної безпеки сталого розвитку суб'єктів господарської діяльності. *Ефективна економіка*. 2019. № 12. DOI: <https://doi.org/10.32702/2307-2105-2019.12.101>
7. Лагодієнко В. В., Суханова А. В. Характеристика економічної безпеки підприємств та оцінка основних індикаторів безпеки. *Вісник Харківського національного технічного університету сільського господарства. «Економічні науки»*. 2019. Вип. 206 С. 318-333. URL: <https://repo.btu.kharkov.ua/handle/123456789/5105>
8. Федонін О.С., Репіна І.М., Олексюк О.І. Потенціал підприємства: формування та оцінка : навч. посіб. Київ : КНЕУ, 2004. 316 с.
9. Хобта В.М., Башкевич Ю.В. Обґрунтування напрямків розвитку потенціалу підприємства. *Вісник Хмельницького національного університету*. 2011. № 6. Т4. С. 306-310. URL: https://journals.khnu.km.ua/vestnik/pdf/ekon/2011_6_4/306-310.pdf
10. Гончарук О.В. Бузовська Г.О. Методичні положення графоаналітичної оцінки економічного потенціалу підприємства. *Глобальні та національні проблеми економіки*. 2016. Випуск 10. С. 302-307. URL: <http://global-national.in.ua/archive/10-2016/64.pdf>
11. Строченко Н.І., Ковальова О.М. Формування напрямів підвищення ефективності використання ресурсного потенціалу на основі його оцінки. *Приазовський економічний вісник*. 2023. Випуск 2(34). С. 42-49. DOI: <https://doi.org/10.32782/2522-4263/2023-2-8>
12. Мешкова-Кравченко Н.В. Тарасюк А.В. Оцінка економічної безпеки підприємства. *Вісник Херсонського національного технічного університету*. 2021. № 1(76). С. 204-212. DOI: <https://doi.org/10.35546/kntu2078-4481.2021.1.26>
13. Офіційний сайт Державної служби статистики України. URL: <https://www.ukrstat.gov.ua/> (дата звернення: 10.08.2025).
14. Офіційний сайт ТДВ «Яготинський маслозавод». URL: <https://milkalliance.com.ua/company/enterprises/yagotinskij-maslozavod/> (дата звернення: 15.08.2025).
15. Офіційний сайт ПрАТ «Молочний альянс». URL: <https://milkalliance.com.ua/> (дата звернення: 15.08.2025).

References

1. Bosa, I. (2022). Etymolohiia poniattia "Ekonomichna bezpeka pidpriyemstva" [Etymology of the concept of economic security of an enterprise]. *Herald of Khmelnytskyi National University. Economic Sciences*, 312(6(2)), 300-308. [https://doi.org/10.31891/2307-5740-2022-312-6\(2\)-50](https://doi.org/10.31891/2307-5740-2022-312-6(2)-50)
2. Kolenda N.V. (2016). Ponyattya sotsial'no-ekonomichnoyi bezpeky pidpriyemstva [The concept of socio-economic security of an enterprise]. *Economy and Society*. 7, 672-678.
3. Kopytko, M., & Malanchuk, A. (2023). Doslidzhennya sutnosti sotsial'no-ekonomichnoyi bezpeky pidpriyemstva v suchasnykh umovakh rozvytku [Study of the essence of socio-economic security of an enterprise in the modern conditions of development]. *Scientific Journal of Lviv State University of Internal Affairs. Economics*. (1), 18-22. <https://doi.org/10.32782/2311-844X/2023-1-3>
4. O.M., Shubalyi & B.Yu., Brovchenko. (2025). Otsinka sotsial'no-ekonomichnoyi bezpeky pidpriyemstv v ukrayini v umovakh nehatyvnykh eksternaliy [Assessment of the socio-economic security of enterprises in Ukraine in the context of negative externalities]. *Economic Sciences, Series: Regional Economics*. (1), 489-495. [https://doi.org/10.36910/2707-6296-2024-21\(83\)-57](https://doi.org/10.36910/2707-6296-2024-21(83)-57)
5. Shylo, Zhanna. (2022). Metodyka kompleksnoyi otsinky rivnya ekonomichnoyi bezpeky pidpriyemstva [Methodology for comprehensive assessment of the level of economic security of an enterprise]. *International Science Journal of Management, Economics & Finance*. 1. 17-25. <https://doi.org/10.46299/j.isjmef.20220104.03>.
6. Subota, M. (2019). Metodolohiya otsynyuvannya sotsial'no-ekonomichnoyi bezpeky staloho rozvytku sub'yektiv hospodars'koyi diyal'nosti [Methodology of assessment of socio-economic provision of sustainable development of enterprises], *Efektivna ekonomika*, [Online], vol. 12. <https://doi.org/10.32702/2307-2105-2019.12.101>
7. Lahodiienko, V.V., Sukhanova, A.V. (2019). Kharakterystyka ekonomichnoyi bezpeky pidpriyemstv ta otsinka osnovnykh indyikatoriv bezpeky [Characteristics of the economic security of enterprises and evaluation of the main

security indicators]. *Visnyk Kharkivs'koho natsional'noho tekhnichnoho universytetu sil's'koho hospodarstva, Seriiia «Ekonomiczni nauky»*. 206, 318-333.

8. Fedonin O.S., Repina I.M. & Oleksiuk O.I. (2004). Potentsial pidpryyemstva: formuvannya ta otsinka [Enterprise potential: formation and assessment]: a textbook. Kyiv: KNEU. [in Ukrainian]

9. Khobta V.M. & Bashkevych Y.V. (2011). Obgruntuvannya napryamkiv rozvytku potentsialu pidpryyemstva [Justification of directions for the development of the enterprise's potential]. *Herald of Khmelnytskyi National University. Economic Sciences*. 6, 306-310.

10. Goncharuk O.V. & Buzovska G.O. (2016). Metodychni polozhennya hrafoanalytychnoyi otsinky ekonomichnoho potentsialu pidpryyemstva [Methodological provisions of graphic-analytical assessment of the economic potential of an enterprise]. *Global and national problems of the economy*. 10, 302-307

11. Strochenko N.I. & Kovaleva O.M. (2023) Formuvannya napryamiv pidvyshchennya efektyvnosti vykorystannya resursnoho potentsialu na osnovi yoho otsinky [Formation of directions for increasing the efficiency of resource potential use based on its assessment]. *Pryazovsky Economic Bulletin*. 2(34), 42-49. <https://doi.org/10.32782/2522-4263/2023-2-8>

12. Mieshkova-Kravchenko, N.V. & Tarasiuk, A.V. (2021). Otsinka ekonomichnoyi bezpeky pidpryyemstva [Enterprise's economic security assessment]. *Visnyk of Kherson National Technical University*. 76, 204-212. <https://doi.org/10.35546/kntu2078-4481.2021.1.26>.

13. State Statistics Service of Ukraine. URL: <https://www.ukrstat.gov.ua/> (accessed on 10.08.2025).

14. Yagotynsky Butter Plant. URL: <https://milkalliance.com.ua/en/company/enterprises/yagotinskij-maslozavod/> (accessed on 15.08.2025)

15. Group of companies "Milk Alliance". URL: <https://milkalliance.com.ua/en/> (accessed on 15.08.2025)

Дата першого надходження рукопису до видання: 27.11.2025

Дата прийнятого до друку рукопису після рецензування: 12.12.2025

Дата публікації: 31.12.2025

НОТАТКИ

ВІСНИК

Херсонського національного технічного університету

Відповідальний за випуск	головний редактор Литвиненко В. І. д.т.н., професор, завідувач кафедри інформатики і комп'ютерних наук, заслужений діяч науки і техніки України
Комп'ютерна верстка	Клименко Т. О.
Відповідальний секретар	Лур'є І. А. к.т.н., доцент, начальник навчально-методичного відділу, доцент кафедри інформатики і комп'ютерних наук

Реєстрація суб'єкта у сфері друкованих медіа: Рішення Національної ради України
з питань телебачення і радіомовлення № 1541 від 09.05.2024 року. Ідентифікатор медіа: R30-04565.

Суб'єкт у сфері друкованих медіа – Херсонський національний технічний університет
(вул. Інститутська, буд. 11, м. Хмельницький, 29016, olena.chepelyuk@kntu.net.ua, тел. +38 (0382) 77-35-65).

Періодичність виходу: 4 рази на рік.

Підписано до друку 31.12.2025.

Формат 60x84/8. Гарнітура Times New Roman. Папір офсет. Цифровий друк.
Ум. друк. арк. 41,39. Замов. № 0126/037. Наклад 100 прим.

Видавництво і друкарня – Видавничий дім «Гельветика»
65101, Україна, м. Одеса, вул. Інглезі, 6/1
Телефон +38 (095) 934 48 28, +38 (097) 723 06 08
E-mail: mailbox@helvetica.ua
Свідоцтво суб'єкта видавничої справи
ДК № 7623 від 22.06.2022 р.