

Національний університет
“Полтавська політехніка імені Юрія Кондратюка”

National University
“Yuri Kondratyuk Poltava Polytechnic”

Системи управління, навігації та зв'язку

Control, navigation and communication systems

Випуск 3 (85)

Issue 3 (85)

Щоквартальне видання

Засноване у 2007 році

У журналі відображені результати наукових досліджень з розробки та удосконалення систем управління, навігації та зв'язку у різних проблемних галузях.

Засновник і видавець:

Національний університет
“Полтавська політехніка імені Юрія Кондратюка”

Телефон:

+38 (050) 302-20-71

E-mail редколегії:

kuchuk56@ukr.net

Інформаційний сайт:

<http://journals.nupp.edu.ua/sunz>

Quarterly

Founded in 2007

Journal represent the research results on the development and improvement of control, navigation and communication systems in various areas

Founder and publisher:

National University
“Yuri Kondratyuk Poltava Polytechnic”

Phone:

+38 (050) 302-20-71

E-mail of the editorial board:

kuchuk56@ukr.net

Information site:

<http://journals.nupp.edu.ua/sunz>

*Затверджений до друку Вченою Радою Національного університету
“Полтавська політехніка імені Юрія Кондратюка” (протокол від 18 вересня 2026 року № 12).*

*Ідентифікатор медіа R30-04135 згідно з рішенням Національної ради України
з питань телебачення і радіомовлення від 25.04.2024 № 1416*

*Включений до “Переліку наукових фахових видань України, в яких можуть публікуватися
результати дисертаційних робіт на здобуття наукових ступенів доктора наук, кандидата наук
та ступеня доктора філософії” до категорії Б з 01.06.2026 року
(в редакції наказу Міністерства освіти і науки України від 19 січня 2026 року № 56).*

Кластер: Інформаційні технології та електроніка.

Шифри спеціальностей: F3, F5, F6, F7, G5, G7.

Полтава • 2026

Редакційна колегія

Головний редактор:

КОСЕНКО Віктор Васильович
(*д-р техн. наук, проф., Полтава, Україна*).

Заступник головного редактора:

ШЕФЕР Олександр Віталійович
(*д-р техн. наук, проф., Полтава, Україна*).

Члени редакційної колегії:

ГАШИМОВ Ельшан Гіяс огли
(*д-р наук, проф., Баку, Азербайджан*);

ГОПЕЕНКО Вікторс
(*д-р інжен. наук, проф., Рига, Латвія*);

КОВАЛЕНКО Андрій Анатолійович
(*д-р техн. наук, проф., Харків, Україна*);

КУЧУК Георгій Анатолійович
(*д-р техн. наук, проф., Харків, Україна*);

ЛЕВЧЕНКО Лариса Олексіївна
(*д-р техн. наук, проф., Київ, Україна*);

МИРОНЦОВ Микита Леонідович
(*д-р фіз.-мат. наук, проф., Київ, Україна*);

СЕМЕНОВ Сергій Геннадійович
(*д-р техн. наук, проф., Краків, Польща*);

ЯНКО Аліна Сергіївна
(*канд. техн. наук, доц., Полтава, Україна*).

Відповідальний секретар:

ЗАХАРЧЕНКО Руслан Володимирович
(*канд. техн. наук, доц., Полтава, Україна*).

Editorial board

Editor-in-Chief:

Viktor KOSENKO
(*Dr. Sc. (Tech.), Prof., Poltava, Ukraine*).

Associate editor:

Oleksandr SHEFER
(*Dr. Sc. (Tech.), Prof., Poltava, Ukraine*).

Editorial board members:

Elshan Giyas oglu HASHIMOV
(*Dr. Sc., Prof., Baku, Azerbaijan*);

Viktors GOPEJENKO
(*Dr. Sc. (Tech.), Prof., Riga, Latvia*);

Andriy KOVALENKO
(*Dr. Sc. (Tech.), Prof., Kharkiv, Ukraine*);

Heorhii KUCHUK
(*Dr. Sc. (Tech.), Prof., Kharkiv, Ukraine*);

Larysa LEVCHENKO
(*Dr. Sc. (Tech.), Prof., Kyiv, Ukraine*);

Mykyta MYRONTSOV
(*Dr. Sc. (Phys. & Math.), Prof., Kyiv, Ukraine*);

Serhii SEMENOV
(*Dr. Sc. (Tech.), Prof., Krakow, Poland*);

Alina YANKO
(*PhD (Tech.), Ass. Prof., Poltava, Ukraine*).

Responsible secretary:

Ruslan ZAKHARCHENKO
(*PhD (Tech.), Ass. Prof., Poltava, Ukraine*).

З М І С Т

НАВІГАЦІЯ ТА ГЕОІНФОРМАЦІЙНІ СИСТЕМИ

<i>Бойчун С. О., Левицький С. М.</i> Автономна пасивна навігація безпілотної наземного транспортного засобу за втрати супутникового сигналу: порівняння мікроелектромеханічного та волоконно-оптичного гіроскопів	5
---	---

УПРАВЛІННЯ В СКЛАДНИХ СИСТЕМАХ

<i>Невлюдов І. Ш., Євсєєв В. В., Чала О. О., Горовий К. Ю.</i> Розроблення інтелектуального метода контролю якості PAD контактів цифрового двійника FEK DELVOTEC 6400 для моделювання процесу клинового зварювання у мікроелектроніці	11
<i>Рудаков І. С.</i> Метод та модель планування маршруту польоту групи безпілотної літальних апаратів для уникнення перешкод у міській забудові на основі технологій машинного навчання	22
<i>Слободяник О. Ю.</i> Метод підвищення ефективності автоматизованого аналізу PQ-даних показників якості електроенергії	31
<i>Сукач С. В., Лашко Є. Є., Ченчева О. О., Резнік Д. В., Петренко І. С.</i> Синтез трифакторної регресійної моделі пилопоглинання у повітряному середовищі виробничих приміщень	37
<i>Усенко Д. В., Биба В. В., Пінчук Н. М.</i> Математичне моделювання динамічної теплової реакції багатопарової стінової конструкції за періодичного та імпульсного теплового навантаження	47

ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

<i>Barkovska O.</i> Model and method of inclusive interaction with visual and textual content based on contextual gesture interpretation	57
<i>Бондаренко М. Е., Іващенко Г. С.</i> Гібридизація методів вилучення ознак у системі голосової ідентифікації користувача	68
<i>Дрючко О. Г., Бунякіна Н. В., Шефер О. В., Галай В. М., Захарченко Р. В.</i> Особливості синтезу і характерні властивості перовскітоподібних шаруватих оксидних матеріалів рідкісноземельних і перехідних елементів як фундаментальні аспекти для практики інноваційних фотокаталітичних рішень	78
<i>Diachenko D., Shostak M.</i> Method for assessing the risk of defect manifestation in .NET microservices in distributed computer systems based on operational metrics	88
<i>Kapiton A., Franchuk T., Tyshchenko D., Zakharov R.</i> Development of decentralized forecasting systems based on multi-cloud infrastructure	100
<i>Коваленко А. А., Кучук Г. А.</i> Метод адаптивної кластеризації мобільних вузлів туманного шару Інтернету речей	106
<i>Lashyn Ya., Voitko O., Fedoriienko V.</i> Hybrid neuro-mathematical model for detecting information threats: integrating large language models and modified diffusion processes	115
<i>Skorlupin O., Podorozhniak A.</i> YOLO family architectures for explosive object detection in embedded systems of autonomous unmanned vehicles: a comparative analysis	123
<i>Шостак А. В.</i> Експериментальне дослідження методів пошуку в АВЛ-дереві з урахуванням ієрархії пам'яті	131

КІБЕРБЕЗПЕКА ТА ЗАХИСТ ІНФОРМАЦІЇ

<i>Hashimov E., Akhundov R., Talibov A., Radovanović M.</i> A risk- and time-based model for processing information flows in the physical protection of special-purpose and critical infrastructure facilities	139
<i>Горносталь О. А., Челак В. В.</i> Диверсифіковані ансамблі класифікаторів для виявлення дефектів програмного коду за статичними метриками якості	150
<i>Derkach T. M., Dmytrenko A. O., Klochko L. A.</i> A comprehensive approach to ensuring the cybersecurity of corporate network infrastructure	161

ЗВ'ЯЗОК, ТЕЛЕКОМУНІКАЦІЇ ТА РАДІОТЕХНІКА

<i>Гічко Ю. Г., Мазор С. Ю., Храновська Т. В., Криховецький Г. Я.</i> Методика експериментального вимірювання основних параметрів антен надвисокочастотного діапазону	168
<i>Zhuchenko O.</i> Discrete integer model of complex time-division signals with a periodic structure	172
<i>Коваленко О. О., Почерняєв В. М., Івко С. О.</i> Модифікований метод неортогонального множинного доступу з парціальним спектральним кодуванням для мобільних вузлових цифрових тропосферних станцій	179

<i>Кучук Н. Г., Пироженко С. С. Методика розрахунку оверхеду live-міграції віртуальних машин на базі концепції ресурсного куба NRC</i>	189
<i>Самборський І. І., Залужний О. В., Лозова Н. Т. Метод логіко-динамічного управління функціональною стійкістю електронної комунікаційної мережі на основі цифрового двійника</i>	195
АЛФАВІТНИЙ ПОКАЖЧИК	204

За достовірність викладених фактів, цитат та інших відомостей відповідальність несе автор

Журнал індексується міжнародними наукометричними базами: Index Copernicus (ICV = 85.62), General Impact Factor, Google Scholar, Academic Resource Index, Scientific Indexed Service

Організації авторів / Authors affiliation

Азербайджан	Azerbaijan
Азербайджанський технічний університет, Баку, Азербайджан	Azerbaijan Technical University, Baku, Azerbaijan
Університет національної оборони, Баку, Азербайджан	National Defense University, Baku, Azerbaijan
Італія	Italy
Політехнічний університет Барі, Барі, Італія	Politecnico di Bari, Bari, Italy
Сербія	Serbia
Університет оборони, Белград, Сербія	University of Defence in Belgrade, Belgrade, Serbia
Швейцарія	Switzerland
Геологічне бюро BEG SA, Швейцарія	Geological Bureau BEG SA, Switzerland
Україна	Ukraine
Військовий інститут телекомунікацій та інформатизації імені Героїв Крут, Київ, Україна	Kruty Heroes Military Institute of Telecommunications and Information Technology, Kyiv, Ukraine
Державний торговельно-економічний університет, Київ, Україна	State University of Trade and Economics, Kyiv, Ukraine
Інститут фізики напівпровідників імені В. Є. Лашкарьова НАН України, Київ, Україна	V. E. Lashkaryov Institute of Semiconductor Physics, NAS of Ukraine, Kyiv, Ukraine
Кременчуцький національний університет імені Михайла Остроградського, Кременчук, Україна	Kremenchuk Mykhailo Ostrohradskyi National University, Kremenchuk, Ukraine
Науково-дослідний інститут воєнної розвідки, Київ, Україна	Research Institute of Military Intelligence, Ukraine
Національний аерокосмічний університет «ХАІ», Харків, Україна	National Aerospace University "KHAI", Kharkiv, Ukraine
Національна академія Служби безпеки України, Київ, Україна	National Academy of the Security Service of Ukraine, Kyiv, Ukraine
Національний технічний університет «Харківський політехнічний інститут», Харків, Україна	National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine
Національний технічний університет України «КПІ імені Ігоря Сікорського», Київ, Україна	National Technical University "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine
Національний університет оборони України, Київ, Україна	National Defence University of Ukraine, Kyiv, Ukraine
Національний університет «Полтавська політехніка імені Юрія Кондратюка», Полтава, Україна	National University "Yuri Kondratyuk Poltava Polytechnic", Poltava, Ukraine
Харківський національний університет радіоелектроніки, Харків, Україна	Kharkiv National University of Radio Electronics, Kharkiv, Ukraine
Український державний університет залізничного транспорту, Харків, Україна	Ukrainian State University of Railway Transport, Kharkiv, Ukraine

Навігація та геоінформаційні системи

УДК: 629.5-052.001.5; 004.896; 681.518.3

doi: 10.26906/SUNZ.2026.3.005

С. О. Бойчун, С. М. Левицький

Інститут фізики напівпровідників імені В. Є. Лашкарьова НАН України, Київ, Україна

АВТОНОМНА ПАСИВНА НАВІГАЦІЯ БЕЗПІЛОТНОГО НАЗЕМНОГО ТРАНСПОРТНОГО ЗАСОБУ ЗА ВТРАТИ СУПУТНИКОВОГО СИГНАЛУ: ПОРІВНЯННЯ МІКРОЕЛЕКТРОМЕХАНІЧНОГО ТА ВОЛОКОННО-ОПТИЧНОГО ГІРОСКОПІВ

Анотація. Актуальність. Безпілотний наземний транспортний засіб (далі – наземний робот) за втрати глобальної навігаційної супутникової системи (Global Navigation Satellite System, GNSS) потребує автономного визначення курсу та сприйняття перешкод. **Предмет дослідження:** навігація наземного робота за поєднання колісної одометрії, мікроелектромеханічного гіроскопа (MEMS-гіроскоп), волоконно-оптичного гіроскопа (ВОГ), камери, пасивних маркерів і далекоміра. **Мета дослідження:** кількісно оцінити точність базового ВОГ та порівняти пасивні й гібридні конфігурації за різної видимості. **Завдання дослідження.** Реалізувати параметричну модель ВОГ; відокремити її від зорового каналу; перевірити статичні, рухомі й нові перешкоди; сформувати відтворювану матрицю запусків. **Методи.** Виконано 1440 проходжень полігона у 48 конфігураціях і 180 додаткових 300-секундних випробувань курсу. Базовий ВОГ задано стабільністю зміщення $1^\circ/\text{год}$, випадковим блуканням кута $0,04^\circ/\sqrt{\text{год}}$ і похибкою масштабного коефіцієнта $250 \cdot 10^{-6}$. **Результати.** ВОГ забезпечив середню середньоквадратичну похибку курсу $0,0568^\circ$ і кінцеву похибку $0,0995^\circ$ проти $9,8981^\circ$ та $17,1355^\circ$ для MEMS-сценарію. За ясної видимості волоконно-оптична конфігурація з маркерами досягла цілі у 46,7% запусків, із них 23,3% – без зіткнень; з далекоміром отримано 76,7 та 40,0% відповідно. **Висновки.** ВОГ є найточнішим пасивним засобом курсової орієнтації у випробуваній моделі й не залежить від видимості. Проте блокування камери позбавляє систему сприйняття перешкод, тому практична архітектура має поєднувати ВОГ із зоровими засобами за сприятливих умов та з окремим резервом сприйняття за туману, снігу або диму.

Ключові слова: безпілотний наземний транспортний засіб; пасивна навігація; волоконно-оптичний гіроскоп; MEMS-гіроскоп; супутникова навігація; оптичні маркери; злиття вимірювань; далекомір; відтворюване моделювання.

1. Вступ

1.1 Актуальність тематики

Навігаційна стійкість наземного робота визначається не лише точністю координат, а й здатністю безпечно продовжити рух після втрати зовнішнього абсолютного орієнтира. У міській забудові, приміщеннях, під навісами та за навислих або природних завад супутникова навігація може стати недоступною. Тоді одометрія та інерціальні вимірювання підтримують безперервність, але їх похибки накопичуються; стохастичні складові гіроскопа доцільно характеризувати дисперсією Аллана [1]. ВОГ вимірює кутову швидкість за ефектом Саньяка і не потребує видимості зовнішньої сцени. Його слід відрізнити від зорової одометрії, яка оцінює рух за послідовністю кадрів. Зорово-інерціальні системи та штучні оптичні орієнтири корисні для корекції положення [2, 3], але дощ, туман, сніг і дим погіршують саме зовнішній зоровий канал [4, 5]. Практична архітектура повинна окремо контролювати точність орієнтації та доступність сприйняття перешкод.

1.2 Аналіз літератури і постановка проблеми

Сучасні дослідження навігації за відсутності GNSS групують доступні підходи у відносно інерціально-одометричне оцінювання, зорову та зорово-інерціальну локалізацію, LiDAR-інерціальну

одометрію, прив'язку до карт і мультимодальне злиття [6]. Для наземних роботів вибір методу визначається не лише середньою точністю, а й стійкістю до слабкої текстури, змін освітлення, динамічної рослинності та інших ознак неструктурованого середовища [7].

Зорово-інерціальні системи VINS-Mono та ORB-SLAM3 показують можливість обмежувати дрейф за рахунок камери та інерціального модуля, але залежать від калібрування, якості ознак і доступності зображення [8, 9]. Набір EuRoC демонструє значення синхронізованих сенсорів і незалежної еталонної траєкторії для перевірки таких систем [3]. Порівняння LIO-, VIO- та резервованих IMU-конфігурацій підтверджує, що зовнішні геометричні оновлення можуть істотно стримувати накопичення інерціальної похибки під час тривалого блокування GNSS, однак їх результат залежить від властивостей сцени та якості калібрування [10].

Для наземних транспортних засобів сучасні адаптивні моделі низьковартісних IMU використовують обмеження руху, псевдовимірювання та оцінювання шуму, щоб продовжити автономну ділянку маршруту [11, 12]. Волоконно-оптичні гіроскопи формують інший клас рішення: вони забезпечують пасивне вимірювання кутової швидкості за ефектом Саньяка й не потребують зовнішньої сцени. Експериментальні роботи 2025 року підтверджують навігаційний клас інтерферометричних ВОГ та прогрес їх термічної

стабільності [13], а огляд 2026 року систематизує архітектури, джерела похибок і застосування ВОГ у сучасних інерціальних системах [14]. Інтеграція оптичного передавача і приймача на кремнієвій фотонній платформі додатково показує напрям мініатюризації навігаційних ВОГ [15].

У відібраному корпусі праць точність курсового інерціального каналу, доступність зовнішнього сприйняття та безпечне завершення маршруту переважно оцінюються окремо. Це залишає відкритим питання, як точний пасивний ВОГ впливає на результат наземного робота, коли камера деградує, а планувальник отримує лише фактично доступні сліди перешкод. Тому в цій роботі ВОГ відокремлено від зорового каналу, а модифікований планувальник за принципом VFH використовує тільки локальні спостереження, а не істинний список об'єктів [16].

1.3 Мета і завдання дослідження

Мета роботи – кількісно оцінити точність ВОГ та порівняти автономні конфігурації наземного робота за повної втрати супутникової навігації, різної видимості й різної якості програмного злиття.

Для досягнення мети поставлено такі завдання:

- реалізувати взаємозамінні модулі оцінювача стану та локального планувальника;
- порівняти одометрію, МЕМС-гіроскоп, ВОГ та їх злиття з пасивними маркерними корекціями;
- відокремити похибки ВОГ від деградації камери за ясної, обмеженої та заблокованої видимості;
- оцінити реакцію на статичні, рухомі та нові перешкоди з далекоміром і без нього;
- зберегти первинні журнали, сесії, цифровий відбиток моделі й перевірити детерміноване відтворення.

2. Матеріали і методи дослідження

2.1 Архітектура симулятора

Розроблено двовимірний (2D) браузерний симулятор без зовнішніх програмних залежностей. Істинний стан містить положення $\mathbf{p} = [x, y]^T$, курс ψ і швидкість v . Оцінювач стану отримує одометрію та вибрані спостереження; планувальник – оцінений стан і локальні сліди перешкод. Такий інтерфейс дає змогу замінити оцінювач або планувальник без зміни моделі світу та журналювання.

2.2 Моделі сенсорів і корекції

Дискретний прогноз положення задається як

$$\hat{\mathbf{p}}_{k+1} = \hat{\mathbf{p}}_k + \hat{v}_k [\cos \hat{\psi}_k, \sin \hat{\psi}_k]^T \Delta t + \boldsymbol{\varepsilon}_p.$$

МЕМС-модель містить масштабну похибку, повільне зміщення та шумові складові. Їх значення є сценарними й не подаються як калібрування конкретного виробу [1]. Для ВОГ кутовий приріст:

$$\Delta \hat{\psi} = \Delta \psi (1 + k_s) + b \Delta t / 3600 + N \sqrt{\Delta t / 3600} \xi,$$

де k_s – масштабна похибка, b – зміщення у градусах за годину, N – коефіцієнт випадкового блукання кута, ξ – детерміновано відтворювана нормальна випадкова величина.

Схема обміну даними між середовищем, каналами спостереження, оцінювачем, планувальником і моделлю руху (архітектура навігації наземного робота) наведена на рис. 1.



Рис. 1. Обмін даними між середовищем, каналами спостереження, оцінювачем, планувальником і моделлю руху

За основу взято цифрову версію ВОГ класу EMCORE EG-120 при 25 °C [17]. Термін «частин на мільйон» (parts per million, ppm) у технічному описі переведено в безрозмірну величину.

Досліджені конфігурації навігаційних засобів наведені у табл. 1, а параметри базового волоконно-оптичного гіроскопа – у табл. 2.

Таблиця 1 – Досліджені конфігурації навігаційних засобів

Конфігурація	Орієнтація	Корекція або сприйняття
Одометрична	колісна одометрія	немає
МЕМС	МЕМС + одометрія	немає
Волоконно-оптична	ВОГ + одометрія	5 пасивних маркерів
Злита	МЕМС + ВОГ + одометрія	5 пасивних маркерів
Злита з далекоміром	як у злитій	маркери + далекомір 1 Гц

Технічний опис указує смугу цифрового ВОГ 250 Гц, тоді як модель використовує 20 Гц через крок інтегрування. Передбачається початкове нульове калібрування; температурний дрейф, вібраційна чутливість, похибка встановлення та час прогріву не моделюються. Отже, параметри відтворюють

сприятливий базовий режим, а не повний температурний паспорт виробу.

Таблиця 2 – Параметри базового волоконно-оптичного гіроскопа

Параметр	Значення	Одиниця	Використання в моделі
Стабільність зміщення	1	°/год	нормальна стала похибка запуску
Випадкове блукання кута	0,04	°/√год	білий шум приросту
Масштабна похибка	$250 \cdot 10^{-6}$	безрозмірна	похибка кутového приросту
Частота моделі	20	Гц	крок 0,05 с

Видимий маркер у полі зору створює корекцію

$$x_c = x_p + K[z - h(x_p)],$$

де x_c – скоригований стан, x_p – прогноз стану, а K – коефіцієнт корекції, заданий профілем програмного забезпечення. Маркерна корекція впливає на курс і положення, але не змінює фізичні параметри ВОГ.

2.3 Видимість та якість програмного забезпечення

За ясної видимості камера має дальність 28 м і поле зору 70°. Обмежена видимість моделює дощ або легкий туман: дальність камери зменшено до 62%, імовірність виявлення маркера – до 0,72, а шум спостереження збільшено. Повне блокування моделює густий туман, сніг або дим: камера і маркери вимкнені, а далекомір має меншу дальність і надійність [4, 5]. ВОГ і МЕМС-гіроскоп у всіх трьох профілях залишаються доступними.

Профілі програмного забезпечення названо базовим, типовим і стійким до похибок. Останній відповідає англійському поняттю robust software. У роботі не використовується калька «робастний»: українське формулювання означає здатність алгоритму зберігати прийнятну роботу за шуму, пропусків і хибних спостережень. Профілі одночасно змінюють пам'ять сліду, силу корекції, відбраковування вимірювань і вагу ВОГ у злитті, але не є оцінкою конкретного комерційного виробу.

2.4 Локальне уникнення і нові перешкоди

На кожному кроці сліди перетворюються на полярну гістограму $H(\theta)$. Позначивши похибки відносно цільового та попереднього курсів як e_g і e_s , серед вільних секторів обирається

$$\theta_{opt} = \operatorname{argmin}_{\theta} [w_1|e_g| + w_2|e_s| + w_3H(\theta)].$$

Критерій підтримує рух до цілі, обмежує різкі зміни керування й відхиляє напрями з великою густиною перешкод [16]. Якщо далекомір тривалий час підтверджує фронтальне блокування, контролер виконує відступ і повторний вибір сектора.

Сценарій містить чотири статичні та дві рухомі перешкоди. Ще дві рухомі перешкоди з'являються на 9-й та 25-й секундах. Реєструються перше зорове або далекомірне виявлення та затримка після появи.

2.5 План експерименту і показники

Полігон становив 100×100 м, швидкість – 8 км/год, крок – 0,05 с, граничний час – 300 с. Основна матриця охопила 5 наборів $\times$ 3 профілі видимості $\times$ 3 профілі ПЗ = 45 конфігурацій із втратою супутникової навігації. Додано три контрольні конфігурації з доступною супутниковою навігацією. Кожну конфігурацію виконано на 30 однакових сесіях, усього 1 440 запусків. Окремий стенд курсу містив 180 запусків: МЕМС, ВОГ і їх злиття, по 30 зерен за ясного та блокованого зорового каналу. Протягом 300 с кутова швидкість дорівнювала $\omega(t) = 1,5 + 6\sin(2\pi t/35)$ °/с; супутникові й маркерні корекції були вимкнені. Основні показники: досягнення цілі; безпечне досягнення без зіткнень; середньоквадратична похибка (СКП) положення і курсу; мінімальна відстань до перешкоди; час і довжина шляху; затримка виявлення. Для часток обчислено 95% інтервали Вілсона [18].

3. Результати

3.1 Точність волоконно-оптичного гіроскопа

На 300-секундному стенді ВОГ мав середню СКП курсу $0,0568 \pm 0,0349^\circ$ і кінцеву похибку $0,0995 \pm 0,0624^\circ$ (рис. 2). Для МЕМС-сценарію відповідні значення становили $9,8981 \pm 5,7123^\circ$ і $17,1355 \pm 9,8844^\circ$. Отже, середня СКП ВОГ була приблизно у 174 рази меншою. Злиття з фіксованою вагою дало $0,5869 \pm 0,3766^\circ$, тобто поступилося окремому ВОГ через домішування МЕМС-похибки.

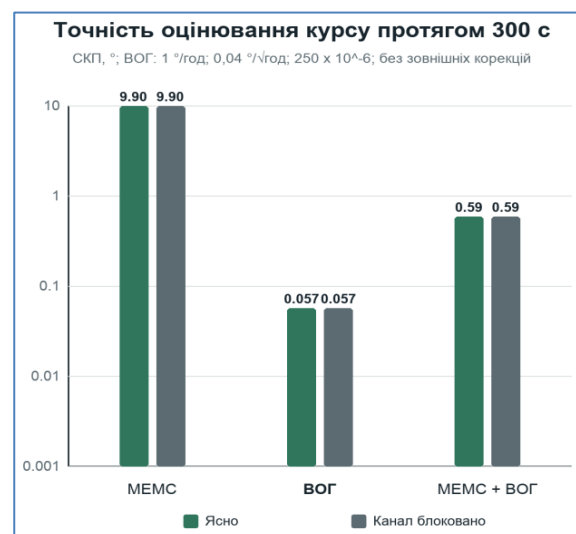


Рис. 2. Середня СКП курсу без зовнішніх корекцій; логарифмічна шкала

Результати ВОГ за ясної та заблокованої видимості збіглися для кожної сесії. Це контрольна ознака правильної моделі: погода впливає на зовнішній зоровий канал, але не на інерціальне вимірювання ВОГ. Водночас така точність не враховує температурний дрейф і похибку встановлення.

3.2 Порівняння на полігоні (рис. 3, табл. 3)

За ясної видимості та стійкого до похибок ПЗ одометрія досягла цілі в 1 із 30 запусків: 3,3% (95% інтервал 0,6–16,7%), МЕМС – у 23,3% (7/30).

Волоконно-оптична конфігурація дала 46,7% (14/30; 95% інтервал 30,2–63,9%) досягнень і 23,3% (7/30) безпечних досягнень; злиття - 43,3% (13/30) і 23,3% (7/30). СКП положення становила 4,02 і 4,13 м відповідно проти 23,44 м для одометрії.

Додавання далекоміра підвищило досягнення до 76,7% (23/30; 95% інтервал 59,1–88,2%), а безпечне досягнення – до 40,0% (12/30). За обмеженої видимості злиття дало 36,7% досягнень і 16,7% безпечних досягнень, а варіант із далекоміром - 56,7% та 30,0%. Прибуття до цілі не можна використовувати як єдиний показник безпеки.

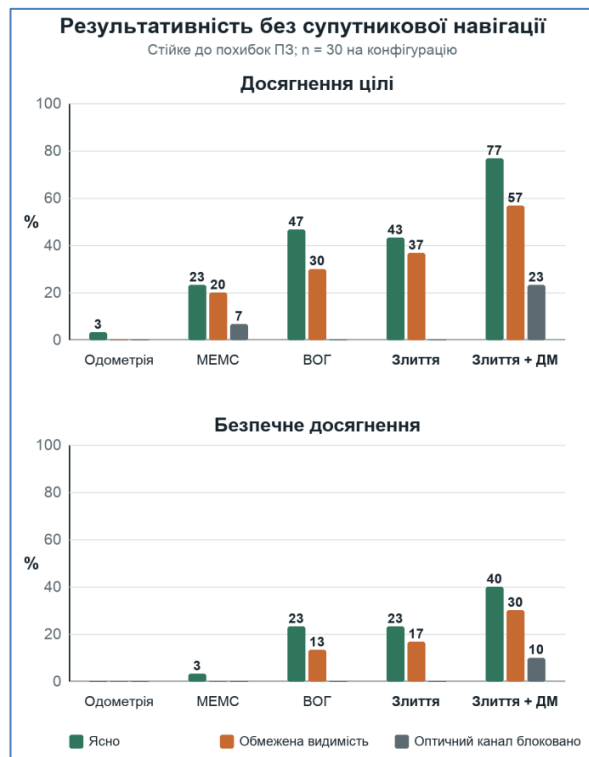


Рис. 3. Досягнення цілі та безпечне досягнення за стійкого до похибок ПЗ; ДМ - далекомір

Таблиця 3 – Злиття та варіант із далекоміром за стійкого до похибок ПЗ

Видимість	Злиття: ціль / безпечно; СКП	З далекоміром: ціль / безпечно; СКП
Ясна	43,3 / 23,3%; 4,13 м	76,7 / 40,0%; 3,12 м
Обмежена	36,7 / 16,7%; 4,37 м	56,7 / 30,0%; 3,76 м
Блокована	0 / 0%; 5,11 м	23,3 / 10,0%; 4,34 м

3.3 Видимість і якість програмного забезпечення (рис. 4)

За блокованої видимості окремий ВОГ зберіг СКП курсу 0,1046° у полігонному сценарії, але не досяг цілі в жодному запуску. Причина полягає не у втраті курсу, а у відсутності зорових слідів перешкод і маркерних корекцій положення. Далекомір підвищив досягнення до 23,3%, а безпечне - до 10,0%.

За ясної видимості стійке до похибок ПЗ підвищило пасивне злиття від 20,0% до 43,3% досягнень, а безпечний результат із далекоміром - від 3,3% до 40,0%. Водночас за блокованої видимості

типовий профіль поступився базовому за окремими показниками. Це наслідок взаємодії пам'яті слідів, корекцій і траєкторії; назви профілів не слід трактувати як ранжування готових програмних продуктів.

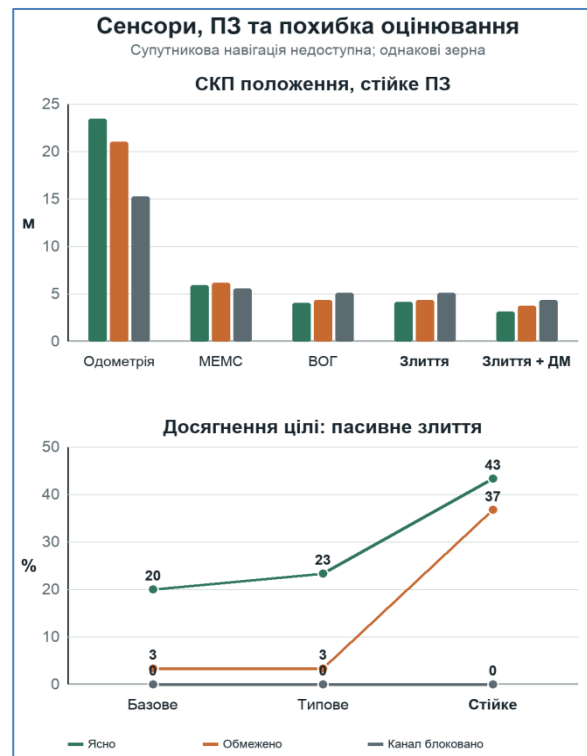


Рис. 4. СКП положення та досягнення цілі залежно від конфігурації й профілю ПЗ

3.4 Виявлення нових перешкод

Стойке пасивне злиття виявляло обидві нові перешкоди у всіх запусках за ясної та обмеженої видимості із середньою затримкою 2,02 і 2,18 с. Далекомір зменшив її до 1,61 і 1,93 с. За заблокованої камери лише далекомір зберіг канал виявлення: у середньому 1,67 із двох об'єктів із затримкою 12,75 с. Вузький промінь 1 Гц не забезпечує суцільного огляду.

4. Обговорення результатів

4.1 Переваги й обмеження засобів

Одометрія проста й безперервна, але її позиційний і курсовий дрейф обмежує самостійне використання. MEMC-гіроскоп компактний і не залежить від видимості, однак зміщення накопичується [1, 19]. ВОГ забезпечує значно точніший курс; для класу використаного технічного опису вказано масу до 0,08 кг і типову потужність 1,8 Вт за кімнатної температури [17]. У моделі енергоспоживання та маса не враховувалися.

ВОГ є пасивним щодо зовнішнього середовища: світло циркулює всередині приладу, але навігаційний сигнал назовні не випромінюється. Водночас він не визначає абсолютне положення і не бачить перешкод. Камера та маркери надають ці обмеження без активного опромінення сцени, але залежать від погоди, освітлення й перекриття [2–5].

Далекомір є керованим активним резервом. Він підвищує частоту досягнення, але вузький промінь,

частота 1 Гц і помилки планувальника залишають зіткнення. Тому безпечність залежить від оглядовості, прогнозування рухомих об'єктів, кінематики та правила безпечної зупинки.

4.2 Найкращий сценарій і практична архітектура

Для суворо пасивної роботи за ясної або обмеженої видимості доцільним ядром є ВОГ, одометрія, камери та пасивні маркери зі стійким до похибок оцінюванням. MEMC слід залишити незалежним резервом і використовувати адаптивну, а не сталу вагу: точніший ВОГ у стенді погіршувався після домішування MEMC-вимірювання. За густого туману, снігу або диму ВОГ зберігає курс, але потрібен окремий засіб сприйняття: пасивний тепловізійний канал, якщо це допускають умови, або активний радар чи далекомір. Якщо достовірного сліду перешкоди немає, безпечна зупинка має вважатися штатним результатом, а не невдачею планувальника.

4.3 Достовірність і загрози валідності

Інтерактивний і пакетний режими використовують той самий код. Матриця задана до агрегації, невдалі запуски не вилучено, зерна спільні, первинні журнали збережено, а 11 672 незалежні автоматизовані перевірки не виявили порушень. Така процедура забезпечує внутрішню відтворюваність порівнянь. Зовнішня достовірність обмежена. Модель двовимірна й кінематична; вона не враховує рельєф, ковзання, вібрацію, температурний цикл ВОГ, затримки, електромагнітні завади та енергоспоживання. Параметри ВОГ узяті з технічного опису, але не перевірено на придбаному зразку. Тому робота підтримує вибір архітектури, але не підтверджує готовність до експлуатації.

5. Висновки

1. Параметрична модель базового ВОГ із 1 °/год, 0,04 °/год і масштабною похибкою $250 \cdot 10^{-6}$ дала середню СКП курсу 0,0568° за 300 с, приблизно у 174 рази меншу за MEMC-сценарій.

2. ВОГ не залежав від ясності зовнішньої сцени, однак блокування камери різко знижувало безпечність через втрату сприйняття перешкод і маркерних корекцій положення.

3. За ясної видимості волоконно-оптична конфігурація з маркерами досягла 46,7% цілей і 23,3% безпечних досягнень; злиття дало 43,3% і

23,3%, а далекомір підвищив ці значення до 76,7% і 40,0%. За обмеженої видимості для злиття та далекоміра отримано 36,7%/16,7% і 56,7%/30,0% відповідно.

4. Рекомендоване пасивне ядро складається з ВОГ, одометрії, камери й маркерів, а MEMC використовується як незалежний резерв. За блокованої видимості потрібен додатковий канал сприйняття та правило безпечної зупинки.

5. Наступний етап має включати апаратно-програмне моделювання із залученням реального обладнання (hardware-in-the-loop, HIL), температурні випробування ВОГ, калібрування MEMC дисперсією Аллана та польову перевірку з незалежною еталонною траєкторією.

Конфлікт інтересів

Автори декларують відсутність конфлікту інтересів фінансового, особистого, авторського або іншого характеру, який міг би вплинути на дослідження та його результати.

Фінансування

Дослідження проводилося без фінансової підтримки.

Доступність даних

Пов'язані дані надано у відтворюваному пакеті: первинні результати CSV/JSON, агрегати, фіксовані зерна, параметри ВОГ, код пакетного запуску, код рисунків і цифровий відбиток моделі. Перед публікацією автори мають розмістити пакет у репозитарії та додати постійне посилання або цифровий ідентифікатор об'єкта (DOI).

Використання засобів штучного інтелекту

Автори повідомляють про використання системи штучного інтелекту (ШІ) OpenAI Codex для допомоги у програмуванні симулятора, підготовці відтворюваного пакетного експерименту, агрегації результатів, побудові рисунків, мовному редагуванні та форматуванні рукопису. Усі числові дані отримано виконанням локального детермінованого коду, перевірено авторами й надано разом із первинними журналами та параметрами. Засіб ШІ не створював фізичних вимірювань і не використовувався як джерело наукових фактів. Автори відповідають за постановку дослідження, перевірку результатів, посилання та висновки.

СПИСОК ЛІТЕРАТУРИ

1. El-Sheimy, N., Hou, H., and Niu, X. (2008), "Analysis and Modeling of Inertial Sensors Using Allan Variance", *IEEE Transactions on Instrumentation and Measurement*, vol. 57, no. 1, pp. 140–149, doi: <https://doi.org/10.1109/TIM.2007.908635>
2. Olson, E. (2011), "AprilTag: A robust and flexible visual fiducial system", *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 3400–3407, doi: <https://doi.org/10.1109/ICRA.2011.5979561>
3. Burri, M., Nikolic, J., Gohl, P., Schneider, T., Rehder, J., Omari, S., Achtelik, M. W., and Siegwart, R. (2016), "The EuRoC micro aerial vehicle datasets", *IJRR*, vol. 35, no. 10, pp. 1157–1163, doi: <https://doi.org/10.1177/0278364915620033>
4. Bijelic, M., Gruber, T., Mannan, F., Kraus, F., Ritter, W., Dietmayer, K., and Heide, F. (2020), "Seeing Through Fog Without Seeing Fog: Deep Multimodal Sensor Fusion in Unseen Adverse Weather", *Proceedings of CVPR*, pp. 11682–11692, URL: <https://www.cs.princeton.edu/~fheide/AdverseWeatherFusion/figures/AdverseWeatherFusion.pdf>
5. Hahner, M., Sakaridis, C., Dai, D., and Van Gool, L. (2021), "Fog Simulation on Real LiDAR Point Clouds for 3D Object Detection in Adverse Weather", *Proceedings of ICCV*, pp. 15283–15292, URL: <https://arxiv.org/pdf/2108.05249>
6. Jarraya, I., Al-Batati, A., Kadri, M. B., Abdelkader, M., Ammar, A., Boulila, W., and Koubaa, A. (2025), "Gnss-denied unmanned aerial vehicle navigation: analyzing computational complexity, sensor fusion, and localization methodologies", *Satellite Navigation*, vol. 6, article 9, doi: <https://doi.org/10.1186/s43020-025-00162-z>

7. Pereira, T., Viegas, C., Soares, S., and Ferreira, N. (2026), “Visual and Visual-Inertial SLAM for UGV Navigation in Unstructured Natural Environments: A Survey of Challenges and Deep Learning Advances”, *Robotics*, vol. 15, no. 2, article 35, doi: <https://doi.org/10.3390/robotics15020035>
8. Qin, T., Li, P., and Shen, S. (2018), “VINS-Mono: A Robust and Versatile Monocular Visual-Inertial State Estimator”, *IEEE Transactions on Robotics*, vol. 34, no. 4, pp. 1004–1020, doi: <https://doi.org/10.1109/TRO.2018.2853729>
9. Campos, C., Elvira, R., Rodriguez, J. J. G., Montiel, J. M. M., and Tardós, J. D. (2021), “ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual-Inertial, and Multimap SLAM”, *IEEE Transactions on Robotics*, vol. 37, no. 6, pp. 1874–1890, doi: <https://doi.org/10.1109/TRO.2021.3075644>
10. Chiang, K.-W., Tsai, S., Huang, C.-H., Lu, Y.-E., Srinara, S., Tsai, M.-L., El-Sheimy, N., and Ai, M. (2026), “Navigation Error Characteristics of LIO-, VIO-, and RIMU-Assisted INS/GNSS Multi-Sensor Fusion Schemes in a GNSS-Denied Environment”, *Sensors*, vol. 26, no. 7, article 2068, doi: <https://doi.org/10.3390/s26072068>
11. Tong, B., Chen, W., and Du, L. (2025), “Adaptive Navigation Strategy for Low-Cost IMU-Assisted Vehicles in GNSS-Denied Traffic Environment”, *Journal of Advanced Transportation*, vol. 2025, doi: <https://doi.org/10.1155/atr/2762711>
12. Huang, G., Liu, J., Zhao, Y., Gong, X., Hyyppä, J., and Zhang, W. (2025), “Vehicle Inertial Localization With Adaptive Noise Estimation and Pseudo-Measurement Constraints in GNSS Denied Environment”, *IEEE Internet of Things Journal*, early access, doi: <https://doi.org/10.1109/JIOT.2025.3593605>
13. Li, M., Sun, Y., Gao, S., Zhao, X., Hui, F., Luo, W., Hu, Q., Chen, H., Wu, H., Wang, Y., Yan, M., and Ding, W. (2025), “Navigation-grade interferometric air-core antiresonant fibre optic gyroscope with enhanced thermal stability”, *Nature Communications*, vol. 16, article 3449, doi: <https://doi.org/10.1038/s41467-025-58381-6>
14. Smailov, N., Tashtay, Y., Komada, P., Nussupov, Y., Zhunussov, K., Batyrgaliyev, A., Naubetov, D., Amir, A., Sekenov, B., and Yerezhap, D. (2026), “Fiber-Optic Gyroscopes in Modern Navigation Systems: A Comprehensive Review”, *Network*, vol. 6, no. 2, article 28, doi: <https://doi.org/10.3390/network6020028>
15. Li, T., Chen, H., Fu, H., Zhou, F., Gao, S., Fan, S., Wang, C., Wu, Y., Liu, H., Zhao, H., Song, J., Jin, L., Cui, N., Feng, J., and Lu, Z. (2025), “Integrated optoelectronic transceiver toward navigation grade fiber optic gyroscope using silicon photonics”, *Optics Letters*, vol. 50, no. 22, pp. 7139–7142, doi: <https://doi.org/10.1364/OL.576264>
16. Borenstein, J., and Koren, Y. (1991), “The Vector Field Histogram-Fast Obstacle Avoidance for Mobile Robots”, *IEEE Transactions on Robotics and Automation*, vol. 7, no. 3, pp. 278–288, doi: <https://doi.org/10.1109/70.88137>
17. EMCORE Corporation (2023), “EG-120 Lithium-Niobate Fiber Optic Gyroscope (FOG)”, datasheet, revision 12.12.2023, URL: https://www.emcore.com/wp-content/uploads/966798_A1_EG-120-FOG.pdf
18. Wilson, E. B. (1927), “Probable Inference, the Law of Succession, and Statistical Inference”, *Journal of the American Statistical Association*, vol. 22, no. 158, pp. 209–212, doi: <https://doi.org/10.1080/01621459.1927.10502953>
19. Moore, T., and Stouch, D. (2016), “A Generalized Extended Kalman Filter Implementation for the Robot Operating System”, *Intelligent Autonomous Systems 13*, pp. 335–348, doi: https://doi.org/10.1007/978-3-319-08338-4_25

Received (Надійшла) 29.06.2026

Accepted for publication (Прийнята до друку) 02.09.2026

Publication date (Дата публікації) 25.09.2026

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Бойчун Сергій Олегович – кандидат політичних наук, молодший науковий співробітник відділу оптоелектроніки, Інститут фізики напівпровідників ім. В.Є. Лашкарьова НАН України, Київ, Україна;

Boychun Serhii – Candidate of Political Sciences, Junior Researcher, Optoelectronics Department, V.E. Lashkaryov Institute of Semiconductor Physics, NAS of Ukraine, Kyiv, Ukraine;

e-mail: s@lambit.tech; ORCID Author ID: <https://orcid.org/0009-0009-7099-1990>.

Левицький Сергій Миколайович – кандидат технічних наук, старший дослідник, старший науковий співробітник відділу оптоелектроніки, Інститут фізики напівпровідників ім. В.Є. Лашкарьова НАН України, Київ, Україна;

Levytskyi Serhii – Candidate of Technical Sciences, Senior Researcher, Senior Scientist, Optoelectronics Department, V. E. Lashkaryov Institute of Semiconductor Physics, NAS of Ukraine, Kyiv, Ukraine;

e-mail: levytskyism@gmail.com; ORCID Author ID: <https://orcid.org/0000-0002-3909-0993>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=25923472700>.

**Autonomous passive navigation of an unmanned ground vehicle during GNSS loss:
MEMS and fiber-optic gyroscope comparison**

Serhii Boychun, Serhii Levytskyi

Abstract. Background. GNSS loss requires an unmanned ground vehicle to maintain orientation and obstacle awareness from onboard sensing. A fiber-optic gyroscope must be distinguished from camera-based visual odometry. **Subject of the study:** navigation using wheel odometry, a MEMS gyroscope, a parameterized fiber-optic gyroscope, passive optical markers, and an optional rangefinder. **Objective:** to quantify baseline fiber-optic gyro precision and compare passive and hybrid configurations under degraded visibility. **Methods.** The study comprised 1,440 obstacle-course runs and 180 separate 300-second heading tests. The baseline fiber-optic gyro used 1 deg/h bias stability, 0.04 deg/sqrt(h) angle random walk, and 250 ppm scale-factor error. **Results.** The fiber-optic gyro achieved 0.0568 deg mean heading RMSE and 0.0995 deg final heading error, versus 9.8981 deg and 17.1355 deg for the MEMS scenario. With clear visibility, the fiber-optic gyro plus markers reached the target in 46.7% of runs with 23.3% safe arrival; adding the rangefinder to the fused stack increased these values to 76.7% and 40.0%. **Conclusions.** The fiber-optic gyro is the most accurate passive heading source in the model and remains available when camera visibility is blocked. However, safe navigation still requires obstacle perception and position correction; blocked visibility therefore needs a separate sensing fallback or a safe-stop policy.

Keywords: unmanned ground vehicle; GNSS-denied navigation; fiber-optic gyroscope; MEMS gyroscope; passive navigation; optical markers; sensor fusion; rangefinder; reproducible simulation.

І. Ш. Невлюдов, В. В. Євсєєв, О. О. Чала, К. Ю. Горовий

Харківський національний університет радіоелектроніки, Харків, Україна

РОЗРОБЛЕННЯ ІНТЕЛЕКТУАЛЬНОГО МЕТОДА КОНТРОЛЯ ЯКОСТІ РАД КОНТАКТІВ ЦИФРОВОГО ДВІЙНИКА FEK DELVOTEC 6400 ДЛЯ МОДЕЛЮВАННЯ ПРОЦЕСУ КЛИНОВОГО ЗВАРЮВАННЯ У МІКРОЕЛЕКТРОНІЦІ

Анотація. Об'єктом дослідження є процес автоматизованого контролю якості контактних площадок у цифровому двійнику обладнання FEK Delvotec 6400 для моделювання технологічного процесу клинового зварювання у мікроелектроніці. Предметом дослідження є методи, математичні моделі та алгоритмічне забезпечення інтелектуального контролю якості контактних площадок цифрового двійника обладнання FEK Delvotec 6400 для моделювання процесу клинового зварювання у мікроелектроніці. Метою роботи є розроблення інтелектуального методу автоматизованого оцінювання стану контактних площадок, що забезпечує підвищення точності позиціонування, достовірності цифрового моделювання та якості прийняття рішень під час виконання технологічного процесу. Для досягнення поставленої мети виконано розроблення математичних моделей формування зображення технічного зору, калібрування координатного простору, нейромережевого розпізнавання контактних площадок і реперних міток, уточнення координат центрів, координатних перетворень, автоматичної корекції траєкторії посадки дроту, оцінювання точності позиціонування та інтегральної перевірки достовірності результатів комп'ютерного зору перед використанням у цифровому двійнику. Методично дослідження ґрунтується на використанні математичного моделювання, сучасних методів комп'ютерного зору, глибокого навчання, згорткових нейронних мереж, детекторів об'єктів, цифрової обробки зображень, автоматизованого аналізу геометричних параметрів контактних площадок і чисельного моделювання. Для експериментальної перевірки сформовано набір аугментованих зображень контактних площадок, виконано навчання нейромережевої моделі та проведено серію експериментів із визначення повноти виявлення контактних площадок, точності локалізації їх центрів, похибки вимірювання геометричних параметрів, достовірності класифікації дефектів, інтегрального індексу валідації цифрового двійника та часу обробки одного зображення. Отримані результати підтвердили високу ефективність запропонованого підходу, забезпечивши повноту виявлення контактних площадок до 100 %, похибку локалізації центрів у межах 1,35–3,05 мкм, похибку визначення геометричних параметрів до 4,7 %, високий рівень достовірності класифікації дефектів та час обробки одного зображення, що не перевищує 50 мс. Практична цінність роботи полягає у можливості інтеграції запропонованого методу до цифрового двійника FEK Delvotec 6400 для автоматизованого контролю якості контактних площадок, підвищення точності моделювання процесу клинового зварювання, своєчасного виявлення дефектів і забезпечення стабільності технологічного процесу у мікроелектронному виробництві.

Ключові слова: цифровий двійник; контактні площадки; клинове зварювання; комп'ютерний зір; нейронні мережі; контроль якості; математичне моделювання; мікроелектроніка; автоматизований контроль; цифрова обробка зображень.

Вступ

Актуальність дослідження

Сучасне виробництво мікроелектронних виробів висуває підвищені вимоги до точності позиціонування, автоматизованого контролю якості контактних площадок та достовірності моделювання процесів клинового зварювання, що обумовлено переходом до концепцій Industry 5.0 і широким впровадженням технології Digital Twin [1, 2]. Одним із найбільш критичних етапів технологічного процесу є своєчасне виявлення дефектів PAD-контактів, оскільки їх геометричні відхилення, забруднення, окислення або механічні пошкодження безпосередньо впливають на якість мікрозварних з'єднань, надійність електронних модулів та ефективність роботи автоматизованих систем wire bonding [3, 4]. Використання традиційних алгоритмів машинного зору не завжди забезпечує необхідну точність і стійкість до змін умов освітлення, шумів CCD-зображень та технологічних варіацій поверхні,

що знижує достовірність цифрового моделювання виробничого процесу [5, 6]. У зв'язку з цим актуальним науково-практичним завданням є розроблення інтелектуального методу контролю якості PAD-контактів на основі інтеграції сучасних нейромережевих моделей комп'ютерного зору, системи CCD/PRU та цифрового двійника FEK Delvotec 6400, що дозволить підвищити точність автоматичного позиціонування, достовірність оцінювання стану контактних площадок і ефективність моделювання процесу клинового зварювання у мікроелектроніці.

Аналіз літературних джерел і визначення проблеми

У роботі [7] запропоновано огляд підходів до підвищення надійності контролю кутових зварних з'єднань із використанням фазованих ультразвукових решіток, що дає можливість покращити виявлення дефектів у складних геометричних зонах і знизити ризик пропуску критичних пошкоджень.

Запропоноване рішення орієнтоване на ультразвуковий неруйнівний контроль металевих зварних з'єднань і не враховує специфіку CCD/PRU-зображень, нейромережевого аналізу PAD-контактів, координатної корекції та інтеграції результатів у цифровий двійник FEK Delvotec 6400.

У статті [8] досліджено інноваційний метод неруйнівного контролю зварних швів вакуумної камери компактного термоядерного реактора, що дає можливість підвищити достовірність перевірки відповідальних зварних конструкцій у складних енергетичних системах. Але, даний метод має обмежену придатність для мікроелектронного виробництва, оскільки не розглядає аналіз контактних площадок, дефекти типу забруднення, окислення або відсутності PAD, а також не містить моделей машинного зору для автоматизованого оцінювання якості перед клиновим зварюванням.

У статті [9] розроблено метод виявлення дефектів плати керування двигуном на основі цифрової обробки зображень, що дає можливість виконувати автоматизовану класифікацію дефектних і справних виробів з точністю понад 99 % та використовувати підхід для потокового контролю електронних плат. Даний метод зосереджений переважно на загальній обробці зображень і пороговому виділенні ознак, тому недостатньо враховує мікронну точність позиціонування PAD-контактів, нейромережеву валідацію CCD/PRU alignment і зв'язок результатів контролю з моделюванням процесу клинового зварювання.

У статті [10] досліджено сучасні застосування цифрових двійників для інтелектуального управління якістю у будівництві, що дає можливість сформувати структуроване уявлення про роль збору даних, цифрової моделі, аналітики та зворотного зв'язку у процесах контролю якості [10]. Хотя робота має високий методологічний інтерес, проте вона належить до будівельної галузі й не деталізує моделі комп'ютерного зору, CNN/YOLO-розпізнавання, калібрування CCD-камери та валідації контактних площадок у задачах мікроелектронного wire bonding.

У статті [11] досліджено метод визначення та аналізу глибини текстури асфальтобетонного покриття на основі цифрової аналітики зображень, що дає можливість виконувати кількісне оцінювання поверхневих параметрів за візуальними даними. Запропонований метод підтверджує ефективність цифрової обробки зображень для аналізу поверхонь, проте не охоплює високоточний контроль мікрооб'єктів, дефекти PAD-контактів, субпіксельне уточнення координат і інтеграцію результатів у цифровий двійник технологічного обладнання для клинового зварювання.

У дослідженні [12] запропоновано нову схемну реалізацію медіанного фільтра на основі атомних кремнієвих квантових точок для цифрової обробки зображень та інтернету речей, що дає можливість зменшити енергоспоживання і площу апаратної реалізації при фільтрації шумів зображень. Робота є корисною для апаратного прискорення попередньої обробки CCD-зображень, проте вона не вирішує

задачі розпізнавання PAD-контактів, класифікації дефектів, координатної корекції bonding points і валідації результатів у структурі цифрового двійника FEK Delvotec 6400.

Загальний аналіз наведених публікацій показує, що сучасні дослідження активно розвивають неруйнівний контроль зварних з'єднань, цифрові двійники, цифрову обробку зображень та апаратні засоби фільтрації, проте вони не формують цілісного рішення для автоматизованого контролю PAD-контактів у мікроелектроніці з урахуванням CCD/PRU-зображень, CNN/YOLO-аналізу, координатного перетворення та інтегральної валідації. Це підтверджує необхідність і актуальність розроблення інтелектуального методу контролю якості PAD-контактів цифрового двійника FEK Delvotec 6400 для підвищення точності автоматизованого оцінювання контактних площадок і достовірності моделювання процесу клинового зварювання у мікроелектроніці.

Постановка проблеми та мета дослідження

Сучасні системи цифрових двійників обладнання для клинового зварювання у мікроелектроніці потребують високоточного автоматизованого контролю якості PAD-контактів, однак традиційні методи машинного зору недостатньо ефективно працюють в умовах шумів CCD-зображень, змін освітлення, технологічних варіацій поверхні та різноманітних дефектів контактних площадок, що знижує достовірність моделювання технологічного процесу та прийняття рішень. У зв'язку з цим актуальною науково-практичною задачею є розроблення інтелектуального методу контролю якості PAD-контактів із використанням сучасних нейромережевих технологій комп'ютерного зору та їх інтеграцією до цифрового двійника FEK Delvotec 6400 для підвищення точності автоматичного позиціонування, оцінювання стану контактних площадок і достовірності моделювання процесу клинового зварювання.

Метою роботи є розробка інтелектуального методу контролю якості PAD-контактів цифрового двійника FEK Delvotec 6400 на основі математичних моделей CCD/PRU, нейромережевого аналізу CNN/YOLO та алгоритмів валідації для підвищення точності автоматизованого оцінювання контактних площадок і достовірності моделювання процесу клинового зварювання у мікроелектроніці.

Розроблення математичного забезпечення метода інтелектуального контролю якості PAD контактів

Обладнані FEK Delvotec 6400 оснащена Charge-Coupled Device (CCD) це ПЗЗ-камера, тобто системою технічного зору на базі світлочутливої матриці, яка використовується для отримання високоточного зображення контактних площадок, кристала та wire bonding-зони [13, 14], яка працює сумісно з Pattern Recognition Unit (PRU) – модуль розпізнавання образів, який виконує аналіз зображення з CCD-камери, знаходить reference marks, pads, координати

bonding points та виконує автоматичне вирівнювання позиції зварювання [15]. Під керуванням процедури Alignment, то б то автоматичного вирівнювання та корекції координат bonding tool відносно реального положення кристала, підкладки або контактних площадок. Технічні характеристики параметрів CCD-камери яка використовується у обладнанні FEK Delvotec 6400, приведені у табл. 1 [15].

Для систем FEK Delvotec CCD – модуль зазвичай інтегрується з системою vision alignment та виконує наступні функції:[15]

- пошуку bonding pad;
- вирівнювання capillary;
- контролю loop height;
- розпізнавання fiducial marks;
- контролю точності посадки дроту.

Для розробки математичного представлення CDD – робочої зони (вигляді з CCD – камери) обрані наступні параметри, які зображені на рис. 1.

Таблиця 1 – Технічні характеристики параметрів CCD-камери

Параметр	Типове значення для CCD FEK Delvotec 6400
Роздільна здатність	640×480, 1024×768 або 1280×1024 <i>px</i>
Тип сенсора	Monochrome CCD
Частота кадрів	30–120 <i>FPS</i>
Dynamic Range	60–72 <i>dB</i>
Signal-to-Noise Ratio	50–60 <i>dB</i>
Sensitivity	~0.1–1 <i>lux</i>
Exposure Time	10 μ s – 100 <i>ms</i>
Точність позиціонування	$\pm 1-3 \mu$ m
Інтерфейс передачі	CameraLink / LVDS
Field of View	1–10 <i>mm</i> залежно від оптики
Оптичне збільшення	5× – 20×
Auto Gain Control	підтримується
Sub-pixel detection accuracy	до 0.1 <i>pixel</i>

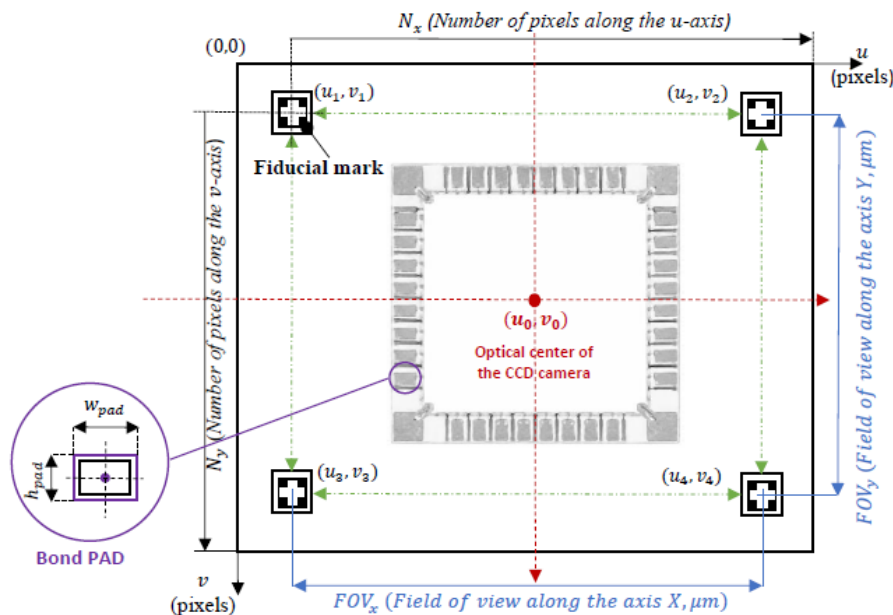


Рис. 1. CDD – зображення робочої зони (вигляді з CCD – камери)

Опишемо систему координат CDD (рис. 1) зображення робочої зони (вигляд з CCD – камери):
 u – вісь зображення по горизонталі [пікселі];
 v – вісь зображення по вертикалі [пікселі];
 $(0,0)$ – початок координат (лівий верхній піксель);
 (u_0, v_0) – оптичний центр CDD – камери (пікселі);
 N_x – роздільна здатність по осі u [пікселі];
 N_y – роздільна здатність по осі v [пікселі];
 FOV_x – поле огляду по осі X [μ m];
 FOV_y – поле огляду по осі Y [μ m].

Відповідно до рисунку 1 масштаб CDD – камери розраховуються відповідно:

$$k_x = \frac{FOV_x}{N_x}; \quad k_y = \frac{FOV_y}{N_y}, \quad (1)$$

де k_x, k_y – масштаб CDD – камери по осі u та v відповідно [μ m/піксель].

Розробимо математичну модель CCD-модуля з vision alignment для Digital Twin FEK Delvotec 6400 з

удосконаленням на базі нейронної мережі для розпізнавання fiducial marks та контролю точності посадки дроту. У FEK Delvotec 6400 PRU використовує сигнал CCD-камери, встановленої на bond head, для високоточного позиціонування, розпізнавання шаблонів після teach-in та компенсації зміщення adjust point через offset, а також має налаштування contrast, brightness, lighting, gain, offset і framegrabber.

На першому кроці розробимо модель формування CCD-зображення робочої зони. Дана модель потрібна для опису того, як реальна координата контактної площадки або fiducial mark у робочій зоні FEK Delvotec 6400 перетворюється у піксельне зображення CCD-камери, яке надалі використовується Pattern Recognition Unit (PRU) або нейронною мережею для вирівнювання координат bonding tool. Вона є основою для Digital Twin, оскільки дозволяє зв'язати фізичну систему координат $X - Y$ з цифровим зображенням, та має такий вигляд:

$$\begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = K \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} + \varepsilon_{cdd}, \quad (2)$$

де u_i, v_i – координати точки на CCD-зображенні у пікселях; x_i, y_i – реальні координати fiducial mark або bonding point у системі $X - Y, [\mu m]$; K – матриця внутрішніх параметрів CCD-оптики; s_x, s_y – коефіцієнти масштабування між μm і пікселями; ε_{cdd} – похибка CCD-зображення, яка враховує шум, фокус, освітлення та дискретизацію.

Для визначення просторової роздільної здатності CCD-модуля, тобто скільки мікрометрів відповідає одному пікселю зображення розробимо модель калібрування масштабу CCD-камера - $X - Y$ стіл. Данна модель у Digital Twin що розроблюється буде використовуватися для перерахунку координат, знайдених нейронною мережею або PRU, у реальні координати bonding program:

$$k_x = \frac{\Delta X_{table}}{\Delta u}; \quad k_y = \frac{\Delta Y_{table}}{\Delta v}, \quad (3)$$

де k_x, k_y – масштабні коефіцієнти, $[\mu m/px]$; $\Delta X_{table}, \Delta Y_{table}$ – контрольне переміщення $X - Y$ столу, $[\mu m]$; $\Delta u, \Delta v$ – відповідна зміна положення об'єкта на CCD-зображенні [пікселі].

Для підвищення надійності vision alignment у Digital Twin FEK Delvotec 6400, то б то для удосконалення класичного PRU за рахунок використання CNN або YOLO-подібної нейронної мережі, яка автоматично визначає fiducial marks, bonding pads та контрольні точки навіть при зміні освітлення, шумі, частковому забрудненні або неоднорідності поверхні пропонується використовувати гібридну схему CNN + YOLO-like detector. YOLO-подібна мережа швидко знаходить fiducial marks, bond pads і контрольні області на CCD-зображенні, а невелика CNN або sub-pixel refinement module уточнює координати центра мітки з мікронною точністю. Запропонований підхід добре узгоджується з принципом роботи PRU у FEK Delvotec 6400, оскільки штатна система вже використовує CCD-камеру на bond head для високоточного позиціонування, розпізнавання шаблонів після teach-in та компенсації зміщення adjust point через offset, та може бути представлена наступною модель:

$$\hat{B}_j, \hat{p}_j = f_{\theta}(I_{cdd}), \quad (4)$$

де I_{cdd} – вхідне CCD-зображення для нейронної мережі; f_{θ} – нейронна мережа з параметрами θ ; $\hat{B}_j = (u_j, v_j, w_j, h_j)$ – знайдена область fiducial mark або bonding pad на зображенні; $\hat{p}_j$ – імовірність правильного розпізнавання об'єкта; j – номер знайденого об'єкта.

Модель вхідного CCD-зображення для нейронної мережі (I_{cdd}) з моделі 4, потрібна для формалізації того, що нейронна мережа отримує не координати напряду, а цифрове CCD-зображення робочої зони, на якому присутні fiducial marks, bond pads, кристал, підкладка та можливі шуми освітлення:

$$I_{cdd} = F_{cam}(X, Y, L, G, O, \eta); \quad (5)$$

де I_{cdd} – вхідне CCD-зображення робочої зони; X, Y – фізичні координати елементів на робочому столі; L – параметри освітлення; G – gain CCD/PRU (коефіцієнт підсилення сигналу в системі технічного зору CCD + PRU, який використовується для регулювання яскравості та контрастності CCD-зображення перед виконанням Pattern Recognition та vision alignment); O – яскравість зображення або зміщення рівня сигналу або корекція базового рівня яскравості framegrabber-модуля; η – шум зображення, викликаний освітленням, відбиттям поверхні та дискретизацією.

Для автоматичного пошуку fiducial marks на CCD-зображенні без ручного вибору шаблону, як у класичному PRU teach-in. Рекомендовано використовувати легку YOLO-подібну архітектуру типу YOLO-nano або YOLO-small або компактний detector, тому що у задачах wire bonding важлива не тільки точність, але й швидкість обробки кадру перед виконанням alignment [16, 17]. Модель YOLO-like detection для fiducial marks представлена нижчі:

$$\{\hat{B}_j, \hat{p}_j, \hat{c}_j\}_{j=1}^N = f_{\theta}^{YOLO}(I_{cdd}), \quad (6)$$

де f_{θ}^{YOLO} – YOLO-подібна нейронна мережа з параметрами θ ; I_{cdd} – вхідне CCD-зображення; $\hat{B}_j = (u_j, v_j, w_j, h_j)$ – знайдена bounding box область fiducial mark або bond pad; $\hat{p}_j$ – рівень впевненості розпізнавання (рекомендоване $\hat{p}_j = 0.98$); $\hat{c}_j$ – клас об'єкта, наприклад fiducial mark, pad, defect, edge; N – кількість знайдених об'єктів на зображенні.

Для підвищення точності після YOLO-детекції, оскільки YOLO добре знаходить область об'єкта, але для wire bonding потрібна не просто рамка, а точний центр мітки або контактної площадки, не обхідно модель CNN-уточнення центра fiducial mark [18]. Для цього з області $\hat{B}_j$ вирізається ROI-фрагмент, який передається у CNN-регресор для уточнення центра:

$$(\Delta u_j, \Delta v_j) = f_{\phi}^{CNN}(ROI_j);$$

$$u_{c,j} = u_j + \Delta u_j, \quad v_{c,j} = v_j + \Delta v_j, \quad (7)$$

де ROI_j – фрагмент CCD-зображення всередині знайденої області $\hat{B}_j$; f_{ϕ}^{CNN} – CNN-регресор з параметрами ϕ ; $\Delta u_j, \Delta v_j$ – уточнення положення центра в пікселях; $u_{c,j}, v_{c,j}$ – уточнений центр fiducial mark у системі координат CCD-зображення.

Наступним кроком розробимо модель перетворення піксельних координат у координати $X - Y$ столу на базі 3, ця модель потрібна для того, щоб результат роботи нейронної мережі можна було використати у Digital Twin як реальні координати робочої зони FEK Delvotec 6400:

$$x_{c,j} = k_x(u_{c,j} - u_o), \quad y_{c,j} = k_y(v_{c,j} - v_o), \quad (8)$$

де $x_{c,j}, y_{c,j}$ – координати fiducial mark $[\mu m]$; $u_{c,j}, v_{c,j}$ – координати центра мітки у пікселях; u_o, v_o – оптичний центр CCD-камери (рис. 1); k_x, k_y – масштабні коефіцієнти $[\mu m/px]$.

Для визначення різниці між еталонним положенням міток, збереженим під час teach-in, та фактичним положенням, знайденим CNN/YOLO-

модулем, [19] пропонується така модель обчислення зміщення vision alignment:

$$\begin{aligned}\Delta X &= \frac{1}{M} \sum_{j=1}^M (x_{c,j} - x_{ref,j}); \\ \Delta Y &= \frac{1}{M} \sum_{j=1}^M (y_{c,j} - y_{ref,j}); \\ \Delta \theta &= \arctan \frac{y_{c,2} - y_{c,1}}{x_{c,2} - x_{c,1}} - \arctan \frac{y_{ref,2} - y_{ref,1}}{x_{ref,2} - x_{ref,1}},\end{aligned}\quad (9)$$

де $\Delta X, \Delta Y$ – лінійне зміщення підкладки; $\Delta \theta$ – кутове зміщення; $x_{c,j}, y_{c,j}$ – координати знайденої fiducial mark; $x_{ref,j}, y_{ref,j}$ – еталонні координати fiducial mark; M – кількість використаних міток.

Модель корекції bonding points після neural vision alignment потрібна для перерахунку координат Source Bond та Destination Bond перед моделюванням $X - Y - Z - P$ motion system у Digital Twin:

$$\begin{bmatrix} \hat{x}_i \\ \hat{y}_i \end{bmatrix} = \begin{bmatrix} \cos \Delta \theta & -\sin \Delta \theta \\ \sin \Delta \theta & \cos \Delta \theta \end{bmatrix} \begin{bmatrix} x_i \\ y_i \end{bmatrix} + \begin{bmatrix} \Delta X \\ \Delta Y \end{bmatrix}, \quad (10)$$

де x_i, y_i – координати bonding point із Wire Bonding Program Input; $\hat{x}_i, \hat{y}_i$ – скориговані координати після neural alignment; $\Delta X, \Delta Y, \Delta \theta$ – знайдені параметри зміщення підкладки.

Для перевірки, чи потрапляє скоригована точка зварювання у допустиму область bond pad, розробимо модель контролю точності посадки дроту:

$$e_i = \sqrt{(\hat{x}_i - x_{pad,i})^2 + (\hat{y}_i - y_{pad,i})^2},$$

$$V_{place}(i) = \begin{cases} 1, & e_i \leq e_{max}; \\ 0, & e_i > e_{max}, \end{cases} \quad (11)$$

де e_i – похибка посадки дроту; $\hat{x}_i, \hat{y}_i$ – скориговані координати точки зварювання; $x_{pad,i}, y_{pad,i}$ – центр контактної площадки; e_{max} – максимально допустима похибка; $V_{place}(i)$ – результат перевірки посадки дроту. Щоб прийняти рішення, чи можна використовувати результати CNN/YOLO-розпізнавання у подальшому моделюванні клинового зварювання [20], розробимо інтегральна модель валідності neural PRU alignment.

$$\begin{aligned}V_{NeuralPRU}(i) &= \\ &= V_{det}(j) \cdot V_{coord}(j) \cdot V_{place}(i) \cdot V_{conf}(j),\end{aligned}\quad (12)$$

де $V_{NeuralPRU}(i)$ – інтегральна валідність neural vision alignment; $V_{det}(j)$ – факт виявлення fiducial mark; $V_{coord}(j)$ – коректність перетворення координат; $V_{place}(i)$ – точність посадки дроту; $V_{conf}(j)$ – перевірка confidence score нейронної мережі.

У межах даних досліджень рекомендовано використовувати YOLO-like detector для первинного пошуку fiducial marks та невелику CNN-регресійну модель для уточнення центра, тому що така комбінація забезпечує швидке розпізнавання, стійкість до змін освітлення, можливість роботи з різними типами підкладок і достатню точність для подальшого перерахунку координат у Digital Twin FEK Delvotec 6400. Структурна схема реалізації моделювання на базі мови Python, представлена на рис. 2.

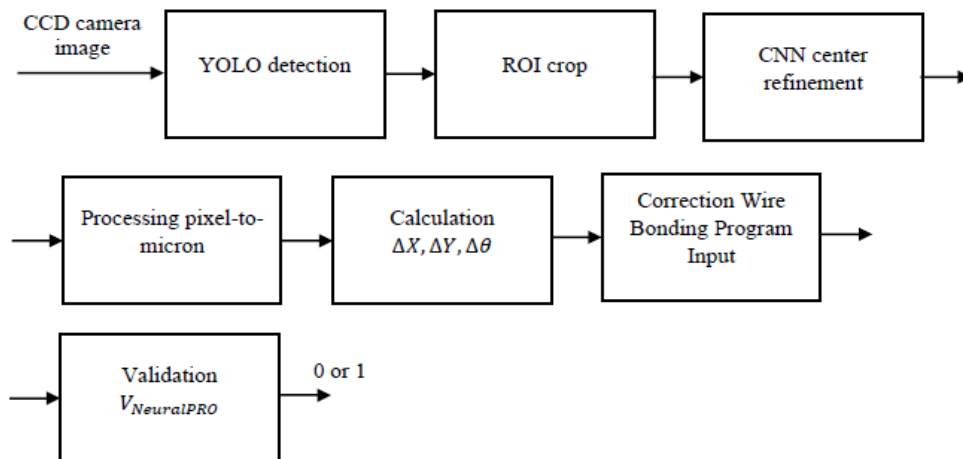


Рис. 2. Структурна схема реалізації моделювання на базі мови Python

Для того, щоб Digital Twin не використовував сумнівні результати комп'ютерного зору, якщо fiducial mark або bonding pad були розпізнані з низькою впевненістю розробимо модель довіри до результату нейромережевого розпізнавання. Результат рішення даної моделі особливо важливо для автоматичного прийняття рішення про запуск bonding process:

$$V_{NeuralPRU}(j) = \begin{cases} 1, & \hat{p}_j \geq p_{min}; \\ 0, & \hat{p}_j < p_{min}, \end{cases} \quad (13)$$

де $V_{NeuralPRU}(j)$ – результат перевірки довіри до нейромережевого розпізнавання; $\hat{p}_j$ – імовірність або

confidence score, отриманий від нейронної мережі; p_{min} – мінімально допустимий рівень довіри; j – номер fiducial mark або bonding pad.

Для об'єднання всіх часткових перевірок у єдиний показник, який визначає, чи можна використовувати скориговані координати для моделювання ultrasonic wedge bonding у Digital Twin FEK Delvotec 6400 розроблена така інтегральна модель валідності CCD + Neural PRU alignment.

$$\begin{aligned}V_{vision}(i) &= \\ &= V_{calib} \cdot V_{NeuralPRU}(j) \cdot V_{place}(i) \cdot V_{offset}(i),\end{aligned}\quad (14)$$

де $V_{vision}(i)$ – інтегральна валідність vision alignment для i -го bonding point; V_{calib} – результат перевірки калібрування CCD-камери; $V_{NeuralPRU}(j)$ – результат перевірки довіри до неймережевого розпізнавання; $V_{place}(i)$ – результат перевірки точності посадки дроту; $V_{offset}(i)$ – результат перевірки допустимості зміщення $\Delta X, \Delta Y, \Delta \theta$.

Розроблена математична модель CCD-модуля з vision alignment та інтеграцією CNN/YOLO-подібної нейронної мережі забезпечує підвищення точності автоматичного позиціонування bonding tool, стійкість до змін освітлення, шумів CCD-зображення та варіацій поверхні контактних площадок під час ultrasonic wedge bonding у FEK Delvotec 6400. Запропонований підхід дозволяє реалізувати adaptive PRU alignment у Digital Twin середовищі, автоматизувати контроль точності посадки дроту та підвищити достовірність моделювання технологічного процесу клинового зварювання у мікроелектроніці та гібридних електронних модулях.

Проведення експериментальних досліджень та аналіз отриманих результатів

Метою експерименту є перевірка працездатності розробленого методу оцінки якості Bond PAD контактних площадок за CCD/PRU-зображеннями FEK Delvotec 6400 у межах створення Digital Twin технологічного процесу клинового зварювання.

Постановка задач експерименту полягає у завантаженні CCD/PRU-зображення мікросхеми, автоматичному виявленні всіх PAD, визначенні їх центрів, розрахунку геометричних параметрів контактних площадок, порівнянні отриманих значень з еталонними межами та класифікації можливих дефектів, таких як забруднення, пошкодження, відсутність або окислення PAD.

Вхідними даними експерименту є зображення робочої зони у вигляді з CCD-камери, масштабна сітка 200 μm , еталонні мінімальні та максимальні розміри PAD (рис. 3), а також навчена модель YOLO26 для розпізнавання класів pad_ok, pad_contamination, pad_damage, pad_missing і pad_oxidation [21].

Отримані результати повинні підтвердити можливість використання програми для кількісного та якісного контролю стану контактних площадок перед виконанням wire bonding, що є важливим етапом підвищення достовірності Digital Twin FEK Delvotec 6400.

На базі представлених Bond PAD мікросхеми CCD/PRU-зображень (рис.3) було розроблений початковий датасет для навчання YOLO26, але в поточному вигляді його буде недостатньо для повноцінного навчання неймережі. Ці зображення можуть використовуватися як еталонні представники класів: рис. 3,а – pad_ok; рис. 3, б – pad_oxidation; рис. 3, в – pad_contamination та рис. 3, г – pad_missing.

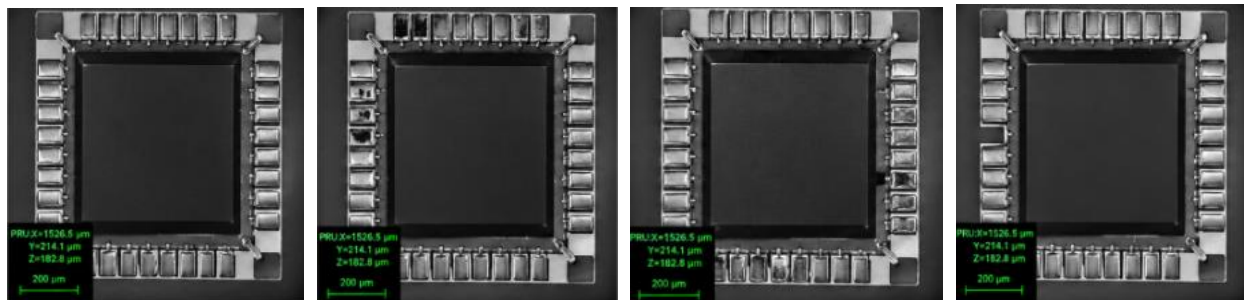


Рис. 3. Приклади Bond PAD мікросхеми:

а – ідеальний PAD; б – забруднення PAD; в – пошкоджений PAD; г – відсутній PAD;

Далі для кожного вихідного зображення було автоматично сформувати набір синтезованих аугментованих даних. Для CCD/PRU-зображень FEK Delvotec 6400 було прийняте рішення, генерувати варіації за рахунок повороту на $\pm 3-5^\circ$, зміни яскравості на $\pm 20\%$, зміни контрасту на $\pm 20\%$, додавання шуму CCD-матриці, зміни параметрів gain, brightness та framegrabber, а також невеликого масштабування й зміщення. З одного вихідного зображення було отримано приблизно 200–500 синтетичних зразків без суттєвої втрати достовірності. Звіт про успішний результат генерації синтетичних зразків для класів представлено на рис. 4.

Далі проводимо процес ініціалізації та запуску навчання неймережевої моделі YOLO26 для задачі автоматизованої оцінки якості Bond PAD контактних площадок у межах Digital Twin FEK Delvotec 6400. На початковому етапі система визначає доступний обчислювальний ресурс (CPU), завантажує архітектуру

моделі yolo26n.pt, формує структуру нейронної мережі та автоматично адаптує вихідний шар детектора під п'ять цільових класів дефектів контактних площадок. Далі виконується побудова повної конфігурації мережі із зазначенням типів шарів, кількості параметрів, міжшарових зв'язків та обчислювальної складності моделі, що необхідно для подальшого навчання на сформованому CCD/PRU-датасеті (рис. 5). Отримана конфігурація є основою для навчання YOLO26 розпізнаванню справних і дефектних PAD, що дозволяє використовувати модель як модуль автоматичного візуального контролю у складі Digital Twin FEK Delvotec 6400 для підвищення достовірності оцінки якості PAD контактів.

Після навчання отримуємо навчений CCD/PRU-датасеті по такому шляху: runs/detect/runs_pad_quality/yolo26_pad_quality/weights/best.pt, на базі якого були проведені експерименти по оцінці якості PAD контактів, результати якого приведені у табл. 2.

Для оцінювання ефективності розробленого інтелектуального методу контролю якості PAD-контактів проведено серію експериментальних досліджень, результати яких дозволяють комплексно оцінити точність розпізнавання, локалізації,

вимірювання геометричних параметрів та придатність запропонованого підходу до використання у цифровому двійнику FEK Delvotec 6400.

Узагальнені результати проведених експериментів наведено на рис. 6.

```
Generating class: pad_ok
100%|██████████| 1000/1000 [01:05<00:00, 15.32it/s]
0%|          | 0/1000 [00:00<?, ?it/s]Generating class: pad_oxidation
100%|██████████| 1000/1000 [01:42<00:00, 9.80it/s]
0%|          | 0/1000 [00:00<?, ?it/s]Generating class: pad_contamination
100%|██████████| 1000/1000 [00:59<00:00, 16.72it/s]
0%|          | 0/1000 [00:00<?, ?it/s]Generating class: pad_missing
100%|██████████| 1000/1000 [01:21<00:00, 12.34it/s]
0%|          | 0/1000 [00:00<?, ?it/s]Generating class: pad_damage
100%|██████████| 1000/1000 [01:03<00:00, 15.83it/s]

DATASET GENERATION COMPLETE
-----
Train images: 3432
Val images: 1023
Test images: 542
YAML file: PAD_DATASET\pad_quality.yaml
-----
Next training command:
yolo detect train data=PAD_DATASET\pad_quality.yaml model=yolo26n.pt imgsz=640 epochs=300 batch=8

Process finished with exit code 0
```

Рис. 4. Звіт про успішний результат генерації синтетичних зразків для класів

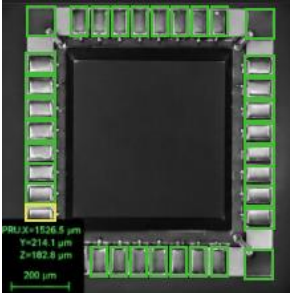
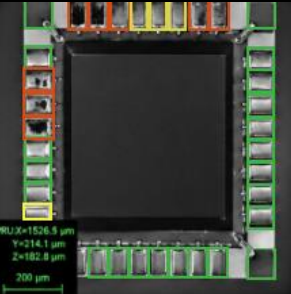
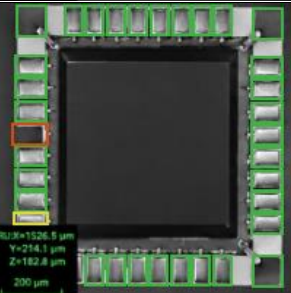
```
Selected device: cpu
Trying to load yolo26n.pt ...
Loaded model: yolo26n.pt
Ultralytics 8.4.60 Python-3.13.3 torch-2.12.0+cpu CPU (12th Gen Intel Core(TM) i7-1265U)
engine\trainer: agnostic_nms=False, amp=True, angle=1.0, augment=False, auto_augment=randaugument, batch=2, bgr=0.0,
Downloading https://ultralytics.com/assets/Arial.ttf to 'C:\Users\vlday\AppData\Roaming\Ultralytics\Arial.ttf': 100%
Overriding model.yaml nc=80 with nc=5

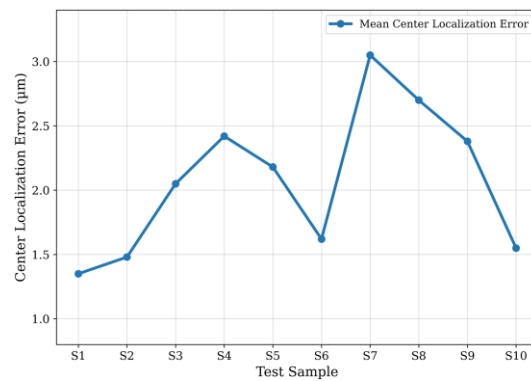
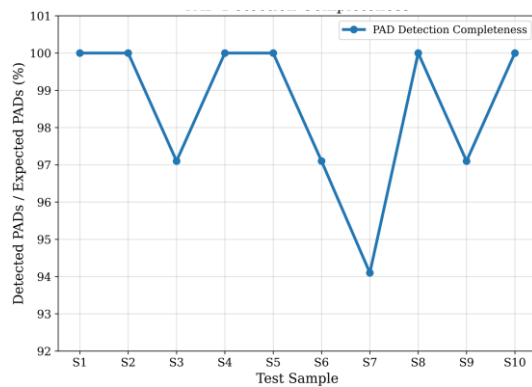
      from  n  params module                                arguments
0         -1  1    464  ultralytics.nn.modules.conv.Conv                [3, 16, 3, 2]
1         -1  1   4672 ultralytics.nn.modules.conv.Conv                [16, 32, 3, 2]
2         -1  1   6640 ultralytics.nn.modules.block.C3k2                  [32, 64, 1, False, 0.25]
3         -1  1  36992 ultralytics.nn.modules.conv.Conv                [64, 64, 3, 2]
4         -1  1  26080 ultralytics.nn.modules.block.C3k2                  [64, 128, 1, False, 0.25]
5         -1  1 147712 ultralytics.nn.modules.conv.Conv                [128, 128, 3, 2]
6         -1  1  87040 ultralytics.nn.modules.block.C3k2                  [128, 128, 1, True]
7         -1  1 295424 ultralytics.nn.modules.conv.Conv                [128, 256, 3, 2]
8         -1  1 346112 ultralytics.nn.modules.block.C3k2                  [256, 256, 1, True]
9         -1  1 164608 ultralytics.nn.modules.block.SPPF                  [256, 256, 5, 3, True]
10        -1  1 249728 ultralytics.nn.modules.block.C2PSA                  [256, 256, 1]
11        -1  1     0  torch.nn.modules.upsampling.Upsample          [None, 2, 'nearest']
12        [-1, 6] 1     0  ultralytics.nn.modules.conv.Concat                [1]
13        -1  1 119808 ultralytics.nn.modules.block.C3k2                  [384, 128, 1, True]
14        -1  1     0  torch.nn.modules.upsampling.Upsample          [None, 2, 'nearest']
15        [-1, 4] 1     0  ultralytics.nn.modules.conv.Concat                [1]
16        -1  1  34304 ultralytics.nn.modules.block.C3k2                  [256, 64, 1, True]
17        -1  1  36992 ultralytics.nn.modules.conv.Conv                [64, 64, 3, 2]
18        [-1, 13] 1     0  ultralytics.nn.modules.conv.Concat                [1]
19        -1  1  95232 ultralytics.nn.modules.block.C3k2                  [192, 128, 1, True]
20        -1  1 147712 ultralytics.nn.modules.conv.Conv                [128, 128, 3, 2]
21        [-1, 10] 1     0  ultralytics.nn.modules.conv.Concat                [1]
22        -1  1 463104 ultralytics.nn.modules.block.C3k2                  [384, 256, 1, True, 0.5, True]
23        [16, 19, 22] 1 243126 ultralytics.nn.modules.head.Detect                  [5, 1, True, [64, 128, 256]]

YOLO26n summary: 260 layers, 2,505,750 parameters, 2,505,750 gradients, 5.8 GFLOPs
```

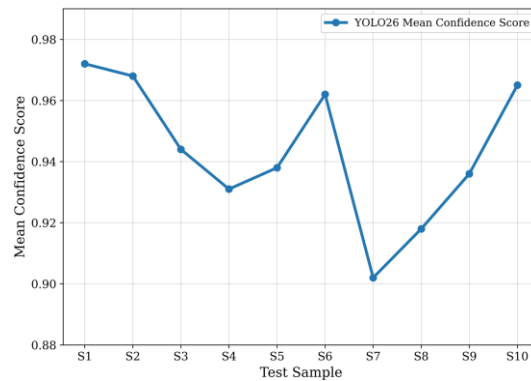
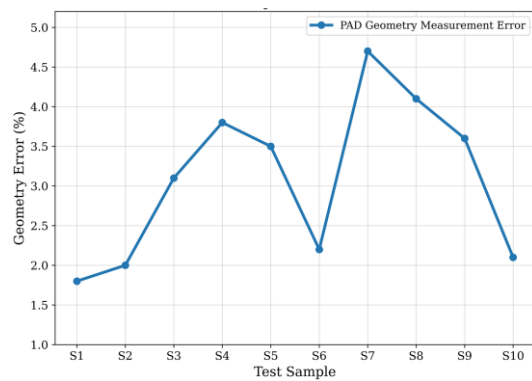
Рис. 5. Фрагмент побудови конфігурації мережі із зазначенням типів шарів, кількості параметрів, міжшарових зв'язків та обчислювальної складності моделі, що необхідно для подальшого навчання на сформованому CCD/PRU-датасеті

Таблиця 2 – Результати інтелектуальної оцінки якості PAD контактів на базі навченого Datasets

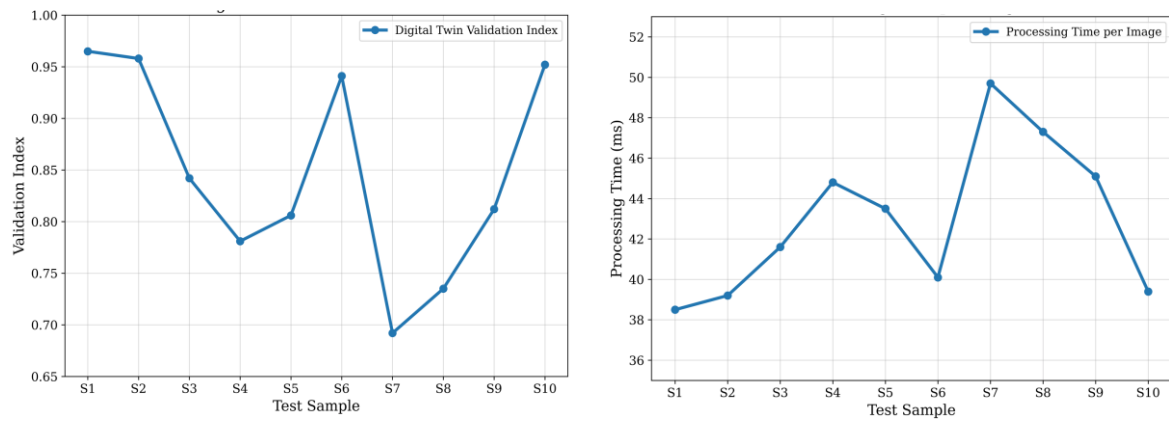
Розпізнані PAD	Результат оцінки якості PAD для Digital Twin FEK DELVOTEC 6400	Клас
	Recognition mode: YOLO26 Scale: 2.9851 μm/px Total number of detected PADs: 34 PAD OK: 34 PAD contamination: 0 PAD damage: 0 PAD missing: 0 PAD oxidation: 0 RISK: 1 FAIL: 0	pad_ok
	Recognition mode: YOLO26 Scale: 2.9851 μm/px Total number of detected PADs: 34 PAD OK: 22 PAD contamination: 4 PAD damage: 0 PAD missing: 0 PAD oxidation: 8 RISK: 4 FAIL: 8	pad_oxidation
	Recognition mode: YOLO26 Scale: 2.9851 μm/px Total number of detected PADs: 34 PAD OK: 33 PAD contamination: 0 PAD damage: 0 PAD missing: 1 PAD oxidation: 0 RISK: 1 FAIL: 1	pad_damage



a 6



B Γ



д е

Рис. 6. Результати проведених експериментів:

а – PAD Detection Precision; б – Center Localization Error;

в – PAD Geometry Measurement Error; г – Vision Confidence Score; д – Digital Twin Validation Index; е – Processing Time

Відповідно до результатів моделювання, наведених на рис. 6а, розроблений метод забезпечує повне виявлення контактних площадок у більшості експериментів, де показник Detection Completeness становить 100 % для зразків S1, S2, S4, S5, S8 та S10. Для зразків із дефектами значення знижується до 97,1 % (S3, S6, S9) та 94,1 % (S7), що свідчить про вплив пошкоджених або відсутніх PAD на ефективність детектування, однак навіть мінімальний показник підтверджує високу надійність запропонованого методу. Як видно з рис. 6б, середня похибка визначення координат центрів контактних площадок змінюється в межах від 1,35 μm до 3,05 μm , причому найменші значення отримано для якісних зразків S1 та S10. Максимальна похибка 3,05 μm спостерігається для зразка S7 із найбільшою кількістю дефектів, що підтверджує стійкість алгоритму локалізації та достатню точність визначення координат PAD для задач цифрового двійника FEK Delvotec 6400. З результатів, наведених на рис. 6в, встановлено, що відносна похибка визначення геометричних параметрів PAD знаходиться у межах від 1,8 % до 4,7 %, причому для більшості коректно сформованих контактних площадок вона не перевищує 2,2 %. Максимальна похибка 4,7 % отримана для зразка S7, однак навіть у цьому випадку точність вимірювання залишається достатньою для автоматизованої оцінки якості контактних площадок перед виконанням процесу клинового зварювання. Відповідно до рис. 6г, середній коефіцієнт впевненості моделі YOLO26 змінюється від 0,902 до 0,972, що свідчить про стабільну роботу нейронної мережі при розпізнаванні PAD різного технічного стану. Найменше значення отримано для зразка S7, тоді як для справних контактних площадок коефіцієнт перевищує 0,96, підтверджуючи високу достовірність класифікації та коректність навченої моделі. Як показано на рис. 6д, інтегральний індекс валідації Digital Twin знаходиться в інтервалі від 0,692 до 0,965, що відображає ступінь відповідності цифрової моделі реальному стану контактних площадок. Для більшості експериментів значення

Validation Index перевищує 0,94, а його зниження до 0,692 для зразка S7 пояснюється одночасною наявністю декількох дефектів, які суттєво впливають на достовірність цифрової репрезентації технологічного процесу. Результати, представлені на рис. 6е, демонструють, що час обробки одного CCD/PRU-зображення змінюється від 38,5 ms до 49,7 ms, причому для більшості експериментів він становить 39–45 ms. Навіть максимальний час 49,7 ms не перевищує 50 ms, що підтверджує можливість використання розробленого інтелектуального методу у складі Digital Twin FEK Delvotec 6400 для автоматизованого контролю PAD контактів у режимі, близькому до реального часу.

За результатами проведеного моделювання підтверджено, що розроблений інтелектуальний метод контролю якості PAD контактів забезпечує високу точність їх автоматичного виявлення, локалізації, визначення геометричних параметрів і класифікації дефектів, зберігаючи стійкість роботи навіть за наявності технологічних відхилень контактних площадок. Отримані результати підтверджують адекватність запропонованих математичних моделей, ефективність навченої нейронної мережі YOLO26 та доцільність інтеграції розробленого методу до цифрового двійника FEK Delvotec 6400 для підвищення достовірності моделювання процесу клинового зварювання й автоматизованого прогнозування якості мікрозварних з'єднань.

Висновки

Розроблений метод інтелектуальний метод контролю якості PAD-контактів поєднує математичні моделі системи технічного зору, нейромережеве розпізнавання, координатні перетворення, автоматичну корекцію положення та інтегральну валідацію результатів, що забезпечує підвищення точності автоматизованого оцінювання стану контактних площадок і достовірності цифрового моделювання процесу клинового зварювання. Результати експериментальних досліджень підтвердили високу ефективність

запропонованого підходу, забезпечивши повноту виявлення контактних площадок $\approx 100\%$, похибку локалізації центрів у межах 1,35–3,05 мкм, похибку визначення геометричних параметрів до 4,7 %, високий рівень достовірності класифікації дефектів та час обробки одного зображення, що не перевищує 50 мс, що підтверджує можливість його використання в режимі, близькому до реального часу. Розроблений метод доцільно використовувати у цифрових двійниках обладнання для клинового та дротяного зварювання, автоматизованих системах контролю якості мікроелектронних виробів, виробництві гібридних інтегральних схем, роботизованих комплексах прецизійного складання та інтелектуальних системах машинного зору мікроелектронного виробництва.

Перспективними напрямками подальших досліджень є розширення методу для тривимірного аналізу процесу зварювання, інтеграція мультимодальних сенсорних даних, використання трансформерних архітектур комп'ютерного зору та розроблення самонавчального цифрового двійника

з можливістю прогнозування якості зварних з'єднань і автоматичної оптимізації технологічних параметрів у реальному часі.

Конфлікт інтересів

Автори декларують, що не мають конфлікту інтересів стосовно даного дослідження, в тому числі фінансового, особистісного характеру, авторства чи іншого характеру, що міг би вплинути на дослідження та його результати, представлені в даній статті.

Фінансування

Дослідження проводилося без фінансової підтримки.

Доступність даних

Рукопис не має пов'язаних даних.

Використання засобів штучного інтелекту

Автори підтверджують, що не використовували технології штучного інтелекту при створенні представленої роботи.

СПИСОК ЛІТЕРАТУРИ

1. Pham, H. S. T., and Li, S. B. (2026), "Towards Industry 5.0: a conceptual model for leading organisational change in digital age", *Journal of Organizational Change Management*, vol. 39(1), pp. 125–161, doi: <https://doi.org/10.1108/JOCM-03-2025-0264>
2. Невлюдов, І. Ш., Євсєєв, В. В., Максимова, С. С. (2026), *Industry 5.0 та колаборативна робототехніка: динамічний опис навколишнього середовища роботів-маніпуляторів з використанням мови Python*, монографія, Видавництво Іванченка І. С., Харків, 279 с., doi: <https://doi.org/10.30837/978-617-8332-95-2>
3. Buchner, C., Riedle, B., Krauß, J., Seidler, C. T., Huber, M. F., Eigenbrod, H., and Schlicht, F. (2026), "Machine learning-driven multi-objective parameter optimization for sustainable, efficient, and high-quality ultrasonic wire bonding", *Journal of Intelligent Manufacturing*, vol. 37(4), pp. 1681–1699, doi: <https://doi.org/10.1007/s10845-025-02615-3>
4. Yuan, T., Zhou, M., Jiang, W., and Yi, D. (2026), "High-curvature bonding wire metrology via hemispherical light homogenization and gradient optimization", *Optics and Laser Technology*, 193, doi: <https://doi.org/10.1016/j.optlastec.2025.114177>
5. Nevliudov, I., Yevsieiev, V., Horovyi, K. (2026), "Analysis of the FEK Delvotec 6400 Wedge Bonding System", *Débats scientifiques et orientations prospectives du développement scientifique*, pp. 89–92, doi: <https://doi.org/10.36074/logos-05.06.2026>
6. Nevliudov, I., Yevsieiev, V., Horovyi, K. (2026), "Analysis of the Technological Process of Wedge Bonding Using FEK Delvotec 6400 in Microelectronics", *Current scientific goals, approaches and challenges*, Dresden, Federal Republic of Germany: International Center of Scientific Research, pp. 117–119, doi: <https://doi.org/10.36074/scientia-05.06.2026>
7. Kumar, S. (2026), "Improving Inspection Reliability of Corner Weld Joints Using Phased Array Ultrasonic Testing: A Review for Failure Prevention", *Journal of Failure Analysis and Prevention*, doi: <https://doi.org/10.1007/s11668-026-02413-w>
8. Li, C., Liu, Z., Ji, H., Ma, J., Wang, Z., Chen, R., and Wu, H. (2026), "Innovative method for the vacuum vessel welds NDT of compact fusion reactor", *Fusion Engineering and Design*, vol. 228, doi: <https://doi.org/10.1016/j.fusengdes.2026.115767>
9. Huang, J., Huang, Z., Li, C., and Liu, J. (2026), "Research on defect detection method of motor control board based on image processing", *Journal of Computational and Applied Mathematics*, vol. 473, doi: <https://doi.org/10.1016/j.cam.2025.116916>
10. Yue, H., Wang, Q., Zhao, M., Yang, Z., and Lu, L. (2026), "Advancements in digital twin applications for intelligent construction quality management", *Journal of Construction Engineering and Management*, vol. 152(1), 03125011, doi: <https://doi.org/10.1061/JCEMD4.COENG-17126>
11. Yu, D., Cao, Y., and Zhao, Q. (2025), "Detection and analysis of asphalt pavement texture depth based on digital image analytics technology", *International Journal of Pavement Research and Technology*, vol. 18(3), pp. 628–637, doi: <https://doi.org/10.1007/s42947-023-00368-x>
12. Ahmadvpour, S. S., Avval, D. B., Navimipour, N. J., Rasmi, H., Heidari, A., Kassa, S., and Patidar, M. (2025), "A new median filter circuit design based on atomic silicon quantum-dot for digital image processing and IoT applications", *IEEE Internet of Things Journal*, doi: <https://doi.org/10.1109/JIOT.2025.3553706>
13. Nevliudov, I., Yevsieiev, V., and Horovyi, K. (2026), "Model of the Technological Process of Wedge Bonding Based on FEK Delvotec 6400 in the Production of Hybrid Electronic Modules", *Proceedings of the XII International Scientific and Theoretical Conference*, International Center of Chicago, USA, pp. 170–173, doi: <https://doi.org/10.36074/scientia-12.06.2026>
14. Martini, G., Paroli, B. and Milani, P. (2025), "Non-linear reconfigurable threshold logic gates based on nanostructured metallic films", *npj Unconv. Comput*, vol. 2(23), doi: <https://doi.org/10.1038/s44335-025-00037-1>
15. (2026), *Bonder 6400 User Guide v3.0*, URL: <https://www.scribd.com/document/827992279/6400GB>
16. Apostolidis, K. D., and Papakostas, G. A. (2025), "Delving into yolo object detection models: Insights into adversarial robustness", *Electronics*, vol. 14(8), 1624, doi: <https://doi.org/10.3390/electronics14081624>
17. Abu-Jassar, A. T., Attar, H., Amer, A., Lyashenko, V., Yevsieiev, V., and Solyman, A. (2025), "Development and investigation of vision system for a small-sized mobile humanoid robot in a smart environment", *International Journal of Crowd Science*, vol. 9(1), pp. 29–43, doi: <https://doi.org/10.26599/IJCS.2023.9100018>

18. Abu-Jassar, A., Attar, H., Hafez, M., Hamdan, M., Yevsieiev, V., and Lyashenko, V. (2025), "Adaptive System for Constructing a Mobile Robot Trajectory Using Neural Network and Stochastic Optimizer", 2025 26th International Arab Conference on Information Technology, IEEE, pp. 422–428, doi: <https://doi.org/10.1109/ACIT68900.2025.11510567>
19. Nguyen, D. T., Nguyen, T. N., Kim, H., and Lee, H. J. (2019), "A high-throughput and power-efficient FPGA implementation of YOLO CNN for object detection", *IEEE Trans. on VLSI Systems*, vol. 27(8), pp. 1861–1873, doi: <https://doi.org/10.1109/TVLSI.2019.2905242>
20. Wei, G., Hou, Y., Cui, Q., Deng, G., Tao, X., and Yao, Y. (2018), "YOLO acceleration using FPGA architecture", 2018 IEEE/CIC Int. Conf. on Communications in China, IEEE, pp. 734–735, doi: <https://doi.org/10.1109/ICCChina.2018.8641256>
21. Sapkota, R., and Karkee, M. (2025), "Ultralytics YOLO evolution: An overview of YOLO26, YOLO11, YOLOv8 and YOLOv5 object detectors for computer vision and pattern recognition", *arXiv*, preprint arXiv:2510.09653, doi: <https://doi.org/10.48550/arXiv.2510.09653>

Received (Надійшла) 15.06.2026

Accepted for publication (Прийнята до друку) 26.08.2026

Publication date (Дата публікації) 25.09.2026

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Невлюдов Ігор Шакирович – доктор технічних наук, професор, завідувач кафедри комп'ютерно-інтегрованих технологій, автоматизації, робототехніки та безпекової інженерії, Харківський національний університет радіоелектроніки, Україна;
Igor Nevliudov – Doctor of Technical Sciences, Professor, Head of the Department of Computer-Integrated Technologies, Automation, Robotics and Safety Engineering, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;
e-mail: igor.nevliudov@nure.ua; ORCID Author ID: <https://orcid.org/0000-0002-9837-2309>;
Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57216434058>.

Євсєєв Владислав В'ячеславович – доктор технічних наук, професор, професор кафедри комп'ютерно-інтегрованих технологій, автоматизації, робототехніки та безпекової інженерії, Харківський національний університет радіоелектроніки, Харків, Україна;
Vladyslav Yevsieiev – Doctor of Technical Sciences, Professor, Professor of the Department of Computer-Integrated Technologies, Automation, Robotics and Safety Engineering, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;
e-mail: vladyslav.yevsieiev@nure.ua; ORCID Author ID: <https://orcid.org/0000-0002-2590-7085>;
Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57190568855>.

Чала Олена Олександрівна – кандидат технічних наук, доцент, доцент кафедри комп'ютерно-інтегрованих технологій, автоматизації, робототехніки та безпекової інженерії, Харківський національний університет радіоелектроніки, Україна;
Olena Chala – Candidate of Technical Sciences, Associate Professor, Associate Professor of the Department of Computer-Integrated Technologies, Automation, Robotics and Safety Engineering, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;
e-mail: olena.chala@nure.ua; ORCID Author ID: <https://orcid.org/0000-0003-2454-3774>;
Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57211027250>.

Горовий Кирило Юрійович – аспірант кафедри комп'ютерно-інтегрованих технологій, автоматизації, робототехніки та безпекової інженерії, Харківський національний університет радіоелектроніки, Харків, Україна;
Kyrylo Horovyi - PhD student, Department of Computer-Integrated Technologies, Automation, Robotics and Safety Engineering, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;
e-mail: kyrylo.horovyi@nure.ua ORCID Author ID: <https://orcid.org/0009-0003-7272-9626>.

Development of an intelligent method for PAD contact quality control in the FEK DELVOTEC 6400 digital twin for modeling the wedge bonding process in microelectronics

Igor Nevliudov, Vladyslav Yevsieiev, Olena Chala, Kyrylo Horovyi

Abstract. The object of the study is the process of automated quality control of bonding pads in the digital twin of the FEK Delvotec 6400 equipment used for modeling the wedge bonding process in microelectronics. The subject of the study comprises the methods, mathematical models, and algorithmic tools for intelligent quality control of bonding pads in the digital twin of the FEK Delvotec 6400 equipment for modeling the wedge bonding process in microelectronics. The aim of the study is to develop an intelligent method for the automated assessment of bonding pad conditions, providing improved positioning accuracy, enhanced reliability of digital modeling, and better decision-making quality during the technological process. To achieve this aim, mathematical models for machine vision image formation, coordinate space calibration, neural network-based recognition of bonding pads and fiducial markers, refinement of center coordinates, coordinate transformations, automatic correction of the wire bonding trajectory, positioning accuracy assessment, and integrated validation of computer vision results prior to their use in the digital twin were developed. Methodologically, the study is based on mathematical modeling, modern computer vision techniques, deep learning, convolutional neural networks, object detection methods, digital image processing, automated analysis of the geometric parameters of bonding pads, and numerical modeling. For experimental validation, an augmented dataset of bonding pad images was created, a neural network model was trained, and a series of experiments was conducted to determine bonding pad detection completeness, the accuracy of center localization, errors in measuring geometric parameters, the reliability of defect classification, the integral validation index of the digital twin, and the processing time per image. The obtained results confirmed the high efficiency of the proposed approach, providing bonding pad detection completeness of up to 100%, center localization errors within 1.35–3.05 μm , errors in determining geometric parameters of up to 4.7%, a high level of defect classification reliability, and an image processing time not exceeding 50 ms. The practical significance of the study lies in the possibility of integrating the proposed method into the FEK Delvotec 6400 digital twin to enable automated quality control of bonding pads, improve the accuracy of wedge bonding process modeling, ensure timely defect detection, and maintain the stability of the technological process in microelectronic manufacturing.

Keywords: digital twin; bonding pads; wedge bonding; computer vision; neural networks; quality control; mathematical modeling; microelectronics; automated control; digital image processing.

І. С. Рудаков, О. В. Коломійцев

Національний технічний університет “Харківський політехнічний інститут”, Харків, Україна

МЕТОД ТА МОДЕЛЬ ПЛАНУВАННЯ МАРШРУТУ ПОЛЬОТУ ГРУПИ БЕЗПІЛОТНИХ ЛІТАЛЬНИХ АПАРАТІВ ДЛЯ УНИКНЕННЯ ПЕРЕШКОД У МІСЬКІЙ ЗАБУДОВІ НА ОСНОВІ ТЕХНОЛОГІЙ МАШИННОГО НАВЧАННЯ

Анотація. Актуальність. Збільшення кількості безпілотних літальних апаратів (БпЛА) з кожним роком призводить до підвищення щільності їх польотів та ймовірності зіткнень, що потребує забезпечення високого рівня безпеки їх польотів. Застосування концепції динамічних коридорів щільності D2-C2 дає змогу знизити частоту виникнення конфліктів, але не усуває повністю ризик непередбачених зближень БпЛА з перешкодами (статичними та динамічними) – за наявності природних явищ та динамічних факторів. Таким чином, удосконалення методу планування траєкторій БпЛА завдяки впровадженню засобів прогнозування можливих конфліктів на основі технологій машинного навчання є актуальною науковою задачею. **Предметом дослідження** в статті є інтеграція концепцію D2-C2 з модулем штучного інтелекту. **Метою дослідження** є підвищення рівня безпеки польотів групи БпЛА у міській забудові. **Завданням дослідження** є розробка моделі та удосконалення методу планування траєкторій БпЛА завдяки впровадженню засобів прогнозування можливих конфліктів. В ході досліджень використано методи математичної статистики та моделювання. **Результатами дослідження** у статті є запропоновано неймережеву модель LSTM, яка навчена на даних про зближення БпЛА. Удосконалено метод планування траєкторій БпЛА з проактивним уникненням конфліктів, який здійснює прогнозне виявлення небезпеки за допомогою ML-моделі. **Висновки.** Запропонована модель апроксимує функцію, яка відображає складну залежність між часовими рядами параметрів руху двох (і більше) БпЛА та бінарним результатом – зіткнення чи ні. Удосконалений метод інтегрує концепцію D2-C2 з модулем штучного інтелекту, що дозволяє завчасно та автоматично виявляти потенційні зіткнення БпЛА між собою та перешкодами. Отримані результати можуть бути у подальшому використані для створення нових поколінь автономних систем керування БпЛА, здатних адаптивно та навчально покращувати свої навички уникнення небезпеки у складних умовах повітряного руху.

Ключові слова: безпілотний літальний апарат, траєкторія польоту, машинне навчання, штучний інтелект, нейронна мережа, безпека, зіткнення, метод, математична модель, алгоритм.

1. Вступ

1.1 Актуальність дослідження

Стрімке розширення сфер застосування безпілотних літальних апаратів (БпЛА) та збільшення інтенсивності їх використання зумовлюють необхідність підвищення рівня автоматизації й безпеки польотів, особливо в умовах складного низьковисотного повітряного простору міської забудови. Урбанізоване середовище характеризується значною кількістю стаціонарних перешкод, складною тривимірною структурою забудови, обмеженою видимістю, можливим погіршенням або втратою сигналів глобальної супутникової навігації, а також наявністю динамічних об'єктів [1]. За таких умов планування безпечного маршруту польоту БпЛА є складною багатокритеріальною задачею, що потребує одночасного врахування геометрії перешкод, динамічних характеристик літальних апаратів, ризиків зіткнення, часу виконання місії тощо [2, 3].

Зростання кількості БпЛА у спільному повітряному просторі закономірно призводить до збільшення ймовірності виникнення потенційно конфліктних ситуацій. Під конфліктною ситуацією доцільно розуміти небезпечне зближення БпЛА між собою або з перешкодами, за якого відстань між об'єктами стає меншою за встановлену мінімально допустиму дистанцію безпеки. На відміну від пілотованої авіації, для якої традиційно застосовуються централізовані процедури управління повітряним рухом, масове використання БпЛА потребує більш гнучких, масштабованих і значною

мірою автоматизованих методів організації та планування руху. Водночас застосування лише заздалегідь визначених маршрутів або традиційних алгоритмів планування не завжди забезпечує необхідний рівень безпеки в умовах невизначеності та швидкої зміни параметрів польотної обстановки [4, 5].

Особливо складною є задача планування траєкторій у міському середовищі. Висотні будівлі та інші об'єкти формують складний тривимірний простір, у якому необхідно визначати маршрути з урахуванням не лише геометричних обмежень, а й навігаційних ризиків. Зокрема, щільна забудова може спричиняти екранування або погіршення якості сигналів GNSS, що збільшує невизначеність визначення координат БпЛА та, відповідно, ризик відхилення від запланованої траєкторії. Сучасні дослідження пропонують враховувати такі ризики безпосередньо під час побудови маршруту [6].

Одним із напрямів підвищення безпеки та впорядкування руху БпЛА є структуризація низьковисотного повітряного простору шляхом формування маршрутів, секторів або коридорів руху. Такий підхід дає змогу зменшити кількість потенційних конфліктів завдяки просторовому розділенню потоків БпЛА та узгодженню напрямків їх руху. Але навіть за умови попереднього розподілу повітряного простору ризик непередбачених зближень повністю не усувається. Причинами цього можуть бути пориви вітру, похибки систем навігації та позиціонування, затримки або втрати зв'язку, відмови окремих бортових систем, а також поява нових перешкод, тому система планування маршруту

повинна забезпечувати не лише формування початкової безпечної траєкторії, а й можливість оперативного прогнозування небезпечних ситуацій та коригування маршруту в реальному часі [7].

Важливого значення набуває застосування технологій машинного навчання, насамперед методів глибокого навчання та навчання з підкріпленням. Їхньою перевагою є здатність формувати стратегії керування та планування на основі даних про поточний стан навколишнього середовища, виявляти закономірності у складних ситуаціях і адаптувати поведінку БпЛА до зміни польотної обстановки. Дослідження останніх років демонструють можливість застосування глибокого навчання з підкріпленням для автономного слідування заданому маршруту, локального уникнення перешкод, навігації у невідомому середовищі та побудови траєкторій у тривимірному просторі [8].

Подальший розвиток цього напрямку пов'язаний із використанням гібридних моделей, у яких глобальна інформація про маршрут поєднується з локальними даними бортових сенсорів. Такий підхід дозволяє забезпечити досягнення кінцевої точки за оптимальною траєкторією, та оперативно реагувати на перешкоди та непередбачувані зміни обстановки. Особливий інтерес становлять моделі, здатні здійснювати прогнозування можливих конфліктів, оцінювати рівень безпеки поточної ситуації та завчасно формувати рішення щодо зміни напрямку або параметрів руху БпЛА. Застосування таких моделей є особливо актуальним для групового польоту, оскільки в цьому випадку необхідно враховувати взаємний вплив траєкторій декількох БпЛА та забезпечувати їх безпечну координацію [9].

Таким чином, аналіз сучасних досліджень свідчить про актуальність розроблення методів і моделей планування маршруту польоту групи БпЛА в умовах міської забудови, які поєднують стратегічне формування маршруту з локальним прогнозуванням статичних і динамічних перешкод. Перспективним напрямом розв'язання цієї задачі є застосування технологій машинного навчання для прогнозування можливих конфліктних ситуацій, оцінювання ризику зіткнення та адаптивного коригування траєкторій польоту, тому необхідно розроблення моделі та вдосконалення методу планування маршруту польоту групи БпЛА, що забезпечуватимуть підвищення безпеки та ефективності функціонування в складному динамічному середовищі міської забудови.

1.2 Аналіз літературних джерел і визначення проблеми

У статті [10] розглянуто особливості застосування технологій машинного навчання для підвищення безпеки польотів групи безпілотних літальних апаратів (БпЛА) в умовах міської забудови. Особливу увагу приділено можливості використання інтелектуальних методів для підвищення автономності прийняття рішень і вдосконалення забезпечення безпеки польоту. Результати дослідження свідчать про перспективність подальшого розвитку цього

напрямку шляхом застосування методів машинного навчання для прогнозування небезпечних ситуацій.

У роботі [11] запропоновано застосування множинних мурашиних алгоритмів для розв'язання задач логістичного забезпечення під час групового використання вантажних БпЛА. Основну увагу приділено розподілу маршрутів між окремими БпЛА та оптимізації виконання завдань групою. Запропонований підхід забезпечує можливість пошуку раціональних маршрутів, однак питання оперативного уникнення перешкод у процесі польоту групи БпЛА не є основним предметом дослідження.

У джерелі [12] досліджено особливості підвищення ефективності управління БпЛА залежно від структури матриці зв'язків під час використання нелінійного псевдовипадкового генератора, а отримані результати спрямовані на вдосконалення процесів формування та передавання керуючої інформації. Але у роботі не розглянуто комплексне розв'язання задачі планування маршруту групи БпЛА з урахуванням перешкод.

У роботі [13] запропоновано систему планування маршрутів для рою БпЛА на основі Q-learning. Підхід орієнтований на визначення траєкторій руху в середовищі з перешкодами та передбачає можливість формування маршрутів без повної попередньої інформації про карту середовища. Перевагою роботи є застосування методів машинного навчання до задачі групового планування. Водночас основна увага приділена покриттю простору та вибору маршрутів, тоді як питання прогнозування конфліктних ситуацій між БпЛА та завчасного коригування їх траєкторій потребує подальшого дослідження.

У дослідженні [14], що присвячене уникненню перешкод і запобіганню зіткненням рою БпЛА на основі вдосконаленого алгоритму Vector Field Histogram (VFH) та стратегії обміну інформацією, запропоновано підхід до організації польоту групи в складному невідомому середовищі. Результати роботи підтверджують важливість координації та обміну даними для забезпечення безпечного групового польоту. Проте запропонований підхід спрямований на локальне реагування на перешкоди, що обумовлює доцільність поєднання таких методів із засобами довгострокового прогнозування конфліктів і глобального планування маршруту.

У роботі [15] розглянуто задачу кооперативної оптимізації планування маршрутів декількох БпЛА з урахуванням обмежень системи зв'язку. Для її розв'язання запропоновано кооперативну модель планування та вдосконалений алгоритм оптимізації рою частинок. Важливою перевагою дослідження є врахування того, що характеристики інформаційного обміну між БпЛА можуть безпосередньо впливати на можливість узгодженого виконання групового завдання. Разом із тим запропонований підхід зосереджений переважно на оптимізації маршрутів і комунікаційних обмеженнях, тоді як комплексне врахування міської забудови, динамічних перешкод і прогнозування ймовірних конфліктів потребує подальшого розвитку.

У роботах [16, 17] запропоновано метод тривимірного динамічного планування маршруту БПЛА на основі вдосконаленої моделі динамічної системи рідини та штучної нейронної мережі. Метод спрямований на підвищення якості сформованого маршруту й зменшення обчислювальних витрат під час польоту в складному середовищі зі статичними та раптово виникаючими динамічними перешкодами. Однак дослідження орієнтоване переважно на одиночний БПЛА, тому задача узгодженого планування та коригування маршрутів групи БПЛА потребує окремого розв'язання.

Отже, проведений аналіз літературних джерел показав, що сучасні дослідження охоплюють окремі аспекти групового застосування БПЛА, зокрема оптимізацію маршрутів, кооперативне планування, уникнення перешкод, запобігання зіткненням, обмін інформацією між літальними апаратами та використання технологій машинного навчання. Водночас існуючі підходи розглядають зазначені завдання окремо. Недостатньо дослідженим залишається комплексний метод планування маршруту польоту групи БПЛА в умовах міської забудови, який би одночасно враховував тривимірну структуру середовища, статичні та динамічні перешкоди, взаємне розташування БПЛА та прогнозування виникнення конфліктних ситуацій.

Таким чином, наукова проблема полягає у необхідності розроблення моделі та вдосконалення методу планування маршруту польоту групи БПЛА в умовах складного динамічного тривимірного середовища міської забудови. Перспективним напрямом її розв'язання є поєднання методів глобального планування маршруту з технологіями машинного навчання для прогнозування можливих конфліктів, оцінювання ризику зіткнення та адаптивного коригування траєкторій польоту. Такий підхід має забезпечити підвищення безпеки та ефективності групового застосування БПЛА.

1.3 Мета і завдання дослідження

Метою дослідження є підвищення рівня безпеки польотів групи БПЛА у міській забудові.

Для досягнення мети були поставлені наступні завдання: розробка моделі та удосконалення методу планування траєкторій БПЛА завдяки впровадженню засобів прогнозування можливих конфліктів. В ході досліджень використано методи математичної статистики та моделювання.

2. Матеріали і методи дослідження

У дослідженні удосконалено метод планування траєкторії польоту БПЛА на основі технологій машинного навчання, завдяки впровадженню засобів прогнозування можливих конфліктів.

Сутність удосконалення полягає у тому, щоб на етапі планування та виконання польоту БПЛА передбачати потенційно небезпечні їх зближення з перешкодами завчасно та динамічно коригувати маршрути польотів для уникнення зіткнення.

Зокрема, необхідно розробити нейромережеву модель, яка за даними про поточний політ БПЛА та

оточуючу обстановку прогнозує ризик зіткнення з перешкодами (як статичними, так й рухомими) на заданому інтервалі вперед.

Для вирішення поставленої задачі її необхідно формалізувати як задачу прогнозування часових послідовностей, яка піддається моделюванню нейронною мережею типу LSTM.

Розглянуто траєкторію польоту БПЛА як дискретну часову послідовність його станів: положення та динамічні параметри на послідовності часових кроків. Зроблено таке припущення:

$$p(t) = (x(t), y(t), z(t)) -$$

просторові координати БПЛА в момент часу t .

Також, розглянуто множину N потенційних перешкод (інших літальних апаратів або стаціонарних об'єктів) з такими координатами:

$$p_i(t) = (x_i(t), y_i(t), z_i(t))$$

для $i=1, 2, \dots, N$.

Введено метрику відстані між БПЛА та i -ою перешкодою:

$$d(t) = \|p(t) - p_i(t)\| = \sqrt{(x(t) - x_i(t))^2 + (y(t) - y_i(t))^2 + (z(t) - z_i(t))^2}, \quad (1)$$

Визначено порогову безпечну дистанцію D_{min} , менше за яку зближення БПЛА з перешкодою вважається небезпечним (конфліктом). Тоді, подію конфлікту можливо описати як існування моменту часу τ у межах прогнозованого горизонту ΔT , що

$$d_i(\tau) < D_{min} \quad (2)$$

для деякого i (тобто БПЛА надто близько наблизиться до перешкоди i).

Задача прогнозування конфлікту полягає у визначенні ймовірності виникнення такої події або факту її виникнення (True/False) до того, як вона станеться у реальності.

У термінах машинного навчання вищевказане завдання можливо подати як задачу прогнозування на основі часових рядів. Тобто, задані послідовності станів БПЛА та перешкод на інтервалі часу $[t_0, t_k]$.

Наприклад, виміряні або заплановані координати та швидкості на кількох останніх кроках. Необхідно передбачити, чи відбудеться конфлікт у інтервалі часу $[t_k, t_k + \Delta T]$ у майбутньому.

Вхідними даними моделі у момент часу t_k є наступна послідовність:

$$X = \{s(t_0), s(t_1), \dots, s(t_k)\}, \quad (3)$$

де $s(t_j)$ – вектор ознак, який характеризує стан системи «БПЛА-перешкоди» у час (t_j) .

До складу $s(t)$ можуть входити, наприклад: власні координати та вектор швидкості БПЛА, параметри його траєкторії (курс, висота та

прискорення); відносні координати найближчих перешкод або параметри зближення (відстань $d_i(t)$, відносна швидкість зближення тощо) для найбільш критичних i ; інші фактори (наприклад, швидкість вітру, ступінь невизначеності позиціонування).

Виходом моделі є прогнозоване значення цільової змінної y – показника наявності конфлікту у найближчому майбутньому. Можливо визначити y як бінарну змінну: $y=1$, якщо конфлікт очікується, та $y=0$ – інакше. Тоді модель реалізує таку функцію:

$$f: X \mapsto \hat{y}, \quad (4)$$

де $y \approx \hat{y}$ – передбачене значення (ймовірність або клас конфліктної ситуації) на основі аналізу послідовності X .

Обрано архітектуру Long short-term memory (LSTM) – різновид рекурентної нейронної мережі (PHM), яка призначена для моделювання залежностей у таких послідовностях X . Мережа LSTM послідовно обробляє вхідні дані за часовими кроками, акумулюючи внутрішній стан, який кодує інформацію про попередні етапи. У запропонованій постановці це дозволяє моделі «навчитися» закономірностям розвитку конфліктних ситуацій у часі та на основі поточної послідовності станів виявляти ознаки неминучого зіткнення ще до того, як відстань стане критично малою.

Таким чином, задача планування траєкторії (маршруту) польоту БПЛА з урахуванням ризику зводиться до задачі прогнозування часової послідовності засобами LSTM-моделі. Тобто, мережа розглядає історію польотів БПЛА та їх взаємного розташування, а також – потенційних перешкод та видає прогноз щодо безпеки продовження поточної траєкторії руху. Подібний підхід застосовується й в суміжних галузях. Наприклад, для автономних автомобілів розроблено LSTM-моделі прогнозування траєкторій сусідніх транспортних засобів (об'єктів), що дає змогу планувати більш безпечний маршрут руху наперед [4]. Запропоновано використання PHM LSTM за таких причин.

1. Наявність часової залежності – ситуація зближення двох БПЛА розвивається у часі та для надійного прогнозу необхідно враховувати послідовність попередніх станів. На відміну від звичайних (не рекурентних) нейронних мереж, які аналізують кожен стан незалежно, PHM мають пам'ять. Вони зберігають вектор прихованого стану h_t , що оновлюється на кожному кроці та несе інформацію про попередні кроки. LSTM спроектована для ефективного запам'ятовування довгострокових залежностей у послідовностях, завдяки введенню механізму керованої пам'яті завдяки гейтам (воротам) [5]. Такий підхід дозволяє їй вирішувати проблему «затухаючого градієнта», яка характерна для класичних PHM та утримувати важливу інформацію на довгих інтервалах.

2. Адаптація до змінної довжини послідовностей – у реальних умовах тривалість маневру зближення або сценарію конфлікту може

варіюватися. LSTM не потребує фіксованої довжини входу. Вона обробляє події послідовно та може працювати як з короткими, так й з довгими серіями даних, продовжуючи оновлювати свій стан. Це важливо для задачі прогнозування.

3. Гнучкість виходу – PHM LSTM можливо налаштувати як для видачі прогнозу після повного перегляду послідовності, так і для формування вихідної послідовності прогнозованих значень (many-to-many, наприклад, прогноз траєкторії руху іншого БПЛА на декілька секунд вперед). Обрано варіант mapu-to-one, де вихідний шар мережі дає оцінку ризику конфлікту на горизонті ΔT .

До складу архітектури LSTM-мережі входять особливі шари пам'яті – комірки LSTM, кожна з яких на кожному кроці виконує декілька операцій. Для чергового часу t та попереднього прихованого стану h_{t-1} і пам'яті C_{t-1} можливо провести розрахунок наступних внутрішніх величин:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f), \quad i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i),$$

$$C_t = f_t \odot C_{t-1} + i_t \odot C_t, \quad o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o),$$

$$C_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C), \quad h_t = o_t \odot \tanh(C_t). \quad (5)$$

де $\sigma(\cdot)$ – сигмоїда; $\tanh(\cdot)$ – гіперболічний тангенс;

символ $\odot$ – покомпонентне множення векторів.

Вектори W_f , W_i , W_C , W_o та відповідні b_f , b_i , b_C , b_o – параметри (вагові коефіцієнти та зміщення) моделі, які налаштовуються під час навчання.

Наведені формули описують роботу трьох внутрішніх гейтів LSTM: forget gate f_t визначає, яку частину попереднього стану пам'яті C_{t-1} слід забути; input gate i_t разом з кандидатним станом C_t визначає, яку інформацію слід додати до пам'яті; output gate o_t вирішує, яка частина оновленого стану C_t буде виведена як новий прихований стан h_t .

Завдяки даній будові, LSTM-комірка може керовано накопичувати інформацію (за i_t та C_t) або забувати її (за f_t), а також видавати релевантні дані (o_t) назовні. Такий підхід робить LSTM надзвичайно ефективною для проведення дослідження: мережа сама навчиться, які патерни входів слід запам'ятовувати (наприклад, поступове зближення курсів двох БПЛА), а які можливо ігнорувати як шум, щоби точно передбачити конфлікт.

Відомі підходи, такі як методи машинного навчання без урахування часової структури (наприклад, дерева рішень, Random Forest, метод опорних векторів тощо) або класичні PHM без гейтів, поступаються LSTM у здатності моделювати довготривалі залежності.

Отже, LSTM поєднує у собі універсальність нелінійної моделі (як і інші нейронні мережі) з «пам'яттю» про попередні стани. Це важливо для

виявлення передумов конфлікту. Іноді небезпечно зближення можливо розпізнати лише проаналізувавши тенденцію зміни відстані та відносної швидкості протягом певного часу, а не з одиничного вимірювання. LSTM як модель прогнозування конфліктів є обґрунтованим вибором,

підтвердженням її успішним застосуванням в суміжних задачах прогнозування траєкторій рухомих об'єктів та запобігання їх зіткненням. Для вирішення задачі прогнозування конфліктів (спроектовано архітектуру LSTM-мережі, яка адаптована до доменної області планування траєкторій БпЛА (рис. 1).

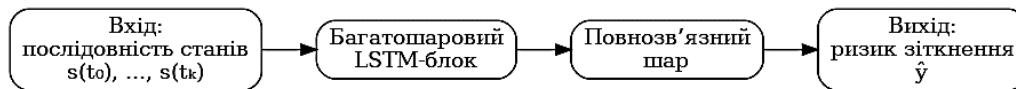


Рис. 1. Структурна схема нейромережевої моделі прогнозування конфліктів на основі LSTM

Вхідний шар моделі сприймає на кожному кроці часу t вектор ознак $s(t)$, що характеризують ситуацію. Вхідні дані нормалізуються та можуть додатково проходити через шар ущільнення (embedding) для зменшення розмірності, якщо кількість ознак велика.

Наступним компонентом є блок LSTM-осередків, які обробляють послідовність вхідних векторів, зберігаючи внутрішній стан.

У запропонованій моделі використано декілька шарів LSTM, які побудовані послідовно (мультишарова архітектура). Кількість шарів та кількість нейронів (осередків) у кожному шарі обрано експериментальним шляхом, забезпечуючи достатню здібність до апроксимації без перенавчання.

Після останнього LSTM-шару формується вихідний шар моделі. У випадку бінарної класифікації конфлікту, це один нейрон, до якого під'єднані виходи останнього прихованого стану h_t (на останньому кроці послідовності T) усіх нейронів передостаннього шару. Даний вихідний нейрон використовує сигмоїдальну активаційну функцію, щоб перетворити сумарний сигнал на ймовірність конфлікту $\hat{p}$.

Значення $\hat{p} > 0.5$ інтерпретується як прогноз наявності конфлікту (небезпечного зближення), $\hat{p} < 0.5$ – відсутності. За умови, якщо розглядати задачу як регресійну (прогноз мінімальної відстані або часу до зближення), то вихідний шар може містити один нейрон лінійної активації (для безперервного значення ризику) або декілька нейронів (для прогнозування траєкторії координат). Запропоновано використання класифікації «конфлікт / нормальний політ», що достатньо для подальшого планування траєкторії польоту БпЛА.

В архітектурі моделі передбачено механізми запобігання перенавчання та підвищення збіжності навчання. Зокрема, між LSTM-шарами впроваджено шари відсіву (Dropout) із коефіцієнтом, який оптимально підібрано експериментальним шляхом для зменшення залежності від окремих нейронів.

Також, використано регуляризацию вагових коефіцієнтів (L2-норма) у помірному режимі. Активация прихованих шарів LSTM – гіперболічний тангенс (вбудований у формули LSTM), що забезпечує значення виходів у діапазоні $[-1; 1]$ та стабілізує градієнти. Параметри моделі (кількість шарів, розмірність прихованого стану тощо) представлено у табл. 1.

Таблиця 1 – Параметри архітектури LSTM-моделі

Параметр	Значення
Тип моделі	RNN LSTM
Кількість LSTM-шарів	2 (послідовні)
Розмірність прихованого стану	64 (кожен LSTM-шар має 64 осередки)
Вихідний шар	1 нейрон (сигмоїдна активація)
Вхідна розмірність ознак	6 (координати і швидкості БпЛА та перешкоди)
Dropout після LSTM-шарів	0,2 (20%)
Регуляризація L2	0,001
Загальна кількість параметрів	~50 тис. (вагових коефіцієнтів)

За отриманими у таблиці 1 даними можливо стверджувати, що нейромережа має відносно невелику складність (порядку декількох десятків тисяч параметрів), що дозволяє здійснювати її навчання та виконання на борту БпЛА або у наземному центрі керування у реальному масштабі часу. Обраний розмір прихованого шару забезпечує баланс між здатністю моделі представляти складні залежності та швидкістю/енергоспоживанням обчислень. Два LSTM-шари дають змогу моделювати взаємозв'язки більш високого рівня (другий шар узагальнює послідовність виходів першого шару), що підвищує точність прогнозу без істотного збільшення обчислювальної складності.

Удосконалений метод планування траєкторії польоту БпЛА з урахуванням ризику зіткнення є поєднанням описаної вище RNN та алгоритму оперативного коригування маршруту. Структурно метод наведено у вигляді послідовності наступних кроків, сутність якого полягає у наступному.

1. Збір даних про стан: на початку кожного кроку планування траєкторії польоту отримуються актуальні дані про стан БпЛА та обстановку. Такі дані формують вектор ознак $s(t)$.

2. Прогноз конфліктів за допомогою LSTM: зібрана послідовність останніх k станів подається на вхід нейромережевої моделі LSTM. Модель обчислює прогноз щодо наявності конфлікту у найближчому майбутньому (на горизонті ΔT). За умови, якщо прогноз перевищує заданий поріг (наприклад, $\hat{p} > 0.5$ для ймовірності конфлікту), то вважається, що існує загроза зіткнення. У протилежному випадку – прогноз безпечний.

3. Аналіз результату та вибір дії – якщо конфлікт не прогнозується, то БпЛА продовжує рухатися за поточною траєкторією (або запланованим маршрутом) без змін. Якщо ж прогнозовано конфлікт, то система ініціює процедуру уникнення, залежно від ситуації може бути обрано різні маневри: зміну висоти, відхилення вбік від маршруту, гальмування або зупинку (для мультикоптера) тощо. Мета маневру – збільшити мінімальну відстань до перешкоди та уникнути перетину траєкторій у критичній точці.

4. Коригування траєкторії – на основі обраного маневру відбувається коригування наступних точок траєкторії БпЛА. Наприклад, якщо прогнозовано зіткнення з іншим БпЛА на перетині курсів, то планується зміна курсу або висоти апарату, аби розійтися у просторі. Корекція може відбуватись у автоматичному режимі автопілотом або надсилатися як рекомендація оператору БпЛА.

5. Циклічність процесу – описані кроки повторюються на кожному такті оновлення даних (наприклад, щосекунди).

Таким чином, метод працює у режимі реального часу, постійно відстежуючи ситуацію та завчасно реагуючи на можливі зміни. Фактично – це система підтримки прийняття рішень пілота, але –

автоматизована та налаштована на специфіку польотів БпЛА. Даний метод поєднує стратегічне планування маршруту польоту БпЛА (наприклад, по динамічним коридорам за концепцією D2-C2) з тактичною превентивною перевіркою майбутнього маршруту на безпеку. Модель LSTM може прогнозувати конфлікти навіть за відсутності явних поточних ознак (наприклад, інший БпЛА ще далеко), спираючись на тенденції – це дає виграв у часі на реакцію. Метод інтегрований у загальну архітектуру системи управління БпЛА або UTM (системи управління безпілотним трафіком). Він отримує дані від підсистем спостереження (радар, ADS-B, обмін даними між БпЛА) та передає команди на підсистему автопілота для виконання маневру ухилення.

Таким чином, досягається замкнений контур: сприйняття – прогноз – дія, що підвищує автономність та безпеку польоту БпЛА.

Розглянутий метод та РНМ тісно пов'язані між собою, утворюючи єдиний комплекс забезпечення безпеки польоту групи БпЛА в міській забудові. Модель LSTM виступає основним компонентом методу – її можливо розглядати як модуль прогнозування у загальному алгоритмі планування.

На структурній схемі (рис. 2) показано, як саме модель вбудована в удосконалений метод.

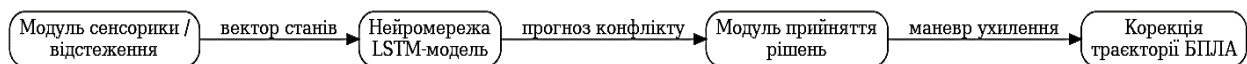


Рис. 3. Взаємозв'язок між нейронною мережею та методом прогнозування конфліктів

Модель LSTM отримує на вході дані від модуля сенсорики/відстеження та передає результати (прогноз конфлікту) до модуля прийняття рішень, що змінює траєкторію польоту БпЛА.

Перед використанням у реальному польоті БпЛА РНМ спочатку навчають на вибірці типових сценаріїв. Після навчання модель вбудовується у програмне забезпечення автопілота або наземної системи керування польотом БпЛА.

Удосконалений метод – це алгоритмічна надбудова, яка організовує роботу моделі та інших компонентів. Під час виконання метод на кожному кроці надає моделі актуальні дані (вектор ознак $s(t)$) та отримує від неї прогноз. Далі, метод визначає потрібну дію та реалізує її за відповідними виконавчими механізмами БпЛА.

Розроблена модель та удосконалений метод утворюють ієрархічну систему: модель LSTM – вирішує вузькоспеціалізовану задачу (оцінка ризику), тоді як метод – відповідає за загальну поведінку польоту БпЛА.

Даний підхід відповідає принципам модульності – модель можливо удосконалювати або перевчити незалежно від решти систем, наприклад, на нових даних, а потім, оновити модуль прогнозування, не змінюючи логіку усього алгоритму.

Ефективність удосконаленого методу напряму залежить від точності та швидкодії розробленої моделі. За умови, якщо модель дає хибні спрацювання або пропуски, то метод може вживати

зайвих маневрів або не зреагувати вчасно. Тому, у взаємозв'язку «модель-метод» реалізовано механізми перевірки достовірності прогнозу (наприклад, введено поріг впевненості та фільтрацію одиничних спрацювань).

Таким чином, розроблена модель та удосконалений метод доповнюють один одного: РНМ забезпечує інтелектуальне прогнозування на основі накопичених даних, а метод – забезпечує реалізацію дій для уникнення прогнозованих загроз. Їх поєднання дає синергійний ефект – підвищує ситуаційну обізнаність системи управління польотом БпЛА та дозволяє реалізувати принцип проактивної безпеки їх польотів у міській забудові.

Розроблена модель та удосконалений метод реалізовані у програмному середовищі для подальших експериментальних досліджень

На етапі реалізації визначено ряд параметрів, наведених нижче у таблицях, які забезпечують відтворюваність та налаштування запропонованого підходу під конкретні умови.

У табл. 2 представлено основні параметри навчання РНМ LSTM та характеристику даних для навчання. Модель навчено на певному наборі сценаріїв зближення БпЛА, згенерованих у симуляційному середовищі. Модель навчалася протягом 50 епох на 5000 симульованих конфліктних та безконфліктних траєкторіях, досягнувши високої точності класифікації ~95%.

Горизонт прогнозування обрано 10 с – цього достатньо, щоб встигнути виконати маневр ухилення

БпЛА на типовій швидкості та уникнути зіткнення. Дані для навчання згенеровано таким чином, щоб включати різні сценарії: зближення на зустрічних курсах, наздоганяння, зближення при маневруванні, обліт перешкод тощо, з варіаціями початкових умов та шумами вимірювань.

Таблиця 2 – Параметри навчання LSTM-моделі та даних

Параметр	Значення
Апаратна платформа БпЛА	NVIDIA Jetson TX2 (GPU) / Raspberry Pi 4 (CPU)
Час виконання прогнозу	~5-10 мс (на Jetson TX2, для однієї пропагації LSTM)
Затримка циклу ухилення	< 1 с (вкл. збір даних, прогноз, прийняття рішення)
Комунікація між БпЛА	обмін координатами кожні 1 с (V2V зв'язок)
Інтеграція в автопілот	PX4 Autopilot (через MAVSDK офлайн планувальник)
Програмна реалізація	Python 3.9, фреймворк TensorFlow 2.6, середовище Ubuntu 20.04

Метод розрахований на роботу в межах концепції D2-C2. Тому, симуляцією враховано наявність динамічних сегментів повітряного простору. Симуляційні випробування проведено в умовному міському просторі розміром 10×10 км з декількома десятками БпЛА, що рухаються за правилами концепції D2-C2 (по динамічних коридорах) із випадковими відхиленнями. Безпечна дистанція 50 м обрана відповідно до рекомендацій U-sparse для малих БпЛА (дронів) у міській забудові. Кожен БпЛА обладнаний системою прогнозування; конфлікт вважається успішно попередженим, якщо мінімальна відстань залишилась більшою за $D_{\min}$, завдяки виконанню маневру.

Для реалізації стратегій ухилення використано простий підхід: БпЛА піднімається або опускається на 20 м (у межах свого сегменту висоти) або змінює горизонтальний шлях на сусідній паралельний коридор, якщо недоступний. Параметри таких маневрів (амплітуда зміни висоти, кут відхилення) налаштовувалися залежно від ситуації.

У табл. 3 представлено технічні параметри реалізації програмно-апаратної частини, а саме на якому обладнанні проведено роботу моделі.

Таблиця 3 – Технічні параметри реалізації системи

Параметр	Значення
Розмір області польотів	10 км $\times$ 10 км (умовне міське повітряне середовище)
Кількість БпЛА в сценарії	20 (одночасно в повітрі)
Безпечна відстань $D_{\min}$	50 м (мінімальний інтервал для уникнення)
Максимальна швидкість БпЛА	15 м/с
Типи перешкод	Інші БпЛА, будівлі
Симулятор	BlueSky ATM Simul., модуль D2-C2
Частота оновлення даних	1 Гц (прогноз кожну секунду)
Стратегія уникнення	Зміна висоти або маршруту на запасний коридор

Отже, навіть на відносно компактних бортових комп'ютерах (типу Nvidia Jetson) обчислення LSTM-прогнозу займають лічені мілісекунди, що цілком прийнятно для задачі з періодом оновлення 1 с.

У випадку використання менш потужного CPU (наприклад, Raspberry Pi) можливе згортання моделі (після навчання) до спрощеної версії або запуск прогнозу на наземному сервері з передачею результатів на БпЛА завдяки телеметрії.

3. Результати розробки і дослідження удосконаленого методу та розробленої моделі планування маршруту БпЛА

3.1 Удосконалений метод планування маршруту БпЛА

Отримав подальший розвиток метод планування траєкторій БпЛА з проактивним уникненням конфліктів, який, на відміну від відомих методів здійснює прогнозне виявлення небезпеки за допомогою моделі машинного навчання. Метод інтегрує концепцію динамічних коридорів із модулем штучного інтелекту, що дозволяє завчасно та автоматично виявляти потенційні зіткнення БпЛА з перешкодою (іншим БпЛА). Наукова новизна полягає у представленні завдання уникнення зіткнень БпЛА з перешкодами (іншими БпЛА) як задачі оптимізації маршруту польоту з урахуванням прогнозованої небезпеки, що розширює відомі (існуючі) підходи постановки задачі обтікання перешкод у динаміці.

3.2 Модель планування маршруту БпЛА

Запропоновано нейромережеву модель довготривалої короткочасної пам'яті для планування траєкторій польотів БпЛА, яка навчена на даних про їх зближення.

Модель апроксимує функцію, яка відображає складну залежність між часовими рядами параметрів руху двох БпЛА та бінарним результатом – зіткнення чи ні. Вона навчається даним залежностям на прикладах, що є новим підходом у теорії конфліктної динаміки польотів БпЛА.

Наукова новизна запропонованої моделі полягає у тому, що вона вперше застосована для прогнозування конфліктів між БпЛА з урахуванням їх просторово-часової динаміки.

4. Обговорення результатів дослідження удосконаленого методу та розробленої моделі планування маршруту БпЛА

Метод інтегрує концепцію D2-C2 (динамічних коридорів) з модулем ШІ, що дозволяє завчасно та автоматично виявляти потенційні зіткнення БпЛА з перешкодами (іншими БпЛА). Така інтеграція стратегічного повітряного розподілу з тактичним машинним прогнозуванням запропоновано вперше.

Метод забезпечує нову якість управління безпілотним трафіком – кількісну оцінку безпеки кожного плану польоту БпЛА у реальному масштабі часу та адаптацію плану на основі цієї оцінки.

Модель апроксимує функцію, яка відображає складну залежність між часовими рядами параметрів

польоту двох (і більше) БПЛА та бінарним результатом – зіткнення чи ні.

Раніше подібні залежності не мали аналітичного виразу та визначались алгоритмічно (завдяки геометричним розрахункам на основі припущень про траєкторії польотів БПЛА). Розроблена модель навчається таким залежностям на прикладах, що є новим підходом у теорії конфліктної динаміки БПЛА.

5. Висновки

Удосконалено метод планування траєкторій польоту БПЛА з проактивним уникненням конфліктів, який, на відміну від відомих методів (які реагують на поточні порушення дистанції), здійснює прогнозне виявлення небезпеки за допомогою ML-моделі. Отже, новим є підхід до представлення завдання уникнення зіткнень БПЛА з перешкодами (БПЛА) як задачі оптимізації маршруту їх руху з урахуванням прогнозованої небезпеки, що розширює класичні постановки задачі обтікання перешкод у динаміці. Запропонована РНМ LSTM, навчена на даних про зближення БПЛА, сама по собі є новим інструментом оцінки конфліктів у повітрі. Наукова новизна моделі полягає у тому, що вона вперше застосована для прогнозування конфліктів між БПЛА з урахуванням їх просторово-часової динаміки.

Крім того, до нових наукових результатів можливо віднести встановлення оптимальної архітектури LSTM для даної задачі та виявлення основних ознак, які необхідно включати у вектор стану для досягнення високої точності прогнозу.

Модель показує здатність виявляти приховані патерни (наприклад, критичні поєднання відстаней та відносних швидкостей), які складно виразити

явними правилами – це підкреслює перевагу методу машинного навчання над основними підходами та є новим знанням для подальших досліджень.

Розроблена модель та удосконалений метод розширюють сучасні уявлення про побудову систем запобігання зіткненням БПЛА. Вони показують, як можливо об'єднати динамічне управління повітряним рухом (D2-C2) з інтелектуальним прогнозуванням на основі нейронних мереж, досягаючи підвищення рівня безпеки польотів БПЛА у міській забудові без втручання людини. Такий результат може бути основою для створення нових поколінь автономних систем керування БПЛА, здатних адаптивно покращувати навички уникнення небезпеки у складних умовах повітряного руху.

Конфлікт інтересів

Автор декларує, що не має конфлікту інтересів стосовно даного дослідження, в тому числі фінансового, особистісного характеру, авторства чи іншого характеру, що міг би вплинути на дослідження та його результати, представлені в даній статті.

Фінансування

Дослідження проводилося без фінансової підтримки.

Доступність даних

Рукопис не має пов'язаних даних.

Використання засобів штучного інтелекту

Автор підтверджує, що не використовував технології штучного інтелекту при створенні представленої роботи.

СПИСОК ЛІТЕРАТУРИ

1. Delamer, J.-A., Watanabe, Y., and Chanel, C.P.C. (2021), "Safe path planning for UAV urban operation under GNSS signal occlusion risk", *Robotics and Autonomous Systems*, vol. 142, 103800, doi: <https://doi.org/10.1016/j.robot.2021.103800>
2. Lin, X., Wang, C., Wang, K., Li, M., and Yu, X. (2021), "Trajectory planning for unmanned aerial vehicles in complicated urban environments: A control network approach", *Transportation Research Part C: Emerging Technologies*, vol. 128, 103120, doi: <https://doi.org/10.1016/j.trc.2021.103120>
3. Zhang, N., Zhang, M., and Low, K.H. (2021), "3D path planning and real-time collision resolution of multirotor drone operations in complex urban low-altitude airspace", *Transportation Research Part C: Emerging Technologies*, vol. 129, 103123, doi: <https://doi.org/10.1016/j.trc.2021.103123>
4. Maw, A.A., Tyan, M., Nguyen, T.A., and Lee, J.-W. (2021), "iADA*-RL: Anytime Graph-Based Path Planning with Deep Reinforcement Learning for an Autonomous UAV", *Applied Sciences*, vol. 11(9), doi: <https://doi.org/10.3390/app11093948>
5. Rubí, B., Morcego, B., and Pérez, R. (2021), "Quadrotor Path Following and Reactive Obstacle Avoidance with Deep Reinforcement Learning", *Journal of Intelligent, and Robotic Systems*, vol. 103, doi: <https://doi.org/10.1007/s10846-021-01491-2>
6. Delamer, J.-A., Watanabe, Y., and Chanel, C. (2021), "Multiple UAVs trajectory generation and waypoint assignment in urban environment based on DOP maps", *Aerospace Science and Technology*, vol. 110, doi: <https://doi.org/10.1016/j.ast.2021.106507>
7. Doukhi, O., and Lee, D.-J. (2021), "Deep Reinforcement Learning for End-to-End Local Motion Planning of Autonomous Aerial Robots in Unknown Outdoor Environments: Real-Time Flight Experiments", *Sensors*, vol. 21(7), 2534, doi: <https://doi.org/10.3390/s21072534>
8. Tu, G.-T., and Juang, J.-G. (2023), "UAV Path Planning and Obstacle Avoidance Based on Reinforcement Learning in 3D Environments", *Actuators*, vol. 12(2), 57, doi: <https://doi.org/10.3390/act12020057>
9. Tang, J., Liang, Y., and Li, K. (2024), "Dynamic Scene Path Planning of UAVs Based on Deep Reinforcement Learning", *Drones*, vol. 8(2), 60, doi: <https://doi.org/10.3390/drones8020060>
10. Коломійцев О.В., Рудаков І.С., Філіппенков О.В., Павлій В.О., Катунін А.М., Коваль В.В., Згода Г.І., Кібальник В.М., Пальчиков В.В., Белоус М.В., Павлій Л.В. (2025), "Особливості застосування технології машинного навчання для підвищення безпеки польотів групи безпілотних літальних апаратів у міській забудові", *ГРААЛЬ НАУКИ*, № 50, С. 537–550, doi: <https://doi.org/10.36074/grail-of-science.21.03.2025>
11. Коломійцев О.В., Рудаков І.С., Дмитрієв О.М., Комаров В.О., Катунін А.М., Руснак В.М., Шеремет І.А., Галінський Д.О. Аналіз наукових досліджень у галузі безпеки польотів групи безпілотних літальних апаратів у міській забудові. *ГРААЛЬ НАУКИ*, № 51, С. 569–582. DOI <https://doi.org/10.36074/grail-of-science.18.04.2025>

12. Рисований О.М., Коломійцев О.В., Дмитрієв О.М., Голубничий Д.Ю., Невиніцин А.М., Третяк В.Ф., Крук Б.М., Романюк А.О., Собора А.І., Старцев В.В., Калачова В.В., Пустоваров В.В. (2023), “Метод підвищення ефективності управління безпілотним літальним апаратом у залежності від особливостей матриці зв’язків при використанні нелінійного псевдовипадкового генератора”, *II International Scientific and Practical Conference. Science and education in progress*, pp. 344–352, available at: <https://archive.interconf.center/index.php/conference-proceeding/article/view/3889>
13. Puente-Castro, A., Rivero, D., Pedrosa, E., Pereira, A., Lau, N., and Fernandez-Blanco, E. (2024), “Q-Learning based system for Path Planning with Unmanned Aerial Vehicles swarms in obstacle environments”, *Expert Systems with Applications*, vol. 235, 121240, doi: <https://doi.org/10.1016/j.eswa.2023.121240>
14. Fu, X., Zhi, C., and D. Wu (2023), “Obstacle avoidance and collision avoidance of UAV swarm based on improved VFH algorithm and information sharing strategy”, *Computers, and Industrial Engineering*, vol. 186, 109761, doi: <https://doi.org/10.1016/j.cie.2023.109761>
15. Xu, L., Cao, X., Du, W., and Li, Y. (2023), “Cooperative path planning optimization for multiple UAVs with communication constraints”, *Knowledge-Based Systems*, vol. 260, 110164, doi: <https://doi.org/10.1016/j.knosys.2022.110164>
16. Niu, Y., Yan, X., Wang, Y., and Niu, Y. (2024) “3D real-time dynamic path planning for UAV based on improved interfered fluid dynamical system and artificial neural network”, *Advanced Engineering Informatics*, vol. 5(9), 102306, doi: <https://doi.org/10.1016/j.aei.2023.102306>
17. Feng S, Liu F, Zhou S, Wong X, and Mai J. (2025), “Dynamic obstacle avoidance strategy for autonomous vehicles based on LSTM model trajectory prediction under urban roads”, *Proceedings of the Institution of Mechanical Engineers, Part D*, doi: <https://doi.org/10.1177/09544070241299283>
18. Fedorchenko, V., Yeroshenko, O., Shmatko, O., Kolomiitsev, O., and Omarov, M. (2024), “Password hashing methods and algorithms on the .net platform”, *Advanced Information Systems*, vol. 8(4), pp. 82–92, doi: <https://doi.org/10.20998/2522-9052.2024.4.11>
19. Shmatko, O., Kolomiitsev, O., Rekova, N., Kuchuk, N., and Matvieiev, O. (2023), “Designing and evaluating dl-model for vulnerability detection in smart contracts”, *Advanced Information Systems*, vol. 7(4), pp. 41–51, doi: <https://doi.org/10.20998/2522-9052.2023.4.05>
20. Gordo, V., Perez-Castan, J.A., Perez Sanz, L., Serrano-Mira, L., and Xu, Y. (2024), “Feasibility of Conflict Prediction of Drone Trajectories by Means of Machine Learning Techniques”, *Aerospace*, vol. 11, doi: <https://doi.org/10.3390/aerospace11121044>

Received (Надійшла) 16.07.2026

Accepted for publication (Прийнята до друку) 02.09.2026

Publication date (Дата публікації) 25.09.2026

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Рудаков Ігор Сергійович – аспірант кафедри комп’ютерної інженерії та програмування, Національний технічний університет «Харківський політехнічний інститут», Харків, Україна;

Rudakov Ihor – Postgraduate student, Department of Computer Engineering and Programming, National Technical University, “Kharkiv Polytechnic Institute”, Kharkiv, Ukraine;

e-mail: ihor.rudakov@cit.khpi.edu.ua; ORCID Author ID: <https://orcid.org/0009-0008-9525-5669>.

Коломійцев Олексій Володимирович – доктор технічних наук, професор, професор кафедри комп’ютерної інженерії та програмування Національного технічного університету «Харківський політехнічний інститут», Харків, Україна;

Oleksii Kolomiitsev – Doctor of Technical Sciences, Professor, Professor of Computer Engineering and Programming Department, National Technical University «Kharkiv Polytechnic Institute», Kharkiv, Ukraine;

e-mail: alexus_k@ukr.net; ORCID Author ID: <https://orcid.org/0000-0001-8228-8404>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57211278112>.

A method and model for planning flight paths for a fleet of unmanned aerial vehicles to avoid obstacles in urban areas using machine learning technologies

Ihor Rudakov, Oleksii Kolomiitsev

Abstract. Background. The increasing number of unmanned aerial vehicles (UAVs) each year leads to a higher flight density and a greater likelihood of collisions, which necessitates ensuring a high level of flight safety. The application of the D2-C2 dynamic density corridor concept makes it possible to reduce the frequency of conflicts, but does not completely eliminate the risk of unforeseen UAV encounters with obstacles (static and dynamic) in the presence of natural phenomena and dynamic factors. Thus, improving the method of UAV trajectory planning through the implementation of tools for predicting potential conflicts based on machine learning technologies is a pressing scientific challenge. **The subject of this article** is the integration of the D2-C2 concept with an artificial intelligence module. **The aim of the study** is to improve the flight safety of UAV fleets in urban environments. **The objective** of the study is to develop a model and refine the method for planning UAV trajectories by implementing tools for predicting potential conflicts. **Methods** of mathematical statistics and modeling were used in the course of the research. The **results** of the study presented in this article include a proposed LSTM neural network model trained on data regarding UAV approach scenarios. The method for planning UAV trajectories with proactive conflict avoidance has been improved; it performs predictive hazard detection using a machine learning model. **Conclusions.** The proposed model approximates a function that reflects the complex relationship between the time series of motion parameters of two (or more) UAVs and a binary outcome—collision or no collision. The improved method integrates the D2-C2 concept with an artificial intelligence module, enabling the early and automatic detection of potential collisions between UAVs and obstacles. **The results obtained** can be used in the future to develop new generations of autonomous UAV control systems capable of adaptively and through learning improving their ability to avoid hazards in complex air traffic conditions.

Keywords: unmanned aerial vehicle; flight path; machine learning; artificial intelligence; neural network; safety; collision; method; mathematical model; algorithm.

О. Ю. Слободяник

Національний технічний університет “Харківський політехнічний інститут”, Харків, Україна

МЕТОД ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ АВТОМАТИЗОВАНОГО АНАЛІЗУ PQ-ДАНИХ ПОКАЗНИКІВ ЯКОСТІ ЕЛЕКТРОЕНЕРГІЇ

Анотація. Актуальність. Стрімкий розвиток цифровізації енергетики та зміни характеру навантажень у сучасних енергосистемах призводить до наступних наслідків: зростання обсягів даних, збільшення кількості аномалій, поширення нелінійних навантажень та економічним збиткам. Відомі методи та підходи при аналізі за жорсткими порогоми не контролюють складні, комбіновані та динамічні аномалії PQ-даних (Power Quality data) показників якості електроенергії. Таким чином, розробка методу підвищення ефективності автоматизованого аналізу PQ-даних показників якості електроенергії є актуальною науковою задачею. **Предметом дослідження** в статті є математичні моделі, методи та процедури інтелектуального аналізу великих масивів даних показників якості електроенергії, спрямовані на виявлення відхилень та типових режимів. **Метою дослідження** є підвищення ефективності автоматизованого аналізу PQ-даних показників якості електроенергії. **Завданням дослідження** є розробка методу підвищення ефективності автоматизованого аналізу PQ-даних показників якості електроенергії. В ході досліджень використано **методи** машинного навчання (для класифікації PQ-подій за ознаками), методи інтелектуального аналізу (для PQ Big Data), графові методи (для відображення топології електромережі та просторових взаємозв'язків), а також методи щільності, локальної близькості та кластеризації. **Результатами дослідження** у статті є розроблений метод підвищення ефективності автоматизованого аналізу PQ-даних показників якості електроенергії на основі інтеграції удосконаленого методу глибинного автоенкодера з кластеризацією ознак для аналізу великих масивів даних показників якості електроенергії та розробленого методу глибокого виявлення аномалій у PQ-даних, зумовлених геофізичними факторами. **Висновки.** Запропоновано розроблений метод підвищення ефективності автоматизованого аналізу PQ-даних показників якості електроенергії. Представлено два взаємопов'язаних методи (автоенкодер + аномальний моніторинг), які разом підвищують ефективність автоматизованого аналізу PQ-даних показників якості електроенергії. Розроблено узагальнену структурну схему процесу автоматизованого аналізу вимірювальних PQ-даних показників якості електроенергії. Отримані **результати** можуть бути основою для розширення переліку об'єктів спостереження, підвищення стійкості до зміни конфігурацій мережі та сезонно-добової нестаціонарності, а також розвитку процедур адаптації моделей та інтеграції аналітики з подіями керування для поглиблення причинно-наслідкової інтерпретації в експлуатаційних умовах.

Ключові слова: PQ-дані; показник якості електроенергії; глибинний автоенкодер з кластеризацією ознак; глибоке виявлення аномалій; метод; математична модель; машинне навчання; інтелектуальний аналіз.

1. Вступ

1.1 Актуальність дослідження

Стрімкий розвиток цифровізації енергетики та зміни характеру навантажень у сучасних енергосистемах призводить до наступних наслідків:

- зростання обсягів даних – впровадження смарт-лічильників та систем Phasor Measurement Unit (PMU – це пристрої синхронізованих векторних вимірювань, що є основними елементами технології Wide Area Measurement Systems (WAMS) для цифрових енергосистем (Smart Grid)) генерує терабайти інформації, яку неможливо обробити вручну;

- збільшення кількості аномалій – масове підключення сонячних та вітрових електростанцій створює нестабільність частоти та напруги електроживлення;

- поширення нелінійних навантажень – побутова електроніка, LED-освітлення та зарядні станції для електромобілів спотворюють «чисту» синусоїду струму;

- економічні збитки – низька якість електроенергії призводить до перегріву трансформаторів, збоїв автоматики тощо та простою виробництва;

- недоліки відомих методів та підходів – аналіз за жорсткими порогоми пропускає складні,

комбіновані та динамічні аномалії даних показників якості електроенергії.

Але відома статистика має низку недоліків:

- прокляття розмірності – PQ-дані (Power Quality data) містять десятки взаємопов'язаних параметрів (статистичні матриці коваріації стають занадто громіздкими та втрачають точність);

- припущення про нормальність – статистика часто вимагає, щоб дані мали нормальний (Гаусовий) розподіл (реальні спотворення напруги мають складні та мультимодальні розподіли);

- жорсткі статичні пороги – класичний аналіз фіксує лише вихід за межі (він пропускає приховані аномалії, які знаходяться у межах норми);

- чутливість до часових зсувів – статистичні методи не якісно вловлюють складні динамічні зміни форми сигналу у часі.

На даний час відома статистика опирається на фіксовані математичні моделі, а використання елементів штучного інтелекту (ШІ), зокрема глибинного навчання без вчителя, дозволяє адаптуватися до реального стану енергосистеми.

Сучасна статистика не обмежується лише фіксованими моделями. Вона активно використовує динамічні, адаптивні та непараметричні методи, баєсівський підхід, а також алгоритми машинного навчання, які самостійно оновлюються й змінюються під час обробки нових даних у реальному часі.

До основних сучасних підходів можливо віднести наступні:

- Баєсівська статистика – яка постійно оновлює ймовірності при появі нових фактів;
- машинне навчання – створює моделі, які навчаються на даних без жорстких попередніх формул;
- непараметричні методи – не вимагають заздалегідь заданих законів розподілу для даних;
- адаптивні системи – змінюють свої параметри автоматично зі зміною умов.

До математичних переваг ІІІ (автоенкодерів) можливо віднести наступні:

- нелінійне стиснення – автоенкодери використовують нелінійні функції активації, що дозволяє знаходити складні залежності між гармоніками, які ігнорує класичний метод головних компонент (РСА);
- автоматичний пошук ознак – ІІІ самостійно виділяє основні паттерни з сирого сигналу (відома статистика вимагає від інженера вручну задавати математичні критерії);
- динамічний простір станів – кластеризація у латентному (стиснутому) просторі автоенкодера групує режими роботи не за абсолютними значеннями, а за внутрішньою структурою сигналу;
- оцінка за помилкою реконструкції – ІІІ навчається відновлювати нормальний сигнал (за умови, якщо на вхід подається аномалія, помилка реконструкції (Mean Squared Error, MSE) різко зростає, що математично сигналізує про дефект без налаштування порогів.

Таким чином, розробка методу підвищення ефективності автоматизованого аналізу PQ-даних показників якості електроенергії є актуальною науковою задачею.

1.2 Аналіз літературних джерел і визначення проблеми

У [1] для покращення виявлення аномалій у розподілених енергосистемах запропоновано автоенкодерно-федеративний підхід. Акцентовано увагу на проблематику відомих методів у масштабуванні та конфіденційності. Розкрито основні переваги глибокого представлення даних.

У [2] розкрито особливості гібридної системи, яка застосовується для ідентифікації порушень якості електроенергії. При цьому, безперервні вейвлет-перетворення та глибокі нейронні мережі ResNet/VGG (Residual Network, ResNet, залишкова мережа та Visual Geometry Group VGG, група візуальної геометрії) комбінуються з аналізом важливості ознак та SVM-класифікатором (Support Vector Machine, метод опорних векторів).

У [3] підкреслено необхідність у формуванні стійких латентних представлень часових фрагментів PQ-даних (сигналів) та виявлення аномальних відхилень без повної розмітки даних.

У [4] наведено вимоги галузевих стандартів (IEC 61000-4-30), у тому числі [5].

Однак, запропоновані рішення орієнтуються на класифікацію за заздалегідь визначеними класами

порушень або потребують ручного узгодження параметрів. В умовах польових вимірювань доводиться працювати з різнорідними сигналами, відсутністю однозначних міток та великою часткою пропущених або зіпсованих даних.

Тому, виникає необхідність у розробці методів, які спроможні забезпечити адаптивність та стійкість до шуму, а також здатність до самоорганізації та інтерпретації результатів [6–8].

При цьому, можливо використовувати інтеграцію двох взаємопов'язаних аналітичних контурів – глибокого автоенкодера з кластеризацією ознак та механізму глибокого виявлення аномалій у спільному латентному просторі представлення PQ-даних [9, 10].

1.3 Мета і завдання дослідження

Метою дослідження є підвищення ефективності автоматизованого аналізу PQ-даних показників якості електроенергії.

Для досягнення мети було поставлено наступне завдання: розробка методу підвищення ефективності автоматизованого аналізу PQ-даних показників якості електроенергії.

2. Матеріали і методи дослідження

Розроблений метод підвищення ефективності автоматизованого аналізу PQ-даних показників якості електроенергії оснований на інтеграції [11]:

- удосконаленого методу глибокого автоенкодера з кластеризацією ознак для аналізу великих масивів даних показників якості електроенергії [12, 13];

- розробленого методу глибокого виявлення аномалій у PQ-даних, зумовлених геофізичними факторами [14, 15].

При цьому, обидва методи працюють у складі одного аналітичного конвеєра, який пов'язує моделювання типової «норми» з моніторингом відхилень [16]. При аналізі даних показників якості електроенергії в умовах цифровізації енергосистем використано методи аналізу. При розробці методу підвищення ефективності автоматизованого аналізу вимірювальних даних електроенергії, удосконаленні методу глибокого автоенкодера з кластеризацією ознак та розробці методу глибокого виявлення аномалій у PQ-даних використано методи машинного навчання (для класифікації PQ-подій за ознаками), методи інтелектуального аналізу (для PQ Big Data), графові методи (для відображення топології електромережі та просторових взаємозв'язків), а також методи щільності, локальної близькості та кластеризації.

3. Результати розробки і дослідження розробленого методу підвищення ефективності автоматизованого аналізу PQ-даних показників якості електроенергії

Сутність розробленого методу підвищення ефективності автоматизованого аналізу PQ-даних

показників якості електроенергії полягає у тому, що вихідні дані – це великі масиви багатоканальних часових рядів показників якості електроенергії (напруга, струм, гармоніки, індекси пульсацій тощо) з відповідними часовими мітками та метаданими (стан пристроїв, умови роботи та оцінка якості прийому). Дані, після попередньої обробки утворюють набір часових вікон спостереження. При цьому, спочатку формуються компактні латентні представлення для типових режимів роботи системи за допомогою автоенкодера, а потім – у тому ж просторі, здійснюється детектування аномалій, які не відповідають нормальному режиму.

На рис. 1 наведено узагальнену структурну схему розробленого методу від попередньої обробки даних до виділення типових режимів та подієвого детектування аномалій.

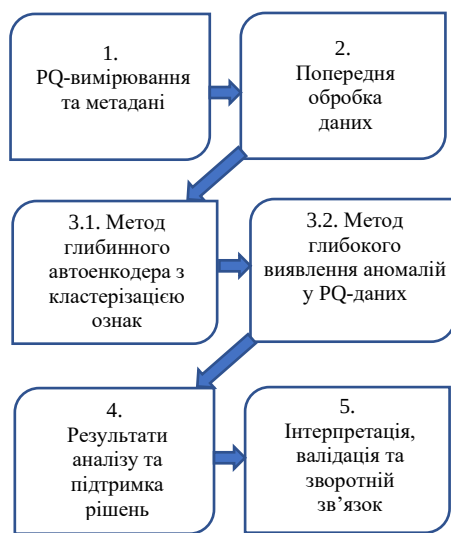


Рис. 1. Узагальнена структурна схема розробленого методу

За рис. 1 зроблено припущення. Отже, X – матриця вхідних даних, яка подана у вигляді множини N часових вікон (фрагментів) довжини T (кількість відліків вікна) та з M каналами (різні показники) наступним чином:

$$X = \{x_i\}_{i=1}^N, x_i \in R^{T \times M}, \quad (1)$$

де x_i – вектор-матриця вимірювань i -го вікна, яка містить значення усіх M показників у T моментах часу.

Тоді, наприклад, за кроком 10 хв протягом доби матриця x_i може мати розмірність $T=144$ та $M=5$ каналів (5 показників якості). Набір X формується після виділення «спокійних» інтервалів (без техногенних аномалій та підвищеної геомагнітної активності) у вихідних PQ-даних.

Відповідно, для коректного навчання автоенкодера важливо вирівняти амплітуди сигналів, щоб модель аналізувала форми, а не абсолютні рівні. Тому, кожен канал даних стандартизується за фіксованим часом доби наступним чином:

$$\tilde{x}_i(t) = \frac{x_i(t) - \mu_t}{\sigma_t}, \quad (2)$$

де μ_t та σ_t – середнє та стандартне відхилення усіх $x_j(t)$ у момент доби t за усіма вікнами j .

Даний підхід знижує вплив дисбалансу та дрейфу потужностей на результат та забезпечує узгодженість масштабів під час кластеризації.

Вихідні дані часто мають фрагментації (добові цикли, зміни навантаження тощо). Тоді, набір добових фрагментів можливо записати таким чином:

$$\{x_i\}_{i=1}^N, x_i = [x_i(1), x_i(2), \dots, x_i(T)], \quad (3)$$

де $x_i(t)$ – значення показника у часовий відлік t (при множинних каналах – відповідний вектор).

Такий крок виконується перед автоенкодером та забезпечує порівнянність форм. Модель навчається на одноманітних послідовностях (наприклад, кожне x_i – це добове коливання гармоніки напруги).

Сутність основних аналітичних блоків методу підвищення ефективності аналізу даних показників якості електроенергії полягає у наступному (рис. 1).

Першим аналітичним блоком розробленого методу є удосконалений метод глибокого автоенкодера з кластеризацією ознак (рис. 1, блок 3.1). Автоенкодер складається з енкодера F_θ та декодера D_η , що послідовно зменшують розмірність входу до компактного латентного вектора, а потім відновлюють сигнал. Процес навчання формалізується наступним чином:

$$z_i = F_\theta(\tilde{x}_i), z_i \in R^d, d \ll T \cdot M, \quad (4)$$

де $\tilde{x}_i$ – нормалізоване вхідне вікно; z_i – його латентний код розмірності d ; θ – параметри енкодера.

Декодер відновлює сигнал наступним чином:

$$\hat{x}_i = D_\eta(z_i), \quad (5)$$

де η – параметри декодера.

Під час навчання зменшується середньоквадратична помилка реконструкції:

$$L_{rec}(\theta, \eta) = \frac{1}{N} \sum_{i=1}^N \|\tilde{x}_i - \hat{x}_i\|_F^2. \quad (6)$$

Одночасно з навчанням автоенкодера задається кількість кластерів K . Після отримання латентних кодів $\{z_i\}$ виконується кластеризація (наприклад, методом k -середніх) у латентному просторі. Кожному з K кластерів відповідає типова група вікон. Для кожного кластера обчислюється центр m_k у просторі z та здійснюється його декодування:

$$x_{repr,k} = D_\eta(m_k), \quad (7)$$

що дає репрезентативний профіль $x_{repr,k}$ – типовий добовий сигнал для даного режиму.

Таке реконструювання профілів та побудова календарних карт режимів (з тривалістю та частотою появи кластерів у часі) забезпечують інтерпретацію результатів (типових шаблонів роботи).

Таким чином, удосконалений метод глибинного автоенкодера з кластеризацією ознак для аналізу великих масивів даних показників якості електроенергії вирішує задачу виділення типових режимів. Він формує компактні ознаки, групує їх у кластери та відновлює репрезентативні шаблони.

Другий аналітичний контур – розроблений метод глибокого виявлення аномалій у PQ-даних, який призначений для моніторингу відхилень від типового режиму (рис. 1, блок 3.2). Він використовує той же простір латентних ознак z_i , але для виявлення часових вікон, що не узгоджуються з «нормою».

Спочатку задається навчання моделі (наприклад, того самого автоенкодера або прогностичної неймережі) лише на нормальних даних (таймові проміжки без буревіїв та техногенних перебоїв). Після цього, для кожного вікна x_i обчислюється аномальний скор s_i , наприклад як відстань MSE між x_i та його реконструкцією $\hat{x}_i$ (або приріст помилки рекурсії):

$$s_i = \| \tilde{x}_i - \hat{x}_i \|_F^2. \quad (8)$$

Далі, обирається поріг τ детектування аномалій, щоб контролювати рівень помилкових спрацювань.

Наприклад, поріг може бути визначений як квантиль α розподілу скорів на тренувальній вибірці (невеликий α дає малу частку хибних спрацювань):

$$\tau = Q_\alpha(\{s_j\}_{j \in \text{train}}). \quad (9)$$

Вікно i позначається аномальним, якщо $s_i > \tau$. Окремі ізольовані спрацювання групуються у події за часовим критерієм (наприклад, суміжні або близькі за часом вікна з $s > \tau$ об'єднуються в один інцидент з піковим скором $s_{\max}$ та тривалістю).

Таким чином, розроблений метод глибокого виявлення аномалій у PQ-даних, зумовлених геофізичними факторами, вирішує задачу виявлення де тимчасові PQ-серії виходять за межі нормального контексту. Він формує безперервний показник

аномальності та агрегацію спрацювань у події з атрибутами (піковий скор, тривалість та зв'язок із GIC). Визначення порога завдяки ROC/PR-аналізу із заданим рівнем хибних спрацювань та зведення метрик Precision/Recall/F1 для оцінки якості детектування дають чисельні критерії.

Блок інтерпретації, валідації та зворотного зв'язку здійснює наступне: перевіряє достовірність PQ-даних результатів, надає пояснення, збирає операторські підтвердження та повертає сигнали для корекції порогів, правил агрегації та циклу супроводження моделей. Така замкнена архітектура відповідає загальній ідеї статті: «типові режими → аномалії → інтерпретація → рішення → контроль якості та адаптація». Блок результатів аналізу та підтримки рішень призначений для перетворення виходу ядра глибинного аналізу без вчителя на форму, яка придатна для оперативного використання в цифровізованому контурі моніторингу PQ-даних. Блок виконує не окрему класифікацію кожного вікна, а агрегування суміжних (близьких у часі) аномальних вікон у завершений інцидент із часовими межами, тривалістю, піковою інтенсивністю та контекстом вузла.

Контур зворотного зв'язку є формалізованим. Отже, оператор (інженер-експерт) підтверджує подію, позначає її як хибне спрацювання (false positive), уточнює клас порушення, додає причину, фіксує вплив на обладнання, а також ініціює огляд порогів або запуск контрольованого перенавчання.

За NIST AI RMF (Artificial Intelligence Risk Management Framework – системи керування ризиками ШІ) – це відповідає вимогам щодо ролей і відповідальностей у конфігураціях взаємодії людини та ШІ (human-AI configurations), документування рівня людського нагляду, безперервного моніторингу, журналювання подій та процедур корекції (remediation), якщо система виходить за межі допустимого ризику.

За рис. 1 такий сигнал зворотного зв'язку надходить не тільки на блок результатів аналізу та підтримки рішень, а й назад до ядра аналізу та правил попередньої обробки (блоки 3.1 та 3.2, рис. 1), якщо виявлено систематичні артефакти або зміну режимів функціонування електроенергетичної системи.

У табл. 1 представлено приклад порівняння вихідних артефактів та повідомлень зворотного зв'язку для узагальненої структурної схеми розробленого методу.

Таблиця 1 – Приклад порівняння вихідних артефактів та повідомлень зворотного зв'язку

Сутність	Тип / формат	Частота	Призначення
Подієве повідомлення / Повідомлення про аномалію	JSON (формат обміну даними) / подія OPC UA (Open Platform Communications Unified Architecture) / повідомлення IEC 61850	Подієво	Оперативне оповіщення SCADA/АСУ про інцидент із severity (рівнем критичності), часовими межами, qc_flag (прапорцем контролю якості), measurement_class (класом вимірювання) та посиланням на waveform (хвильову форму)
Черга ранжованих інцидентів	REST/API (програмний інтерфейс) / список інформаційної панелі (dashboard list)	Квазі-реальний час	Пріоритизація роботи оператора за критичністю об'єкта, інтенсивністю події та довірою моделі
KPI-знімок моніторингу	Віджет інформаційної панелі (dashboard widget), витяг CSV (comma-separated values) / Parquet	Годинно / добово	Контроль частоти спрацювань, FAR (частоти хибних спрацювань), затримки виявлення, дрейфу даних і скорів

Закінчення табл. 1

Сутність	Тип / формат	Частота	Призначення
Періодичний RQ-звіт	PDF / вебзвіт (Web report) / експорт CSV	Добово / тижнево / місячно	Аналітика інцидентів, тяжкість sags/swells/transients (провалів/перенапруг/перехідних процесів), location/impact (місця виникнення та впливу), звітність і аудит
Пакет кейсу для експерта	Архів запису, waveform (хвильова форма), metadata (метадані), explanation payload (пакет пояснення)	За запитом	Поглиблена верифікація, розбір причин, навчання персоналу, ескалація
Підтвердження оператора	ack/close/comment message (повідомлення підтвердження/закриття/коментаря)	Подієво	Фіксація результату human review (експертної перевірки людиною) та завершення життєвого циклу інциденту
Повід. false positive	feedback message (повідомлення зворотного зв'язку)	Подієво	Поповнення контрольованої вибірки, перегляд класу події, корекція шаблонів
Переналаштування	config-change / threshold-review (перегляд конфігурації та порогів)	За KPI-тригером	Перегляд порогів, правил агрегації, ваг пріоритету
Trigger на перенавчання	retrain request / MLOps job (запит на перенавчання / задача керування життєвим циклом ML-моделі)	Епізодично	Реакція на drift (дрейф даних), зміну режимів, накопичення підтверджених нових кейсів

4. Обговорення результатів дослідження удосконаленого методу та розробленої моделі планування маршруту БПЛА

Таким чином, для кількісної оцінки ефективності обох блоків запропоновано використовувати узгоджені метрики, зокрема: для блоку кластеризації – середній силует (silhouette) та стійкість кластерів при різних ініціалізаціях/підвибірках; для блоку детектування – метрики якості детекції подій на рівні епізодів, частка False Alarm Rate (FAR), а також медіанна затримка (lead time) виявлення; значущість зв'язку подій з геомагнітними індексами – оцінюється за p-value блочними перестановками.

5. Висновки

Таким чином, розроблено метод підвищення ефективності автоматизованого аналізу RQ-даних показників якості електроенергії на основі інтеграції удосконаленого методу глибинного автоенкодера з кластеризацією ознак для аналізу великих масивів даних показників якості електроенергії та розробленого методу глибокого виявлення аномалій у RQ-даних, зумовлених геофізичними факторами.

Представлено два взаємопов'язаних методи (автоенкодер + аномальний моніторинг), які разом підвищують ефективність автоматизованого аналізу даних показників якості електроенергії.

Розроблено узагальнену структурну схему процесу автоматизованого аналізу вимірювальних даних електроенергії, яка поєднує:

– автоматичне нелінійне узагальнення даних у компактний простір та кластеризацію режимів;

– безперервну оцінку аномальності з керованим налаштуванням порогів;
– агрегування спрацювань у події з інженерно змістовними характеристиками;
– статистично обґрунтовану верифікацію виявлених подій за геомагнітними індикаторами.

Отриманий науковий результат може бути основою для розширення переліку об'єктів спостереження, підвищення стійкості до зміни конфігурацій мережі та сезонно-добової нестационарності, а також розвитку процедур адаптації моделей та інтеграції аналітики з подіями керування для поглиблення причинно-наслідкової інтерпретації в експлуатаційних умовах.

Конфлікт інтересів

Автор декларує, що не має конфлікту інтересів стосовно даного дослідження, в тому числі фінансового, особистісного характеру, авторства чи іншого характеру, що міг би вплинути на дослідження та його результати, представлені в даній статті.

Фінансування

Дослідження проводилося без фінансової підтримки.

Доступність даних

Рукопис не має пов'язаних даних.

Використання засобів штучного інтелекту

Автор підтверджує, що не використовував технології штучного інтелекту при створенні представленої роботи.

СПИСОК ЛІТЕРАТУРИ

- Kea, K., Han, Y., and Kim, T-K (2023), "Enhancing anomaly detection in distributed power systems using autoencoder-based federated learning", *PLoS ONE*, vol. 18(8), doi: <https://doi.org/10.1371/journal.pone.0290337>
- Peruman, P.M., and Ayyar, K. (2025), "Power quality disturbance identification using hybrid deep learning in renewable energy systems", *Scientific Reports*, vol. 15, 44592, doi: <https://doi.org/10.1038/s41598-025-28291-0>
- Коломійцев, О.В., & Слободяник О.Ю. (2026), "Аналіз даних показників якості електроенергії в умовах цифровізації енергосистем", *ГРААЛЬ НАУКИ*, № 68. С. 1046-1062, doi: <https://doi.org/10.36074/grail-of-science>
- (2015), *IEC 61000-4-30:2015. Electromagnetic compatibility (EMC), Testing and measurement techniques – Power quality measurement methods*, IEC, 2015, available at: <https://webstore.iec.ch/en/publication/21844>

5. (2022), *EN 50160:2022. Voltage characteristics of electricity supplied by public electricity networks*, Brussels, CENELEC, 2022 Dugan R. C., McGranaghan M. F., Santoso S., Beaty H. W. *Electrical Power Systems Quality*. 3rd ed. New York: McGraw-Hill, 2012, available at: <https://standards.iteh.ai/catalog/standards>
6. Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., and Bouchachia, A. (2014), "A Survey on Concept Drift Adaptation", *ACM Computing Surveys*, vol. 46, no. 4, art. 44, doi: <https://doi.org/10.1145/2523813>
7. Pimentel, M. A. F., Clifton, D. A., Clifton, L., and Tarassenko, L. (2014), "A review of novelty detection", *Signal Processing*, vol. 99, pp. 215–249, doi: <https://doi.org/10.1016/j.sigpro.2013.12.026>
8. Jain, A. K. (2010), "Data clustering: 50 years beyond K-means", *Pattern Recognition Letters*, vol. 31, no. 8, pp. 651–666, doi: <https://doi.org/10.1016/j.patrec.2010.08.023>
9. Machlev, R., Heistrene, L., Perl, M., Levy, K. Y., Belikov, J., Mannor, S., and Levron, Y. (2022), "Explainable Artificial Intelligence (XAI) techniques for energy and power systems: Review, challenges and opportunities", *Energy and AI*, vol. 9, art. 100169, doi: <https://doi.org/10.1016/j.egyai.2022.100169>
10. Ribeiro, M. T., Singh, S., and Guestrin, C. (2016), "Why Should I Trust You?": Explaining the Predictions of Any Classifier", *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1135–1144, doi: <https://doi.org/10.1145/2939672.2939778>
11. Коломійцев, О.В., & Слободяник, О.Ю. (2026), "Метод підвищення ефективності автоматизованого аналізу вимірювальних даних електроенергії на основі глибинного навчання без вчителя", *Moderní aspekty vědy: LXIX. Díl mezinárodní kolektivní monografie* str. 672, doi: <https://doi.org/10.52058/69-2026>
12. Коломійцев, О.В., & Слободяник, О.Ю. (2026), "Удосконалений метод глибинного навчання для автоматизованого виявлення типових шаблонів у PQ-даних показників якості електроенергії", *ГРААЛЬ НАУКИ*, № 69, С. 1163–11816 doi: <https://doi.org/10.36074/grail-of-science>
13. Коломійцев, О.В., & Слободяник, О.Ю. (2026), "Метод глибокого виявлення аномалій у вимірювальних PQ-даних (показників якості електроенергії), зумовлених геофізичними факторами", *ГРААЛЬ НАУКИ*, № 70, С. 831–850, doi: <https://doi.org/10.36074/grail-of-science.31.05.2026>
14. Mahela, O.P., and Shaik, A.G. (2015), "A critical review of detection and classification of power quality events", *Renewable and Sustainable Energy Reviews*, vol. 41, pp. 549–556, doi: <https://doi.org/10.1016/j.rser.2014.08.070>
15. Fedorchenko, V., Yeroshenko, O., Shmatko, O., Kolomiitsev, O., and Omarov, M. (2024), "Password Hashing Methods and Algorithms on the .NET Platform", *Advanced Information Systems*, vol. 8, no. 4, pp. 82–92, doi: <https://doi.org/10.20998/2522-9052.2024.4.11>
16. Kozina, O. A., Panchenko, V. I., Kolomiitsev, O. V., Usik, V. V., Stratiienko, N. K., Safoshkina, L. V., & Kucherenko, Y. F. (2024), "Data consistency protocol for multicloud systems", *International Journal of Cloud Computing*, vol. 13(1), pp. 42–61, doi: <https://doi.org/10.1504/IJCC.2024.136284>

Received (Надійшла) 07.06.2026

Accepted for publication (Прийнята до друку) 26.08.2026

Publication date (Дата публікації) 25.09.2026

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Слободяник Олег Юрійович – аспірант кафедри комп'ютерної інженерії та програмування, Національний технічний університет «Харківський політехнічний інститут», Харків, Україна;

Oleg Slobodanyk – Postgraduate student, Department of Computer Engineering and Programming, National Technical University, "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;

e-mail: Oleh.Slobodanyk@cs.khpi.edu.ua; ORCID Author ID: <https://orcid.org/0009-0003-2886-7116>.

A method for improving the efficiency of automated analysis of pq data on electric power quality indicators

Oleg Slobodanyk

Abstract. Background. The rapid development of digitalization in the energy sector and changes in the nature of loads in modern power systems lead to the following consequences: an increase in data volumes, a rise in the number of anomalies, the proliferation of nonlinear loads, and economic losses. Existing methods and approaches based on hard thresholds fail to detect complex, combined, and dynamic anomalies in PQ data (power quality indicators). Thus, developing a method to improve the efficiency of automated analysis of PQ data is a pressing scientific challenge. The subject of this study is mathematical models, methods, and procedures for the intelligent analysis of large datasets of power quality indicators, aimed at identifying deviations and typical operating modes. The purpose of this study is to improve the efficiency of automated analysis of PQ data on power quality indicators. The objective of this study is to develop a method for improving the efficiency of automated analysis of PQ data on power quality indicators. The research utilized machine learning methods (for classifying PQ events based on characteristics), intelligent analysis methods (for PQ Big Data), graph-based methods (for representing the topology of the power grid and spatial relationships), as well as density, local proximity, and clustering methods. The results of the research presented in this article include a method for improving the efficiency of automated analysis of PQ data on power quality indicators, based on the integration of an improved deep autoencoder method with feature clustering for the analysis of large datasets of power quality indicators, and a newly developed method for deep anomaly detection in PQ data caused by geophysical factors. **Conclusions.** A method for improving the efficiency of automated analysis of PQ data on power quality indicators is proposed. Two interrelated methods (autoencoder + anomaly monitoring) are presented, which together improve the efficiency of automated analysis of PQ data on power quality indicators. A generalized block diagram of the process of automated analysis of measured PQ data on power quality indicators has been developed. The results obtained can serve as a basis for expanding the list of monitored objects, improving resilience to changes in network configurations and seasonal and diurnal variability, as well as for developing procedures for model adaptation and integrating analytics with control events to deepen the cause-and-effect interpretation under operational conditions.

Keywords: PQ data; power quality index; deep autoencoder with feature clustering; deep anomaly detection; method; mathematical model; machine learning; intelligent analysis.

С. В. Сукач, Є. Є. Лашко, О. О. Ченчева, Д. В. Резнік, І. С. Петренко

Кременчуцький національний університет імені Михайла Остроградського, Кременчук, Україна

СИНТЕЗ ТРИФАКТОРНОЇ РЕГРЕСІЙНОЇ МОДЕЛІ ПИЛОПОГЛИНАННЯ У ПОВІТРЯНОМУ СЕРЕДОВИЩІ ВИРОБНИЧИХ ПРИМІЩЕНЬ

Анотація. Актуальність. Забезпечення нормативної якості повітря у виробничих приміщеннях є важливою умовою промислової, техногенної безпеки, безпеки праці, оскільки дрібнодисперсний пил здатний тривалий час перебувати у зваженому стані, ресуспендуватися з поверхонь і формувати підвищене інгаляційне навантаження на працівників. **Предмет дослідження:** процеси пилопоглинання, седиментації та ресуспензії дрібнодисперсних частинок у внутрішньому повітряному середовищі за змінних параметрів мікроклімату. **Мета дослідження:** розроблення науково обґрунтованої трифакторної регресійної моделі коефіцієнта пилопоглинання з урахуванням температури, відносної вологості та масової концентрації пилу для предиктивного керування якістю повітря у приміщеннях. **Завдання дослідження.** Систематизувати фізичні механізми осадження та вторинного зависання дрібнодисперсного пилу; встановити аналітичну залежність концентрації пилу від відносної вологості; синтезувати та статистично перевірити регресійну модель пилопоглинання; визначити області стабілізації та сухої деструкції пилової хмари й обґрунтувати можливість практичного використання моделі. **Методи дослідження.** Теоретичний аналіз, узагальнення експериментальних даних, нелінійний метод найменших квадратів, трифакторне рототабельне планування другого порядку, регресійний і дисперсійний аналіз, перевірка коефіцієнтів за критерієм Стюдента, перевірка адекватності за критерієм Фішера, аналіз поверхонь відгуку й ізоліній. **Результати дослідження.** У роботі узагальнено сучасні уявлення про вплив вологості, температури, адгезії частинок, гіроскопічного росту та механічної активності людини на формування аерозольного навантаження. На основі експериментального масиву побудовано експоненційну залежність концентрації від відносної вологості повітря, а також синтезовано поліноміальну трифакторну модель другого порядку для коефіцієнта пилопоглинання. За результатами математичного моделювання встановлено домінуючу роль відносної вологості, визначено зони стабілізації пилової хмари та зону сухої деструкції, у якій зростає інтенсивність вторинного зависання частинок. **Висновки.** Запропонована модель із коефіцієнтом детермінації 94 % може бути використана як алгоритмічна основа для Smart HVAC, аспіраційних і зволожувальних систем промислових приміщень, а також для оцінювання ризиків впливу дрібнодисперсного пилу на персонал.

Ключові слова: якість повітря у приміщеннях, пилопоглинання, ресуспензія, відносна вологість, температура, мікроклімат, регресійна модель, Smart HVAC.

1. Вступ

1.1 Актуальність тематики

Забезпечення нормативних показників якості повітряного середовища у приміщеннях (IAQ) належить до пріоритетних завдань у галузях промислової, техногенної безпеки й охорони праці. В умовах сучасної урбанізації людина перебуває у закритих просторах різного призначення до 80–90 % добового часу. Внутрішнє повітряне середовище часто характеризується вищим рівнем забруднення порівняно з атмосферним через акумуляцію та постійну рециркуляцію аерозолів антропогенного та природного походження [1, 2].

Особливу небезпеку становлять дрібнодисперсні зважені частинки фракцій $PM_{2.5}$ та PM_{10} . Результати епідеміологічних і токсикологічних досліджень ВООЗ свідчать, що тривалий інгаляційний вплив таких частинок спричиняє розвиток хронічних обструктивних захворювань легень, серцево-судинних патологій, а також зниження когнітивної та фізичної працездатності [3, 4]. На промислових об'єктах накопичення пилу на поверхнях із його подальшою ресуспензією створює ризики формування зон підвищеної небезпеки [5].

Відомі методи оцінювання якості повітря зазвичай ґрунтуються на порівнянні миттєвих концентрацій із гранично допустимими значеннями [6]. Проте такий підхід є пасивним і не враховує динаміку фізико-хімічних процесів у закритому

просторі під впливом параметрів мікроклімату [7]. Розвиток систем кондиціювання й інтелектуальної вентиляції (Smart HVAC) обумовлює необхідність переходу до предиктивних систем керування якістю повітряного середовища. Одним із ключових параметрів у таких системах має стати коефіцієнт пилопоглинання внутрішнього середовища k_{ds} – інтегральний показник, який описує здатність гетерогенної системи приміщення зв'язувати, коагулювати та виводити зважені частинки з аеродисперсного стану під дією термодинамічних чинників.

1.2 Аналіз літературних джерел і визначення проблеми

Аналіз наукових праць свідчить, що концентрація пилу у приміщенні визначається динамічною рівновагою між процесами генерації, дифузії, гравітаційної седиментації та вентиляційного видалення. Зокрема, у дослідженнях [8, 9] розглянуто механізми вторинного пилоутворення внаслідок рухової активності людини у приміщенні. Авторами встановлено, що кінетична енергія механічного контакту є основним чинником руйнування адгезійних зв'язків, проте інтенсивність цього процесу суттєво залежить від вологовмісту середовища.

Вплив відносної вологості на адгезію в системі «частинка – поверхня» детально описано у роботі [10]. Встановлено, що за умови перевищення

порогу вологості 50–60 % виникає капілярна конденсація водяної пари, що посилює зв'язування пилу на поверхнях і перешкоджає його ресуспензії. Крім того, згідно з дослідженням [11], у вологому середовищі спостерігається гігроскопічне укрупнення аерозолів ($>2,5$ мкм), що прискорює їх седиментацію за законом Стокса.

Вплив температурного градієнта повітряного середовища на аеродисперсні системи має складну подвійну природу. Підвищення температури повітря призводить до збільшення його в'язкості, що сповільнює гравітаційне осадження частинок. Водночас температурні коливання інтенсифікують конвективні потоки та прискорюють випаровування вологи з капілярів, що сприяє відриву сухих частинок від поверхонь [12, 13].

Незважаючи на значний обсяг експериментальних даних, представлених зокрема у праці [14], на сьогодні відсутній уніфікований аналітичний апарат для комплексного опису залежності масової концентрації пилу від температури та вологості через інтегральну функцію пилопоглинання. Існуючі моделі переважно обмежені ізотермічними умовами чи мають вузькоспеціалізований статистичний характер [15, 16].

На основі теоретичного аналізу висунуто гіпотезу, що коефіцієнт пилопоглинання є нелінійною функцією термодинамічного стану повітряного середовища у приміщенні. У цій моделі вологість виступає депресором пилової фази за рахунок капілярних ефектів і гігроскопічного росту, тоді як температура регулює кінетичну енергію системи та стабільність адгезійних зв'язків, що визначає специфіку поведінки різних фракцій пилу.

1.3 Мета і завдання дослідження

Метою роботи є розробка науково обґрунтованого підходу до математичного моделювання коефіцієнта пилопоглинання $PM_{2,5}$ у повітряному середовищі приміщень з урахуванням комплексного впливу температури, відносної вологості та поточної масової концентрації пилу як базових параметрів предиктивного керування якістю внутрішнього повітря.

Для досягнення мети були поставлені такі завдання:

- систематизувати теоретичні положення щодо формування, седиментації та вторинної ресуспензії дрібнодисперсного пилу $PM_{2,5}$ у повітряному середовищі виробничих приміщень;
- сформувані репрезентативний масив експериментальних даних і встановити аналітичну залежність масової концентрації $PM_{2,5}$ від відносної вологості повітря;
- синтезувати регресійну модель сукупного впливу температури, відносної вологості та початкової концентрації $PM_{2,5}$ на коефіцієнт пилопоглинання;
- виконати статистичну ідентифікацію та перевірку адекватності моделі;
- інтерпретувати поверхні відгуку й ізолінії, локалізувати області стабілізації та сухої деструкції

пилової хмари, а також обґрунтувати умови інтегрування одержаної моделі в алгоритми предиктивного керування Smart HVAC.

2. Матеріали і методи дослідження

Об'єктом дослідження є процес формування пилопоглинальної здатності повітряного середовища виробничого приміщення дрібнодисперсної фракції $PM_{2,5}$ за сукупної дії температури, відносної вологості та початкової масової концентрації пилу. Предмет дослідження становлять кількісні закономірності впливу зазначених факторів і їхніх взаємодій на інтегральний коефіцієнт пилопоглинання k_{ds} , який характеризує результуюче вилучення частинок із повітряної фази внаслідок гігроскопічного росту, коагуляції, адгезійного закріплення та седиментації з урахуванням протилежного процесу вторинної ресуспензії.

Наукова проблема полягає у відсутності компактної, статистично обґрунтованої та фізично інтерпретованої моделі, придатної для прогнозування зміни пилопоглинальної здатності внутрішнього повітря за одночасної дії кількох мікрокліматичних факторів. Контроль лише поточної концентрації $PM_{2,5}$ має реактивний характер і не дає змоги завчасно виявляти режими, за яких послаблюється природне осадження та зростає ймовірність повторного зависання осілого пилу. Основна гіпотеза дослідження полягає в тому, що k_{ds} є нелінійною функцією T , φ і C , причому вологісний фактор має домінуючий стабілізуювальний вплив, температура модифікує цей вплив через конвективні та випаровувальні процеси, а зростання концентрації обмежує відносну ефективність самоочищення повітряного об'єму.

Дослідження виконано як теоретико-розрахункове узагальнення лабораторних даних. Матеріалом слугував масив із 21 значення масової концентрації $PM_{2,5}$ у діапазоні відносної вологості 10–90 % із кроком 4 %, сформований на основі результатів досліджень в ізольованій аеродисперсійній симуляційній камері [14] та доповнений даними про варіативність ресуспензії за умови механічного збурення [8, 9]. Такий діапазон обрано для охоплення як сухих режимів, потенційно сприятливих для повторного зависання, так і режимів підвищеної вологості, за яких посилюються капілярне зв'язування та седиментація. Рівномірне розміщення реперних точок забезпечує порівнювану інформаційну щільність у всій області вологи і надає змогу ідентифікувати кривину однофакторної залежності.

У дослідженні прийнято такі припущення та спрощення: повітряне середовище у межах розрахункового об'єму вважається квазіоднорідним за температурою, вологістю та концентрацією; домінантною контрольованою фракцією є $PM_{2,5}$; упродовж окремого вимірювального інтервалу мінералогічний склад, густина та форма частинок не змінюються; швидкість повітря, кратність вентиляції, шорсткість поверхонь й активність працівників не вводяться як окремі фактори, а їхній усереднений

вплив частково відображається у варіації k_{ds} ; фазові перетворення води без утворення макроскопічного конденсату не моделюються; взаємодія факторів у дослідженій області може бути адекватно апроксимована поліномом другого порядку. Ці припущення забезпечують ідентифікованість моделі за наявного обсягу даних, але одночасно визначають межі її подальшого застосування.

Для незалежного відтворення експериментальної частини необхідні герметизована камера з контрольованим повітрообміном; генератор або дозатор тестового пилу; оптичний аналізатор масової концентрації $PM_{2.5}$ із можливістю реєстрації часових рядів; референтний гравіметричний канал для періодичної перевірки оптичного сигналу; калібрований термогігрометр; система керованого зволоження й осушення; пристрій відтворюваного механічного збурення поверхневої пилової фази та реєстратор даних. Застосування одночасного оптичного та референтного вимірювання є необхідним через залежність світлорозсіювання від гігроскопічного росту частинок [16, 17]. У цьому дослідженні конкретні марки приладів не задавалися, оскільки математичний етап ґрунтувався на узагальненому масиві опублікованих лабораторних даних; критеріями вибору обладнання для подальшої натурної валідації мають бути метрологічна простежуваність, достатня часова роздільна здатність і відома невизначеність у заданому діапазоні зміни вологості.

На першому етапі оцінювали однофакторну залежність $C(\varphi) = a \cdot \exp(b\varphi)$. Параметри a і b визначали нелінійним методом найменших квадратів, а якість апроксимації – за коефіцієнтом детермінації R^2 й аналізом відповідності знака параметра b фізичному механізму вологості депресії аерозольної фази. Експоненційну форму обрано тому, що вона відтворює інтенсивне зменшення концентрації на початковій ділянці зі зниженням граничного приросту ефекту за високої вологості, тобто відповідає очікуваному насиченню механізмів гігроскопічного росту та капілярної адгезії.

На другому етапі застосовано трифакторне рототабельне планування другого порядку. Незалежними факторами прийнято T , φ і C , функцією відгуку – k_{ds} . Фактори кодували відносно центральних рівнів, після чого оцінювали вільний член, лінійні та квадратичні ефекти, а також парні взаємодії. Вибір рототабельного плану обґрунтований необхідністю отримати приблизно однакову дисперсію прогнозу для точок, рівновіддалених від центра факторного простору, і водночас ідентифікувати нелінійність без проведення повного багаторівневого перебору всіх комбінацій. Поліном другого порядку обрано як оптимальний між фізичною інтерпретованістю, кількістю оцінюваних параметрів і здатністю відтворити кривину поверхні відгуку.

Статистичну значущість коефіцієнтів оцінювали за двобічним критерієм Стюдента при рівні значущості $\alpha=0,05$; незначущі члени послідовно вилучали зі збереженням ієрархічної та фізичної узгодженості моделі. Адекватність

скороченого рівняння перевіряли за критерієм Фішера, описову здатність – за R^2 , а відносний внесок факторів – за стандартизованою картою Парето. Для фізичної інтерпретації будували двовимірні поверхні відгуку й ізолінії при фіксації третього фактора на центральному рівні. Така сукупність методів обрана тому, що поєднує формальну статистичну перевірку з візуальним визначенням областей екстремуму, стабілізації та сухої деградації.

Обчислення коефіцієнтів регресії, стандартизованих ефектів, критеріїв значущості, поверхонь відгуку й ізоліній виконували у програмному середовищі STATGRAPHICS 18. Табличну підготовку та контроль повноти масиву здійснювали засобами електронних таблиць. Вибір STATGRAPHICS 18 обумовлений наявністю реалізованих процедур планування експерименту, регресійного та дисперсійного аналізу, а також узгодженого формування карти Парето та графічних поверхонь за однією моделлю. Критерієм прийнятності результату вважали одночасне виконання статистичних умов адекватності та фізичної несуперечності встановлених ефектів.

3. Результати розробки і дослідження методу керування обчислювальними ресурсами туманного шару

Послідовність подання результатів відповідає етапам перевірки гіпотези: спочатку встановлено часткову залежність концентрації $PM_{2.5}$ від відносної вологості, далі ідентифіковано трифакторне рівняння k_{ds} , після чого визначено статистичну ієрархію ефектів і параметричні області характерних режимів пилопоглинання.

Для побудови регресійної моделі та визначення аналітичної залежності масової концентрації пилу від відносної вологості повітря використано результати лабораторних досліджень [14]. Зазначені експерименти проводилися в ізолюванні аеродисперсійній симуляційній камері з реєстрацією кінетики вторинного зважування частинок. З метою підвищення репрезентативності даних і врахування варіативності аерозольного фону, вихідний масив точок було доповнено результатами вимірювань із суміжних джерел, зокрема щодо динаміки ресуспензії пилу за умов механічного збурення середовища [8, 9]. Сформована вибірка із 21 реперної точки (табл. 1) враховує природну дисперсію фракції $PM_{2.5}$ і похибки вимірювальних приладів. На основі цих даних встановлено залежність масової концентрації пилу від відносної вологості повітря у приміщенні (рис. 1).

Застосовуючи нелінійний метод найменших квадратів до розширеного масиву із 21 експериментальних точок, синтезовано базове математичне рівняння першого етапу:

$$C(\varphi) = 0,8491e^{-0,014\varphi}, \quad (1)$$

де C – масова концентрація дрібнодисперсного пилу фракції $PM_{2.5}$, мг/м³; φ – відносна вологість внутрішнього повітряного середовища, %; 0,8491 –

асимптотичний коефіцієнт, який характеризує гранично можливу масову концентрацію сухого зваженого пилу при H наближається до 0 %; -0,014 – коефіцієнт депресії пилоутворення, який безпосередньо відображає інтенсивність капілярного зв'язування та седиментації для зазначеної дисперсної системи.

Таблиця 1 – Експериментальні значення масової концентрації пилу $PM_{2,5}$ залежно від відносної вологості

№ точки	Відносна вологість повітря ϕ , %	Фактична масова концентрація $C_{факт}$, мг/м ³
1	10	0,675
2	14	0,692
3	18	0,621
4	22	0,645
5	26	0,602
6	30	0,541
7	34	0,558
8	38	0,492
9	42	0,515
10	46	0,452
11	50	0,468
12	54	0,412
13	58	0,425
14	62	0,381
15	66	0,362
16	70	0,371
17	74	0,332
18	78	0,280
19	82	0,218
20	86	0,152
21	90	0,113

Коефіцієнт детермінації отриманої моделі становить $R^2=0,911$, що згідно зі шкалою Чеддока свідчить про функціональний, надзвичайно високий рівень адекватності та зв'язку між параметрами.

На основі проведених досліджень та встановленого взаємозв'язку температури та відносної вологості [6] для визначення коефіцієнта пилопоглинання $PM_{2,5}$ запропоновано застосування методу планування експерименту, який дозволяє варіювати одночасно всі фактори, отримувати кількісні оцінки основних ефектів й ефектів парних взаємодій, визначати параметр пилопоглинання $PM_{2,5}$ за відомими залежностями шуканого показника окремо для кожного з факторів.

Застосування методів планування експерименту порівняно із традиційними методами дають змогу підвищити ефективність наукових досліджень у 2–10 разів. При цьому результати моделювання можуть бути використані при практичній реалізації у системах керування рівнем пилопоглинання $PM_{2,5}$ у робочій зоні виробничих приміщень.

В основу побудови моделі покладені визначений діапазон температури, відносної вологості та концентрації пилу $PM_{2,5}$, зазначені у нормативних документах.

Трифакторна модель пилопоглинання $PM_{2,5}$ для врахування мікрокліматичних параметрів і концентрації пилу та результатів їх взаємодії має вигляд:

$$k_{ds} = b_0 + b_1T + b_2\phi + b_3C + b_{12}T\phi + b_{23}\phi C + b_{31}CT + b_{11}T^2 + b_{22}\phi^2 + b_{33}C^2 + b_{123}T\phi C, \quad (2)$$

де T – температура, °C; ϕ – відносна вологість, %; C – концентрації пилу $PM_{2,5}$, мг/м³; b – коефіцієнти регресійної моделі.

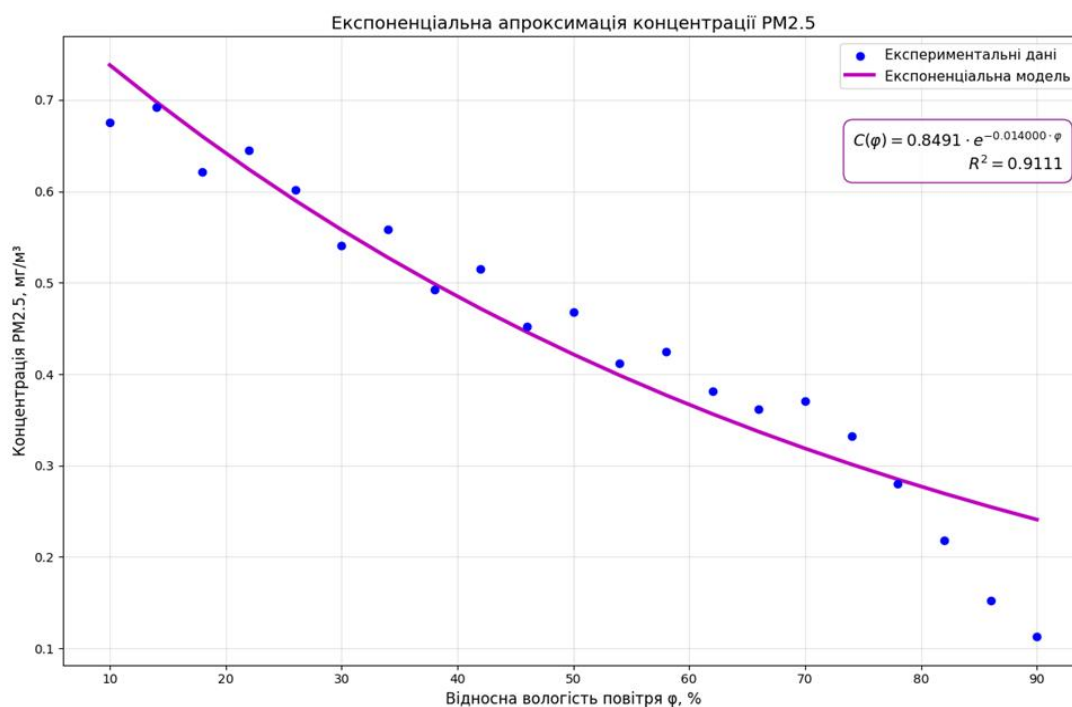


Рис. 1. Залежність концентрації пилу $PM_{2,5}$ від відносної вологості у приміщенні

Прийняті області варіювання параметрів при проведенні експерименту наведені у табл. 2.

Таблиця 2 – Області варіювання параметрів при проведенні експерименту

Фактори	Позначення	Область досліду
Температура, °C	T	12–30
Відносна вологість повітря, %	φ	10–90
Концентрації пилу $PM_{2,5}$, мг/м ³	C	0,15–0,85

Математична обробка отриманих даних дозволила встановити спільний вплив температури, вологості та концентрації пилу повітря у приміщенні на коефіцієнт пилопоглинання. Як спосіб отримання поліноміальної моделі застосовано трифакторне рототабельне планування другого порядку. Функцією відгуку є коефіцієнт пилопоглинання. Фактори у кодованому вигляді й інтервали їх варіювання наведені у табл. 3.

Таблиця 3 – Матриця планування експерименту

T , в.о.	φ , в.о.	C , в.о.	T , °C	φ , %	C , мг/м ³	K , в.о.
-1	-1	-1	14	30	0,25	0,42
+1	-1	-1	22	30	0,25	0,35
-1	+1	-1	14	70	0,25	0,68
+1	+1	-1	22	70	0,25	0,59
-1	-1	+1	14	30	0,75	0,14
+1	-1	+1	22	30	0,75	0,08
-1	+1	+1	14	70	0,75	0,51
+1	+1	+1	22	70	0,75	0,44
-1,682	0	0	12	50	0,50	0,56
+1,682	0	0	30	50	0,50	0,38
0	-1,682	0	18	10	0,50	0,12
0	+1,682	0	18	90	0,50	0,64
0	0	-1,682	18	50	0,15	0,73
0	0	+1,682	18	50	0,85	0,18
0	0	0	18	50	0,50	0,49

При обробці результатів розрахунків експерименту шляхом виключення статистично незначущих ефектів за критерієм Стюдента було отримано рівняння регресії другого порядку, адекватність якого підтверджена за критерієм Фішера.

Остаточно регресійна модель залежності коефіцієнта пилопоглинання дрібнодисперсної фракції $PM_{2,5}$ у кодованих значеннях має вигляд:

$$k_{ds} = 0,49 - 0,045T + 0,145\varphi - 0,152C - 0,022T^2 - 0,038\varphi^2 - 0,015C^2 + 0,012T\varphi - 0,025\varphi C, \quad (3)$$

де T – температура повітря в приміщенні, °C; φ – відносна вологість повітря, %; C – масова концентрація пилу $PM_{2,5}$ мг/м³.

Перевірка статистичної значущості параметрів здійснювалася за допомогою стандартизованої

Парето-карти. Адекватність моделі перевірена за коефіцієнтом детермінації $R^2=0,942$ що свідчить про високу якісну характеристику зв'язку коефіцієнтів системи та можливість її використання для предиктивного керування автоматизованими комплексами підтримки комфортних умов праці.

Для визначення ступеня та характеру впливу досліджуваних чинників на цільову функцію відгуку (коефіцієнт пилопоглинання k_{ds}) у програмному комплексі «STATGRAPHICS Plus» було виконано процедуру стандартизації ефектів. Оцінка статистичної значущості лінійних, квадратичних і парних взаємодій отриманого поліноміального рівняння регресії здійснювалася за допомогою стандартизованої карти Парето (критерій Стюдента, рівень значущості $\alpha=0,05$).

Аналіз отриманого розподілу свідчить, що домінуючий внесок у формування величини коефіцієнта пилопоглинання має відносна вологість повітря, яка має додатне значення стандартизованого ефекту. Це підтверджує раніше висунуту гіпотезу про те, що інтенсифікація процесів вологої седиментації та мінімізація вторинного зависання пилу за рахунок запобігання сухій деструкції мікрозв'язків відіграють ключову роль у дезактивації пилової хмари фракції $PM_{2,5}$.

Другим за значущістю фактором є вхідна масова концентрація пилу, яка має негативний знак ефекту, що свідчить про поступове насичення ємності самоочищення повітряного об'єму при критичних рівнях запилення. Температурний чинник і квадратичний ефект вологості (φ^2) також визнані статистично значущими, оскільки їх абсолютні значення перевищують критичну межу Стюдента ($t_{крит}=2,57$), тоді як ефекти потрібної взаємодії ($T\cdot\varphi\cdot C$) та квадратичний ефект концентрації (C^2) виявилися нижчими за поріг значущості та були виключені при оптимізації моделі.

Для детального вивчення взаємного впливу термодинамічних параметрів і концентрації на пилопоглинальне кумулятивне поле зафіксовано один із факторів на базовому (нульовому) рівні та побудовано двовимірні поверхні відгуку та їхні ізолінії (лінії рівного рівня).

На рис. 2 представлена поверхня відгуку та лінії рівня залежності коефіцієнта пилопоглинання k_{ds} від температури та відносної вологості повітря при фіксованій середній концентрації пилу $PM_{2,5}$ ($C=0,50$ мг/м³).

Аналіз геометричного профілю поверхні (рис. 2) показує чітко виражений нелінійний екстремальний характер взаємодії. Максимальна ефективність утримання й осадження дрібнодисперсного аерозолі ($k_{ds}\geq 0,60$) досягається у зоні низьких температур (12–16 °C) та високої відносної вологості ($\varphi=75-90$ %). Фізично це пояснюється зниженням кінетичної енергії термічної конвекції повітряних мас й одночасним досягненням максимальної стабільності водних мікрозв'язків між частинками пилу та поверхнею підлоги.

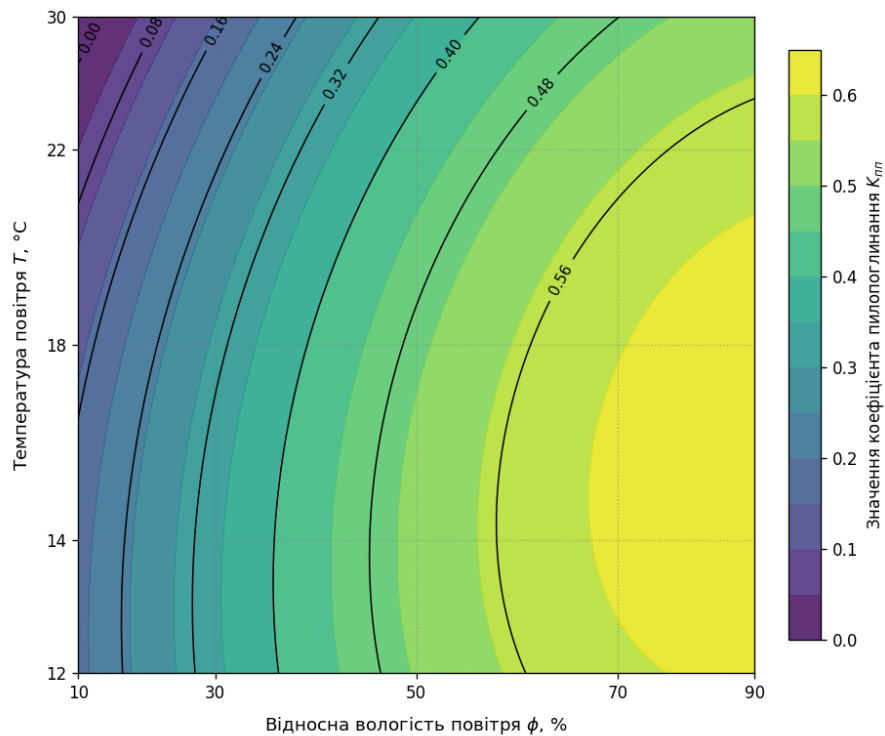


Рис. 2. Залежності температури та відносної вологості повітря при фіксованій середній концентрації пилу $PM_{2.5}$ ($C=0.50$ mg/m^3)

Зі зростанням температури понад $24^\circ C$ та зниженням вологості нижче 30% спостерігається стрімке падіння коефіцієнта пилопоглинання до критичного мінімуму ($k_{ds} \leq 0.15$). Зазначена область характеризується як зона сухої деструкції, де повне випаровування води призводить до руйнування капілярних сил адгезії, внаслідок чого зважений пил

$PM_{2.5}$ під дією слабких механічних збурень (конвективних потоків або руху персоналу) інтенсивно ресуспендується (зависає), практично не піддаючись природному осадженню.

На рис. 3 наведено ізолінії взаємодії відносної вологості та концентрації пилу при нормативній базовій температурі середовища ($T=18^\circ C$).

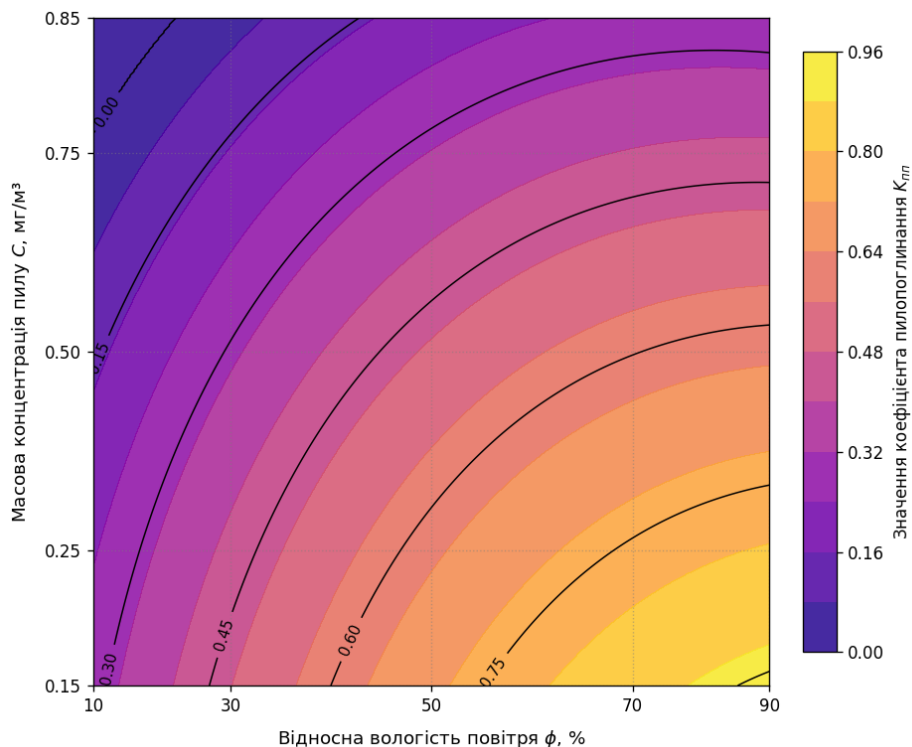


Рис. 3. Залежності відносної вологості та концентрації пилу при нормативній базовій температурі середовища ($T=18^\circ C$)

Характер ізоліній на рис. 3 підтверджує, що навіть за умов високого первинного запилення робочої зони ($C \geq 0,70$ мг/м³), штучне підвищення відносної вологості повітря до рівня 65–80 % дозволяє утримувати коефіцієнт пилопоглинання на стабільно високому рівні ($k_{ds}=0,45\text{--}0,50$). Це свідчить про високу компенсаторну здатність вологісного фактора та вказує на доцільність інтеграції отриманої поліноміальної моделі в алгоритми предиктивного автоматизованого керування аспіраційними та зволожувальними інженерними системами промислових приміщень.

Отже, сукупність результатів підтвердила нелінійність функції відгуку та переважання вологісного фактора. Модель відокремлює область підвищеної стабілізації пилової фази від області сухої деструкції та забезпечує кількісну основу для подальшої інтерпретації механізмів, оцінювання практичної придатності та визначення обмежень застосування.

4. Обговорення результатів дослідження синтезу трифакторної регресійної моделі пилопоглинання у повітряному середовищі виробничих приміщень

Отримані результати підтверджують основну гіпотезу дослідження: коефіцієнт пилопоглинання є нелінійною функцією спільного впливу температури, відносної вологості та початкової концентрації РМ_{2,5}. Значення $R^2=0,942$ свідчить, що у межах прийнятого факторного простору модель пояснює переважну частину варіації k_{ds} . Принципово важливо, що статистично значущі коефіцієнти мають фізично несуперечливі знаки й інтерпретацію. Це надає підстави розглядати рівняння не лише як формальну апроксимацію, а як спрощене інженерне представлення динамічного балансу між гігроскопічним ростом, коагуляцією, адгезійним закріпленням, седиментацією та ресуспензією частинок.

Встановлене експоненційне зменшення C зі зростанням φ ($R^2=0,911$) узгоджується з результатами [10], які показали залежність ресуспензії від вологості та властивостей системи «частинка–поверхня», а також із дослідженням [11], де вологість впливала на осадження та коагуляцію аерозольних частинок. Дані [12] і огляд [13] додатково підтверджують, що швидкість осадження істотно залежить від поверхні та пристінного повітряного потоку. На відміну від зазначених праць, у яких окремі механізми аналізуються переважно автономно, запропонована модель інтегрує їхній результуючий прояв у показнику k_{ds} і пов'язує його із трьома керованими параметрами.

Переважаючий вплив відносної вологості пояснюється її одночасним впливом на кілька ланок аерозольної динаміки. Адсорбція водяної пари збільшує ефективний діаметр і масу гігроскопічних частинок, що підвищує швидкість гравітаційного осадження. Формування капілярних містків збільшує силу зчеплення з поверхнею та зменшує ймовірність повторного відриву. Водночас значущість

квадратичного члена φ^2 свідчить про насичення цього ефекту: після певного рівня подальше зволоження забезпечує менший приріст k_{ds} . Саме нелінійність пояснює перевагу поліноміальної моделі другого порядку над лінійною та дозволяє визначати не просто напрям зміни, а й область раціонального керування.

Температурний ефект формується через зміну в'язкості повітря, інтенсивності природної конвекції, швидкості випаровування адсорбованої вологи та стійкості капілярних зв'язків. За помірно низьких температур послаблюється конвективне перемішування та зберігається волога у контактній зоні, що сприяє осадженню та фіксації частинок. Підвищення температури прискорює висушування поверхневої пилової фази та створює умови для відриву частинок. Тому температура не є ізольованим монотонним регулятором: її вплив реалізується через взаємодію з вологістю та станом поверхневих відкладень, що підтверджується кривою одержаної поверхні відгуку.

Негативний ефект початкової концентрації інтерпретується як прояв обмеженої асиміляційної ємності повітряного об'єму та поверхонь осадження. Зі збільшенням пилового навантаження зростає абсолютна кількість частинок, що залишається у зваженому стані, а накопичений на підлозі й обладнанні пил утворює вторинне джерело. Дослідження [5, 8, 9] показали, що ходіння й інша механічна активність можуть істотно збільшувати внутрішню концентрацію частинок. Це пояснює, чому k_{ds} не може вважатися сталою властивістю пилу: він є характеристикою конкретної системи «аерозоль–повітря–поверхня–режим збурення».

Локалізація області $T=12\text{--}16$ °C і $\varphi=75\text{--}90$ % як зони максимального k_{ds} має фізичний зміст, оскільки відповідає поєднанню зниженої термічної конвекції, гігроскопічного укрупнення та стабілізації капілярних контактів. Протилежна область $T \geq 24$ °C і $\varphi \leq 30$ %, визначена як зона сухої деструкції, характеризується послабленням адгезійних сил і підвищенням імовірності ресуспензії. Виявлення цієї зони є ключовим результатом: воно перетворює якісне уявлення про небезпеку сухого мікроклімату на кількісну межу, яку можна використати як попереджувальний індикатор у системі керування.

Запропоноване рішення вирішує проблему реактивного контролю завдяки переходу від реєстрації вже сформованого перевищення РМ_{2,5} до прогнозування зміни пилопоглинальної здатності. Це забезпечується трьома властивостями моделі: використанням синхронізованого вектора стану (T, φ, C), урахуванням нелінійних і парних ефектів і виділенням області сухої деструкції. Завдяки цьому керовальна система може реагувати не лише на фактичну концентрацію, а й на наближення мікроклімату до режиму, в якому природне вилучення частинок слабшає, а ризик вторинного зависання зростає.

Основною сферою застосування результатів є автоматизовані системи Smart HVAC, локальна аспірація, кероване зволоження, виробничий

моніторинг IAQ та ризик-орієнтоване управління охороною праці. На вході алгоритму мають використовуватися синхронізовані вимірювання $PM_{2,5}$, T і φ ; на виході – прогноз k_{ds} і сигнал про наближення до несприятливої області. Керувальні дії можуть включати зміну продуктивності аспірації, кратності загальнообмінної вентиляції, інтенсивності локального зволоження та режиму безпилового прибирання. Потенційно очікуваними ефектами є зменшення амплітуди та тривалості піків $PM_{2,5}$, скорочення часово зваженої інгалаційної експозиції працівників, раціональніше використання вентиляційного обладнання та зниження ймовірності формування вторинних пилових хмар.

Умовою практичного застосування є попередня локальна ідентифікація моделі для конкретного типу пилу, технологічного процесу, вентиляційної схеми та поверхонь приміщення. Інтервал $\varphi=75\text{--}90\%$, у якому розраховано максимальні значення k_{ds} , не варто безпосередньо трактувати як рекомендований режим постійної експлуатації. Надмірна вологість може спричинити конденсацію, корозію, деградацію матеріалів і погіршення теплового комфорту. Практично доцільно визначати мінімально достатній рівень φ , який забезпечує необхідне зниження $PM_{2,5}$ за обмежень щодо точки роси, санітарних норм, енергоспоживання та технічного стану обладнання.

Важливою умовою є також метрологічна коректність вхідних даних. Гігроскопічний ріст змінює інтенсивність світлорозсіювання, тому оптичний сенсор може реєструвати зміну сигналу навіть без пропорційної зміни сухої маси пилу. Дослідження [16, 18] й огляд [17] підтверджують необхідність калібрування, введення температурно-вологісних поправок і контролю дрейфу. Отже, безперервний оптичний канал доцільно періодично звіряти з референтним методом, а невизначеність C , T і φ враховувати під час формування прогностичного інтервалу k_{ds} .

Перевагами запропонованого підходу є фізична інтерпретованість факторів, можливість оперативного обчислення та безпосередня придатність до алгоритмізації. Рототабельний план забезпечує оцінювання кривини та взаємодій за меншої кількості комбінацій, ніж повний багаторівневий експеримент. Порівняно із суто емпіричним пороговим керуванням модель формує причинно обґрунтовану прогностичну ознаку. Порівняно з тривимірним CFD-моделюванням вона потребує значно менше обчислювальних ресурсів і може працювати у реальному часі, хоча не відтворює просторову неоднорідність повітряного потоку.

До обмежень проведеного дослідження належать невеликий обсяг вихідного масиву, використання узагальнених вторинних лабораторних даних, відсутність незалежної зовнішньої вибірки та явних предикторів повітрообміну, швидкості потоку, шорсткості поверхні, хімічного складу пилу й активності працівників. Припущення про квазі-однорідність не дозволяє описати локальні зони біля джерел, припливних струменів або робочих місць. Значення R^2 характеризує відповідність у межах

навчального масиву, але саме по собі не може гарантувати прогностичної точності поза дослідженими діапазонами. Недоліком є також можливе змішування реальної зміни сухої концентрації з вологісною чутливістю оптичного вимірювання.

Подальший розвиток дослідження має передбачати натурну валідацію на незалежних виробничих об'єктах; повторні серії у центральних точках плану; стратифікацію за складом і дисперсністю пилу; введення швидкості повітря, кратності вентиляції, поверхневого пилового навантаження й інтенсивності руху працівників як додаткових факторів; оцінювання розширеної невизначеності; аналіз залишків на нормальність, гомоскедастичність і автокореляцію; перехресну перевірку та порівняння з моделлю «grey-box», узагальненими адитивними та CFD-моделями [19, 20]. Завершальним етапом доцільно зробити замкнений виробничий експеримент, у якому керування вентиляцією та зволоженням здійснюється за прогностичним k_{ds} , а ефективність оцінюється за зменшенням експозиції $PM_{2,5}$ та стабільністю нормативних параметрів мікроклімату.

5. Висновки

У представленому дослідженні виконано теоретико-експериментальне узагальнення та математичне моделювання спільних термодинамічних і седиментаційних процесів у повітряному середовищі приміщень. Отримані результати дозволили сформулювати такі основні висновки:

1. Математично обґрунтовано та вперше синтезовано цілісну трифакторну регресійну модель другого порядку, яка комплексно описує нелінійну залежність інтегрального коефіцієнта пилопоглинання дрібнодисперсної фракції $PM_{2,5}$ від сумісного впливу базових термодинамічних параметрів мікроклімату (температури і відносної вологості) та запилення повітряного середовища промислових приміщень. Високе значення коефіцієнта детермінації моделі ($R^2=0,942$) підтверджує її статистичну адекватність за критерієм Фішера та гарантує високу точність предиктивного аналізу кумулятивних властивостей повітряного об'єму.

2. Експериментально виявлено та графічно інтерпретовано істотно екстремальний характер взаємодії досліджуваних чинників, що дозволило чітко розмежувати просторові зони стабілізації та деструкції пилової хмари. Шляхом топографічного аналізу двовимірних ізоліній відгуку локалізовано область максимальної ефективності природного осадження аерозолі ($k_{ds}\geq 0,60$) в інтервалі температур ($12\text{--}16\text{ }^\circ\text{C}$) та високої відносної вологості ($\varphi=75\text{--}90\%$), а також математично доведено існування «зони сухої деструкції» ($T\geq 24\text{ }^\circ\text{C}$, $\varphi\leq 30\%$), де нівелювання капілярних сил адгезії через дефіцит вологи призводить до інтенсивного вторинного ресуспендування (зависання) частинок $PM_{2,5}$ під дією конвективних потоків.

3. Доведено високу компенсаторну здатність вологісного чинника, яка дозволяє нівелювати

негативний вплив критичного первинного запилення робочої зони ($C \geq 0,70$ мг/м³) і стабілізувати коефіцієнт пилопоглинання на рівні 0,45–0,50 за рахунок штучного локального збільшення відносної вологості повітря до 65–80%. Одержана закономірність разом із розробленою математичною моделлю становить алгоритмічний базис для побудови предиктивних систем автоматизованого керування комплексами промислової аспірації та зволоження повітря, забезпечуючи досягнення нормативної якості внутрішнього повітря (IAQ).

Конфлікт інтересів

Автори декларують, що не мають конфлікту інтересів стосовно даного дослідження, в тому числі

фінансового, особистісного характеру, авторства чи іншого характеру, що міг би вплинути на дослідження та його результати, представлені в даній статті.

Фінансування

Дослідження проводилося без фінансової підтримки.

Доступність даних

Рукопис не має пов'язаних даних.

Використання засобів штучного інтелекту

Автори підтверджують, що не використовували технології штучного інтелекту при створенні представленої роботи.

СПИСОК ЛІТЕРАТУРИ

1. Spengler, J.D., McCarthy, J.F. and Samet, J.M. (2000), *Indoor Air Quality Handbook*, 1st ed., McGraw-Hill, 1488 p., available at: <https://www.accessengineeringlibrary.com/content/book/9780074455494>
2. Петренко І. С., Ченчева О. О., Лашко Є. Є., Чеберячко Ю. І., Чеберячко С. І., Білоусова К. П. Дослідження впливу пилового навантаження та оцінка ризику на здоров'я працівників. *Проблеми охорони праці в Україні*. 2024. Т. 40, № 1–2. С. 112–120. <https://doi.org/10.36804/nndipbop.40-1-2.2024.52-56>
3. (2018), *Housing and health guidelines*. World Health Organization. Retrieved July 14, 2026, available at: <https://iris.who.int/server/api/core/bitstreams/be014865-921a-482c-89e4-9191c7476e13/content>
4. Ченчева О. О., Бурдейна Н. Б., Лашко Є. Є., Шевченко В. Г., Петренко І. С. Вплив пилоутворення при механічному обробленні карбон-карбонових композитів на ризик виникнення професійних захворювань. *Проблеми охорони праці в Україні*. 2022. Т. 38, № 3–4. С. 25–33. <https://doi.org/10.36804/nndipbop.38-3-4.2022.25-33>
5. Ferro, A. R., Kopperud, R. J., and Hildemann, L. M. (2004), "Source strengths for indoor human activities that resuspend particulate matter", *Environmental Science & Technology*, vol. 38, no. 6, pp. 1759–1764, doi: <https://doi.org/10.1021/es0263893>
6. Запорожець О. І., Сукач С. В., Галаган О. Г., Козловська Т. Ф. Визначення параметрів оптимальної комфортності у робочій зоні приміщення за показниками повітряного середовища. *Вісник Кременчуцького національного університету ім. М. Остроградського*. 2017. № 102. С. 17–21, URL: https://visnikkrmu.kdu.edu.ua/statti/2017_1_17-21_1-2017.pdf
7. Nazaroff, W. W. (2004), "Indoor particle dynamics", *Indoor Air*, vol. 14, pp. 175–183, doi: <https://doi.org/10.1111/j.1600-0668.2004.00286.x>
8. Tian, Y., Sul, K., Qian, J., Mondal, S., and Ferro, A. R. (2014), "A comparative study of walking-induced dust resuspension using a consistent test mechanism", *Indoor air*, vol. 24, pp. 592–603, doi: <https://doi.org/10.1111/ina.12107>
9. Rosati, J. A., Thornburg, J., and Rodes, C. (2008), "Resuspension of particulate matter from carpet due to human activity", *Aerosol Science and Technology*, vol. 42, pp. 472–482, doi: <https://doi.org/10.1080/02786820802187069>
10. Kim, Y., Wellum, G., Mello, K., Strawhecker, K. E., Thoms, R., Giaya, A., and Wysloulzil, B. E. (2016), "Effects of relative humidity and particle and surface properties on particle resuspension rates", *Aerosol Science and Technology*, vol. 50, pp. 339–352, doi: <https://doi.org/10.1080/02786826.2016.1152350>
11. Wang, Y., Chen, L., Chen, R., Tian, G., Li, D., Chen, C., Ge, X., and Ge, G. (2017), "Effect of relative humidity on the deposition and coagulation of aerosolized SiO₂ nanoparticles", *Atmospheric Research*, vol. 194, pp. 100–108, doi: <https://doi.org/10.1016/j.atmosres.2017.04.030>
12. El Hamdani, S., Limam, K., Abadie, M. O., and Bendou, A. (2008), "Deposition of fine particles on building internal surfaces", *Atmospheric Environment*, vol. 42, pp. 8893–8901, doi: <https://doi.org/10.1016/j.atmosenv.2008.09.005>
13. Costa, D., Malet, J., and Gehin, E. (2022), "Dry aerosol particle deposition on indoor surfaces: Review of direct measurement techniques", *Aerosol Science and Technology*, vol. 56, pp. 261–280, doi: <https://doi.org/10.1080/02786826.2021.2013431>
14. Kim, M., and Park, J. (2024), "Study on size distribution and characteristics of particulate matter suspension in indoor space depending on relative humidity", *Particle And Aerosol Research*, vol. 20, no. 2, pp. 25–33, doi: <https://doi.org/10.11629/jpaar.2024.20.2.025>
15. Duarte, R. M. B. O., Gomes, J. F. P., Querol, X., Cattaneo, A., Bergmans, B., Saraga, D., Maggos, T., Di Gilio, A., Villanueva, F., and Mihucz, V. G. (2022), "Particulate matter indoors: A strategy to sample and monitor size-selective fractions", *Applied Spectroscopy Reviews*, vol. 57, pp. 675–704, doi: <https://doi.org/10.1080/05704928.2022.2088554>
16. Bousiotis, D., Alconcel, L.-N. S., Beddows, D. C. S., Harrison, R. M., and Pope, F. D. (2023), "Monitoring and apportioning sources of indoor air quality using low-cost particulate matter sensors", *Environment International*, vol. 174, 107907, doi: <https://doi.org/10.1016/j.envint.2023.107907>
17. Lowther, S. D., Jones, K. C., Wang, X., Whyatt, J. D., Wild, O., and Booker, D. (2019), "Particulate matter measurement indoors: A review of metrics, sensors, needs, and applications", *Environmental Science & Technology*, vol. 53, pp. 11644–11656, doi: <https://doi.org/10.1021/acs.est.9b03425>
18. Lunderberg, D. M., Liang, Y., Singer, B. C., Apte, J. S., Nazaroff, W. W., and Goldstein, A. H. (2023), "Assessing residential PM_{2.5} concentrations and infiltration factors with high spatiotemporal resolution using crowdsourced sensors", *Proceedings of the National Academy of Sciences*, vol. 120, no. 50, e2308832120, doi: <https://doi.org/10.1073/pnas.2308832120>

19. González-Martín, J., Kraakman, N. J. R., Pérez, C., Lebrero, R., and Muñoz, R. (2021), "A state-of-the-art review on indoor air pollution and strategies for indoor air pollution control", *Chemosphere*, vol. 262, 128376, doi: <https://doi.org/10.1016/j.chemosphere.2020.128376>
20. Zhang, L., Ou, C., Magana-Arachchi, D., Vithanage, M., Vanka, K. S., Palanisami, T., Masakorala, K., Wijesekara, H., Yan, Y., Bolan, N., and Kirkham, M. B. (2021), "Indoor particulate matter in urban households: Sources, pathways, characteristics, health effects, and exposure mitigation", *International Journal of Environmental Research and Public Health*, vol. 18, no. 21, 11055, doi: <https://doi.org/10.3390/ijerph182111055>

Received (Надійшла) 29.06.2026

Accepted for publication (Прийнята до друку) 02.09.2026

Publication date (Дата публікації) 25.09.2026

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Сукач Сергій Володимирович – доктор технічних наук, професор, завідувач кафедри цивільної безпеки, охорони праці, геодезії та землеустрою, Кременчуцький національний університет ім. М. Остроградського, Кременчук, Україна;
Serhii Sukach – Doctor of Technical Sciences, Professor, Head of the Department of Civil and Labour Safety, Geodesy and Land Management, Kremenchuk Mykhailo Ostrohradskyi National University, Kremenchuk, Ukraine;
 e-mail: sergvs69@ukr.net; ORCID Author ID: <https://orcid.org/0000-0002-6834-0197>;
 Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57202153691>.

Лашко Євгеній Євгенович – кандидат технічних наук, доцент, доцент кафедри цивільної безпеки, охорони праці, геодезії та землеустрою, Кременчуцький національний університет ім. М. Остроградського, Кременчук, Україна;
Yevhenii Lashko – Candidate of Technical Sciences, Associate Professor, Associate Professor of Department of Civil and Labour Safety, Geodesy and Land Management, Kremenchuk M. Ostrohradskyi National University, Kremenchuk, Ukraine;
 e-mail: evgeny.lashko.lj@gmail.com; ORCID Author ID: <http://orcid.org/0000-0001-9691-4648>;
 Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57203623830>.

Ченчева Ольга Олександрівна – доктор технічних наук, доцент, доцент кафедри цивільної безпеки, охорони праці, геодезії та землеустрою, Кременчуцький національний університет ім. М. Остроградського, Кременчук, Україна;
Olha Chencheva – Doctor of Technical Sciences, Associate Professor, Associate Professor of Department of Civil and Labour Safety, Geodesy and Land Management, Kremenchuk M. Ostrohradskyi National University, Kremenchuk, Ukraine;
 e-mail: chenchevaolga@gmail.com; ORCID Author ID: <https://orcid.org/0000-0002-5691-7884>;
 Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57203619235>.

Резнік Дмитро Володимирович – кандидат технічних наук, доцент, доцент кафедри цивільної безпеки, охорони праці, геодезії та землеустрою, Кременчуцький національний університет ім. М. Остроградського, Кременчук, Україна;
Dmytro Rieznik – Candidate of Technical Sciences, Associate Professor, Associate Professor of Department of Civil and Labour Safety, Geodesy and Land Management, Kremenchuk M. Ostrohradskyi National University, Kremenchuk, Ukraine;
 e-mail: 2411dimareznik@gmail.com; ORCID Author ID: <https://orcid.org/0000-0003-1258-6136>;
 Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=58493079400>.

Петренко Іван Сергійович – доктор філософії, асистент кафедри цивільної безпеки, охорони праці, геодезії та землеустрою, Кременчуцький національний університет імені Михайла Остроградського, Кременчук, Україна;
Ivan Petrenko – Ph.D., Assistant of Department of Civil and Labour Safety, Geodesy and Land Management, Kremenchuk Mykhailo Ostrohradskyi National University, Kremenchuk, Ukraine;
 e-mail: ivanpetrenko95@outlook.com; ORCID Author ID: <https://orcid.org/0000-0002-9846-3737>;
 Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=58096873600>.

Synthesis of a three-factor regression model of dust suppression in the air environment of industrial premises

Serhii Sukach, Yevhenii Lashko, Olha Chencheva, Dmytro Rieznik, Ivan Petrenko

Abstract. Background. Ensuring the regulatory quality of air in industrial premises is an important condition for industrial, technogenic, and occupational safety, since fine dust can remain suspended for a long time, become resuspended from surfaces, and form an increased inhalation load on workers. **Subject of the study:** the processes of dust suppression, sedimentation, and resuspension of fine particles in the indoor air environment under variable microclimate parameters. **Objective:** to develop a scientifically substantiated three-factor regression model of the dust suppression coefficient, taking into account temperature, relative humidity, and mass concentration of dust for predictive control of indoor air quality. **Research tasks.** To systematize the physical mechanisms of fine-particle deposition and resuspension; establish the analytical relationship between PM_{2.5} concentration and relative humidity; synthesize and statistically validate a second-order polynomial model; identify dust-cloud stabilization and dry-destruction regions; and substantiate practical model applications. **Methods.** Theoretical analysis, synthesis of experimental data, nonlinear least squares, three-factor rotatable second-order design, regression and analysis of variance, Student's t-test for coefficients, Fisher's F-test for model adequacy, and response-surface and contour analysis. **Results.** The paper summarizes modern concepts regarding the influence of humidity, temperature, particle adhesion, hygroscopic growth, and human mechanical activity on the formation of aerosol load. Based on an experimental dataset, an exponential dependence of concentration on relative air humidity was constructed, and a second-order polynomial three-factor model for the dust suppression coefficient was synthesized. According to the results of mathematical modeling, the dominant role of relative humidity was established, the zones of dust cloud stabilization and the zone of dry destruction were determined, in which the intensity of secondary particle suspension increases. **Conclusions.** The proposed model, with a coefficient of determination of 94 %, can be used as an algorithmic basis for Smart HVAC, aspiration, and humidification systems in industrial premises, as well as for assessing the risks of fine dust exposure to personnel.

Keywords: indoor air quality; dust suppression; resuspension; relative humidity; temperature; microclimate; regression model; Smart HVAC.

Д. В. Усенко¹, В. В. Биба¹, Н. М. Пінчук²

¹ Національний університет «Полтавська політехніка імені Юрія Кондратюка», Полтава, Україна

² Politecnico di Bari, Bari, Italy

МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ ДИНАМІЧНОЇ ТЕПЛОВОЇ РЕАКЦІЇ БАГАТОШАРОВОЇ СТІНОВОЇ КОНСТРУКЦІЇ ЗА ПЕРІОДИЧНОГО ТА ІМПУЛЬСНОГО ТЕПЛОВОГО НАВАНТАЖЕННЯ

Анотація. Актуальність. Стационарний термічний опір не дає вичерпної характеристики багатошарових огорожень за змінних температурних впливів, оскільки не враховує просторового розподілу теплової інерції та режиму збудження. **Предмет дослідження:** динамічна тепла реакція стаціонарно еквівалентної тришарової стінової конструкції за періодичного та скінченного імпульсного теплового навантаження. **Мета дослідження:** оцінити вплив порядку шарів, температурозалежних теплофізичних властивостей і компенсованого міжшарового термічного опору за незмінного сумарного стаціонарного термічного опору. **Завдання дослідження.** Сформулювати сценарії В, О, N, C5 і C10; визначити режимно відповідні показники; перевірити чисельну стійкість і роздільність ефектів. **Методи дослідження.** Використано одновимірну модель нестационарної теплопровідності, граничні умови Робіна, метод скінчених об'ємів, неявну схему Ейлера, ітерації Пікара, енергетичні й міжшарові діагностики та просторово-часове уточнення. **Результати дослідження.** Для всіх конфігурацій забезпечено $R_{tot}=3,65315 \text{ м}^2 \cdot \text{К}/\text{Вт}$. Просторово-часова чисельна невизначеність становила 0,325166 % за порогу 0,5 %. Для пари В/О ефект змінив знак від $-0,905363\%$ у періодичному режимі до $+0,56625\%$ в імпульсному; відношення ефекту до невизначеності становили 79,2365 і 88,85. Для N періодичний ефект $-2,00484\%$ був розділений, а імпульсний $-0,282142\%$ залишився нерозділеним. Ефекти C5/C10 також не були чисельно розділені. **Висновки.** Контрольована стаціонарна еквівалентність разом з оцінкою чисельної невизначеності дає змогу розділяти режимно залежні ефекти; зміна знака встановлена лише для дослідженої пари В/О та використаних показників.

Ключові слова: нестационарна теплопровідність; багатошарова стінова конструкція; динамічна тепла реакція; термічний опір; порядок шарів; температурозалежні властивості; міжшаровий контактний опір; періодичне теплове навантаження; імпульсне теплове навантаження.

1. Вступ

Актуальність тематики. Підвищення енергоефективності будівель потребує не лише зменшення стаціонарних тепловтрат через огорожувальні конструкції, а й коректного врахування їх реакції на змінні температурні впливи. Чисельне моделювання дає змогу досліджувати такі процеси без зведення їх до одного інтегрального показника та створює основу для обґрунтованого вибору конструктивних рішень у задачах енергоефективності й модернізації будівель [1, 2]. Традиційне теплотехнічне оцінювання огорожувальної конструкції значною мірою ґрунтується на її сумарному термічному опорі R або коефіцієнті теплопередачі U . Ці характеристики визначають стаціонарну теплопередачу, проте не містять повної інформації про просторовий розподіл теплового опору й теплоємності. Тому конструкції з однаковим R або U можуть мати різну нестационарну реакцію. Така відмінність підтверджується як теоретичними дослідженнями багатошарових систем, так і порівняннями стінових конструкцій з однаковим коефіцієнтом теплопередачі [3, 4]. Зокрема, для систем з однаковим U у літературі показано відмінності часових температурних характеристик, пов'язані з теплофізичними властивостями та розподілом теплової маси. Отже, стаціонарна еквівалентність є необхідною умовою коректного порівняння конструкцій за тепловтратами, але недостатньою для висновку про еквівалентність їхньої динамічної поведінки. Особливого значення це набуває за періодичних

температурних коливань і короткочасних теплових впливів, коли результат визначається не лише величиною сумарного опору, а й розподілом теплофізичних властивостей у просторі та характерним часовим масштабом процесу.

2. Аналіз літературних джерел і визначення проблеми

Вплив внутрішньої будови багатошарових огорожень на їхню динамічну теплову поведінку досліджується протягом тривалого часу. Встановлено, що розташування теплоізоляції, розподіл теплової маси та взаємне положення шарів можуть істотно змінювати часові характеристики конструкції навіть без пропорційної зміни її стаціонарної ефективності [5, 6]. Сучасні дослідження розподілу теплоізоляційних і несучих шарів підтверджують, що за незмінної загальної кількості матеріалів їх просторове розташування впливає на запізнення теплової хвилі та ступінь її загасання [7, 8]. Водночас такі результати здебільшого отримано для одного заданого періоду температурних коливань або в межах конкретної оптимізаційної задачі. Зокрема, в одному з найближчих за постановкою досліджень перерозподіл однакової кількості теплоізоляційного та несучого матеріалу аналізувався за фіксованого 24-годинного періоду.

Динамічну теплопередачу можна досліджувати як у частотній області, так і безпосередньо в часовій. Частотні методи дають змогу визначати реакцію конструкції на гармонічне збудження та оцінювати загасання й часовий зсув теплової хвилі. Для різних

перестановок тих самих шарів уже показано суттєві відмінності таких характеристик за добового періодичного навантаження [9]. Поряд із цим методи функцій відгуку та теплового імпульсу дають змогу характеризувати нестационарну поведінку системи в часовій області, визначати характерні часові масштаби та оцінювати теплові властивості за реакцією на контрольоване збудження [10–12]. Таким чином, у літературі сформовано розвинені, але методично різні підходи до аналізу періодичної та перехідної теплової реакції.

Для періодичного режиму використовують коефіцієнти загасання, періодичну теплопередачу та часові або фазові характеристики [13, 14]. Їх величина залежить від співвідношення між характерним часом теплового процесу та періодом зовнішнього збудження. Водночас реакція на скінченний тепловий імпульс характеризує іншу часову постановку задачі. Сучасні методи, засновані на функціях відгуку, дають змогу пов'язувати імпульсні випробування зі стаціонарними та періодичними характеристиками [15, 16], однак наявність такого математичного зв'язку сама по собі не означає тотожності відносно поведінки різних багатошарових конфігурацій у цих режимах. Danovska та співавтори, зокрема, використовують єдиний імпульсний експеримент для одержання стаціонарних і 24-годинних періодичних характеристик, що підтверджує методичну цінність такого зв'язку, але не досліджує зміну ранжування стаціонарно еквівалентної пари за порядком шарів.

Додаткове ускладнення виникає через те, що реальні теплофізичні властивості матеріалів можуть залежати від температури. Нелінійні моделі нестационарної теплопровідності дають змогу враховувати таку залежність [17], а експериментальні дослідження матеріалів підтверджують зміну їх теплофізичних характеристик із температурою [18]. Проте ці залежності є матеріалоспецифічними й справедливими лише в досліджених температурних межах, тому їх не можна переносити на довільні матеріали або температурні діапазони. Відповідно, вплив температурної нелінійності на динамічну реакцію має оцінюватися окремо від впливу структури шарів.

Ще одним чинником є недосконалість теплового контакту на межах матеріалів. Міжшаровий контактний опір призводить до температурного стрибка за неперервності теплового потоку, а аналітичні розв'язки нестационарної теплопровідності дають можливість перевіряти чисельне відтворення таких умов [19, 20]. Однак звичайне введення додаткового контактного опору одночасно змінює сумарний стаціонарний опір системи. Унаслідок цього без спеціальної компенсації неможливо однозначно відокремити власне динамічний вплив межі контакту від наслідку зміни R_{total} .

Отже, попередні дослідження достатньо переконливо встановили окремі закономірності: відмінність динамічної реакції за однакових стаціонарних показників, вплив розташування й

розподілу шарів, залежність теплової реакції від характерного часу, роль температурозалежних властивостей та міжшарового опору. Ці окремі ефекти не є новизною цієї роботи. Невирішеною залишається інша методична проблема: *як роздільно оцінити внесок зазначених чинників у стаціонарно еквівалентних багатошарових конструкціях та простежити, чи зберігаються напрям і доказовість відповідного ефекту під час переходу від періодичного до скінченного імпульсного теплового навантаження.*

Для розв'язання цієї проблеми потрібна контрольована система порівнянь, у якій сумарний стаціонарний термічний опір зберігається незмінним, а порядок шарів, температурна залежність властивостей і міжшаровий контактний опір вводяться як окремі досліджувані чинники. Така постановка дає змогу уникнути змішування причин і відрізнити фізичний ефект від зміни стаціонарної теплопередачі. Водночас висновок про можливу відмінність поведінки між режимами повинен випливати з власних чисельних результатів.

3. Мета і завдання дослідження

Метою дослідження є чисельне визначення та міжрежимне порівняння впливу порядку шарів, температурозалежних теплофізичних властивостей і компенсованого міжшарового контактного опору на динамічну теплову реакцію стаціонарно еквівалентної багатошарової стінової конструкції за періодичного та скінченного імпульсного теплового навантаження.

Для досягнення поставленої мети необхідно розв'язати такі завдання:

1. Сформулювати контрольовану систему стаціонарно еквівалентних розрахункових сценаріїв В, О, N, C5 і C10, у якій зміна досліджуваного чинника не супроводжується зміною сумарного термічного опору конструкції.

2. Розробити та перевірити чисельну модель нестационарної теплопровідності багатошарової конструкції, яка дає змогу враховувати порядок шарів, температурну залежність теплофізичних властивостей і міжшаровий контактний опір, а також окремо визначати показники періодичної та імпульсної реакції.

3. Кількісно оцінити й зіставити вплив досліджуваних чинників у двох режимах теплового навантаження, не ототожнюючи фазове запізнення періодичного процесу з часовими характеристиками скінченного імпульсного відгуку.

4. Перевірити чисельну роздільність і стійкість установлених ефектів за даними просторово-часової збіжності та сформулювати лише ті висновки, величина яких перевищує відповідний рівень чисельної невизначеності.

Науково-методичний внесок роботи полягає не в повторному встановленні відомого впливу порядку шарів, теплової інерції, температуро-залежних властивостей або міжшарового опору окремо, а у контрольованій схемі чисельного розділення цих чинників за однакового сумарного стаціонарного

термічного опору та їх міжрежимного оцінювання. У межах цієї схеми для дослідженої пари В/О чисельними результатами встановлено стійку зміну знака відносного ефекту порядку шарів під час переходу від періодичного до імпульсного режиму. Цей результат стосується саме дослідженої пари конфігурацій і заданих режимів навантаження та не розглядається як універсальна закономірність для довільних багатошарових конструкцій.

4. Матеріали і методи дослідження

4.1. Об'єкт дослідження та математична постановка

Розглянуто одновимірний нестационарний теплоперенос у тришаровій плоскій стіновій конструкції загальною товщиною 0,34 м. У базовій конфігурації шари розташовано в послідовності А–В–С: шар А – легкий бетон на керамзитовому заповнювачі завтовшки 0,10 м, шар В – теплоізоляційний матеріал завтовшки 0,12 м, шар С – цегла завтовшки 0,12 м. Вихідні теплофізичні властивості матеріалів отримано з нормалізованих джерельних записів бази даних; геометрія конструкції є параметром сформованої розрахункової постановки. Саме ці записи та товщини закріплено у конфігурації моделі. Координату x відрховано від зовнішньої поверхні до внутрішньої. Додатний напрям теплового потоку прийнято від зовнішнього середовища до приміщення. Теплоперенос у кожному шарі описується рівнянням нестационарної теплопровідності:

$$\rho_i(T)c_{p,i}(T)\frac{\partial T_i}{\partial t} = \frac{\partial}{\partial x}\left[k_i(T)\frac{\partial T_i}{\partial x}\right], \quad (1)$$

де $T_i = T_i(x, t)$ – температура i -го шару, °C; x – координата, м; t – час, с; k_i – коефіцієнт теплопровідності, Вт/(м·K); $c_{p,i}$ – питома теплоємність, Дж/(кг·K); ρ_i – густина, кг/м³.

Рівняння (1) виражає локальний закон збереження енергії для теплопровідності. Загальна програмна реалізація допускає температурну залежність теплофізичних характеристик, однак у фактично використаному сценарії N температурні закони активовано лише там, де вони мають підтвердження з джерел. Для матеріалу шару А використовуються залежності $k(T)$ і $c_p(T)$, отримані з нормалізованих експериментальних даних [17, 18]. Густина для цієї конфігурації використовується як джерельно визначене опорне значення; окремий закон $\rho(T)$ у поточній моделі для цього матеріалу не активований. Таким чином, рівняння (1) подано в загальній формі, але в поточному дослідженні температурна нелінійність не вводиться для параметрів, для яких відсутня джерельна залежність.

Локальний тепловий потік визначається законом Фур'є

$$q_i = -k_i(T)\frac{\partial T_i}{\partial x}, \quad (2)$$

де q_i – густина теплового потоку, Вт/м².

На зовнішній та внутрішній поверхнях задаються конвективні граничні умови Робіна:

$$q_{(out)} = h_{(out)}[T_e(t) - T_{(s,out)}(t)],$$

$$q_{(in)} = h_{(in)}[T_{(s,in)}(t) - T_{(in)}], \quad (3)$$

де $h_{out} = 10$ Вт/(м²·K) і $h_{in} = 7$ Вт/(м²·K) – коефіцієнти конвективного теплообміну; $T_e(t)$ – температура зовнішнього повітря; $T_{s,out}$ і $T_{s,in}$ – температури зовнішньої та внутрішньої поверхонь; $T_{in} = 20^\circ\text{C}$ – температура внутрішнього повітря.

Для ідеального контакту температура та тепловий потік на межі суміжних шарів є неперервними. Для недосконалого контакту з термічним опором $R_{c,jR_{\{c,j\}}R_{c,j}}$ виконується

$$q_j^- = q_j^+ = q_j, T_j^- - T_j^+ = q_j R_{(c,j)}, \quad (4)$$

де $R_{c,j}$ має розмірність м²·K/Вт. За $R_{c,j} = 0$ умова переходить у випадок ідеального контакту. Таке представлення узгоджується з аналітичними постановками нестационарної теплопровідності через недосконалі межі [19, 20]. У самій чисельній реалізації додатково контролюється залишок співвідношення $\Delta T - qR_c$ для кожної міжшарової межі.

4.2. Контроль стаціонарної еквівалентності та система сценаріїв

Для відокремлення динамічних ефектів від звичайної зміни теплоізоляційної здатності всі основні конфігурації порівнювали за однакового контрольного сумарного стаціонарного термічного опору. При опорній температурі $T_{ref} = 20^\circ\text{C}$ $T_{ref} = 20^\circ\text{C}$ він визначається як

$$R_{tot} \frac{1}{h_{out}} + \sum_{i=1}^n \frac{L_i}{k_i(T_{ref})} + \sum_{j=1}^{n-1} R_{c,j} + \frac{1}{h_{in}}, \quad (5)$$

де L_i – товщина i -го шару, м; інші позначення наведено вище.

Для всіх досліджуваних конфігурацій В, О, N, С5 і С10 отримано $R_{tot} = 3,65315$ м²·K/Вт, а зафіксована відносна похибка стаціонарної еквівалентності дорівнює нулю.

Порівнювали п'ять причинно розділених конфігурацій:

В – базова конфігурація А–В–С зі сталими теплофізичними властивостями та ідеальними контактами;

О – ті самі три шари у зворотній послідовності С–В–А; змінюється лише просторовий порядок шарів;

N – базова послідовність А–В–С та ідеальні контакти, але для шару А активуються джерельно обґрунтовані температурні залежності $k(T)$ і $c_p(T)$;

С5 – базова послідовність А–В–С, у якій 5 % вихідного об'ємного термічного опору контролювано перерозподілено до двох міжшарових контактів;

С10 – аналогічна конфігурація з перерозподілом 10 % об'ємного термічного опору.

У С5 і С10 міжшаровий опір не додається поверх вихідного R_{tot} . Загальний заданий R_c симетрично розподіляється між двома фізичними

межами, а опір шару B відповідно зменшується шляхом масштабування його теплопровідності без зміни товщини. Після компенсації модель повторно перевіряє рівність сумарного опору заданому базовому значенню. Контактні опори C5/C10 є спеціально сформованими параметрами

контрольованого чисельного експерименту, а не вимірними характеристиками реальної конструкції. Ця побудова дає змогу відокремити вплив локалізації опору на міжшарових межах від тривіального ефекту збільшення загального стаціонарного термічного опору.

Таблиця 1 – Теплофізичні характеристики та стаціонарна еквівалентність контрольованих багатшарових конфігурацій

Конф.	Порядок шарів	Модель в-востей	Частка R_c , %	$R_{bulk}, M^2K / Bm$	$R_c, M^2K / Bm$	$R_{tot}, M^2K / Bm$	$C_A, кДж / (м^2K)$
B	A-B-C	Сталі	0	3.4103	0	3.653	354.6
O	C-B-A	Сталі	0	3.4103	0	3.653	354.6
N	A-B-C	Температурозалежні	0	3.4103	0	3.653	354.6
C5	A-B-C	Компенсовані сталі	5	3.2398	0.2	3.653	354.6
C10	A-B-C	Компенсовані сталі	10	3.0693	0.3	3.653	354.6

4.3. Режими теплового навантаження та показники динамічної реакції

Кожну з п'яти конфігурацій розраховували у двох режимах, тобто загальний набір містив десять основних розрахункових сценаріїв. У періодичному режимі температура зовнішнього повітря задається гармонічним законом:

$$T_e(t) = \bar{T}_e + A_e \sin \left[\frac{2\pi}{P} (t - t_\phi) \right], \quad (6)$$

де $P = 24 \text{ год}$ – період; $\bar{T}_e = 30^\circ\text{C}$ – середня температура; $A_e = 10 \text{ K}$ – амплітуда; $t_\phi = 0$ – часовий зсув зовнішнього сигналу. Таким чином, задана зовнішня температура змінюється в межах $20\text{--}40^\circ\text{C}$. Ці параметри безпосередньо закріплені канонічною конфігурацією розрахункового пакета. Основними температурними характеристиками усталеного періодичного режиму є коефіцієнт загасання:

$$f_T = \frac{A_{(s,in)}}{A_e},$$

$$\Delta t_\phi = \frac{[(\varphi_e - \varphi_{(s,in)}) \bmod 2\pi]}{\omega}, \quad \omega = \frac{2\pi}{P}, \quad (7)$$

де $A_{s,in}$ – амплітуда температури внутрішньої поверхні, K ; φ_e і $\varphi_{s,in}$ – фази зовнішнього температурного сигналу та температури внутрішньої поверхні; ω – кутова частота, рад/с ; Δt_ϕ – фазове запізнення, s або год після перерахунку. Для визначення гармонічних складових використовувалось узгоджене з відомими методами періодичного аналізу представлення [13].

Періодичний розрахунок продовжується до досягнення усталеного циклу. Після щонайменше трьох циклів порівнюються температурна історія внутрішньої поверхні та внутрішній тепловий потік у двох послідовних циклах. Для обох нормованих максимальних відмінностей встановлено критерій 10^{-4} ; гранична кількість циклів становить 30. Другий режим представлено скінченим трикутним температурним імпульсом. Перед його початком конструкція приводиться до відповідного

стаціонарного стану. Базова зовнішня температура становить 20°C , початок імпульсу – через 1 год , амплітуда – 20 K , тривалість лінійного наростання – 3 год , тривалість лінійного спаду – 3 год . Отже, максимальна зовнішня температура також становить 40°C , що зберігає N -сценарій у межах $0\text{--}50^\circ\text{C}$, для яких використані температурозалежні властивості мають обґрунтування у відповідних джерелах. Максимальне вікно спостереження становить 240 год . Основною амплітудною характеристикою імпульсного режиму є коефіцієнт передавання піку

$$r_{(peak)} = \frac{\max_t [T_{(s,in)}(t) - T_{(s,in,0)}]}{\Delta T_p}, \quad (8)$$

де $T_{s,in,0}$ – температура внутрішньої поверхні до початку імпульсу, $^\circ\text{C}$; $\Delta T_p = 20 \text{ K}$ – амплітуда зовнішнього температурного імпульсу.

Додатково визначаються часова затримка максимуму та затримка центроїда температурного відгуку. Їх розраховують відносно відповідних характеристик вхідного імпульсу. Вони характеризують перехідний часовий процес і не трактуються як фазове запізнення. У програмній реалізації коефіцієнт передавання піку визначається відносно базового передімпульсного стану, а центроїд обчислюється за додатною частиною відповідної реакції. Методична доцільність імпульсного опису теплової реакції та його зв'язку з функціями відгуку обґрунтована в [15, 16].

Повнота імпульсного розрахунку перевіряється за залишковою реакцією наприкінці вікна спостереження. Відношення залишкового температурного відхилення та залишкового теплового потоку до їхніх відповідних максимальних значень має не перевищувати 10^{-3} . Якщо ця умова не виконується, розрахунок не приймається як завершений.

4.4. Чисельний метод і контроль збіжності

Рівняння теплопровідності дискретизується консервативним методом скінчених об'ємів на

сітці, межі контрольних об'ємів якої узгоджені з межами матеріальних шарів. Такий підхід забезпечує безпосередній баланс теплових потоків та дозволяє вводити міжшаровий опір як складову опору відповідної грані.

Інтегрування в часі здійснюється неявною схемою Ейлера. Для сценарію з температурозалежними властивостями на кожному часовому кроці виконуються ітерації Пікара: теплофізичні коефіцієнти уточнюються за поточним температурним полем, після чого повторно формується і розв'язується лінійна система. Максимальна кількість ітерацій становить 50; абсолютний критерій збіжності – 10^{-7} °C, відносний – 10^{-8} ; коефіцієнт релаксації дорівнює 1.

Метод скінченних об'ємів, неявна схема Ейлера та ітерації Пікара використовуються як стандартні чисельні засоби і не розглядаються як складові наукової новизни. Для основних розрахунків застосовано 40 контрольних об'ємів на кожний шар і крок часу 300 с. Перевірка залежності результатів від дискретизації охоплювала чотири рівні:

40 комірок на шар, $\Delta t = 300$ с – базовий;

80 комірок на шар, $\Delta t = 300$ с – просторове уточнення;

40 комірок на шар, $\Delta t = 150$ с – часове уточнення;

80 комірок на шар, $\Delta t = 150$ с – комбіноване уточнення.

Сумарна оцінена просторово-часова чисельна невизначеність фінального пакета становить $U_{grid+time} = 0,325166\%$, що нижче наперед прийнятого критерію 0,5 %.

Для центрального порівняння В/О додатково виконано цілове уточнення, в якому контролювали не лише величину, а й збереження знака відносного ефекту на всіх чотирьох рівнях дискретизації.

Чисельна діагностика охоплювала також дискретний енергетичний баланс і виконання міжшарових умов. Максимально допустимий відносний залишок енергетичного балансу встановлено на рівні 10^{-5} . Для меж шарів контролювалися тепловий потік, реконструйований температурний стрибок та залишок умови $\Delta T = qR_c$.

4.5. Верифікація та межі доказовості моделі

Перевірку чисельної реалізації виконано незалежно від подальшої інтерпретації основних сценаріїв. Верифікаційний ланцюг містив три рівні. В1 перевіряв лінійну багатошарову нестационарну теплопровідність і коректність базової реалізації. В2 був призначений для перевірки недосконалого міжшарового контакту, зокрема неперервності теплового потоку й температурного стрибка за скінченного R_c , з використанням аналітичних постановок [19, 20]. В3 перевіряв помірно нелінійну нестационарну теплопровідність із температурозалежними властивостями [17]. Загальний результат верифікаційного ланцюга має статус *PASS WITH RESTRICTIONS*. Обмеження В2 полягає в тому, що опубліковане джерело зберігає інформацію про похибки та необхідні фізико-математичні співвідношення, але не містить повного набору

точкових еталонних температур для незалежного відтворення всієї температурної історії. Для В3 опубліковано межу розбіжності з розрахунком методом скінченних елементів менше 4%, однак повний набір табличних для нового точкового зіставлення відсутній. У цій роботі термін *верифікація* означає перевірку правильності чисельної реалізації відносно незалежних аналітичних, чисельних і опублікованих еталонних задач. *Валідація* означала б кількісне підтвердження адекватності повної фізичної моделі за незалежними експериментальними даними для відповідного об'єкта й режиму. *Калібрування* передбачало б підбір параметрів моделі за експериментальними даними. Калібрування поточної моделі не виконувалося, а наявні дослідження повномасштабних стінових комірок і двовимірні випробування в гарячій камері [21] використовуються лише як зовнішні свідчення фізичної правдоподібності, а не як калібрувальні дані. Отже, модель вважається *чисельно верифікованою з обмеженнями*, але не повністю експериментально валідованою.

5. Результати дослідження

5.1 Контроль стаціонарної еквівалентності

Для конфігурацій В, О, N, C5 і C10, параметри яких наведено в табл. 1, підтверджено однаковий сумарний стаціонарний термічний опір $R_{tot} = 3,65315$ м²·K/Вт; відносна похибка стаціонарної еквівалентності дорівнювала нулю. У C5 і C10 введення міжшарового контактного опору компенсувалося відповідним зменшенням об'ємної складової опору, тому загальний R_{tot} залишався незмінним.

5.2 Періодична тепла відповідь

Для базової конфігурації В коефіцієнт температурного загасання становив $f_T = 0,00754845$, а фазове запізнення внутрішньої поверхні $\Delta t_{pf} = 10,5805$ год.

Відносно В зміна основної періодичної метрики становила $-0,905363$ % для О та $-2,00484$ % для N; для C5 і C10 вона була значно меншою: $-0,0167601$ % і $-0,0395018$ % відповідно.

Часові відхилення О і N від базової відповіді наведено на рис. 1.

5.3 Відповідь на скінченний тепловий імпульс

Для В коефіцієнт передавання температурного піку становив $r_{peak} = 0,00545139$, затримка максимуму $-10,000$ год, затримка центроїда $-17,9752$ год. Відносно В зміна r_{peak} становила $+0,56625$ % для О та $-0,282142$ % для N; для C5 і C10 вона дорівнювала лише $+0,00250643$ % і $+0,00444627$ %.

Часову структуру відповіді О і N наведено на рис. 2.

Величини Δt_{peak} і Δt_c є часовими затримками скінченного імпульсного процесу, а не фазовим запізненням.

5.4 Режимно залежна зміна знака ефекту В/О

Для О встановлено зміну знака відносного ефекту: $-0,905363$ % у періодичному режимі та $+0,56625$ % за скінченного імпульсу (рис. 3).

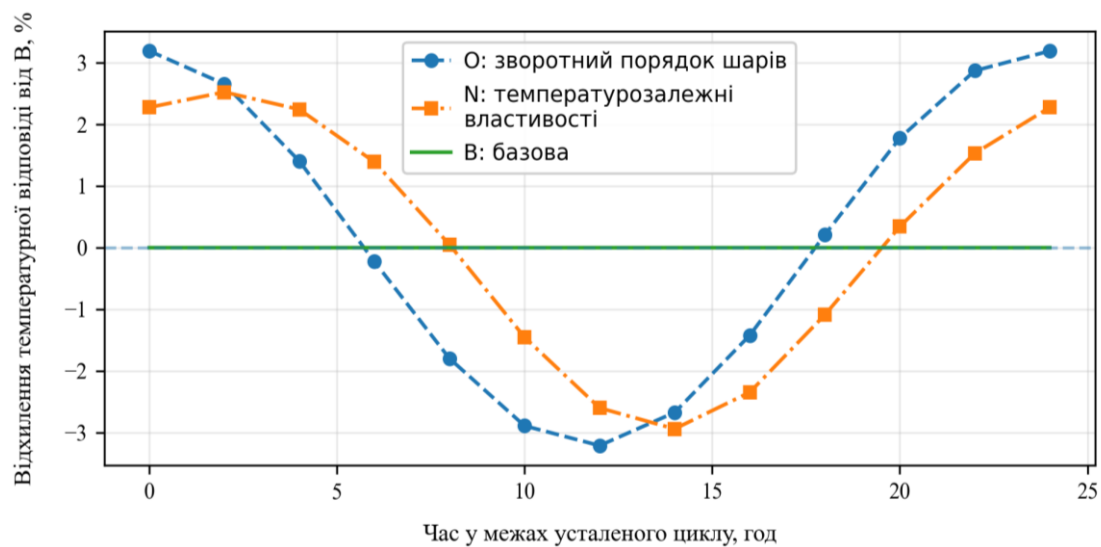


Рис. 1. Вплив порядку шарів і температурозалежних властивостей на періодичну теплову відповідь відносно базової конфігурації

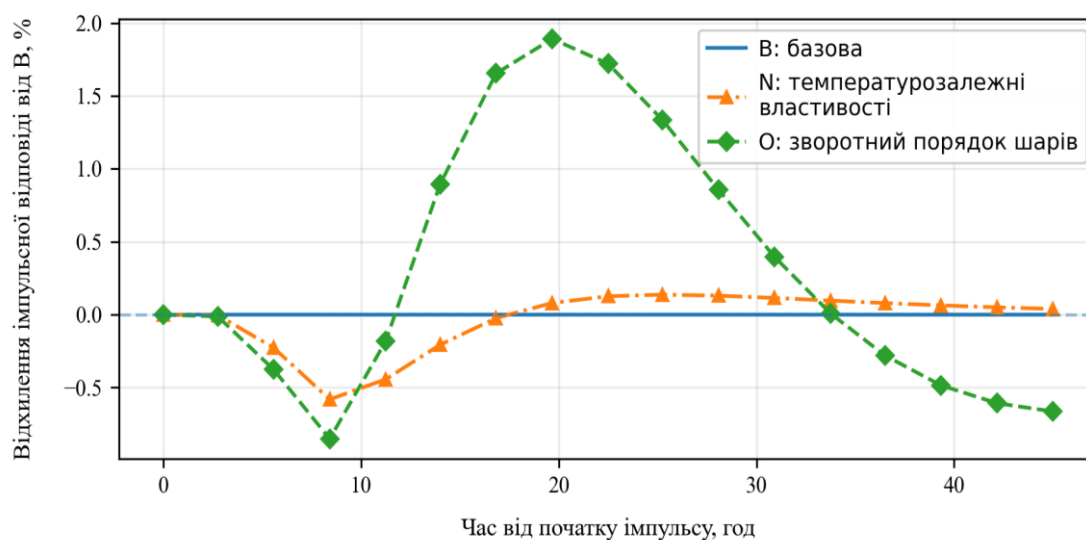


Рис. 2. Вплив порядку шарів і температурозалежних властивостей на теплову відповідь за скінченного імпульсного збудження

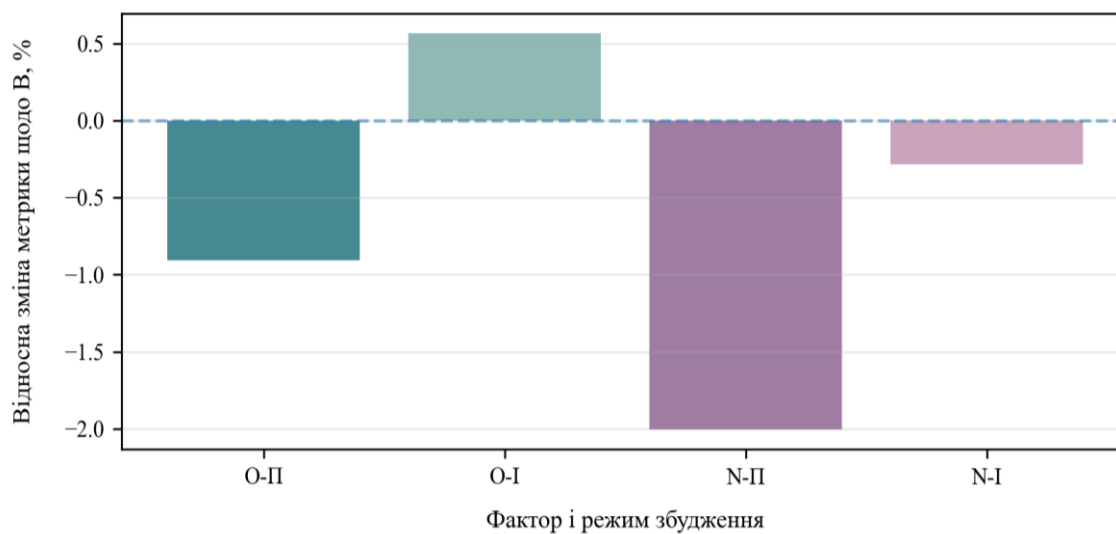


Рис. 3. Режимно залежний вплив порядку шарів і температурозалежних властивостей на основні метрики динамічної теплової відповіді

Цільове просторово-часове уточнення підтвердило стійкість знака; відповідні відношення $|Δ|/u$ становили 79,2365 і 88,85. Отже, обидва ефекти чисельно розділені.

Результат стосується лише дослідженої пари В/О, заданих режимів і відповідних їм основних метрик.

5.5 Вплив температурозалежних властивостей N

Для N періодичний ефект становив $-2,00484\%$ при $|Δ|/u = 6,16561$ і був чисельно розділеним. За скінченного імпульсу номінальний ефект $-0,282142$

% мав $|Δ|/u = 0,867687 < 1$, тому залишився чисельно нерозділеним.

5.6 Результат для компенсованого контактного опору C5/C10

Для C5 відносні зміни основних періодичної та імпульсної метрик становили $-0,0167601\%$ і $+0,00250643\%$, а для C10 – $-0,0395018\%$ і $+0,00444627\%$. Відповідні $|Δ|/u$ перебували в межах $0,0077 - 0,1215$, тобто були значно меншими за одиницю. Отже, C5/C10 дали слабкі чисельно нерозділені результати. Багатометрикову чутливість показано на рис. 4.

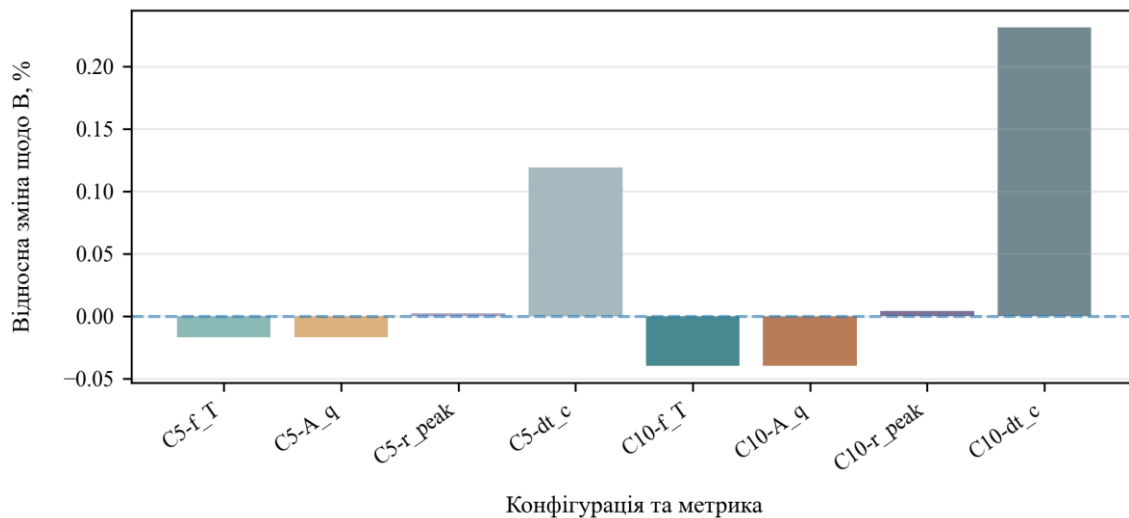


Рис. 4. Чутливість нестационарної теплової відповіді до компенсованого міжшарового контактного опору за незмінного загального стаціонарного термічного опору

5.7 Підсумок чисельної роздільності результатів

Сумарна просторово-часова чисельна невизначеність становила $0,325166\%$, що нижче прийнятого порога $0,5\%$. За критерієм $|Δ|/u > 1$

обидва ефекти O були сильно розділеними, періодичний ефект N – розділеним, тоді як імпульсний ефект N та всі ефекти C5/C10 залишилися нерозділеними. Кількісні показники наведено в табл. 2.

Таблиця 2 – Кількісні показники

Конф.	$f_T, 1$	$Δt_p, \text{год}$	$r_{peak}, 1$	$Δt_{peak}, \text{год}$	$Δt_c, \text{год}$	$ Δ /u (Π), 1$	$ Δ /u (I), 1$
В	0.007548	10.581	0.005451	10	17.975		
О	0.00748	10.697	0.005482	10.25	17.665	79.24	88.85
N	0.007397	10.651	0.005436	10.083	18.014	6.17	0.87
C5	0.007547	10.604	0.005452	10.083	17.997	0.05	0.01
C10	0.007545	10.626	0.005452	10.083	18.017	0.12	0.01

6. Обговорення результатів дослідження

Контроль однакового сумарного стаціонарного термічного опору є принциповою умовою проведеного порівняння. Попередні дослідження вже показали, що однакові R або U не забезпечують тотожної нестационарної реакції, а розташування теплоізоляції та теплової маси впливає на динамічні характеристики стін [3–8]. Тому новизна цієї роботи полягає не в повторному встановленні впливу

порядку шарів, а в його відокремленні від зміни загальної теплоізоляційної здатності. Та сама логіка застосована до C5/C10: контактний опір не додавався до R_{tot} , а компенсувався зменшенням об'ємної складової опору.

Для дослідженої пари В/О встановлено різний напрям відносного ефекту в періодичному та імпульсному режимах. Це узгоджується із загальним уявленням про залежність динамічної реакції від розташування шарів та характерного часу теплового

процесу [9–15], але сама зміна знака є результатом власного чисельного експерименту, а не висновком із літератури. Періодичний режим оцінювався за f_T , а скінченний імпульс – за r_{peak} ; тому йдеться про порівняння режимно відповідних показників. Затримка піку та центроїда імпульсної відповіді не є фазовим запізненням.

Зміну знака B/O не можна поширювати на довільні багат шарові конструкції. Вона встановлена для конкретної тришарової пари, 24-годинного періодичного збудження, заданого трикутного імпульсу та використаних метрик. Дослідження частотних, початкових перехідних і режимно залежних процесів також показують чутливість теплової поведінки до структури та часових умов [14, 24–26], однак не встановлюють універсального правила зміни ранжування.

Для температурозалежної конфігурації N важливим є різний доказовий статус, а не лише різниця номінальних значень. Періодичний ефект перевищив оцінену чисельну невизначеність і був розділеним, тоді як імпульсний залишився нижче відповідного порога. Джерела [17, 18] підтверджують фізичну обґрунтованість температурної залежності теплофізичних властивостей, але не визначають величину ефекту в дослідженій конструкції. Тому нерозділений імпульсний результат не означає відсутності фізичного впливу, а лише неможливість надійно відокремити його від чисельної невизначеності за прийнятої постановки.

Аналогічно слід трактувати $C5/C10$. Теорія недосконалих контактів підтверджує фізичне значення міжшарового термічного опору [19, 20], проте в дослідженому компенсованому діапазоні його додатковий динамічний ефект залишився чисельно нерозділеним. Це негативний, але змістовний результат: ненульова номінальна різниця не повинна автоматично перетворюватися на фізичний висновок. Саме включення чисельної невизначеності до правила інтерпретації дозволило розмежувати стійкий B/O -ефект, розділений періодичний N -ефект та нерозділені N -pulse і $C5/C10$ результати.

Експериментальні та чисельно-експериментальні дослідження реальних огорожень підтверджують фізичну значущість теплової інерції, коефіцієнта загасання й часових затримок [21–23], однак не є прямою калібровкою цієї моделі. Її область застосовності обмежена одновимірною сухою теплопровідністю, конкретними матеріалами й температурними законами, дослідженою B/O -парою, заданими режимами збудження та сценаріями $C5/C10$. Не враховано вологоперенос і дво- чи тривимірні теплові мости. Крім того, перевірочні задачі $B2$ і $B3$ мають обмеження, пов'язані з доступністю вихідних еталонних даних, тому загальний статус залишається *PASS WITH RESTRICTIONS*, а *whole-cell* і *2D hot-box* результати слід розглядати лише як зовнішні свідчення фізичної правдоподібності. Основним методичним внеском є контрольована схема розділення чинників за

незмінного R_{tot} із режимно специфічним та орієнтованим на чисельну невизначеність оцінюванням їхнього впливу.

7. Висновки

1. Сформовано контрольовану систему стаціонарно еквівалентних конфігурацій B , O , N , $C5$ і $C10$ із однаковим сумарним термічним опором $R_{tot} = 3,65315 \text{ m}^2\text{K/W}$. Це дозволило роздільно оцінювати вплив порядку шарів, температурозалежних властивостей та компенсованого міжшарового термічного опору без змішування цих чинників зі зміною загальної стаціонарної теплоізоляційної здатності.

2. Для дослідженої пари B/O встановлено чисельно стійку режимно залежну зміну знака ефекту порядку шарів. Відносна зміна основної періодичної метрики становила $-0,905363\%$, тоді як для скінченного імпульсу вона становила $+0,56625\%$. Відношення ефекту до чисельної невизначеності дорівнювало відповідно 79,2365 та 88,85, що підтверджує стійкість обох результатів. Цей висновок стосується лише дослідженої B/O -пари, заданих режимів навантаження та відповідних їм метрик.

3. Урахування температурозалежних теплофізичних властивостей дало різний доказовий результат у двох режимах. Для періодичного навантаження відносний ефект N становив $-2,00484\%$ при $|A|/u = 6,16561$, тому його визнано чисельно розділеним. Для скінченного імпульсу номінальна зміна $-0,282142\%$ мала $|A|/u = 0,867687$, отже цей ефект залишився нерозділеним.

4. Для компенсованого міжшарового термічного опору сильного додаткового динамічного ефекту в дослідженому діапазоні не встановлено. Для $C5$ відносні зміни основних періодичної та імпульсної метрик становили приблизно $-0,0167601\%$ та $+0,00250643\%$, для $C10$ – $-0,0395018\%$ та $+0,00444627\%$; усі відповідні відношення $|A|/u$ були значно меншими за одиницю. Тому $C5/C10$ класифіковано як слабкі та чисельно нерозділені результати.

5. Сумарна просторово-часова чисельна невизначеність становила $0,325166\%$, що нижче прийнятого порога $0,5\%$. Загальний статус перевірки чисельної реалізації – *PASS WITH RESTRICTIONS* через обмеження вихідних даних перевірочних задач $B2$ і $B3$. Отриманий рівень перевірки не є повною незалежною експериментальною валідацією моделі.

6. Основний науково-методичний внесок роботи полягає у *контрольованій схемі розділення та оцінювання структурних і теплофізичних чинників за незмінного сумарного стаціонарного термічного опору, із застосуванням режимно відповідних динамічних показників та явного порівняння величини ефекту з чисельною невизначеністю*.

Така схема дає змогу однаково коректно фіксувати як розділені ефекти, так і негативні або нерозділені результати, не перетворюючи номінальні числові відмінності на необґрунтовані фізичні висновки.

Конфлікт інтересів

Автори декларують, що не мають конфлікту інтересів стосовно даного дослідження, в тому числі фінансового, особистісного характеру, авторства чи іншого характеру, що міг би вплинути на дослідження та його результати, представлені в даній статті.

Фінансування

Дослідження проводилося без фінансової підтримки.

Доступність даних

Дані та матеріали розрахунків, що підтверджують результати цього дослідження, можуть бути надані авторами за обґрунтованим запитом.

Використання засобів штучного інтелекту

Автори підтверджують, що не використовували технології штучного інтелекту при створенні представленої роботи.

СПИСОК ЛІТЕРАТУРИ

1. Ichanska, N.V., Usenko, I.S., Usenko, D.V., and Pinchuk, N.M. (2025), "Numerical Modeling of the Impact of Complex Loads on Building Materials", *Modern Construction and Architecture*, vol. 14, pp. 33–45, doi: <https://doi.org/10.31650/2786-6696-2025-14-33-45>
2. Усенко І.С., Усенко Д.В. Комплексні ресурсощадні та енергоефективні технології для відновлення та модернізації захисних споруд цивільного захисту. *Екологія. Довкілля. Енергозбереження – 2025: колективна монографія* / за ред. О.Е. Ілляш. Полтава: Національний університет «Полтавська політехніка імені Юрія Кондратюка», 2025. С. 157–168, available at: <https://reposit.nupp.edu.ua/item/19423>
3. Kossecka, E. (1996), "The Effect of Structure on Dynamic Thermal Characteristics of Multilayer Walls", *Archives of Civil Engineering*, vol. 42, no. 3, pp. 351–369.
4. Yüksel, A., Arıcı, M., and Karabay, H. (2021), "Comparison of thermal response times of historical and modern building wall materials", *Journal of Thermal Engineering*, vol. 7, no. 6, pp. 1506–1518, doi: <https://doi.org/10.18186/thermal.991093>
5. Bond, D.E.M., Clark, W.W., and Kimber, M. (2013), "Configuring wall layers for improved insulation performance", *Applied Energy*, vol. 112, pp. 235–245, doi: <https://doi.org/10.1016/j.apenergy.2013.06.024>
6. Al-Sanea, S.A., and Zedan, M.F. (2011), "Improving thermal performance of building walls by optimizing insulation layer distribution and thickness for same thermal mass", *Applied Energy*, vol. 88, no. 9, pp. 3113–3124, doi: <https://doi.org/10.1016/j.apenergy.2011.02.036>
7. Gori, P., Guattari, C., Evangelisti, L., and Asdrubali, F. (2016), "Design criteria for improving insulation effectiveness of multilayer walls", *International Journal of Heat and Mass Transfer*, vol. 103, pp. 349–359, doi: <https://doi.org/10.1016/j.ijheatmasstransfer.2016.07.077>
8. Lu, Y., Hu, J., and Zhong, K. (2022), "Optimization of insulation layer location and distribution considering maximum time lag and damping factor", *Case Studies in Thermal Engineering*, vol. 30, art. 101766, doi: <https://doi.org/10.1016/j.csite.2022.101766>
9. Ma, L., Long, E., and Liu, Q. (2021), "An frequency domain analysis method of thermal parameters unsteady-state detection of building wall", *Journal of Asian Architecture and Building Engineering*, vol. 20, no. 6, pp. 718–728, doi: <https://doi.org/10.1080/13467581.2020.1838292>
10. Rasooli, A., and Itard, L. (2019), "In-situ rapid determination of walls' thermal conductivity, volumetric heat capacity, and thermal resistance, using response factors", *Applied Energy*, vol. 253, art. 113539, doi: <https://doi.org/10.1016/j.apenergy.2019.113539>
11. Rasooli, A., and Itard, L. (2019), "Properties of the triangular excitation pulse and the 3D heat transfer effects in the excitation pulse method", *E3S Web of Conferences*, vol. 111, art. 04018, doi: <https://doi.org/10.1051/e3sconf/201911104018>
12. Petojević, Z., Gospavić, R., and Todorović, G. (2018), "Estimation of thermal impulse response of a multi-layer building wall through in-situ experimental measurements in a dynamic regime with applications", *Applied Energy*, vol. 228, pp. 468–486, doi: <https://doi.org/10.1016/j.apenergy.2018.06.083>
13. Yuan, J. (2018), "Impact of Insulation Type and Thickness on the Dynamic Thermal Characteristics of an External Wall Structure", *Sustainability*, vol. 10, no. 8, art. 2835, doi: <https://doi.org/10.3390/su10082835>
14. Lu, Y., Wang, L., He, J., Yang, R., and Yuan, L. (2024), "Dimensionless resolutions for heat flux decrement factor and time lag of the wall during cyclic variations in outdoor air temperature", *Case Studies in Thermal Engineering*, vol. 60, art. 104678, doi: <https://doi.org/10.1016/j.csite.2024.104678>
15. Danovska, M., Cassol, D., Giongo, I., and Prada, A. (2026), "Experimental assessment of wall thermal properties using an integrated response factor approach", *Energy and Buildings*, vol. 350, art. 116661, doi: <https://doi.org/10.1016/j.enbuild.2025.116661>
16. Peavy, B.A. (1978), *Determination and Verification of Thermal Response Factors for Thermal Conduction Applications*, NBSIR 77-1405, National Bureau of Standards, Washington, DC. <https://doi.org/10.6028/NBS.IR.77-1405>
17. Wang, H., Chen, X., Koenders, E., Dai, Y., Huang, X., Ai, Q., and Yuan, Y. (2023), "Transient Nonlinear Heat Conduction in Concrete Structures: A Semi-Analytical Approach", *Entropy*, vol. 25, no. 4, art. 583, doi: <https://doi.org/10.3390/e25040583>
18. Shafigh, P., Hafez, M.A., Che Muda, Z., Beddu, S., Zakaria, A., and Almkahal, Z. (2022), "Influence of Different Ambient Temperatures on the Thermal Properties of Fiber-Reinforced Structural Lightweight Aggregate Concrete", *Buildings*, vol. 12, no. 6, art. 771, doi: <https://doi.org/10.3390/buildings12060771>
19. Chiba, R. (2018), "An Analytical Solution for Transient Heat Conduction in a Composite Slab with Time-Dependent Heat Transfer Coefficient", *Mathematical Problems in Engineering*, vol. 2018, art. 4707860, doi: <https://doi.org/10.1155/2018/4707860>

20. D'Alessandro, G., and de Monte, F. (2020), "Multi-Layer Transient Heat Conduction Involving Perfectly-Conducting Solids", *Energies*, vol. 13, no. 24, art. 6484, doi: <https://doi.org/10.3390/en13246484>
21. del Coz-Díaz, J.J., Álvarez-Rabanal, F.P., Alonso-Martínez, M., and Martínez-Martínez, J.E. (2021), "Thermal Inertia Characterization of Multilayer Lightweight Walls: Numerical Analysis and Experimental Validation", *Applied Sciences*, vol. 11, no. 11, art. 5008, doi: <https://doi.org/10.3390/app11115008>
22. Toure, P.M., Dieye, Y., Gueye, P.M., Sambou, V., Bodian, S., and Tiguampo, S. (2020), "Experimental Determination of Time Lag and Decrement Factor", *Case Studies in Construction Materials*, vol. 12, art. e00298, doi: <https://doi.org/10.1016/j.cscm.2019.e00298>
23. Belarbi, Y.E., Poullain, P., Othmen, I., Issaadi, N., Bonnet, S., and Leklou, N. (2025), "Experimental evaluation of the time lag and decrement factor of a raw earth wall for various climates: Comparison of different methods", *Journal of Building Engineering*, vol. 113, art. 113996, doi: <https://doi.org/10.1016/j.jobbe.2025.113996>
24. Lu, Y., He, J., Gu, Y., Zhou, H., and Yuan, L. (2024), "Optimizing wall structure for enhanced thermal performance: A layer-specific dimensionless parameter analysis in multi-layer walls", *Journal of Building Engineering*, vol. 94, art. 109889, doi: <https://doi.org/10.1016/j.jobbe.2024.109889>
25. Xu, L., Ding, Y., and Du, C. (2024), "Thermal performance assessment of exterior building walls under intermittent air-conditioning operation in China's hot summer and cold winter zone", *Journal of Building Engineering*, vol. 95, art. 110204, doi: <https://doi.org/10.1016/j.jobbe.2024.110204>
26. Al-Sanea, S.A., and Zedan, M.F. (2001), "Effect of Insulation Location on Initial Transient Thermal Response of Building Walls", *Journal of Thermal Envelope and Building Science*, vol. 24, no. 4, pp. 275–300, doi: <https://doi.org/10.1106/07E7-FGCJ-MFF7-974W>

Received (Надійшла) 21.07.2026

Accepted for publication (Прийнята до друку) 09.09.2026

Publication date (Дата публікації) 25.09.2026

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Усенко Дмитро Валерійович – доктор філософії (PhD), доцент, доцент кафедри вищої і прикладної математики та фізики, Національний університет «Полтавська політехніка імені Юрія Кондратюка», Полтава, Україна;

Dmytro Usenko – PhD, Associate Professor, Associate Professor of the Department of Higher and Applied Mathematics and and Physics, National University "Yuri Kondratyuk Poltava Polytechnic", Poltava, Ukraine;
e-mail: usenko_phd@ukr.net; ORCID Author ID: <https://orcid.org/0000-0001-7133-0638>;
Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57202630040>.

Біба Володимир Васильович – кандидат технічних наук, доцент, доцент кафедри менеджменту і логістики, Національний університет «Полтавська політехніка імені Юрія Кондратюка», Полтава, Україна;

Volodymyr Byba – PhD in Technical Sciences, Associate Professor, Associate Professor of the Department of Management and Logistics, National University "Yuri Kondratyuk Poltava Polytechnic", Poltava, Ukraine;
e-mail: Bibavv1@ukr.net; ORCID Author ID: <https://orcid.org/0000-0002-0949-206X>;
Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57189046983>.

Пінчук Наталія Михайлівна – кандидат технічних наук, доцент, Політехнічний університет Барі, Барі, Італія;

Nataliia Pinchuk – PhD in Technical Sciences, Associate Professor, Politecnico di Bari, Bari, Italy;
e-mail: nataliia.pinchuk@poliba.it; ORCID Author ID: <https://orcid.org/0000-0002-1720-5497>;
Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57202629137>.

Mathematical modeling of the dynamic thermal response of a multilayer wall assembly under periodic and pulsed thermal loading

Dmytro Usenko, Volodymyr Byba, Nataliia Pinchuk

Abstract. Background. Steady-state thermal resistance does not provide a complete characterization of multilayer building envelopes under time-varying thermal conditions because it does not account for the spatial distribution of thermal inertia or the excitation regime. **Subject of the study:** the dynamic thermal response of a stationarily equivalent three-layer wall assembly under periodic and finite-pulse thermal loading. **Objective:** to assess the effects of layer order, temperature-dependent thermophysical properties, and compensated interfacial thermal resistance at unchanged total steady-state thermal resistance. **Research tasks.** To define the *B*, *O*, *N*, *C5*, and *C10* scenarios; determine regime-appropriate response metrics; and assess the numerical robustness and resolvability of the resulting effects. **Methods.** A one-dimensional transient heat-conduction model with Robin boundary conditions, finite-volume discretization, implicit Backward Euler time integration, Picard iteration, energy and interface diagnostics, and grid/time refinement was used. **Results.** All configurations were constrained to $R_{tot} = 3.65315 \text{ m}^2\text{K/W}$. The combined grid/time numerical uncertainty was 0.325166% against the adopted 0.5% threshold. For the investigated *B/O* pair, the relative effect changed sign from -0.905363% under periodic loading to $+0.56625\%$ under finite-pulse loading, with effect-to-uncertainty ratios of 79.2365 and 88.85, respectively. For *N*, the periodic effect of -2.00484% was numerically resolved, whereas the finite-pulse effect of -0.282142% remained unresolved. The *C5/C10* effects were also numerically unresolved. **Conclusions.** Controlled stationary equivalence combined with numerical-uncertainty assessment enables regime-dependent effects to be separated; the observed sign reversal is restricted to the investigated *B/O* pair and the adopted regime-specific metrics.

Keywords: transient heat conduction; multilayer wall assembly; dynamic thermal response; thermal resistance; layer order; temperature-dependent properties; interfacial thermal resistance; periodic thermal loading; pulsed thermal loading.

MODEL AND METHOD OF INCLUSIVE INTERACTION WITH VISUAL AND TEXTUAL CONTENT BASED ON CONTEXTUAL GESTURE INTERPRETATION

Abstract. Background. Contactless gesture control, as an auxiliary channel of interaction with an interface, has now become widely used in various areas of life, as it provides a convenient, safe, and, if necessary, inclusive way of interacting with visual, graphical, and textual content. **The subject of the study** is the process of contactless control of basic navigation scenarios for working with visual, textual, and web content using hand tracking methods and contextual interpretation of gesture events. **The objective** of the study is to develop and experimentally validate a method of contactless interaction with visual and textual content based on RGB hand tracking and contextual interpretation of gestures for implementing practical user scenarios. The main research task of the study is to develop a conceptual model of context-adaptive inclusive user interaction with an information system and the external environment. **Methods.** Functional modeling, computer vision methods, analysis of hand keypoints, geometric rules for gesture recognition, and experimental testing were used. **Results.** A method is proposed in which a gesture event is converted into a command only after determining the type and analyzing the context, namely the technical and scenario context. The experimental validation was carried out using five basic gestures and three scenarios. **Conclusions.** The obtained results confirmed the operability of the method as an auxiliary channel of contactless interaction with visual and textual content.

Keywords: contactless interaction, tracking, gesture recognition, human-computer interaction, Media Pipe Hands, gesture control, interface, model, method.

1. Introduction

1.1 Relevance of the topic

The digitization of education, public services, everyday environments, and professional activities is leading to an increase in the number of scenarios in which users interact with visual and textual content via electronic devices (Human-Computer Interaction, HCI). At the same time, there is a growing need not only to create digital content but also to develop convenient, inclusive, and secure ways to interact with it, with so-called contactless technologies being in particularly high demand.

According to estimates by the World Health

Organization, about 1.3 billion people – or approximately 16% of the world's population – have significant disabilities [1]. For some of these users, traditional input methods may be inconvenient, inaccessible, or physically limited. In this context, gesture-based contactless interaction can be used as an auxiliary method of interface control, especially when a user is able to perform a limited set of hand movements but cannot effectively use a mouse, keyboard, or touchpad. In addition, gesture-based interfaces hold promise for rehabilitation, training, and adaptive environments, where hand movement can simultaneously serve as both a command input and a means of assessing motor activity (Table 1).

Table 1 – An Overview of the Applications of Contactless Human-Machine Interaction Interfaces

Field of application	Situations in which contactless methods are most in demand
Health Care and Surgical Environments	sterile viewing of CT/MRI/X-ray images, interface control without a mouse or keyboard, working while wearing gloves
Rehabilitation	fine motor skills training, hand movement assessment, gamified exercises, recovery from injuries or neurological disorders
Education and Inclusive Learning	lectures, presentations, interactive whiteboards, LMS, maintaining student attention, learning for students with special educational needs
Public Information Displays and Kiosks	public Service Centers, hospitals, universities, museums, exhibitions, transportation hubs, self-service terminals
Everyday and smart-home scenarios	controlling media, screens, and household devices when your hands are full or you don't want to use your voice
Automotive Industry	control volume, calls, and navigation without having to search for elements on the touchscreen
Industry, Manufacturing, and Robotics	operating robots, manipulators, and equipment in hazardous, contaminated, or hard-to-reach environments
VR/AR and Multimedia Environments	interaction with virtual objects, educational simulators, professional XR training simulators, controller-free interaction
Special Needs and Assistive Technologies	assistive computer control, communication aids for people with speech or hearing impairments, adaptive workstations, speech therapy or motor skills training devices

It is worth noting the relevance of this issue for educational and presentation environments. A prime example of its use is digital learning materials, presentations, video lectures, interactive whiteboards, distance learning systems, and public information screens, which are becoming widespread tools in the educational process. In such situations, teachers or users often need to navigate the material without constantly having to use a keyboard, mouse, or touchscreen – to switch slides, scroll through text, open interface elements, confirm selections, or zoom in or out on a visual object.

Therefore, contactless interaction with content can be viewed as an additional control channel that enhances the flexibility and ergonomics of working with educational and informational materials.

A separate group of scenarios relates to public and healthcare settings, where reducing direct physical contact with shared surfaces is crucial. In healthcare facilities, the issue of touching shared surfaces is particularly sensitive because, according to the WHO, among every 100 patients in emergency departments, at least 7 patients in high-income countries and 15 patients in low- and middle-income countries may contract at least one healthcare-associated infection [2]. At the same time, proper hand hygiene can prevent up to 50% of infections that can be avoided during the provision of medical care [3]. Although contactless interfaces do not replace hygiene measures, they can reduce contact with shared surfaces and increase the acceptance of digital services in environments with heightened hygiene requirements.

The limitations of existing solutions include the sensitivity and hardware characteristics of the recording devices used for subsequent hand tracking. The quality of gesture recognition depends on lighting, background, video stream resolution, distance from the camera, hand angle, partial finger occlusion, and movement speed. Review studies on hand gesture recognition show that, despite advances in computer vision and deep learning, challenges related to robustness to environmental changes, gesture complexity, computational efficiency,

and adaptation to different users remain unresolved. Therefore, for practical interaction with content, it is not enough to simply identify which gesture the user is performing. It is also necessary to determine exactly how this gesture should be interpreted in the system's current operating mode and within a specific interaction scenario.

1.2 Literature Review and Problem Definition

This subsection analyzes literature sources and recent studies devoted to hand gesture tracking approaches that differ in the type of input data, hand position estimation accuracy, hardware cost, ease of use, and suitability for large-scale application. The analyzed approaches include:

- tracking using RGB cameras;
- tracking using depth sensors;
- tracking using wearable sensors.

The analysis made it possible to identify several local unresolved issues, which can be systematized as follows:

- RGB-based tracking is accessible and easy to implement, but it depends on illumination, hand orientation, distance to the camera, and video stream resolution;
- most studies treat the recognized gesture as the final result, rather than as an intermediate gesture event that must be further transformed into a control command;
- in applied systems, gestures are often tied to one narrow scenario and a fixed set of commands;
- insufficient attention is paid to the fact that the same gesture may have different meanings depending on the active system mode, content type, hand role, and previous interaction state;
- complete scenarios of interaction with visual and textual content, such as file operations, web navigation, and drag-and-drop, remain insufficiently studied;
- the influence of external recording conditions on gesture recognition accuracy and the stability of applied interaction scenarios requires separate experimental verification (Fig.1).

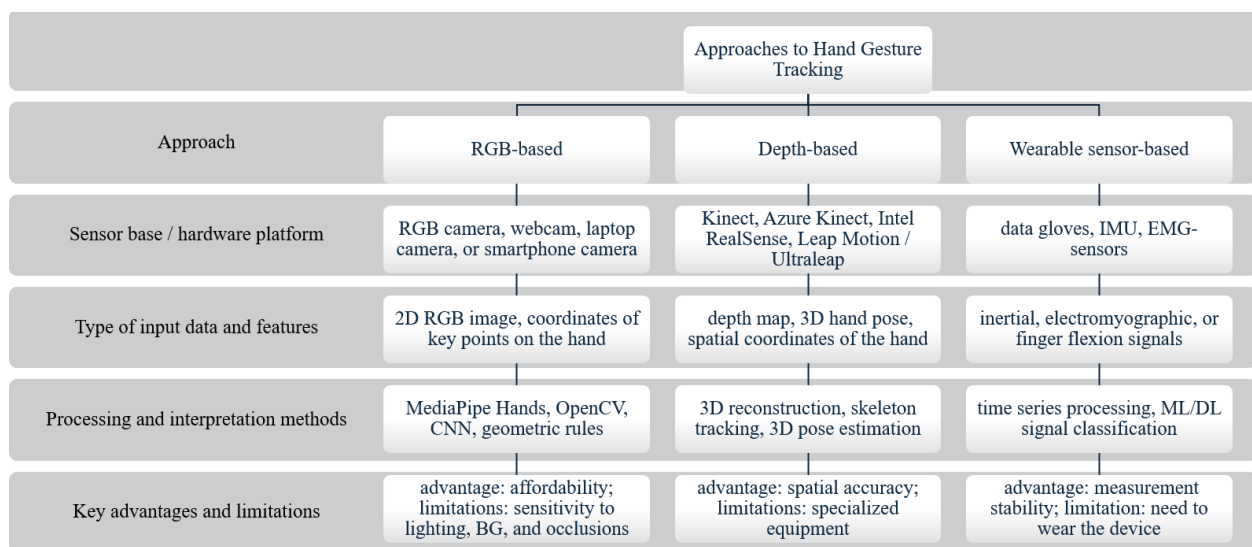


Fig. 1. Approaches to contactless hand gesture tracking

This makes it possible to formulate the general problem in the field of contactless gesture-based content control in human-machine interaction: the system should not only recognize a gesture, but also determine which interface action this gesture corresponds to in the current operating mode, for a specific type of content, and within a particular interaction scenario. This unresolved problem determines the aim of the present study.

In [4], it is shown that depth sensors, in particular Leap Motion, provide sufficiently accurate estimation of the spatial position of the hand and fingers based on three-dimensional hand geometry. However, this approach requires specialized and relatively expensive equipment, which is more difficult to integrate into typical user environments. Therefore, such an approach is less suitable for large-scale implementation in educational, public, or everyday interfaces.

In [5], an approach based on wearable inertial sensors is considered. Its strengths include lower dependence on illumination, background, and hand occlusion in the camera frame. At the same time, the user has to wear an additional device, which complicates the use of such systems in public or short-term interaction scenarios. For this reason, wearable sensors are not always convenient for fast contactless interaction with information content.

In contrast to these approaches, works [6, 7] demonstrate the feasibility of using a standard RGB camera and methods for detecting hand keypoints. In particular, MediaPipe Hands makes it possible to obtain the coordinates of 21 hand keypoints without additional sensors. This approach is inferior to depth sensors in terms of spatial accuracy, but it has important advantages: accessibility, low cost, ease of integration, and the possibility of use on a typical computer or laptop. Therefore, RGB-based hand tracking is appropriate for developing an accessible method of contactless interaction with visual and textual content.

At the same time, even within the RGB-based approach, a number of unresolved issues remain. In [8] and [9], real-time gesture recognition solutions are proposed using MediaPipe, YOLOv11, Inception-v3, and LSTM. The main focus of these studies is the recognition of a gesture as a class, while the further transformation of the recognized gesture into an interface action is considered insufficiently. This is because these studies are mainly aimed at improving classification accuracy, rather than developing a complete cycle of user interaction with digital content.

The practical significance of contactless gesture-based control methods should also be emphasized. The literature describes a wide range of application areas, including the use of gesture analysis in educational environments to recognize teacher signals and support students' attention [10]; interaction with public web displays, where gestures are used for contactless navigation through information content [11]; inclusive offline learning platforms for hands-free interaction with educational materials [12]; rehabilitation and assistive systems for tracking hand movements during rehabilitation exercises [13] and controlling rehabilitation exergame systems [14]; assistive interaction for users with

disabilities in smart-home environments [15]; control of household or IoT devices using local or edge-based gesture processing, command execution, and feedback generation [16, 17]; and human-robot interaction, where gesture recognition is required to form a navigation action or an operation in the environment [18].

Thus, the analysis of literature sources shows that gesture-based approaches are applied in education, public information interfaces, inclusive learning platforms, rehabilitation, smart-home systems, IoT systems, and human-robot interaction. However, most of these solutions are focused on a narrow application scenario and a fixed set of commands.

The weak point of the analyzed studies is insufficient attention to the problem of contextual gesture interpretation. In most cases, the gesture is either recognized as a separate class or directly linked to a fixed command. However, in interaction with visual and textual content, the same gesture may correspond to different actions depending on the current system mode, content type, hand role, previous interaction state, and user scenario. Therefore, the unresolved problem is the development of a method that combines RGB-based hand tracking with contextual interpretation of gesture events for contactless interaction with visual and textual content. This problem defines the aim of the further research.

1.3 Aim and Objectives of the Study

The aim of the study is to develop and experimentally validate a method of contactless interaction with visual and text content based on RGB hand tracking and contextual interpretation of gestures for implementing practical user scenarios.

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

- develop a conceptual model of context-adaptive, inclusive user interaction with an information system and the external environment;
- develop a structural model of a contactless interaction method that encompasses capturing an RGB video stream, determining key points of the hand, calculating geometric features, recognizing gestures, contextually interpreting gesture events, and generating interface control actions;
- experimentally evaluate the impact of external conditions on gesture recognition accuracy, specifically the distance to the camera, lighting level, hand angle, and video stream resolution;
- analysis of the obtained results.

2. Materials and Research Methods

The object of research is the process of contactless management of basic navigation scenarios, related to work with visual, textual and web content by means of one-handed, multi-handed or scripted gesture control based on RGB tracking of the hand and interpretation of gesture events depending on the context and scenario.

The input data for the system is hand movements recorded by a regular RGB camera, and the output data consists of high-level control actions: hovering, selecting an element, moving an element, scrolling the page, confirming, canceling, or executing a script command.

The main hypothesis of the study is that the implementation of two levels of contextual interpretation of gestural events, which determine the states and conditions of receiving input data, will increase the practical suitability of contactless interaction with visual and textual content.

The selected context levels include the following:

- technical context of gesture recognition (illumination, camera distance, hand angle, video flow resolution, occlusion, background complexity, etc.);
- scenic context of gesture interpretation (system state - active or passive; type of operating mode - one-handed, two-handed, pointer mode, sequence mode, calibration; the class of the involved hand (dominant or non-dominant); preliminary data on the state of the system; the selected type of content for interaction is text, web, graphic, OS element; admissibility of action in the current state of the interface).

In the development and experimental validation of the method, the following assumptions and simplifications are adopted.

Condition 1. The interaction is carried out using a conventional RGB camera without the use of depth sensors, wearable devices, instrumented gloves or additional markers on the user's hand.

Condition 2. The user's hand is in the camera's field of view for unimpeded detection and presentation as a set of key points of the hand and geometric relationships between them.

Condition 3. A gestural event is not seen as a completed command. Only after analyzing the context, the gestural event is converted into a control action.

Condition 4. In the work, experimental evaluation is limited to a fixed set of gestures (OPEN_PALM, PINCH, FIST, PINCH_MIDDLE, THUMBS_UP) and user scenarios - basic system and navigation actions, work with text or web content, management of graphic objects.

The principal methods employed in the present study include functional modeling (applied in the construction of the models), computer vision techniques, geometric analysis of hand skeletal keypoints, and experimental validation.

For the acquisition of the input video stream, an RGB camera was utilized as a cost-effective and scalable solution for implementing contactless human-computer interaction (HCI). As substantiated in prior investigations [9, 19], gesture-based interaction with computing systems in educational, clinical, domestic, mobile, and assistive settings represents a rational approach, since it obviates the need for dedicated wearable hardware or complex sensory infrastructures.

The retrieval of 21 keypoint coordinates for the comprehensive description of the palm and phalanges per frame is accomplished via computer vision algorithms, thereby enabling real-time performance even on mobile computational platforms.

The viability of the MediaPipe Hands framework and the landmark-based methodology is corroborated by the findings reported in [14, 20].

The adoption of the landmark-based paradigm is justified by its capacity to shift the analytical focus from

holistic image processing to a compact feature-based hand representation, consequently reducing the dimensionality of the input data through the isolation of a precisely defined region of interest [9].

The methodological choices delineated in [9, 14, 20] are deemed essential, as they ensure sensor-independent operation, facilitate straightforward system integration, and render the approach applicable across conventional educational, domestic, public, and assistive environments. Nevertheless, it has been emphasized in the literature [14, 16, 20] that the practical efficacy of such systems remains contingent upon a number of factors, including ambient conditions, computational resource constraints, illumination levels, camera-to-user distance, real-time processing robustness, and the ergonomic suitability of the chosen gesture vocabulary. These considerations have been duly addressed within the framework of the proposed system and are effectively compensated for by the proposed context-aware interpretation mechanism.

To verify the effectiveness of the proposed method, three groups of test scenarios have been developed, which allow for the evaluation of the proposed method not only in terms of the accuracy of recognizing individual gestures, but also in terms of its ability to perform complete user actions while interacting with the operating system, text, web, and graphical content (scenarios are listed in order of increasing complexity):

- basic system and navigation actions (testing the ability to control operating system elements without using a mouse or keyboard – pointing, moving the cursor with the leading hand, clicking, and opening folders using basic gestures – THUMBS_UP, PINCH, PINCH_MIDDLE);
- working with text or web content (testing interaction with information content – hovering over a page element, selecting, scrolling, and navigating to another part of the content using basic gestures – THUMBS_UP, PINCH, FIST);
- manipulating graphical objects (testing the stability of control when selecting, holding, and moving an object – hovering over the object, triggering the selection action, moving the object to the desired area, and completing the action using basic gestures – THUMBS_UP, PINCH).

The experimental study of the proposed method takes into account the conditions that affect the quality of RGB hand tracking and the stability of gesture event generation. Since the system uses a video stream from a standard RGB camera, its performance depends on:

- the visibility of the hand in the frame;
- lighting;
- background;
- the distance to the camera;
- the angle of the hand;
- the resolution of the video stream.

The main limitations include partial obstruction of the hand or fingers, the hand extending beyond the frame, varying stability of individual gestures, and the transmission of commands at the operating system event level without considering the internal state of a specific application.

External factors include lighting, background, distance, and the angle of the hand, while internal factors include threshold settings, hand detection parameters, system operating mode, camera resolution, and computer performance. Controlling these factors allows for an accurate assessment of the method's performance and helps distinguish errors in the algorithmic interpretation of gestures from errors caused by shooting conditions or the hardware and software environment.

The basic conditions for conducting the study were defined as a constant camera resolution, a consistent distance between the user and the camera, a neutral background, a frontal hand position, and a fixed set of gestures. A video stream resolution of 1280×720 was used as the baseline setting, as it provides a balance between image detail and processing speed.

In a controlled series of measurements, the user held their hand approximately 0.5 m from the camera, the illuminance was about 153 lx, and the hand was positioned primarily facing the camera head-on.

First, the recognition accuracy of individual gestures was determined under baseline conditions; thereafter, individual influencing factors – illuminance, distance from the camera, hand orientation (Fig. 2), and video stream resolution – were systematically varied.

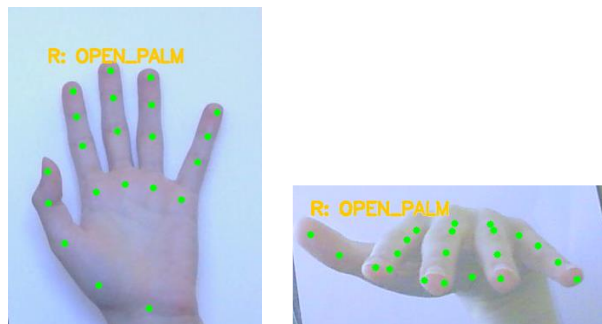


Fig. 2. An example of testing the method's capability when changing the perspective for the open-palm gesture

Next, the system's performance was tested in practical scenarios involving interaction with the operating system, as well as text, web, and graphical content. For each series, recognition accuracy, the number of omissions and errors, frame rate, response time, and scenario execution time were recorded. This methodology made it possible to compare the system's performance under standard and modified conditions, identify the most stable gestures, and determine the factors that most significantly reduce the accuracy and speed of the proposed method. Each scenario was run 10 times using gesture control and 10 times using traditional mouse and keyboard control. The table shows the average execution time, the total number of errors, and the number of critical errors.

3. Results of Development and Experimental Study of the Proposed Contactless Interaction Method

Within the scope of this study, contactless gesture interaction is considered not as an isolated method for recognizing individual hand positions, but as one of the channels for context-adaptive user interaction with an information system. The general logic of such interaction is presented as a conceptual model that takes into account the user's input request, individual needs, the application scenario, the user's context, the interpretation of input data, the selection of an assistance strategy, and the generation of feedback through a multimodal interface.

The proposed conceptual model (Fig. 3) demonstrates that effective interaction with an information system depends not only on the method of entering a command, but also on the context in which it is used. Such context may include the type of information content, the level of accessibility of the environment, the presence of obstacles, the current task scenario, the user's sensory or motor capabilities, as well as the available feedback channels.

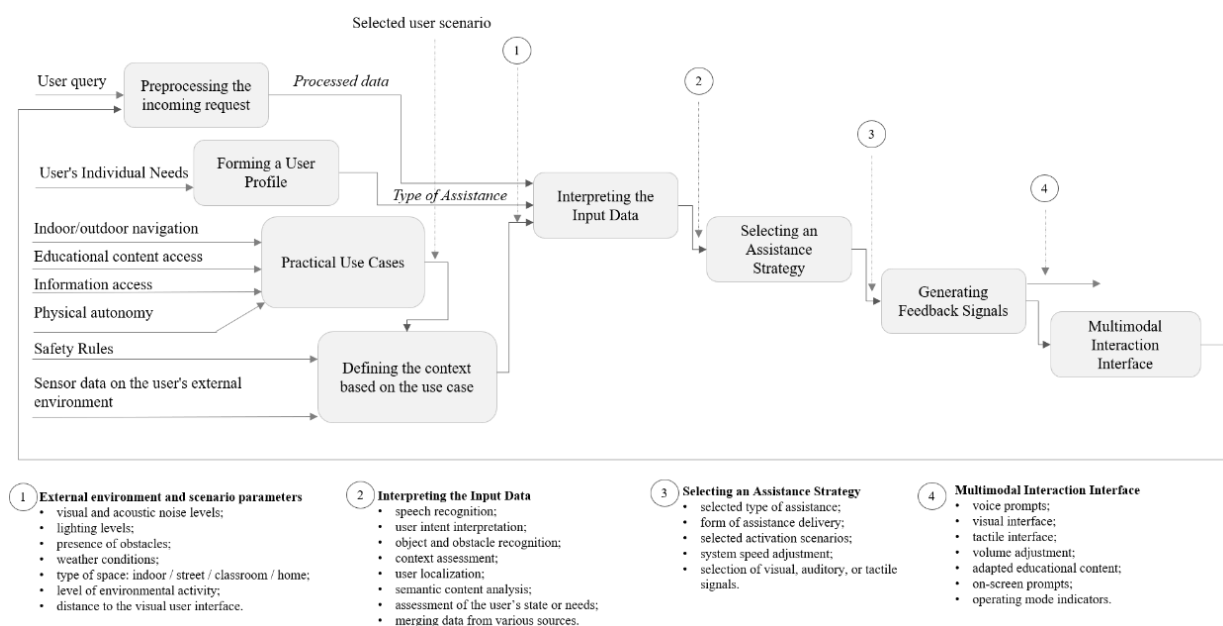


Fig. 3. A Conceptual Model of Context-Adaptive, Inclusive User Interaction with an Information System and the External Environment

In such a system, gesture control can serve as one of the input channels, complementing voice, visual, auditory, or tactile means of interaction.

Within the framework of the conceptual model presented, this article focuses on one of its functional components – the contactless input channel, in which the user's hand movements are converted into commands for interacting with visual and text content. Unlike the general model of inclusive interaction, where various types of sensory data and feedback signals can be used, this study examines RGB hand tracking, the generation of basic gesture events, and their contextual interpretation as interface control commands.

Thus, the proposed method specifies the block responsible for interpreting input data and generating feedback control actions in a context-adaptive interaction system.

Its purpose is to convert hand movements captured by an RGB camera into high-level interface actions: pointing, selecting, scrolling, confirming, canceling, or executing a scripted command.

The model proposed in this paper for a method of non-contact interaction with visual and text content, based on RGB hand tracking and contextual

interpretation of gestures, is presented as a functional diagram consisting of five functional blocks:

- RGB frame acquisition / preprocessing;
- hand detection / keypoint identification;
- gesture event formation;
- mode- and context-based interpretation of the gesture event;
- generation of a remote-control command / OS action.

The model's primary function is to convert the user's movements, captured by an RGB camera, into interface control actions, taking into account the active operating mode, the role of the hand, the type of content, and the current interaction scenario.

A key distinction of the proposed model is the presence of a mode-context interpretation block, in which an elementary gesture event is mapped to a control action not directly, but taking into account the system's active state, operating mode, previous state, the presence of confirmation, content type, and interaction scenario.

The proposed IDEF0 model illustrates the sequence of steps for converting an RGB video stream into a high-level remote-control command for visual and text content (Fig. 4).

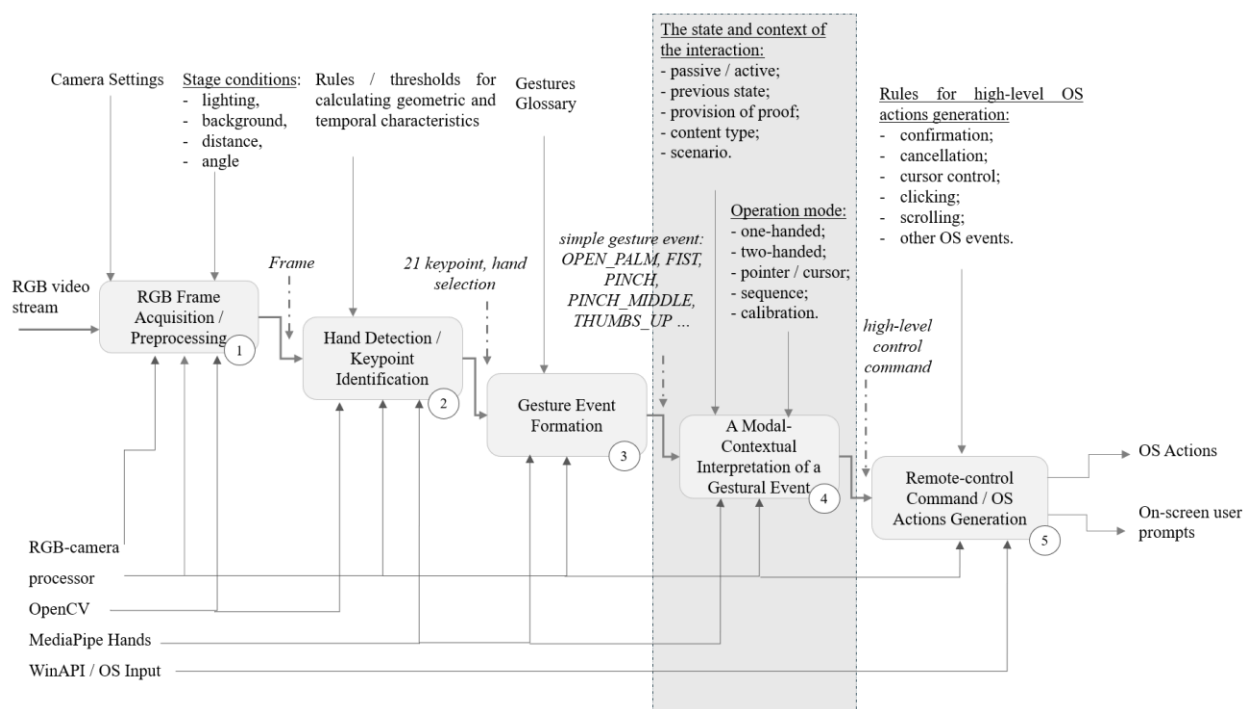


Fig. 4. A Generalized Model of a Non-Contact Interaction Method with Visual and Textual Content Based on RGB Hand Tracking and Contextual Gesture Interpretation

Block 1, “RGB Frame Acquisition / Preprocessing,” acquires a frame from an RGB camera and prepares it for further analysis. At this stage, the camera configuration and scene conditions – including lighting, background, distance to the user, and brush angle – are taken into account. The result of this block's operation is a prepared RGB frame, which is passed on to the brush detection stage.

Block 2, “Hand Detection / Keypoint Identification,” detects the hand in the frame and determines the coordinates of the hand's keypoints. The

output of this block is a compact representation of the hand in the form of a set of keypoints, and it also selects or identifies the hand for further processing. This stage is an intermediate step between image analysis and the generation of a gesture event.

Block 3, “Gesture Event Formation,” converts the coordinates of keypoints into a basic gesture event. This is done using rules and thresholds for calculating geometric and temporal features, as well as a dictionary of basic gestures. Such gesture events include, in particular, OPEN_PALM, FIST, PINCH,

PINCH_MIDDLE, THUMBS_UP, and other valid hand positions or combinations (Fig. 5). At this stage, the system does not yet generate an operating system command, but only determines which basic gesture event was detected.

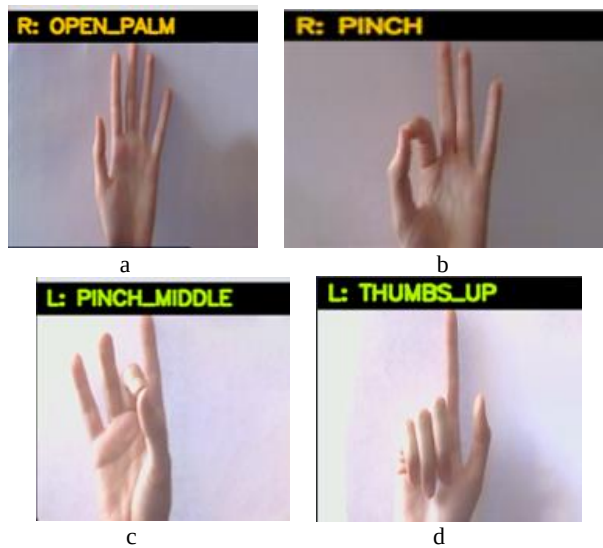


Fig. 5. Basic hand gestures: a – open palm (OPEN_PALM); b – pinch with the thumb and index finger (PINCH); c – pinch with the thumb and middle finger (PINCH_MIDDLE); d – thumbs up (THUMBS_UP)

A key element of the model is Block 4, “Mode-Contextual Interpretation of a Gesture Event.” It is in this block that a basic gesture event is matched to a possible control command – not directly, but taking into account the state and context of the interaction. These factors include the system’s passive or active state, the previous state, the presence of confirmation, the content type, and the current operating scenario. Additionally, the active operating mode is taken into account: one-handed, two-handed, pointer control mode, sequence input mode, or calibration mode.

Within block 4, the same basic gesture event can take on different functional meanings depending on the active mode and context. For example, a gesture can be

interpreted as confirmation, cancellation, a sequence element, activation of pointer mode, or part of a content control command. If the system is in a passive state or the confirmation conditions are not met, no control action is generated. This helps reduce the number of accidental activations and increase the predictability of the interaction.

Block 5, “Remote-control Command / OS Actions Generation,” converts a high-level control command – received after mode- and context-based interpretation – into a specific action of the operating system or interface. Such actions include switching operating modes, buffering sequences, confirmation, cancellation, cursor control, clicking, scrolling, and other operating system events. The model’s outputs are operating system actions and on-screen user prompts, which provide feedback during interaction.

The model’s control inputs include the camera configuration, scene operating conditions, feature calculation rules and thresholds, the gesture vocabulary, the state and context of the interaction, operating modes, and the rules for generating high-level actions of the interface or operating system. The execution mechanisms include an RGB camera, a computer, OpenCV, MediaPipe Hands, a geometric rules module, a contextual interpreter, and a mechanism for transmitting events via WinAPI/OS Input. The proposed model describes not only the technical process of gesture recognition but also the full cycle of remote, contactless content control. Its key feature is mode-contextual interpretation, thanks to which a gesture event is converted into a control action only after checking the active mode, previous state, confirmation, content type, and interaction scenario.

Within the proposed model, passive mode corresponds to a standby state in which the system analyzes the video stream but does not generate OS control events (Table 2). The transition to active mode occurs only after a valid gesture event or activation command is detected. This approach reduces the likelihood of false triggers, since random hand positions are not directly converted into interface actions.

Table 2 – Operating Modes of the Contactless Interaction System

Mode	Purpose	Typical Actions
Passive	Anticipation, frame analysis without forming commands	clearing temporary states, updating the OSD
Active (one-handed)	Executing simple commands with one hand	select, cancel, confirm, single gesture
Active (two-handed)	Assigning roles to each hand	dominant hand – positioning, non-dominant hand – confirm/click
Mouse/Pointer Mode	Cursor control	cursor movement, PINCH, PINCH_MIDDLE, FIST
Sequence Mode	Entering more complex commands	gesture accumulation, confirm, cancel
Calibration	Customizing thresholds for the user	updating threshold values

Active mode, in turn, is divided into one-handed mode, two-handed mode, pointer control mode, sequence mode, and calibration mode. As a result, the same basic event can be interpreted in different ways depending on the current interaction scenario.

Since Block 4 is responsible not for image processing but for decision-making logic, it is advisable to present its decomposition in the form of a flowchart. This allows us to illustrate the system activity check, the selection of the current mode, the mode-based

interpretation of the gesture, context-based alignment, and the generation of a high-level command or the reset of temporary states (Fig. 6).

The block diagram details the operational logic of Block A4, “Mode-Context-Based Interpretation of a Gesture Event.” The input for this stage is a basic gesture event, generated based on the coordinates of key points

on the hand and geometric features. In the first step, the system checks whether it is active and whether the received event is valid.

If these conditions are not met, a high-level command is not generated, temporary states are reset, and then the on-screen prompts, event log, and internal system state are updated.

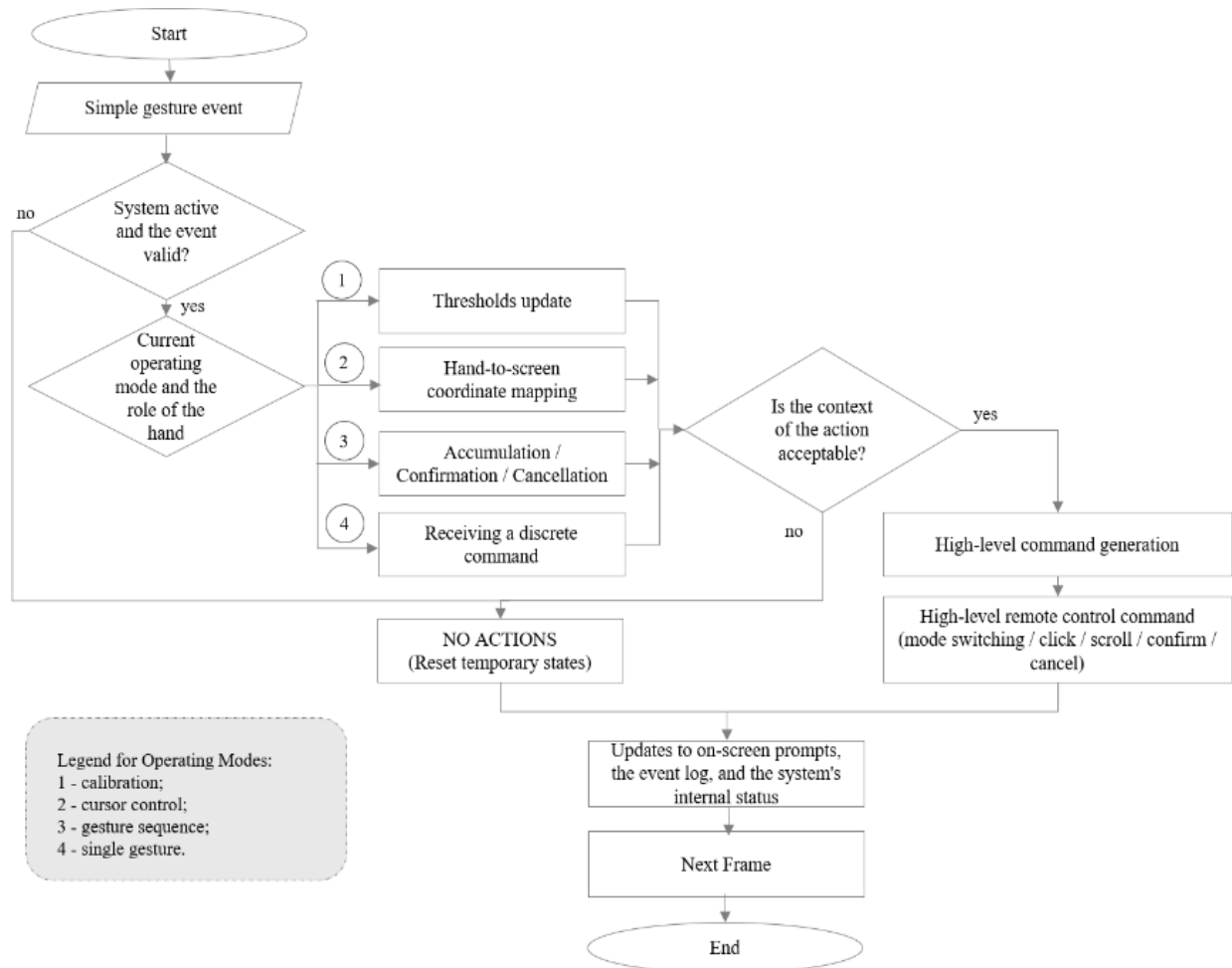


Fig. 6. Flowchart of the Mode-Context Interpretation of a Gesture Event

If a gesture event is valid, the system determines the current operating mode and the role of the hand. The proposed method provides for four main modes: calibration, pointer control, gesture sequence, and single gesture. In calibration mode, threshold values are updated; in pointer control mode, hand coordinates are mapped to screen coordinates; in gesture sequence mode, gestures are used to accumulate, confirm, or cancel a command; and in single-gesture mode, a discrete command is generated.

After the mode interpretation, the system checks whether the action is valid in the current context. At this stage, the system takes into account the previous state of the system, the role of the hand, the active mode, the content type, the interaction scenario, and the presence of confirmation. If the action does not match the current context, the system does not generate a control command. If the action is valid, a high-level remote-control command is generated, such as mode switching, clicking, scrolling, confirmation, or cancellation.

The final step is to update the on-screen prompts, the event log, and the system's internal state, after which the system proceeds to process the next frame. Thus, the flowchart shows that the result of the method is not simply a recognized gesture, but a contextually appropriate remote-control command.

The data presented in Table 3 confirm that the proposed method is operational under the selected experimental conditions and can support the execution of applied interaction scenarios.

The results demonstrate that the system can process visual input, detect gesture events, interpret them according to the current mode and scenario, and generate corresponding control actions.

Therefore, the conducted experiments validate the main idea of the study: contactless interaction with visual and textual content should be based not only on gesture recognition, but also on contextual interpretation of the recognized gesture event.

Table 3 – Summary of the results of the empirical validation of the proposed method

Evaluation Group	Condition / Scenario	Key Indicators
Gesture Recognition Accuracy	Base conditions, 1280×720	Average accuracy – 0,873
	Distance to the camera: 2 m	Average accuracy – 0,673
	Low light, 20 lx	Average accuracy – 0,800
	Change in the hand angle	Average accuracy – 0,513
	Resolution 320×180	Average accuracy – 0,854
	Resolution 1920×1080	Average accuracy – 0,927
Consistency of Individual Gestures	OPEN_PALM	The range of accuracy under different conditions: 0,933 – 0,967, average $\approx$ 0,956
	PINCH	The range of accuracy under different conditions: 0,300 – 0,867, average $\approx$ 0,578
	PINCH_MIDDLE	The range of accuracy under different conditions: 0,333 – 0,867, average $\approx$ 0,650
	FIST	The range of accuracy under different conditions: 0,167 – 1,000, average $\approx$ 0,800
	THUMBS_UP	The range of accuracy under different conditions: 0,733 – 0,967, average $\approx$ 0,878
Scenario Testing	Basic system and navigation actions, using a file operation as an example – creating and opening a folder	Time required for gesture control: 48,692 ms, time spent using the mouse and keyboard – 15,127 ms, errors – 1, critical – 0.
	Working with text or web content using web navigation as an example – opening a browser and navigating to a page	Time required for gesture control – 65129 ms, time spent using the mouse and keyboard – 16206 ms, errors – 0, critical – 0.
	Managing Graphic Objects: An Example Using the Copy and Drag-and-Drop Operations	Time required for gesture control – 80317 ms, time spent using the mouse and keyboard – 13401 ms, errors – 3, critical – 0.
The Effect of Resolution on The Scenarios	320×180, web scenario	Time – 80433 ms, errors – 5, FPS – 22,2
	1280×720, web scenario	Time – 65129 ms, errors – 0, FPS – 19,7
	1920×1080, web scenario	Time – 59379 ms, errors – 2, FPS – 14,4
	1920×1080, drag-and-drop	Time – 58862 ms, errors – 6, critical – 1, FPS – 14,4

This makes it possible to reduce the ambiguity of gesture commands and adapt the system response to the current interaction scenario.

4. Discussion of the Experimental Results of the Contactless Interaction Method

An analysis of the overall results shows that the proposed method provides the most stable performance under standard shooting conditions and when the video stream has sufficient detail. The highest average accuracy was achieved at a resolution of 1920×1080; however, this mode is associated with a higher computational load and a lower frame rate. Therefore, for practical use, the 1280×720 mode is recommended, as it provides an acceptable balance between accuracy, performance, and the number of errors.

Among the individual gestures, OPEN_PALM proved to be the most robust. This is because it has a distinct hand shape and does not require precise determination of small distances between the fingers. The PINCH and PINCH_MIDDLE gestures are more sensitive to image quality, camera angle, and distance from the camera, since their recognition depends on the correct determination of the position of the fingertips. The greatest decrease in accuracy is observed when the

hand's angle changes, which confirms the critical importance of finger overlap for RGB-based tracking.

The scenario testing showed that all the actions considered can be performed using gesture controls. The most stable use case was web browsing, for which no errors were recorded in basic mode. The most challenging scenario turned out to be drag-and-drop, as it requires not only the recognition of a specific gesture but also consistent control while moving an object.

A comparison with traditional input methods shows that gesture control is slower than a mouse and keyboard in terms of execution time. This is to be expected, since the proposed method involves additional steps: hand detection, gesture event recognition, mode verification, contextual interpretation, and command generation. At the same time, the absence of critical errors in basic scenarios confirms the method's viability as a supplementary channel for contactless interaction with visual and text content.

5. Conclusions

This paper proposes a conceptual model of context-adaptive, inclusive user interaction with an information system and the external environment, which takes into account the user's input request, individual needs, the

application scenario, the user's context, the interpretation of input data, the selection of an assistance strategy, and the generation of feedback via a multimodal interface.

This paper presents a method for contactless interaction with visual and text content based on RGB hand tracking and contextual gesture interpretation. The proposed method ensures the sequential conversion of an RGB video stream into an interface control action by identifying key points on the hand, calculating geometric features, forming a basic gesture event, interpreting it in terms of mode and context, and generating a high-level remote-control command.

A functional model of the method has been developed, in which the key element is the block for mode-contextual interpretation of a gesture event. Unlike a direct "gesture-command" mapping, the proposed approach evaluates a gesture event by taking into account the active mode, the system's previous state, the role of the hand, the presence of confirmation, the type of content, and the interaction scenario. This reduces the number of false triggers and increases the predictability of command generation.

The effect of external conditions on gesture recognition accuracy was experimentally evaluated. Under baseline conditions, at a resolution of 1280×720, the average accuracy was 0.873. The highest average accuracy was obtained at a resolution of 1920×1080 – 0.927. H

however, this mode involves a higher computational load; therefore, for practical use, the 1280×720 mode is preferable as a compromise between accuracy, speed, and the number of errors.

It was found that the most critical external factor is a change in the hand's angle, which reduces the average accuracy to 0.513. Increasing the distance to the camera to 2 m also negatively affects recognition quality, reducing the average accuracy to 0.673. Reduced lighting to 20 lux had a more moderate effect: the average accuracy was 0.800. This confirms that for the RGB-based approach, hand visibility, a frontal hand position, and sufficient image detail are particularly important.

A comparison of the stability of individual gestures showed that the OPEN_PALM gesture is the most stable,

with its accuracy varying between 0.933 and 0.967 under different conditions. The PINCH and PINCH_MIDDLE gestures proved to be more sensitive to image quality, distance, and viewing angle, since their recognition depends on the precise determination of the position of the fingertips.

Scenario testing confirmed the effectiveness of the proposed method during basic system operations, web navigation, and control of graphical objects. All basic scenarios were executed without critical errors, confirming the method's suitability as an auxiliary channel for contactless interaction. At the same time, gesture control is slower than a mouse and keyboard in terms of execution time, which is due to the additional steps involved in hand tracking, gesture event recognition, mode verification, contextual interpretation, and command generation.

Further research should focus on improving resistance to changes in hand orientation, enhancing the recognition of subtle and dynamic gestures, adaptively adjusting thresholds for different users, expanding testing scenarios, and deeper integration of gesture control with applications.

Conflict of interest

The authors declare that they have no conflict of interest regarding this research, including financial, personal, authorship-related, or any other type of conflict that could affect the research and its results presented in this article.

Funding

This study was conducted with financial support from the research project "Multimodal Indoor Navigation and Assistive Support System for Users with Special Needs" (state registration number: 0126U002094).

Data availability

The manuscript has no associated data.

Use of artificial intelligence tools

The authors confirm that artificial intelligence technologies were not used in the preparation of this work.

REFERENCES

1. World Health Organization (2023), "Disability", available at: <https://www.who.int/news-room/fact-sheets/detail/disability-and-health>
2. World Health Organization (n.d.), "Hand hygiene", available at: <https://www.who.int/teams/integrated-health-services/infection-prevention-control/hand-hygiene>
3. World Health Organization (2021), "Key facts and figures. World Hand Hygiene Day 2021", available at: <https://www.who.int/campaigns/world-hand-hygiene-day/2021/key-facts-and-figures>
4. Vysocký, A., Grushko, S., Oščádal, P., Kot, T., Babjak, J., Jánoš, R., Sukop, M., and Bobovský, Z. (2020), "Analysis of Precision and Stability of Hand Tracking with Leap Motion Sensor", *Sensors*, vol. 20, no. 15, article 4088, doi: <https://doi.org/10.3390/s20154088>
5. Jeon, H., Choi, H., Noh, D., Kim, T., and Lee, D. (2022), "Wearable Inertial Sensor-Based Hand-Guiding Gestures Recognition Method Robust to Significant Changes in the Body-Alignment of Subject", *Mathematics*, vol. 10, no. 24, article 4753, doi: <https://doi.org/10.3390/math10244753>
6. Yarovy, V., Barkovska, O., Maksymov, D., Ni, Y., and Raptanov, D. (2026), "Methods for recognising hand gestures using computer vision to assist people with movement or speech impairments", *Control, Navigation and Communication Systems*, no. 1(83), pp. 135–141, doi: <https://doi.org/10.26906/SUNZ.2026.1.135>
7. Barkovska, O., Ruban, I., Tymoshenko, D., Holovchenko, O., and Yankovskyi, O. (2025), "Research on mobile machine learning platforms for human gesture recognition in human-machine interaction systems", *Technology Audit and Production Reserves*, vol. 2, no. 2(82), pp. 6–14, doi: <https://doi.org/10.15587/2706-5448.2025.325423>

8. Alsharif, B., Alalwany, E., Ibrahim, A., Mahgoub, I., and Ilyas, M. (2025), "Real-Time American Sign Language Interpretation Using Deep Learning and Keypoint Tracking", *Sensors*, vol. 25, no. 7, article 2138, doi: <https://doi.org/10.3390/s25072138>
9. Yaseen, Kwon, O.-J., Kim, J., Jamil, S., Lee, J., and Ullah, F. (2024), "Next-Gen Dynamic Hand Gesture Recognition: MediaPipe, Inception-v3 and LSTM-Based Enhanced Deep Learning Model", *Electronics*, vol. 13, no. 16, article 3233, doi: <https://doi.org/10.3390/electronics13163233>
10. Chen, I.D.S., Yang, C.-M., Wu, S.-S., Yang, C.-K., Chen, M.-J., Yeh, C.-H., and Lin, Y.-H. (2024), "Continuous Recognition of Teachers' Hand Signals for Students with Attention Deficits", *Algorithms*, vol. 17, no. 7, article 300, doi: <https://doi.org/10.3390/a17070300>
11. Gena, C., Vernerio, F., Mattutino, C., Persichella, G., and Lomanto, L. (2019), "Experimenting Touchless Gestural Interaction for a University Public Web-based Display", *Adjunct Proceedings of the 13th Biannual Conference of the Italian SIGCHI Chapter (CHIItaly 2019)*, Padova, Italy, 5 p., available at: <https://arxiv.org/abs/2011.09749>
12. Drăgoi, M.V., et al. (2025), "An Inclusive Offline Learning Platform Integrating Gesture Recognition and Local AI Models", *Biomimetics*, vol. 10, no. 10, article 693, doi: <https://doi.org/10.3390/biomimetics10100693>
13. Şen, Z., Çil, E.T., and Serif, T. (2024), "Web Camera-Based Gamified Rehabilitation System for Improving Fine Motor Skills: Pilot Study", *International Journal of Digital Waste Engineering*, vol. 1, no. 1, pp. 15–22, doi: <https://doi.org/10.5281/zenodo.14580508>
14. Husna, R., Brata, K.C., Anggraini, I.T., Funabiki, N., Rahmadani, A.A., and Fan, C.-P. (2025), "An Investigation of Hand Gestures for Controlling Video Games in a Rehabilitation Exergame System", *Computers*, vol. 14, no. 1, article 25, doi: <https://doi.org/10.3390/computers14010025>
15. Panagiotou, C., Faliagka, E., Antonopoulos, C.P., and Voros, N. (2025), "Multidisciplinary ML Techniques on Gesture Recognition for People with Disabilities in a Smart Home Environment", *AI*, vol. 6, no. 1, article 17, doi: <https://doi.org/10.3390/ai6010017>
16. Hobbs, T., and Ali, A. (2025), "Smart Home Control Using Real-Time Hand Gesture Recognition and Artificial Intelligence on Raspberry Pi 5", *Electronics*, vol. 14, no. 20, article 3976, doi: <https://doi.org/10.3390/electronics14203976>
17. Lopes, M.C., Silva, P.A., Marengo, L., Vilas Boas, E.C., de Carvalho, J.G.A., Ferreira, C.A., Carvalho, A.L.O., Guimarães, C.V.R., Aquino, G.P., and de Figueiredo, F.A.P. (2026), "GECO: A Real-Time Computer Vision-Assisted Gesture Controller for Advanced IoT Home System", *Sensors*, vol. 26, no. 1, article 61, doi: <https://doi.org/10.3390/s26010061>
18. Wang, X., Shen, H., Yu, H., Guo, J., and Wei, X. (2022), "Hand and Arm Gesture-based Human-Robot Interaction: A Review", *Proceedings of the 6th International Conference on Algorithms, Computing and Systems, Association for Computing Machinery*, New York, article 12, pp. 1–7, doi: <https://doi.org/10.1145/3564982.3564996>
19. Yun, G., Kwak, H., and Kim, D.H. (2024), "Single-Handed Gesture Recognition with RGB Camera for Drone Motion Control", *Applied Sciences*, vol. 14, no. 22, article 10230, doi: <https://doi.org/10.3390/app142210230>
20. Marques, P., Váz, P., Silva, J., Martins, P., and Abbasi, M. (2025), "Real-Time Gesture-Based Hand Landmark Detection for Optimized Mobile Photo Capture and Synchronization", *Electronics*, vol. 14, no. 4, article 704, doi: <https://doi.org/10.3390/electronics14040704>

Received (Надійшла) 14.07.2026

Accepted for publication (Прийнята до друку) 02.09.2026

Publication date (Дата публікації) 25.09.2026

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Барковська Оlesia Юрійвна – кандидат технічних наук, доцент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Olesia Barkovska – Candidate of Technical Sciences, Associate Professor at the Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: olesia.barkovska@nure.ua; ORCID Author ID: <http://orcid.org/0000-0001-7496-4353>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=24482907700>.

Модель та метод інклюзивної взаємодії з візуальним і текстовим контентом на основі контекстної інтерпретації жестів

О. Ю. Барковська

Анотація. **Актуальність.** Безконтактне жестове керування, як допоміжний канал взаємодії з інтерфейсом, на сьогодні набуло широкого поширення у різних сферах життя, оскільки забезпечує зручний, безпечний і, за необхідності, інклюзивний спосіб взаємодії з візуальним, графічним та текстовим контентом. **Предметом** дослідження є процес безконтактного керування базовими навігаційними сценаріями для роботи з візуальним, текстовим і веб-контентом з використанням методів трекінгу кисті та контекстної інтерпретації жестових подій. **Мета** дослідження полягає в розробленні та експериментальній перевірці методу безконтактної взаємодії з візуальним і текстовим контентом на основі RGB-трекінгу кисті та контекстної інтерпретації жестів для реалізації практичних користувацьких сценаріїв. **Головне завдання** дослідження полягає у розробці концептуальної моделі контекстно-адаптивної інклюзивної взаємодії користувача з інформаційною системою та зовнішнім середовищем. **Методи.** Використано функціональне моделювання, методи комп'ютерного зору, аналіз ключових точок кисті, геометричні правила розпізнавання жестів та експериментальне тестування. **Результати.** Запропоновано метод, у якому жестова подія перетворюється на команду лише після визначення типу та аналізу контексту (технічного та сценарного). Експериментальна перевірка проводилась на п'яти базових жестах і трьох сценаріях. **Висновки.** Отримані результати підтвердили працездатність методу як допоміжного каналу безконтактної взаємодії з візуальним і текстовим контентом.

Ключові слова: безконтактна взаємодія, трекінг, розпізнавання жестів, людино-комп'ютерна взаємодія, MediaPipe Hands, жестове керування, інтерфейс, модель, метод.

М. Е. Бондаренко, Г. С. Іващенко

Харківський національний університет радіоелектроніки, Харків, Україна

ГІБРИДИЗАЦІЯ МЕТОДІВ ВИЛУЧЕННЯ ОЗНАК У СИСТЕМІ ГОЛОСОВОЇ ІДЕНТИФІКАЦІЇ КОРИСТУВАЧА

Анотація. Актуальність. Практичне застосування систем голосової ідентифікації обумовлює необхідність високої точності розпізнавання, обчислювальної ефективності та стійкості до синтезованих підробок голосу. Статистичні підходи, такі як акустичний аналіз та спектральне порівняння, є невимоглими до апаратних ресурсів, проте поступаються за точністю сучасним глибинним підходам на основі машинного навчання. Штучні нейронні мережі, попри вищу точність, характеризуються значними обчислювальними витратами та чутливістю до викривлень. Одним із можливих шляхів вирішення є гібридизація методів на етапі вилучення ознак, що передбачає поєднання статистичного акустичного аналізу та нейромережевого підходу в єдиному векторному просторі. **Об'єкт дослідження:** гібридизація методів вилучення ознак у системі голосової ідентифікації на основі поєднання статистичного акустичного аналізу та штучної нейронної мережі. **Предмет дослідження:** метод гібридного вилучення ознак, що поєднує MFCC-вектори та вектори ознак нейронної мережі в єдиному векторному просторі, із порівняльним аналізом способів їх поєднання. **Мета дослідження:** розробка та дослідження методу гібридного вилучення ознак для підвищення точності голосової ідентифікації. **Завдання дослідження.** Реалізувати запропоновані підходи гібридизації статистичних та нейромережевих методів вилучення ознак у єдиному векторному просторі; експериментально оцінити точність ідентифікації у режимах closed-set та open-set; дослідити обчислювальну ефективність і масштабованість запропонованих варіантів. **Методи дослідження.** Статистичний акустичний аналіз на основі мел-частотних кепстральних коефіцієнтів, згорткова нейронна мережа на основі архітектури VGGish, метод головних компонент для зниження розмірності, навчання проєкційного шару за критерієм кутового відступу, експериментальне порівняння отриманих результатів на наборі даних VoxCeleb1. **Результати дослідження.** Запропоновано та досліджено три варіанти гібридизації на рівні вилучення ознак, такі як конкатенація нормованих векторів, зважена лінійна комбінація після зниження розмірності та злиття через навчений проєкційний шар. Отримані результати демонструють скорочення EER та підвищення точності ідентифікації порівняно з використанням кожного методу окремо. **Висновки.** Результати дослідження підтверджують доцільність застосування гібридизації методів на етапі вилучення ознак для підвищення точності ідентифікації мовця. Встановлено, що гібридизація шляхом злиття через навчений проєкційний шар забезпечує найвищу точність, тоді як конкатенація нормованих векторів не потребує додаткового навчання.

Ключові слова: біометричні системи голосової ідентифікації; гібридизація методів; вилучення ознак; MFCC; штучні нейронні мережі; згорткова мережа VGGish; навчання проєкційного шару; злиття ознак.

1. Вступ

1.1 Актуальність тематики

Голосова ідентифікація є актуальним способом контролю доступу, що використовується при створенні персоналізованих інтерфейсів, зокрема для розблокування авто, пристроїв, доступу до приміщень чи комп'ютерних систем [1–3].

Практичні умови застосування таких систем, зокрема мінливе акустичне середовище, фізіологічні варіації голосу, наявність фонового шуму та загрози підробки за допомогою синтезованого аудіо, передбачають суперечливі вимоги щодо точності ідентифікації, стійкості до міжсесійних варіацій та обчислювальної ефективності [3, 4].

Спосіб отримання векторного представлення мовця є основним компонентом архітектури системи біометричної ідентифікації, оскільки саме від якості вилучення ознак залежить подальша точність розрізнення користувачів [5]. Статистичні методи акустичного аналізу, такі як мел-частотні кепстральні коефіцієнти (MFCC), забезпечують інтерпретовані, відтворювані та обчислювально ефективні представлення, але усереднюють характеристики в часі та не відображають складну нелінійну динаміку сигналу [3, 6]. Глибокі нейронні мережі здатні автоматично виявляти складні закономірності, проте є вимоглими до апаратних

ресурсів та потребують великих навчальних вибірок і мають обмежену інтерпретованість [7–9].

Головними проблемними питаннями при цьому стають такі: необхідність поєднання переваг актуальних підходів до вилучення ознак без суттєвого ускладнення обчислювальної архітектури та відсутність універсальних методів їхньої інтеграції [10, 11].

Крім того, слід зазначити, що обчислювальні ресурси пристроїв для біометричної ідентифікації є обмеженими [12]. Це визначає необхідність пошуку методів ранньої інтеграції ознак, які б дозволили формувати компактні та інформативні вектори представлення мовця без суттєвого збільшення обчислювального навантаження.

Тому дослідження, присвячені формуванню ефективних гібридних представлень мовленнєвого сигналу шляхом інтеграції статистичних та глибинних ознак, є актуальним [13, 14]. Вирішення цього питання дозволить підвищити точність розпізнавання мовців в системах біометричної голосової ідентифікації за умов обмежених обчислювальних ресурсів.

1.2 Аналіз літературних джерел і визначення проблеми

У сучасних системах голосової ідентифікації дедалі частіше використовується поєднання різних

методів вилучення ознак з метою компенсації обмежень кожного з них [15-21].

Дослідження [15] розглядає голосову ідентифікацію як багатоетапний процес, що починається з перетворення акустичного сигналу шляхом екстракції 25 статистичних ознак. Цей етап охоплює розрахунок параметрів Гйорта, енергії Тігера та центральних моментів для детального опису морфології голосу мовця. Наступний рівень передбачає селекцію ознак на основі коефіцієнта кореляції Спірмена, що завершується етапом класифікації з використанням моделей на основі методів Decision Tree, Support Vector Machine (SVM) або Ensemble Bagged Tree для прийняття остаточного рішення. Такий підхід до обробки даних дозволив авторам досягти точності 90.16% для моделі SVM та 95.14% для алгоритму Decision Tree. Однак переважання морфологічних параметрів у часовій області обмежує здатність системи враховувати спектрально-динамічні властивості сигналу, що ускладнює часову сегментацію.

У дослідженні [16] пропонується багаторівнева гібридна модель розпізнавання голосу, що поєднує природно-інспіровані алгоритми оптимізації (такі як оптимізація роєм часток (PSO) та генетичні алгоритми (GA)) із системами нечіткої логіки [9]. Рішення ґрунтовано на первинному вилученні акустичних характеристик сигналу з подальшим їх використанням у текстонезалежному режимі. У підході використовуються такі ознаки мовного сигналу, як енергія сигналу (EE), частота перетину нуля (ZCR) та короткочасна енергія (STE), які формують вхідний вектор ознак для класифікатора. Це дозволяє ефективно ідентифікувати особу незалежно від змісту вимовлюваних фраз. Гібридизація методів дозволила досягти точності розпізнавання до 88% при середньоквадратичній похибці (MSE) на рівні 10%. Проте розширення простору ознак ускладнює оптимізацію та збіжність моделі [16].

У роботі [17] продемонстровано модель для виявлення голосового спуфінгу, яка поєднує попередньо навчену модель ШНМ глибокого навчання (VGGish) для автоматичного вилучення ознак та довготривалу короткочасну пам'ять (LSTM) для подальшої класифікації. Архітектура VGGish-LSTM подана як гібридна, оскільки вона поєднує різні архітектури ШНМ, що застосовуються послідовно для розв'язання окремих підзадач – вилучення ознак і класифікації. Запропонована гібридна модель продемонструвала високу ефективність, досягнувши точності 99.11% та рівня рівної помилки (EER) 1.18% під час тестування на наборі даних ASV Spoof 2019 LA. Водночас в роботі зазначено, що обробка великих обсягів даних потребує використання складних попередньо навчених моделей (YAMNet), що спричиняє дефіцит обчислювальних ресурсів ще на етапі вилучення ознак, а застосування рекурентних нейронних мереж для довгих послідовностей ускладнює процес навчання через проблему зникнення градієнта.

У дослідженні [18] розглядається гібридна екстракція ознак шляхом поєднання статистичних

MFCC-параметрів із векторами ознак, відібраними згортковою мережею (CNN). Автори проводять порівняльний аналіз кількох конфігурацій гібридизації. Однак оцінка ефективності підходу обмежена єдиним невеликим набором даних і не враховує стійкості системи до міжсесійних варіацій голосу та її обчислювальної складності при масштабуванні.

Робота [19] пропонує інтеграцію механізму уваги для динамічного зважування кількох груп ознак у задачах верифікації мовця та виявлення піддробленого голосу (SASV). Запропонована архітектура здійснює спільне навчання ASV- та антиспуфінг-гілок для підвищення точності розпізнавання [22, 23]. Однак висока складність моделі та необхідність сумісного тренування обох компонентів ускладнюють інтерпретацію внеску окремих джерел ознак і суттєво підвищують обчислювальні витрати.

У дослідженні [20] пропонується злиття без залучення нейромережових екстракторів фізичних часових і частотних параметрів (спектральний потік та нахил, тон, інтенсивність) з їхніми статистичними варіаціями. Такий підхід орієнтований на реалізацію текстонезалежної ідентифікації мовця на рівні ознак низького порядку. Однак відсутність порівняння з сучасними глибинними архітектурами (VGGish, x-vector, ECAPA-TDNN) звукує застосовність отриманих висновків для сценаріїв із високою міжсесійною варіативністю голосового сигналу.

Дослідження [21] розглядає архітектуру AMDF-CNN, яка забезпечує увага-кероване злиття (attention-guided feature fusion) часових, спектральних та кепстральних вимірів в єдиній згортковій структурі. Застосування багатовимірної моделі дозволило знизити частку помилкової класифікації мовця на 2.52% порівняно з базовими підходами. Оскільки автори не наводять порівняння обчислювальних витрат із простішими методами злиття, оцінити практичну ефективність такого ускладнення архітектури наразі неможливо.

Аналіз розглянутих актуальних досліджень свідчить про доцільність переходу від послідовного використання окремих методів до їх гібридизації. Це пов'язано з тим, що ізолювані алгоритми вже не забезпечують достатнього приросту ефективності, тому все частіше застосовуються їх поєднання в межах єдиної системи. У сучасних рішеннях статистичні методи та ШНМ використовуються разом як частини одного підходу. Класичні алгоритми з виділенням ознак і традиційними класифікаторами (SVM, дерева рішень) ефективні при однорідності даних і не потребують складних обчислень [23], але не підтримують роботу зі складними варіаціями мовлення та великими наборами даних. Натомість моделі на основі штучних нейронних мереж (LSTM, CNN) виявляють складні закономірності в аудіосигналі та враховують його часову структуру, але характеризуються високими вимогами до ресурсів, мають проблему зникаючого градієнта на довгих послідовностях та потребують великі масиви даних для навчання. Принципова

різниця між їх використанням полягає в тому, на якому етапі відбувається обмін інформацією між компонентами системи і чи формується в результаті єдине спільне представлення сигналу.

Отже на основі проведеного аналізу робіт [15-21] є підстави стверджувати, що доцільним є проведення дослідження, спрямованого на підвищення точності текстонезалежної ідентифікації мовця шляхом гібридизації методів вилучення акустичних ознак голосового сигналу.

1.3 Мета і завдання дослідження

Метою роботи є розробка та дослідження методу гібридного вилучення ознак для задач голосової ідентифікації, що поєднує MFCC-представлення та векторні ознаки нейромережі на основі VGGish. Передбачається формування спільного гібридного векторного опису мовленнєвого сигналу з подальшим порівнянням трьох підходів до злиття ознак: конкатенації, зваженої лінійної комбінації та злиття через навчений проєкційний шар [24]. Для досягнення мети були поставлені такі завдання:

- розробити підходи гібридизації методів вилучення статистичних та нейромережових ознак у єдиному векторному просторі;
- визначити критерій та показники якості процесу ідентифікації мовця;
- оцінити точність ідентифікації мовця за допомогою запропонованих способів гібридизації у режимах closed-set та open-set;
- дослідити обчислювальну ефективність і масштабованість запропонованих підходів при збільшенні банку користувачів.

2. Матеріали і методи дослідження

Об'єктом дослідження є гібридизація методів вилучення ознак у системі голосової ідентифікації на основі поєднання статистичного акустичного аналізу та штучної нейронної мережі. Реалізовано ідентифікацію мовця шляхом порівняння з еталонними профілями гібридного вектора ознак, який формується злиттям (через конкатенацію, зважену суму або використання проєкційного шару) MFCC- та VGGish-представлень мовленнєвого сигналу.

Основна гіпотеза дослідження: інтеграція статистичних та нейромережових ознак у єдиному векторному просторі на етапі вилучення ознак дозволить зменшити рівень рівної помилки (ERR) та підвищити точність ідентифікації порівняно з ізольованим застосуванням кожного з підходів. Гіпотеза ґрунтується на припущенні про взаємодоповнювальний характер цих представлень: ознаки, що вилучені за допомогою статистичних методів, описують явні спектральні відмінності між мовцями, тоді як ознаки, вилучені за допомогою ШНМ, відображають нелінійні закономірності, виявлені в процесі навчання [25].

Для реалізації підходу передбачається формування спільного векторного опису голосового сигналу з використанням різних стратегій злиття

ознак, що впливають на структуру фінального представлення.

Розробка та експериментальне оцінювання підходів гібридного вилучення ознак здійснювалися за дотримання таких крайових умов та обмежень:

Умова 1. Мовленнєвий сигнал розглядається як одноканальний із частотою дискретизації 16 кГц, що відповідає формату використаного набору даних.

Умова 2. Профіль користувача формується з фіксованої кількості висловлень, а еталонний вектор профілю обчислюється шляхом усереднення векторів цих висловлень.

Умова 3. Розмірності вихідних представлень є фіксованими: 240 ознак для статистичного методу та 128 ознак для нейромережового.

Умова 4. При навчанні шару злиття ваги базової нейромережової архітектури залишаються сталими, що запобігає деградації попередньо сформованих представлень.

Умова 5. Вплив адитивного акустичного шуму та багатомовності мовлення в межах даного дослідження не моделюється, а міжсесійна варіативність враховується лише через кількість висловлень у профілі користувача.

Процес голосової ідентифікації складається з кількох послідовних етапів, таких як попередня обробка вхідного аудіо сигналу $x(t)$, вилучення ознак, формування векторного представлення мовленнєвого сигналу, а також етап порівняння з еталонними профілями користувачів і прийняття рішення щодо ідентифікації.

На етапі вилучення ознак виконується перетворення мовленнєвого сигналу у компактний числовий вектор, що зберігає індивідуальні характеристики голосу та є стійким до акустичних варіацій між сесіями.

Вибір MFCC як статистичного методу вилучення ознак зумовлений обчислювальною ефективністю, можливостями інтерпретації та відтворення представлення мовного сигналу, не потребує навчання і враховує особливості слухового сприйняття людини [26].

На MFCC реалізовані класичні системи ідентифікації мовця на основі гауссових сумішей, а їх властивості та обмеження докладно проаналізовані в оглядових роботах [27]. Це дозволяє застосовувати вилучення ознак MFCC в умовах обмежених ресурсів та використовувати результати для порівняльного аналізу.

Сигнал розбивається на короткі перекривні кадри тривалістю 25 мс із кроком 10 мс [28]. Для кожного кадру застосовується вікно Геммінга для зменшення спектральних викривлень. Далі виконується швидке перетворення Фур'є та фільтрація через банк із 40 мел-фільтрів у діапазоні 80-7600 Гц.

Після цього обчислюється логарифм енергій фільтрів і застосовується дискретне косинусне перетворення, у результаті чого отримуються MFCC-коефіцієнти. Додатково обчислюються їх перші та другі похідні, які описують зміну ознак у часі.

$$D_t = \frac{\sum_{n=1}^N n \cdot (c_{t+n} - c_{t-n})}{2 \sum_{n=1}^N n^2} t, \quad (1)$$

де c_t – значення статичного MFCC-коефіцієнта у кадри t , а N – розмір часового вікна ($N = 2$).

Для кожного кадру формується вектор ознак, а для отримання фіксованого представлення всього сигналу, обчислюються середнє значення та стандартне відхилення по всіх кадрах:

$$\phi M(x) = [\mu_M; \sigma_M] \in \mathbb{R}^{240}, \quad (2)$$

де μ_M та σ_M – вектори середнього значення та стандартного відхилення ознак кадру по всій тривалості сигналу.

Головним обмеженням підходу є втрата часової структури сигналу внаслідок усереднення кадрів. Як наслідок, порядок появи локальних динамічних змін не враховується, а різні за часовою структурою сигнали отримують подібні представлення, що погіршує розрізнення мовців [16, 29].

Сучасною альтернативою опису сигналу через усереднення спектральних характеристик (зокрема MFCC) є використання згорткових нейронних мереж, які виконують автоматичне вилучення ознак безпосередньо з мел-спектрограми. У роботі у якості нейромережевого екстрактору застосовано модель на основі архітектури VGGish [30]. Завдяки попередньому навчанню на наборі аудіоподій, ця модель формує інформативні високорівневі представлення. Додатковими перевагами є компактність вихідного вектора розмірністю 128. Нейронну мережу було донаведено на цільовому наборі даних для її специфічної адаптації до задачі ідентифікації мовців.

Вхідні дані для мережі формуються у вигляді мел-спектрограми за схемою конвеєра попередньої обробки, наведеною на рис. 1. Аудіосигнал спочатку розбивається на короткі перекривні фрейми з застосуванням віконної функції, після чого для кожного фрейму обчислюється спектр за допомогою швидкого перетворення Фур'є.

На наступному етапі виконується проекція спектральної енергії на мел-шкалу з використанням набору фільтрів, після чого застосовується логарифмування отриманих значень. У результаті формується двовимірне представлення сигналу у вигляді матриці «час/частота». Для подальшої обробки мел-спектрограма фреймується на патчі розміром 96 часових кадрів на 64 мел-смуги, що відповідає інтервалу приблизно 0.96 с.

Архітектура мережі включає чотири згорткові блоки з ядрами 3×3 та кількістю фільтрів 64, 128, 256 і 256 відповідно. Після кожного згорткового блоку застосовується операція максимального об'єднання 2×2 для поступового зменшення просторової розмірності ознак.

Далі використовуються два повнозв'язні шари розмірністю 4096 і 128 нейронів.

Вихід останнього шару формує 128-вимірний вектор ознак:

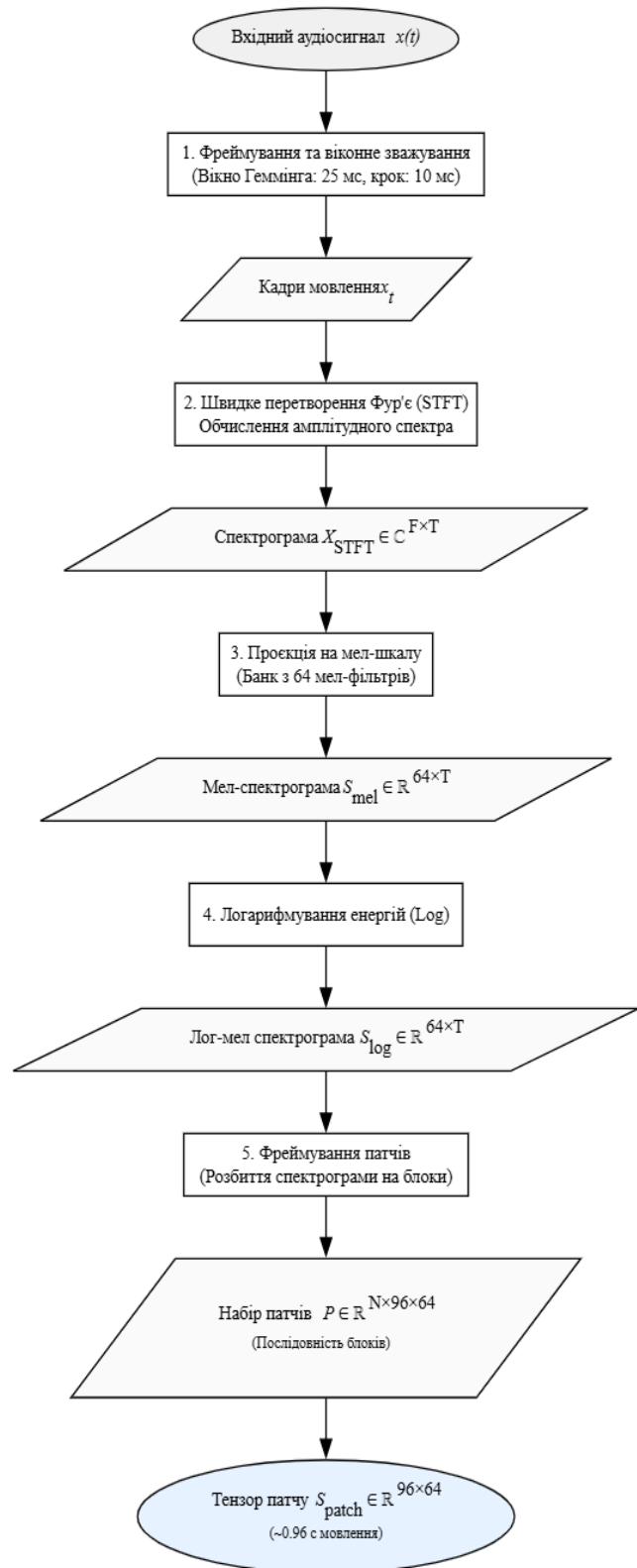


Рис. 1. Блок-схема конвеєра попередньої обробки аудіосигналу та формування мел-спектрограмних патчів

$$\phi V(x) = fVGGish(x; \theta) \in \mathbb{R}^{128}, \quad (3)$$

де $fVGGish$ – нелінійне перетворення вхідного сигналу x нейронною мережею з параметрами θ .

Навчання мережі виконується з використанням функції втрат на основі кутового відступу (Angular

Margin Loss), що обґрунтовано специфікою задачі ідентифікації, у якій рішення приймається за мірою подібності між векторними представленнями, а не за належністю до фіксованого класу. Це забезпечує формування розділених кластерів у просторі ознак для різних мовців і збільшує міжкласову відстань, що безпосередньо сприяє зниженню EER [10, 12]. Для сигналів довільної тривалості вихідний аудіопотік розбивається на патчі з перекриттям, після чого для кожного з них обчислюється вектор ознак, а фінальне представлення отримується шляхом усереднення по часовій осі [11].

На відміну від статичного методу вилучення ознак, запропонований метод на основі ШНМ VGGish дозволяє моделювати складні нелінійні залежності між спектральними компонентами та враховувати часову динаміку сигналу. Водночас такий підхід є обчислювально більш затратним і потребує достатнього обсягу навчальних даних для стабільної роботи моделі.

Методи вилучення ознак формують вектори різного типу: класичний метод описує локальні спектральні характеристики та їхню динаміку, нейромережовий моделює складні закономірності, виявлені під час навчання. Завдяки цьому описані методи є взаємодоповнюваними: на основі MFCC чутливий до явних спектральних відмінностей між мовцями, а на основі VGGish здатний виявляти малопомітні особливості мовлення та артикуляції, які складно формалізувати аналітично [3, 25, 29].

Для приведення представлень вилучених ознак різної розмірності до єдиного простору застосовано метод головних компонент, що виконує лінійне перетворення з мінімальною втратою дисперсії вихідних даних і не потребує навчання з учителем, завдяки чому вплив зниження розмірності можна відокремити від впливу самого способу злиття ознак у спільний простір.

Експериментальну реалізацію гібридної системи виконано за допомогою фреймворку PyTorch в інтерактивному середовищі Apache Zeppelin для структурованого виконання сценаріїв та забезпечення відтворюваності результатів експериментів. Обчислення проводилися на серверній платформі на основі процесора AMD EPYC 9354, графічного прискорювача NVIDIA A100 (80 ГБ VRAM) та 256 ГБ DDR5 ECC.

Як експериментальний набір даних обрано VoxCeleb1 (1251 мовець, понад 145 000 висловлювань) [31, 32]. Вибір обумовлено наявністю записів мовців у неконтрольованих акустичних умовах, що дає змогу оцінити стійкість методів до міжсесійних варіацій голосу, а також його ефективність у задачах ідентифікації мовця для забезпечення порівняльності результатів. Дані розділено на дві вибірки: 1151 мовець для навчання та формування профілів (по 3 висловлювання на користувача), і 100 мовців – для тестування.

Вибір показників FAR, FRR та EER у якості критеріїв обґрунтований тим, що FAR і FRR характеризують два протилежні типи помилок системи залежно від обраного порога прийняття

рішення, тоді як EER є інваріантним критерієм для порівняння конфігурацій.

Оцінювання у режимах closed-set (оцінка точності розпізнавання) та open-set (оцінка здатності відхиляти невідомих мовців) зумовлене необхідністю аналізу роботи системи у різних практичних сценаріях застосування. Додатково вимірювалися експлуатаційні характеристики: середній час обробки запиту, тривалість виконання окремих модулів, обсяг використаної VRAM, залежність часу ідентифікації від розміру банку користувачів та вплив кількості векторів у профілі на підсумкові показники.

3. Результати дослідження

3.1 Підходи гібридизації методів вилучення ознак

Об'єднання традиційних акустичних (MFCC) та нейромережових (VGGish) представлень у єдиний гібридний вектор $\phi H(x)$ дає змогу ефективно використати переваги обох підходів, поєднуючи класичні спектрально-часові характеристики з високорівневими абстрактними ознаками сигналу, вилученими за допомогою ШНМ [17, 30].

Запропоновано три підходи інтеграції векторів ознак сигналу, що відрізняються обчислювальною складністю: конкатенація нормованих векторів, зважена лінійна комбінація та злиття через навчений проєкційний шар.

Конкатенація нормованих векторів передбачає об'єднання векторів після їх нормування. Перед конкатенацією обидва вектори приводяться до одиничної норми за евклідовою метрикою, що дає змогу сформувати цілісний вектор розширеної розмірності без втрати первинної інформації з кожного джерела:

$$\begin{aligned}\hat{\phi}M &= \frac{\phi M}{\|\phi M\|_2}, \\ \hat{\phi}V &= \frac{\phi V}{\|\phi V\|_2},\end{aligned}\quad (4)$$

У подальшому конкатенуються в вектор:

$$\hat{\phi}H = [\hat{\phi}M; \hat{\phi}V] \in R^{368}. \quad (5)$$

Нормування перед конкатенацією запобігає домінуванню вектора з більшою нормою при обчисленні міри подібності та пригніченню внеску іншого джерела ознак. Після нормування обидва вектори мають рівноцінний вплив на формування гібридного представлення. Підсумковий вектор містить 240 статистичних та 128 нейромережових ознак, що забезпечує формування цілісного опису сигналу без втрати інформації з жодного із джерел. Запропонований підхід не потребує навчання та не вносить додаткових апроксимаційних похибок [33].

Спосіб зваженої лінійної комбінації передбачає приведення обох векторів до спільного простору однакової розмірності з подальшим їх зваженим поєднанням. Статистичний вектор має розмірність 240, тоді як нейромережовий 128. Для приведення їх

до спільного простору виконується лінійне перетворення статистичного вектора до 128 вимірів за допомогою матриці головних компонент, побудованої на навчальній вибірці:

$$\varphi' M = P \cdot \varphi M \in R^{128}, \quad (6)$$

де P – матриця проєкції розмірністю 128×240 , стовпці якої відповідають 128 головним компонентам з найбільшою дисперсією.

Після цього обидва вектори нормуються, і формується їх зважена лінійна комбінація:

$$\varphi H = \alpha \cdot \varphi' M + (1 - \alpha) \cdot \varphi V \in R^{128}, \quad (7)$$

де $\alpha \in [0, 1]$ ваговий коефіцієнт (від тільки нейромережевого при $\alpha = 0$ до проєктованого статистичного при $\alpha = 1$), який визначає відносний внесок джерел ознак і підбирається на валідаційній вибірці за критерієм мінімізації EER.

Спосіб злиття через навчений проєкційний шар передбачає нелінійне поєднання ознак за допомогою двошарової штучної нейронної мережі з навченими параметрами. На відміну від попередніх підходів, система самостійно навчається оптимальному способу інтеграції двох джерел ознак без використання припущень щодо їх відносної ваги або структури.

Обидва вектори спочатку конкатенуються у проміжний вектор розмірністю 368, який подається на вхід першого повнозв'язного шару:

$$h = \tanh(W_1 \cdot [\varphi M; \varphi V] + b_1), \quad (8)$$

$$W_1 \in R^{256 \times 368},$$

$$b_1 \in R^{256}.$$

Ваги базової архітектури залишаються сталими, що запобігає деградації попередньо сформованих представлень.

Отриманий у кожному з підходів гібридний вектор $\varphi H(x)$ використовується для формування профілів користувачів та їх порівняння з вхідним сигналом у банку ознак. На основі цих представлень виконується обчислення міри подібності між тестовими та еталонними векторами, що дозволяє здійснювати ідентифікацію мовця.

3.2 Критерій та показники якості процесу ідентифікації мовця

Для виконання оцінки ефективності розпізнавання сформовано систему кількісних показників якості, що ґрунтуються на вимірюванні кількості помилок першого та другого роду. Основними критеріями оцінки обрано метрики помилка хибного прийняття (FAR) та помилка хибного відхилення (FRR), які відображають залежність помилок системи ідентифікації від обраного порога прийняття рішення. Для забезпечення порогово-інваріантного порівняння різних підходів та їхніх конфігурацій визначено інтегральний показник ERR, який відповідає точці

рівноймовірної помилки ($FAR = FRR$). Система показників якості додатково враховує режими closed-set (для визначення загальної точності ідентифікації) та open-set (для аналізу стійкості до міжсесійних варіацій голосу та оцінки здатності системи відхиляти незареєстрованих мовців).

Окрім метрик точності, до системи оцінювання залучено експлуатаційні характеристики: середній час обробки одного запиту $T_{проц}$, тривалість роботи окремих модулів (статистичного та нейромережевого вилучення ознак), обсяг використаної відеопам'яті (VRAM), а також залежність часу ідентифікації від розміру банку користувачів. На основі виділених показників сформовано загальний критерій оцінки процесу ідентифікації мовця, який полягає у мінімізації сукупного рівня помилок EER за умови дотримання часового регламенту обробки сигналу:

$$\min_{\theta} EER(\theta) \text{ при } T_{проц}(\theta) \leq T_{макс}, \quad (9)$$

де θ – вектор параметрів конфігурації системи (метод гібридизації ознак, розмірність векторного представлення, поріг ухвалення рішення), $T_{проц}$ – середній час обробки одного запиту, $T_{макс}$ – гранично допустимий час відповіді системи.

3.3 Оцінка точності ідентифікації в режимах closed-set та open-set

Для оцінки впливу способу гібридизації класичного статичного та нейромережевого підходів на точність ідентифікації користувачів проведено експериментальне порівняння (табл. 1) розглянутих конфігурацій в умовах closed-set постановки. Оцінювання здійснювалося шляхом порівняння записаних тестових висловлювань мовців із заздалегідь сформованими еталонними профілями користувачів у закритій множині мовців.

Таблиця 1 – Порівняння методів вилучення ознак в умовах closed-set постановки

Конфігурація	FAR, %	FRR, %	ERR, %	Точність ідентифікації, %
Статистичний метод (MFCC)	9.2	11.2	10.5	87.6
Нейромережевий метод (VGGish)	4.1	5.1	4.8	93.8
Гібридизація (конкатенація)	3.6	4.8	4.2	94.7
Гібридизація (при $\alpha=0.35$)	3.8	3.1	4.4	94.3
Гібридизація (FC-fusion)	2.9	3.8	3.3	95.9

Основними критеріями порівняння є точність розпізнавання та рівень рівноймовірної помилки

EER. Додатково аналізувалася стабільність роботи системи ідентифікації при кожній з конфігурацій при зміні тривалості тестових записів, що дозволило встановити архітектуру гібридного вектора для практичного застосування.

За результатами експерименту найнижчі значення FAR (2.9%) та EER (3.3%), а також найвищу точність ідентифікації (95.9%) зафіксовано для гібридизації через злиття за допомогою навченого проєкційного шару.

Найнижче значення FRR (3.1%) отримано для зваженої лінійної комбінації при ваговому коефіцієнті $\alpha = 0.35$.

Для статистичного методу зафіксовано найвищі значення FAR (9.2%), FRR (11.2%) та EER (10.5%).

Експеримент (табл. 2) дозволяє оцінити стабільність роботи конфігурацій системи біометричної голосової ідентифікації в умовах open-set постановки при змінах голосових характеристик користувачів між сесіями запису, а також за появи незареєстрованих мовців.

Таблиця 2 – Результати оцінки стійкості системи до міжсесійних варіацій у open-set постановці

К-сть сесій	Метод	EER, %	FRR, %	Точність ідентифікації, %
1	Статистичний	5.5	89.8	90.2
1	Нейромережевий	4.9	91.5	92.1
1	Гібрид (оцінки)	4.6	92.3	92.8
1	Гібрид (рішення)	5.0	91.0	91.7
3	Статистичний	4.8	91.2	91.6
3	Нейромережевий	3.5	94.0	94.8
3	Гібрид (оцінки)	3.1	95.4	95.9
3	Гібрид (рішення)	3.4	94.6	95.1
5	Статистичний	4.5	92.6	92.9
5	Нейромережевий	3.0	95.8	96.3
5	Гібрид (оцінки)	2.7	96.9	97.2
5	Гібрид (рішення)	2.9	96.2	96.7

Найвищу точність ідентифікації (97.2%) зафіксовано для гібридизації на рівні оцінок при п'яти сесіях у профілі.

3.4 Оцінка обчислювальної ефективності та масштабованості

В процесі дослідження (табл. 3) аналізується вплив розміру банку користувачів на час обробки запиту та використання GPU-пам'яті для нейромережевого та гібридних методів.

За результатами експерименту при збільшенні банку користувачів з 50 до 1000 осіб час обробки

одного запиту зростає з 42 до 341 мс для нейромережевого методу, з 55 до 450 мс для гібридизації на рівні ознак та з 50 до 390 мс для гібридизації на рівні оцінок.

Обсяг використаної відеопам'яті змінювався в межах від 812 до 910 МБ.

Таблиця 3 – Порівняння обчислювальної ефективності залежно від розміру банку користувачів

Кількість користувачів	ШНМ VGGish, мс	Гібрид (ознаки), мс	Гібрид (оцінки), мс	GPU memory, МБ
50	42	55	50	812
100	57	75	68	820
200	89	120	105	835
500	176	230	205	860
1000	341	450	390	910

4. Обговорення результатів дослідження

Реалізовано підходи гібридизації на рівні вилучення ознак для задач голосової ідентифікації, що ґрунтуються на інтеграції статистичного акустичного аналізу та нейромережевого підходу в єдиному векторному просторі.

Запропонований гібридний підхід передбачає формування спільного представлення, в якому обидва джерела вилучених ознак присутні одночасно та рівноправно впливають на результат ідентифікації.

Порівняльний аналіз показав, що статистичний метод на основі MFCC продемонстрував найнижчу точність ідентифікації (87.6%) та найвищі значення FAR (9.2%), FRR (11.2%) і EER (10.5%) (табл. 1).

Застосування нейромережевої архітектури VGGish забезпечило підвищення точності до 93.8% та істотне зниження рівня EER до 4.8%, тоді як використання гібридизації дозволило додатково покращити результати: злиття через проєкційний шар (FC-fusion) забезпечило максимальну точність 95.9% та мінімальний EER 3.3% (табл. 1).

Експериментальні дослідження стійкості методів до міжсесійних варіацій показали, що збільшення кількості сесій у профілі з 1 до 5 забезпечує стабільність розглянутих конфігурацій (табл. 2).

При 1 сесії використання статистичного підходу вилучення ознак демонструє точність 90.2%, а використання нейромережевого методу на основі VGGish – 92.1%, гібридизація підвищує цей показник до 92.8%.

При розширенні профілю до 5 сесій гібридний метод дозволяє досягти мінімального рівня EER 2.7% та максимальної точності ідентифікації 97.2% (табл. 2), що є найкращими показниками серед усіх

розглянутих методів у контексті open-set задач. Оцінка обчислювальної ефективності свідчить, що зі збільшенням кількості профілів у банку з 50 до 1000 осіб час обробки зростає майже лінійно для всіх конфігурацій системи голосової ідентифікації (табл. 3).

Мінімальний час відповіді забезпечує нейромережовий метод (від 42 до 341 мс) завдяки відсутності додаткових етапів обробки. Гібридизація на рівні ознак вимагає найбільших обчислювальних ресурсів (55-450 мс) через раннє об'єднання векторів, тоді як гібридизація на рівні оцінок поступається нейромережовому підходу за часом обробки на 15-25% (50-390 мс) (табл. 3), проте є обчислювально ефективнішою при збереженні високої точності. Найбільшу частку ресурсів використовує нейромережовий компонент, про що свідчить незначне коливання GPU-пам'яті в межах 812-910 МБ при розширенні бази з 50 до 1000 користувачів (табл. 3).

Отримані результати підтвердили, що гібридизація на рівні вилучення ознак перевершує окреме використання будь-якого з методів (табл. 1, 2). Це пояснюється взаємодоповненням статистичних та нейромережових представлень: явні спектральні характеристики та глибинні закономірності разом формують більш інформативний опис голосового сигналу мовця, ніж кожен підхід окремо. Перевага гібридних методів зростає зі збільшенням міжсесійної варіативності, що є важливим для практичного застосування систем голосової ідентифікації в біометричних комплексах аутентифікації та контролю доступу.

Застосування запропонованих рішень має обмеження: нейромережовий компонент потребує GPU-прискорення, а при розширенні бази до 1000 користувачів час обробки зростає до 390-450 мс (табл. 3), що важливе для високонавантажених сервісів реального часу.

Крім того, у роботі не досліджено вплив викривлень, зокрема, каналних завад (наприклад, при стисненні сигналу в телефонних мережах). Подальші дослідження передбачають впровадження алгоритмів швидкого векторного пошуку для масштабування бази мовців та стисненні ШНМ для її використання без GPU.

5. Висновки

1. Розроблено три підходи гібридизації статистичного (MFCC) та нейромережового (VGGish) методів вилучення ознак: на основі

конкатенації, зваженої комбінації (з PCA) та злиття через навчений шар (FC-fusion). На відміну від послідовного аналізу мовного сигналу, обидва джерела ознак використовуються одночасно у спільному векторному просторі. Гібридизація через конкатенацію не потребує навчання, зважена комбінація зберігає компактність вектора, а злиття через проєкційний шар забезпечує адаптацію для врахування нелінійних взаємозв'язків.

2. Визначено критерій та показники якості ідентифікації мовця на основі метрик FAR, FRR та інтегрального показника EER у режимах closed-set й open-set. За критерій якості обрано мінімізацію сукупного рівня помилок EER за умови дотримання часового регламенту обробки ($T_{\max} \leq T_{\text{проц}}$).

3. Оцінено точність ідентифікації в режимах роботи closed-set та open-set. Гібридизація методів вилучення ознак демонструє вищу точність, ніж при ізольованому використанні через взаємодоповнення спектральних та глибинних ознак. У режимі роботи closed-set найкращий результат ідентифікації забезпечує злиття через проєкційний шар (EER 3.3%, точність 95.9%), а в режимі open-set при 5 сесіях у профілі – гібридизація на рівні оцінок (EER 2.7%, точність 97.2%).

4. Досліджено обчислювальну ефективність та масштабованість роботи при розширенні банку мовців до 1000 осіб. Гібридизація на рівні оцінок збільшила час обробки лише на 15-25% у порівнянні з методом на основі VGGish при сталому споживанні GPU-пам'яті (до 910 МБ).

Конфлікт інтересів

Автори декларують, що не мають конфлікту інтересів стосовно даного дослідження, в тому числі фінансового, особистісного характеру, авторства чи іншого характеру, що міг би вплинути на дослідження та його результати, представлені в даній статті.

Фінансування

Дослідження проводилося без фінансової підтримки.

Доступність даних

Рукопис не має пов'язаних даних.

Використання засобів штучного інтелекту

Автори підтверджують, що не використовували технології штучного інтелекту при створенні представленої роботи.

СПИСОК ЛІТЕРАТУРИ

1. Chung, J.S., Huh, J., Mun, S., Lee, M., Heo, H.-S., Choe, S., Ham, C., Jung, S., Lee, B.-J., and Han, I. (2020), "In Defence of Metric Learning for Speaker Recognition", *Interspeech 2020*, Shanghai, China, pp. 2977–2981, doi: <https://doi.org/10.21437/Interspeech.2020-1064>
2. Chen, Z. et al. (2022), "Large-Scale Self-Supervised Speech Representation Learning for Automatic Speaker Verification", *ICASSP 2022 – 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Singapore, pp. 6147–6151, doi: <https://doi.org/10.1109/ICASSP43922.2022.9747814>
3. Ramzan, H.A., Zahra, S., Ramzan, S., Musa, S.M.A., Kalsum, T., and Jalil, R. (2024), "Speaker Identification Using Convolutional Neural Networks and Mel-Frequency Cepstral Coefficients", *2024 5th International Conference on Innovative Computing (ICIC)*, Lahore, Pakistan, pp. 1–6, doi: <https://doi.org/10.1109/ICIC63915.2024.11116334>

4. Hershey, S., Chaudhuri, S., Ellis, D.P.W., Gemmeke, J.F., Jansen, A., Moore, R.C., Plakal, M., Platt, D., Saurous, R.A., Seybold, B., Slaney, M., Weiss, R.J., and Wilson, K. (2017), "CNN architectures for large-scale audio classification", *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, New Orleans, LA, USA, pp. 131–135, doi: <https://doi.org/10.1109/ICASSP.2017.7952132>
5. Snyder, D., Garcia-Romero, D., Sell, G., Povey, D., and Khudanpur, S. (2018), "X-Vectors: Robust DNN Embeddings for Speaker Recognition", *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Calgary, AB, Canada, pp. 5329–5333, doi: <https://doi.org/10.1109/ICASSP.2018.8461375>
6. Aliaskar, M., Mazakov, T., Mazakova, A., Jomartova, S., and Shormanov, T. (2022), "Human voice identification based on the detection of fundamental harmonics", *2022 IEEE 7th International Energy Conference (ENERGYCON)*, Riga, Latvia, pp. 1–4, doi: <https://doi.org/10.1109/ENERGYCON53164.2022.9830471>
7. Okabe, K., Koshinaka, T., and Shinoda, K. (2018), "Attentive Statistics Pooling for Deep Speaker Embedding", *Interspeech 2018*, Hyderabad, India, pp. 2252–2256, doi: <https://doi.org/10.21437/Interspeech.2018-993>
8. Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., and Chintala, S. (2019), "PyTorch: An Imperative Style, High-Performance Deep Learning Library", *Advances in Neural Information Processing Systems 32 (NeurIPS 2019)*, pp. 8024–8035, URL: <https://papers.nips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library>
9. Snyder, D., Garcia-Romero, D., Povey, D., and Khudanpur, S. (2017), "Deep Neural Network Embeddings for Text-Independent Speaker Verification", *Interspeech 2017*, Stockholm, Sweden, pp. 999–1003, doi: <https://doi.org/10.21437/Interspeech.2017-620>
10. Бондаренко М.Є., Іващенко Г.С. Багатовекторна голосова ідентифікація користувача з урахуванням фонетичних варіацій голосу. *Системи управління, навігації та зв'язку*. Полтава: ПНТУ, 2026. Вип. 1 (83). С. 40–44. doi: <https://doi.org/10.26906/SUNZ.2026.1.040>
11. Бондаренко М.Є., Іващенко Г.С. Методи оновлення профіля користувача в системах багатовекторної голосової ідентифікації. *Вісник Хмельницького національного університету. Технічні науки*. 2026. № 2 (363). С. 371–378. doi: <https://doi.org/10.31891/2307-5732-2026-363-51>
12. Бондаренко М.Є., Іващенко Г.С. Використання послідовності методів попередньої обробки в системах голосової ідентифікації. *Системи управління, навігації та зв'язку*. Полтава: ПНТУ, 2025. Вип. 2 (80). С. 90–96. doi: <https://doi.org/10.26906/SUNZ.2025.2.090>
13. Nassif, A.B., Shahin, I., Attili, I., Azzeh, M., and Shaalan, K. (2019), "Speech Recognition Using Deep Neural Networks: A Systematic Review", *IEEE Access*, vol. 7, pp. 19143–19165, doi: <https://doi.org/10.1109/ACCESS.2019.2896880>
14. Jahangir, R., Teh, Y.W., Memon, N.A., Mujtaba, G., Zareei, M., Ishtiaq, U., Akhtar, M.Z., and Ali, I. (2020), "Text-Independent Speaker Identification Through Feature Fusion and Deep Neural Network", *IEEE Access*, vol. 8, pp. 32187–32202, doi: <https://doi.org/10.1109/ACCESS.2020.2973541>
15. Bora, B., Emanet, A.E., Elmacı, E., Kandaz, D., and Uçar, M.K. (2023), "Hybrid AI-based Voice Authentication", *Turkish Journal of Forecasting*, vol. 7, no. 2, pp. 17–22, doi: <https://doi.org/10.34110/forecasting.1260073>
16. Al-Tekreeti, N., and Ibrahim, A. A. (2020), "Speaker Voice Recognition Using A Hybrid PSO/Fuzzy Logic System", *2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, Istanbul, Turkey, pp. 1–5, doi: <https://doi.org/10.1109/ISMSIT50672.2020.9254309>
17. Shahzad, K., Farhan, S., Yasin-Ul-Haq, Sana, R., and Pathan, S. (2025), "Enhancing Voice Spoofing Detection: A Hybrid Approach With VGGish-LSTM Model for Improved Security in Automatic Speaker Verification Systems", *IEEE Access*, vol. 13, pp. 40682–40702, doi: <https://doi.org/10.1109/ACCESS.2025.3544639>
18. Hasan Abdulqader, A., Al-Haddad, S.A.R., Abdo, S., Abdulghani, A., and Natarajan, S. (2022), "Hybrid Feature Extraction MFCC and Feature Selection CNN for Speaker Identification Using CNN: A Comparative Study", *2022 2nd International Conference on Emerging Smart Technologies and Applications (eSmarTA)*, Marib, Yemen, doi: <https://doi.org/10.1109/eSmarTA56775.2022.9935422>
19. Shen, Q., Guo, M., and Huang, Y. et al. (2024), "Attentional multi-feature fusion for spoofing-aware speaker verification", *International Journal of Speech Technology*, vol. 27, pp. 377–387, doi: <https://doi.org/10.1007/s10772-024-10112-w>
20. Shah, Z., Jang, G., and Farooq, A. (2024), "Feature Fusion for Performance Enhancement of Text Independent Speaker Identification", *ICCK Transactions on Intelligent Systematics*, vol. 2, no. 1, pp. 27–37, doi: <https://doi.org/10.62762/TIS.2024.649374>
21. Karthikeyan, V., Priyadharsini, S.S., and Balamurugan, K. (2025), "Attention-based multi dimension fused-feature convolutional neural network framework for speaker recognition", *Multimedia Tools and Applications*, vol. 84, pp. 38927–38957, doi: <https://doi.org/10.1007/s11042-025-20694-5>
22. Variani, E., Lei, X., McDermott, E., Lopez-Moreno, I., and Gonzalez-Dominguez, J. (2014), "Deep neural networks for small footprint text-dependent speaker verification", *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Florence, Italy, pp. 4052–4056, doi: <https://doi.org/10.1109/ICASSP.2014.6854363>
23. Todisco, M., Wang, X., Vestman, V., Sahidullah, Md., Delgado, H., Nautsch, A., Yamagishi, J., Evans, N., Kinnunen, T., and Lee, K. A. (2019), "ASVspoof 2019: Future Horizons in Spoofed and Fake Audio Detection", *Interspeech 2019*, Graz, Austria, pp. 1008–1012, doi: <https://doi.org/10.21437/Interspeech.2019-2249>
24. Tak, H., Patino, J., Todisco, M., Nautsch, A., Evans, N., and Larcher, A. (2021), "End-to-End anti-spoofing with RawNet2", *2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Toronto, ON, Canada, pp. 6369–6373, doi: <https://doi.org/10.1109/ICASSP39728.2021.9414234>
25. Semenov, S., Davydov, V., and Grynov, D. (2024), "Study of Machine Learning Methods for Voice Biometric Identification in Data Protection Systems", *2024 14th International Conference on Dependable Systems, Services and Technologies (DESSERT)*, Athens, Greece, pp. 1–6, doi: <https://doi.org/10.1109/DESSERT65323.2024.11122211>
26. Hsu, W.-N., Bolte, B., Tsai, Y.-H. H., Lakhotia, K., Salakhutdinov, R., and Mohamed, A. (2021), "HuBERT: Self-Supervised Speech Representation Learning by Masked Prediction of Hidden Units", *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 3451–3460, doi: <https://doi.org/10.1109/TASLP.2021.3122291>

27. Wang, H., Wang, Y., Zhou, Z., Ji, X., Gong, D., Zhou, J., Li, Z., and Liu, W. (2018), "CosFace: Large Margin Cosine Loss for Deep Face Recognition", *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, UT, USA, pp. 5265–5274, doi: <https://doi.org/10.1109/CVPR.2018.00552>
28. Zeghidour, N., Teboul, O., de Chaumont Quitry, F., and Tagliasacchi, M. (2021), "LEAF: A Learnable Frontend for Audio Classification", *International Conference on Learning Representations (ICLR 2021)*, 16 p., available at: <https://openreview.net/pdf?id=jM76BCb6F9m>
29. Desplanques, B., Thienpondt, J., and Demuynck, K. (2020), "ECAPA-TDNN: Emphasized Channel Attention, Propagation and Aggregation in TDNN Based Speaker Verification", *Interspeech 2020*, Shanghai, China, pp. 3830–3834, doi: <https://doi.org/10.21437/Interspeech.2020-2650>
30. Kurada, S., and Kurada, A. (2020), "Poster: VGGish Embeddings Based Audio Classifiers to Improve Parkinson's Disease Diagnosis", *2020 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*, Crystal City, VA, USA, pp. 9–11, doi: <https://doi.org/10.1145/3384420.3431775>
31. Nagrani, A., Chung, J. S., and Zisserman, A. (2017), "VoxCeleb: A Large-Scale Speaker Identification Dataset", *Interspeech 2017*, Stockholm, Sweden, pp. 2616–2620, doi: <https://doi.org/10.21437/Interspeech.2017-950>
32. Chung, J.S., Nagrani, A., and Zisserman, A. (2018), "VoxCeleb2: Deep Speaker Recognition", *Interspeech 2018*, Hyderabad, India, pp. 1086–1090, doi: <https://doi.org/10.21437/Interspeech.2018-1929>

Received (Надійшла) 30.06.2026

Accepted for publication (Прийнята до друку) 26.08.2026

Publication date (Дата публікації) 25.09.2026

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Івашенко Георгій Станіславович – кандидат технічних наук, доцент, доцент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Heorhii Ivashchenko – Candidate of Technical Sciences, Associate Professor at the Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: heorhii.ivashchenko@nure.ua; ORCID Author ID: <http://orcid.org/0000-0003-1027-5262>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57217030807>.

Бондаренко Максим Едуардович – аспірант кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Maksym Bondarenko – PhD student at the Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: maksym.bondarenko@nure.ua; ORCID Author ID: <http://orcid.org/0000-0002-2500-7626>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57216159508>.

Hybridization of feature extraction methods in a user voice identification system

Maksym Bondarenko, Heorhii Ivashchenko

Abstract. Background. The practical application of voice identification systems necessitates meeting exclusive criteria: identification accuracy, computational efficiency and resistance to synthesised voice forgeries. Statistical approaches, such as acoustic analysis and spectral comparison, place low demands on hardware resources but are inferior in accuracy to modern deep learning architectures. At the same time, deep neural networks, despite their higher accuracy, significantly increase computational costs and sensitivity to distortions. One possible way of reconciling these requirements is to hybridise methods at the feature extraction stage, which involves combining statistical acoustic analysis and a neural network approach within a single vector space. **Object of the study:** the hybridisation of feature extraction methods in a voice identification system based on the combination of statistical acoustic analysis and a neural network. **Subject of the study:** the hybrid feature extraction method that combines MFCC vectors and neural network feature vectors within a single vector space, with a comparison of different ways of merging them. **Mera Objective:** to develop and investigate a hybrid feature extraction method that combines MFCC vectors and neural network feature vectors in a single vector space, with a view to improving the accuracy of voice identification. **Research tasks** are to implement the proposed approaches for hybridising statistical and neural network features within a single vector space; to experimentally evaluate the identification accuracy of the proposed variants in closed-set and open-set modes; to investigate the computational efficiency and scalability of the proposed variants. **Research methods.** Statistical acoustic analysis based on mel-frequency cepstral coefficients, a convolutional neural network based on the VGGish architecture, the principal component analysis method for dimensionality reduction, training of the projection layer using the angular deviation criterion, and an experimental comparison of the results obtained on the VoxCeleb1 dataset. **Results.** Three variants of hybridisation at the feature extraction stage were proposed and investigated, namely concatenation of normalised vectors, a weighted linear combination following dimensionality reduction, and merging via a trained projection layer. Results demonstrate a reduction in EER and an improvement in identification accuracy compared with the use of each method individually. **Conclusions.** The results of the study confirm the feasibility of applying method hybridisation at the feature extraction stage to improve speaker identification accuracy. It has been established that merging via a trained projection layer provides the highest accuracy, whilst concatenation of normalised vectors is a practically effective solution without the need for additional training.

Keywords: biometric systems of voice identification; hybridisation of methods; feature extraction; MFCC; artificial neural networks; VGG-style convolutional network; projection layer training; feature fusion.

О. Г. Дрючко, Н. В. Бунякіна, О. В. Шефер, В. М. Галай, Р. В. Захарченко

Національний університет «Полтавська політехніка імені Юрія Кондратюка», Полтава

ОСОБЛИВОСТІ СИНТЕЗУ І ХАРАКТЕРНІ ВЛАСТИВОСТІ ПЕРОВСКІТОПОДІБНИХ ШАРУВАТИХ ОКСИДНИХ МАТЕРІАЛІВ РІДКІСНОЗЕМЕЛЬНИХ І ПЕРЕХІДНИХ ЕЛЕМЕНТІВ ЯК ФУНДАМЕНТАЛЬНІ АСПЕКТИ ДЛЯ ПРАКТИКИ ІННОВАЦІЙНИХ ФОТОКАТАЛІТИЧНИХ РІШЕНЬ

Анотація. Актуальність. Використання перовскітоподібних шаруватих оксидних фаз та їх попередників комплексних нітратів лантанодів відкриває інноваційний підхід для керованого синтезу високоактивних наноконструкцій перехідних і рідкісноземельних елементів. **Об'єкти дослідження:** Прекурсори – лужні координаційні нітрати рідкісноземельних елементів. Цільові системи – одношарові та шаруваті наноструктуровані оксидні високовпорядковані системи лантанодів і перехідних елементів (зокрема, покриття на основі високодисперсного діоксиду титану анатазної модифікації). **Мета статті:** Узагальнення відомостей про умови утворення, будову та властивості координаційних нітратів лантанодів та розроблення на їхній основі методології фізико-хімічного модифікування центрів окиснення дифузного шару наноструктурованих оксидних зразках покриттів. **Висновки.** Одержані дані (як первинна інформація) є основою для виявлення, ідентифікації, контролю фазового стану об'єктів перероблення у підготовчих стадіях, вибору критеріїв сумісності складових при формуванні одношарових і шаруватих наноструктурованих оксидних композиційних систем лантанодів і перехідних елементів широкого призначення, з каталітичною і фотокаталітичною активністю, покриття здатного самоочищатися з гідрофільними властивостями; розроблення різних комбінованих способів їх активації та встановлення технологічно-функціональних залежностей; керованого модифікування властивостей одержуваних цільових продуктів.

Ключові слова: лужні координаційні нітрати лантанодів; умови утворення; кристалічна будова сполук; характерні властивості; перетворення з фізичною активацією; оксидні високовпорядковані системи лантанодів і перехідних елементів.

Вступ

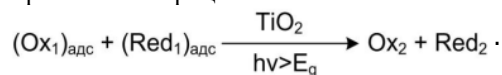
Нині продовжується пошук методів і комплексних технологій по створенню нових й удосконаленню існуючих регламентів одержання досконалих багатофункціональних оксидних матеріалів перехідних і рідкісноземельних елементів зі структурою дефектного перовскіта, граната із відтворюваними властивостями низькотемпературними методами «м'якої хімії» та з використанням нітратних прекурсорів. Вони мають складну будову і у науковому й технологічному відношенні становлять собою непрості об'єкти, що інтенсивно досліджуються [1, 2], в тому числі з участю авторів [3, 4]. Тому сучасне матеріалознавство, яке базується на їх основі, потребує регламентних рішень простих за конфігурацією, малостадійних, енергоефективних, таких, що характеризуються масштабністю, з можливістю відтворення продуктів із заданими однорідністю, стабільністю, комплексом наперед заданих характеристик.

Існує низка методик синтезу вказаних оксидних багатокомпонентних матеріалів [5, 6], заснованих на різних фізичних і хімічних принципах. Вибір використання методу залежить від хімічної природи отримуваних сполук, розміру і морфологічних особливостей частинок продуктів синтезу, умов і способу формування останніх; матеріалу і стану

поверхні, форми зразків, на які наноситься покриття; можливостей наявного технічного оснащення тощо. Ці методи використовують як незалежно один від одного, так і в комбінації.

Останнім часом діоксид титану привертає особливу увагу у зв'язку з новими унікальними перспективами його застосування у формі наноструктурованих матеріалів і наноконструкцій з контрольованими морфологічними, фізико-хімічними та оптичними властивостями. TiO_2 , який володіє високою хімічною і термічною стабільністю, а також домішковими рівнями в електронній структурі матеріалу, створюваними за рахунок заданого типу легування, є унікальним для побудови на його основі нових ефективних функціональних матеріалів, що застосовуються у фотокаталізі і фотовольтаїці, сенсоріці, каталізі, для рідинної хроматографії та інших сферах.

Суть фотокаталітичних властивостей TiO_2 полягає в тому, що в об'ємі напівпровідникової частинки під впливом електромагнітного випромінювання генеруються електрон – діркові пари, які при виході на поверхню частинки TiO_2 вступають в окислювально-відновні реакції з адсорбованими молекулами. Для діоксиду титану характерний такий процес:



У ході перетворень частина електронів і дірок може здійснити рекомбінацію у масі або на поверхні TiO_2 . Для ефективного перебігу фотокаталітичних процесів необхідно, щоб окислювально-відновлювальні реакції за участю пари електрон-дірка були ефективнішими, ніж процеси рекомбінації.

Добре відомі переваги використання такого фотокаталітичного способу очищення:

1) можливість окислити практично будь-які органічні речовини і також ряд неорганічних, таких як CO , H_2S , HCN , NH_3 , NO_x та ін.;

2) метод працює при кімнатній температурі і атмосферному тиску;

3) можна окислювати навіть невеликі концентрації забруднювачів, очищення від яких іншими методами є економічно не вигідним;

4) для реалізації методу фотокаталітичного очищення не потрібні додаткові реагенти, тому що окислювачем є кисень повітря.

При створенні й тепловому обробленні матеріалів на основі діоксиду титану потрібно враховувати можливість виникнення поверхневих і об'ємних дефектів кристалічної решітки TiO_2 (в наслідок існування різноманіття фаз в інтервалі співвідношень $\text{O} / \text{Ti} < 2$ на фазовій діаграмі системи $\text{Ti}-\text{O}$ [7]) та виявлення їх значного впливу на фотохімічні властивості і фотоіндуковану гідрофільну здатність продуктів синтезу [8].

Наявні відомості щодо стану й можливих напрямів удосконалення технологій створення таких матеріалів, чинні вимоги до їх стабільності й відтворюваності властивостей, розширення сфер їх використання [1–8], прояв високої активності малозакристалізованими частками структурних компонентів, сформованими термолізом розчинника [9], нові відомості щодо реакційної здатності й перетворення шаруватих перовскітоподібних оксидів, стабілізації фотокаталітично- і сенсорно-активної кристалічної модифікації анатаза за рахунок іонів NO_3^- [10], допірування Ln_2O_3 [11] при одержанні TiO_2 із розчинів [12] ініціювали продовження нашого дослідження за цією тематикою. Сьогодні з'ясовуються способи управління технічними параметрами цільових продуктів через вибір складу, умов синтезу та способу оброблення.

Одним із найбільш перспективних класів складних оксидних матеріалів рідкісноземельних елементів і титану є наноструктуровані шаруваті перовскітоподібні сполуки і тверді розчини на їхній основі. Залежно від складу і структури, вони мають широкий спектр фізико-хімічних властивостей. Представлені в даній роботі перовскітоподібні шаруваті титанати належать гомологічному ряду $(\text{Me}, \text{Ln})_{n+1}\text{Ti}_n\text{O}_{3n+1}$, де $\text{Ln} - \text{La} - \text{Nd}$, $\text{Me} - \text{Li} - \text{Cs}$, n – число наночарів перовскіта (фаз Руддлессен-Поппера; з товщиною одного шару приблизно 0,5 нм). Відповідно, MeLnTiO_4 у своїй структурі містить один наночар перовскіту, $\text{Me}_2\text{Ln}_2\text{Ti}_3\text{O}_{10}$ – три. Перовскітоподібні наночаруваті титанати були одержані і досліджені авторами робіт [13, 14].

Шарувата структура, що складається з пластинчастих частинок, зберігається продуктом при витримуванні у вологому повітрі і підтверджена методом скануючої електронної мікроскопії [15]. Виявлена стабільність систем рекомендує їх як перспективні фотокаталізатори в умовах близьких до їх використання та є важливою для інших інноваційних сфер їх застосувань.

Аналіз публікацій останніх років засвідчує, що діоксид титану має основне застосування саме у тонко плівковому виді, у якому найефективніше реалізуються його властивості, необхідні для фотокаталізу, сонячної енергетики, сенсорики, самоочищення покриттів тощо.

Нині завдяки технологічним прийомам реакцій «м'якої хімії» з'явилася можливість створення речовин із різноманітними структурними особливостями, отримання метастабільних сполук шляхом послідовності низькотемпературних топомічних синтезів. Такі реакції зі зміною будови і морфології частинок протікають при невисоких температурах зі збереженням основних структурних особливостей у перовскітоподібних шаруватих оксидних сполуках. Залежно від природи і стехіометрії катіонів, що входять до їх складу, вони можуть проявляти різноманітні фізичні і хімічні властивості: надпровідність, колосальний магнітоопір, сегнетоелектрику, каталітичну і фотокаталітичну активність, здатність до іонного обміну в розчинах і розплавах, здатність до гідратації міжшарового простору та інші. Тому вивчення особливостей перетворень проміжних попередників – лужних координаційних нітратів РЗЕ, їх реакційної здатності в ході синтезу шаруватих перовскітоподібних оксидних фаз безпосередньо впливає на можливі сфери подальшого застосування останніх.

Для шаруватих перовскітоподібних сполук до таких процесів сучасні дослідники відносять, зокрема, іонний обмін [16], інтеркаляцію і деінтеркаляцію [15], різні процеси заміщення й конденсації [17], процеси розщеплення [18] і взаємні перетворення однієї структури на іншу [19] (наприклад, перехід із фаз Руддлессен-Поппера у фази Діона-Якобсона; перехід у межах одного типу фаз зі збільшенням або зменшенням числа шарів).

Для ефективного управління властивостями одержуваних продуктів необхідне глибоке розуміння фізико-хімічних процесів, явищ, що відбуваються в ході їх формування. Їх комплексне дослідження із застосуванням сучасних фізико-хімічних методів дозволяє удосконалювати наші знання про характерні особливості швидкоплинних процесів, стадії еволюції структури й мікроструктури технологічних об'єктів.

Мета і завдання дослідження

Мета дослідження: провести дослідження кооперативних процесів взаємодії між структурними компонентами у системах нітратних попередників представників РЗЕ і елементів ІА підгрупи періодичної системи в умовах аналогічних

до регламентних при створенні багатофункціональних оксидних матеріалів різного призначення з використанням фотокаталітично активного TiO_2 з метою формування достовірних уявлень і одержання об'єктивних знань про особливості перетворень і сукупну поведінку складових елементів у підготовчих стадіях перероблення технологічних об'єктів з тепловою активацією (25-1000 °C), необхідних для удосконалення й розроблення методологій і регламентів сучасних технологій їх виготовлення.

Для досягнення поставленої мети у роботі поетапно вирішувалися наступні завдання:

1) вивчення механізмів перетворень у системах координаційних нітратних прекурсорів РЗЕ і лужних металів у ході формування поліфункціональних фотокаталітично активних шаруватих оксидних матеріалів;

2) розроблення методології і виготовлення зразків фотокаталізаторів на основі TiO_2 анатазної модифікації і композиційних зі структурою тришарового титанату $\text{K}_2\text{Ln}_2\text{Ti}_3\text{O}_{10}$ з їх двостадійним нанесенням і формуванням на структурованих металевих носіях;

3) вивчення кінетичних закономірностей реакцій тестового фотокаталітичного окислення парів органічних речовин (на прикладі етанолу) в статичному реакторі;

4) встановлення технологічно-функціональних залежностей; керованого модифікування властивостей одержуваних цільових продуктів.

Методологія експерименту

Для оцінки можливості керування процесами багатостадійного формування складно-оксидних композицій з поліфункціональними властивостями й обґрунтування механізмів фазоутворення як модельні із використанням комплексу фізико-хімічних методів вивчені водно-сольові системи нітратів $\text{MeNO}_3 - \text{Ln}(\text{NO}_3)_3 - \text{H}_2\text{O}$, ($\text{Me} - \text{Li, Na, K, Rb, Cs; Ln} - \text{La-Lu}$) при 25-100 °C. Вибір складу об'єктів дослідження, температурні перерізи зумовлені рядом чинників.

По-перше, серед елементів рідкісноземельного ряду вищу комплексоутворюючу здатність виявляють представники церієвої підгрупи; серед них найбільші зміни складу, структури, властивостей їх сполук – елементи його середини, Pr і Nd. Вибрані компоненти систем задають технічні характеристики цільового продукту чи є модифікаторами його властивостей. А наявність великої кількості для використання потенціальних електронних аналогів (представників природних рядів рідкісноземельних, лужних елементів) зумовлює значну варіативність і широту діапазону модифікування їх характеристик. Температурні перерізи обумовлені областями існування кристалогідратних форм вихідних компонентів.

По-друге, за результатами досліджень авторів [15] трьох шарові калієві титанати $\text{K}_2\text{Ln}_2\text{Ti}_3\text{O}_{10}$ ($\text{Ln} - \text{La, Nd}$) вищевказаних складових елементів, одержані за керамічною технологією, у

суспензійному вигляді у водно-спиртових розчинах під дією УФ випромінювання виявляють найвищу серед відомих видів шаруватих фаз фото каталітичну активність по розкладанню H_2O і виділенню водню. Цей факт пов'язують з особливостями морфології їх частинок і здатністю оборотно інтеркалювати молекули води в міжшаровий простір, що може призводити як до збільшення ефективної питомої поверхні фотокаталізатора, так і сприяти просторовому розподілу центрів окиснення-відновлення.

Для з'ясування характеру хімічної взаємодії і фазових рівноваг у водно-сольових системах досліджуваних нітратів (прекурсорів багатокомпонентних оксидних поліфункціональних матеріалів) у повних концентраційних співвідношеннях в температурному діапазоні існування розчинів використаний метод добавок, описаний в [20, 21] і заснований на дослідженні розчинності як однієї з властивостей найбільш „чутливих“ до виявлення фазових перетворень у системах, що є одночасно параметром їхнього стану, і до того ж простими доступними нині експериментальними методами. Метод дозволяє знайти межі саморозвитку, до яких прямує у конкретних умовах в рівноважному стані ізольована система заданого складу.

Рівновага фаз досягалася протягом 2-3 діб. У якості вихідних солей використовували гідратовані і безводні нітрати вказаних елементів марки „ч.д.а.“.

Хімічний аналіз рідких і твердих фаз, „залишків“ проводили на вміст Ln^{3+} й азоту. Вміст Ln^{3+} визначали трилонометрично у присутності ксиленолового помаранчевого у якості індикатора (ацетатний буферний розчин, $\text{pH} = 5-6$), азоту – методом відгонки [22]; іонів Me^+ – розрахунком по різниці, виходячи із загального вмісту нітратів і частково по сухому залишку.

Одержані дані для окремих іонів перераховувались на сольовий вміст і згідно принципу відповідності наносились на діаграми розчинності.

Графічне відображення складу твердих фаз, що утворюються у системі, проводили за Скрейнемакерсом [20, 21], підтвердження їх індивідуальності і охарактеризування – хімічним, кристалооптичним, рентгенофазовим, рентгеноструктурним, ІЧ-спектроскопічним, термографічним, іншими методами.

На етапі дослідження фотокаталітичного окислення парів органічних субстратів для формування зразків композиційних фотокаталізаторів був використаний Titanium Oxide Micro Powder (TiO_2 , Anatase, 1500 nm, 99,9 %) US Research Nanomaterials, Inc.

Для вивчення впливу сформованих зразків-фотокаталізаторів на кінетику утворення газоподібних інтермедіатів проведено дослідження фотокаталітичного окислення парів органічних речовин (на прикладі етанолу) статичним методом у розробленій герметизованій камері-контейнері, обладнаній змінними навісним тримачем зразків

досліджуваних об'єктів, внутрішнім випарним дозатором і зовнішнім реактором для малих і залпових доз-ін'єкцій субстратів, вентилятором – перемішувачем внутрішнього газового середовища, додатковим внутрішнім нагрівником, датчиком вимірювання концентрації CO_2 , навісними досліджуваними «пасивними» пластинами-адсорберами, системою освітлення.

Для вимірювання концентрації CO_2 в досліджуваному середовищі використаний портативний багатофункціональний електронний газоаналізатор AZ 7755 (AZ Instrument Corp., Тайвань), що дозволяє одночасне вимірювання температури, відносної вологості та має можливість приєднання до зовнішнього інтерфейсу.

У роботі в якості освітлювачів використані: люмінесцентна лампа низького тиску потужністю 8 Вт і бактерицидна лампа тієї ж потужності з довжиною хвилі 254 нм.

Результати дослідження та їх обговорення

Узагальнені і важливі для практичного використання відомості про лужні координаційні нітрати рідкісноземельних елементів – прекурсори перспективних сучасних багатофункціональних матеріалів – щодо умов їх утворення й існування, природи хімічного зв'язку, складу, будови, типу координації ліганд, існування ізотипних рядів по стехіометрії складу, структурі, виявляємим властивостям систематизовані авторами в [23, 24] й у найбільш наочній формі наведені в табл. 1–4. Вибір такої форми представлення даних найбільш

інформативний і корисний під час розроблення інноваційних проєктів дозволяє прогнозувати причинно-наслідкові фундаментальні закономірності поведінки структурних компонентів в аналогічних виробничих процесах, правильно вибирати режими, стабільність, способи формування й одержання цільових продуктів з відтворюваними структуро-чутливими характеристиками.

У потрійних РЗЕ-вмісних системах нітратних прекурсорів, які є невід'ємними компонентами більш складних багатокомпонентних систем, обмінні перетворення починаються з моменту розчинення складових у воді. З'ясовано, що Ln^{3+} церієвої підгрупи в досліджуваних умовах є активними комплексоутворювачами, утворюють аніонні координаційні сполуки з Me^+ усіх лужних металів, а їх стійкість і комплекс притаманних властивостей виявляються сильнодіючими технологічними факторами, що суттєво впливають на характер перетворень в системах як на проміжних стадіях, так і на результати процесів в цілому.

Одержані відомості дозволяють моделювати поведінку структурних компонентів на підготовчих стадіях формування сучасних поліфункціональних фотокаталітичноактивних матеріалів за інноваційними технологічними регламентами з використанням нітратних РЗЕ-вмісних прекурсорів.

Наявні виявлені тенденції фазоутворення в модельних досліджуваних системах є термодинамічно найбільш вірогідними межами перетворень в технологічних об'єктах в умовах формування й отримання цільових продуктів.

Таблиця 1 – Склад і температурний інтервал утворення координаційних нітратів РЗЕ й елементів ІА групи, виявлених у водно-солевих системах $\text{MeNO}_3 - \text{Ln}(\text{NO}_3)_3 - \text{H}_2\text{O}$

МРЗЕ	La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu	Y
Li	Li ₃ [Ln ₂ (NO ₃) ₉]·3H ₂ O кубіч.; P2 ₁ 3					евт.		евт. 25-100 °C								
	100 °C	100 °C	100 °C	65-100 °C		25-100 °C										
Na	Na ₂ [Ln(NO ₃) ₅]·H ₂ O моноклін.; P2 ₁ /a							евт. 25-100 °C								
	25-100 °C		50-100 °C													100 °C
K	K ₂ [Ln(NO ₃) ₅ (H ₂ O) ₂] ромбіч.; Fdd2							K[Tb(NO ₃) ₄ (H ₂ O) ₂] приміт. ромбіч.; P2 ₁ cn 50-100 °C								
	25-50 °C															
	K ₃ [Ln ₂ (NO ₃) ₉]·H ₂ O кубіч.; P4 ₃ 32															
	100 °C		50-100 °C													
Rb	Rb ₂ [Ln(NO ₃) ₅ (H ₂ O) ₂] моноклін.; Cc		Rb ₅ [Ln ₂ (NO ₃) ₁₁]·H ₂ O моноклін.; C2/c					Rb[Ln(NO ₃) ₄ (H ₂ O) ₂] ·H ₂ O моноклін. ; P2 ₁ /n 25-50 °C								
	25-100 °C		25-50 °C													50-100 °C
	Rb ₃ [Ln ₂ (NO ₃) ₉] ·H ₂ O кубіч.; P4 ₃ 32															
	50-100 °C															
Cs	Cs ₂ [Ln(NO ₃) ₅ (H ₂ O) ₂] моноклін.; C2/c							Cs[Ln(NO ₃) ₄ (H ₂ O) ₂] ·H ₂ O моноклін. ; P2 ₁ /n 25-50 °C								
	25-50 °C		25-100 °C													
			Cs[Ln(NO ₃) ₄ (H ₂ O) ₃] триклін.; P1													
			50 °C	50-65 °C												

А можливі реальні відхилення в подібних системах зумовлюються неоднорідністю реакційного середовища за складом, вмістом реагуючих компонентів, за умовами їх знаходження, скінченністю швидкості перетворень, дифузійними особливостями, теплоємністю, в'язкістю, природою перетворень на границях утворюваних гетеро фаз, використанням застосованих принципів і способів їх активації та іншими конкретними чинниками. А виявлені процеси комплексоутворення у водних розчинах нітратів сприяють гомогенізації систем структурних компонентів на молекулярному рівні при комплексному чи комбінованому переробленні.

Аналіз одержаних даних вказує на протікання в системах у найближчому оточенні Ln^{3+} -комплексоутворювача конкуруючих процесів заміщення молекул води на нітрат-іони. Ступінь повноти заміщення залежить від природи Ln^{3+} , присутніх Me^+ , властивостей електронно-донорних атомів кисню і просторової будови лігандів, концентрації аніонів, кількості розчинника. Виявлено значний вплив на ці процеси температурного фактору. Спостерігаються відмінності у комплексоутворюючій здатності елементів церієвої і

ітрієвої підгруп, Y, а також серед РЗЕ у середині першої підгрупи. Одержані результати вказують на постадійність процесів комплексоутворення. Наявність певних значень температури початку виділення у тверду фазу комплексних сполук – на існування енергетичного бар'єру і необхідності деякої енергії активації для здійснення таких перетворень.

При утворенні нітратних комплексів у великій мірі виконуються вимоги симетрії і планарний малий розмірами ліганд NO_3^- є „зручним” для утворення високосиметричного оточення іона Ln^{3+} . Для лантанодів характерна схильність утворювати три типи координації NO_3^- -лігандів. Це веде до утворення, як ізольованих комплексів, так і їх полімеризацію у двоядерні, ланцюжки, каркаси.

Всі встановлені координаційні сполуки були синтезовані в монокристалічному вигляді і охарактеризовані комплексом фізико-хімічних методів. В таблиці 2 приведені рентгеноструктурні характеристики новоутворюваних фаз для можливості їх ідентифікації та виявлення в ході перероблення (ΔT – температурний діапазон знаходження).

Таблиця 2 – Вплив типу симетрії на можливість виникнення фізичних властивостей у кристалах координаційних нітратів рідкісноземельних елементів

Сполуки	Сингонія	Точкова група	Просторова група	Властивості	(ΔT , °C)	Характер розчинності
$\text{Li}_3[\text{Ln}_2(\text{NO}_3)_9] \cdot 3\text{H}_2\text{O}$ $\text{Ln} - \text{La} \div \text{Sm}$	кубіч.	23	$P2_13$		65-100	конгр.
$\text{Na}_2[\text{Nd}(\text{NO}_3)_5] \cdot \text{H}_2\text{O}$ $\text{Ln} - \text{La} \div \text{Sm}$	монокл.	2/m	$P2_1/a$		50-100	інконгр.
$\text{M}_3[\text{Ln}_2(\text{NO}_3)_9] \cdot \text{H}_2\text{O}$ $\text{M} - \text{K}, \text{Rb}, \text{NH}_4^+ \text{Ln} - \text{La} \div \text{Sm}$	кубіч.	432	$P4_332$		50-100	конгр.
$\text{K}_2[\text{Ln}(\text{NO}_3)_5(\text{H}_2\text{O})_2]$ $\text{Ln} - \text{La} \div \text{Sm}$	ромбіч.	mm2	Fdd2		50	інконгр.
$\text{K}[\text{Ln}(\text{NO}_3)_4(\text{H}_2\text{O})_2]$ $\text{Ln} - \text{Y}, \text{Gd} \div \text{Lu}$	ромбіч.	mm2	$P2_1cn$		50-100	конгр.
$\text{Rb}_2[\text{Ln}(\text{NO}_3)_5(\text{H}_2\text{O})_2]$ $\text{Ln} - \text{La}, \text{Ce}$	монокл.	m	Cc		25-100	інконгр.
$\text{Rb}_5[\text{Ln}_2(\text{NO}_3)_{11}] \cdot \text{H}_2\text{O}$ $\text{Ln} - \text{Pr} \div \text{Sm}$	монокл.	2/m	C2/c		25-50	інконгр.
$\text{Rb}[\text{Ln}(\text{NO}_3)_4(\text{H}_2\text{O})_2] \cdot \text{H}_2\text{O}$ $\text{Ln} - \text{Y}, \text{Gd} \div \text{Lu}$	монокл.	2/m	$P2_1/n$		25-50	конгр.
$\text{Cs}_2[\text{Ln}(\text{NO}_3)_5(\text{H}_2\text{O})_2]$ $\text{Ln} - \text{Ce} \div \text{Nd}$	монокл.	2/m	C2/c		25-100	інконгр.
$\text{Cs}[\text{Ln}(\text{NO}_3)_4(\text{H}_2\text{O})_3]$ $\text{Ln} - \text{Pr} \div \text{Sm}$	трикл.	$\bar{1}$	$P\bar{1}$		50-60	інконгр.
$\text{Cs}[\text{Ln}(\text{NO}_3)_4(\text{H}_2\text{O})_2] \cdot \text{H}_2\text{O}$ $\text{Ln} - \text{Y}, \text{Gd} \div \text{Lu}$	монокл.	2/m	$P2_1/n$		25-50	конгр.

Авторами роботи проведено кристалохімічний аналіз сполук лужних рідкісноземельних нітратів, який базується на результатах власних досліджень [23], так і на літературних даних [25]. Особливу увагу звернено на будову координаційних поліедрів Ln, яка багато в чому визначає основні властивості відповідних сполук.

Аналізуючи форму, симетрію й інші властивості поліедрів, ми отримуємо можливість

наблизитися до розуміння індивідуальних особливостей РЗЕ.

При узагальненні структурних даних лужних координаційних нітратів РЗЕ звертає на себе увагу надзвичайно обмежений асортимент координаційних поліедрів для таких високих КЧ = 12. У випадках, коли вода не входить до складу координаційної сфери атома Ln, поліедри побудовані абсолютно однаково.

Їх еквівалентність полягає в тому, що укорочені ребра O–O (загальні з NO₃[–]–лігандами) займають одні й ті самі положення. Виняток становлять по одному з 2 незалежних поліедрів Ln у двох ізоструктурних сполуках Li₃[Ln₂(NO₃)₉]·3H₂O (Ln – La, Nd). Якщо ж координаційне насичення відбувається за участю молекул води, форма 12-вершинників дещо змінюється. Топологічно це все той же ікосаедр. Однак розподіл вкорочених ребер у цьому випадку інший. Аналіз розглянутих нами Ln-поліедрів дає підставу відзначити тенденцію РЗЕ в даному класі сполук до організації симетричного координаційного оточення.

Приведені нижче дані з термографічного вивчення лужних координаційних нітратів рідкісноземельних елементів (табл. 3) з'ясовують характер і закономірності теплових перетворень цих сполук в

температурному діапазоні 25–1000 °С, встановлюють їхню термостійкість, інтервали існування, стадійність процесів, фазоутворення в залежності від складу, вмісту, природи складових компонентів, способу упаковки координаційних поліедрів у просторову побудову, умов і способу оброблення тощо.

Це дає можливість прогнозувати поведінку застосованих потенціальних попередників в реальних технологічних системах в аналогічних умовах. Для порівняння й аналізу там же приводяться відомості про малі стабільність і термостійкість калієвих координаційних нітратів Y, Gd – Lu складу K[Ln(NO₃)₄(H₂O)₂], що обумовлює обмеження в застосуванні такого виду прекурсорів у технологічних перетвореннях з метою модифікування властивостей цільових продуктів.

Таблиця 3 – Характерні властивості представників ізоструктурних груп координаційних нітратів РЗЕ, ітрію

Сполуки; сингонія, просторова група кристалів	Представ- ники	Дегідра- ція	Плавлення у кристалізаційній воді	Поліморфні переходи	Плавлення безводної форми	Примітка
Li ₃ [Ln ₂ (NO ₃) ₉]·3H ₂ O кубіч.; P2 ₁ 3	La–Sm	65 183 216	183	–	274	Температурні дані для координаційної сполуки Nd
Na ₂ [Ln(NO ₃) ₅]·H ₂ O моноклін.; P2 ₁ /a	La–Sm	81 148 236	–	271	328	Температурні дані для координаційної сполуки Nd
K ₂ [Ln(NO ₃) ₅ (H ₂ O) ₂] ромбіч.; Fdd2	La–Sm	95 111	95	219	314	Температурні дані для координаційної сполуки Nd
K ₃ [Ln ₂ (NO ₃) ₉]·H ₂ O кубіч.; P4 ₃ 32	La–Sm	126	–	–	347	Температурні дані для координаційної сполуки Nd
K[Ln(NO ₃) ₄ (H ₂ O) ₂] приміт. ромбіч.; P2 ₁ cn	Y, Gd–Lu	138 172	138	–	–	Температурні дані для координаційної сполуки Gd
Rb ₂ [Ln(NO ₃) ₅ (H ₂ O) ₂] моноклін.; Cc	La, Ce	93 144	93	–	327	Температурні дані для координаційної сполуки La
Rb ₅ [Ln ₂ (NO ₃) ₁₁]·H ₂ O моноклін.; C2/c	Pr–Sm	172	172	–	–	Температурні дані для координаційної сполуки Nd
Rb ₃ [Ln ₂ (NO ₃) ₉]·H ₂ O кубіч.; P4 ₃ 32	La–Sm	136	–	–	302	Температурні дані для координаційної сполуки Nd
Rb[Ln(NO ₃) ₄ (H ₂ O) ₂]·H ₂ O моноклін.; P2 ₁ /n	Y, Gd–Lu	98 165 231	98	–	–	Температурні дані для координаційної сполуки Yb
Cs ₂ [Ln(NO ₃) ₅ (H ₂ O) ₂] моноклін.; 12/a	La	99 148	148	–	263	Температурні дані для координаційної сполуки La
Cs ₂ [Ln(NO ₃) ₅ (H ₂ O) ₂] моноклін.; C2/c	Ce–Nd	122 156	156	–	235	Температурні дані для координаційної сполуки Nd
Cs[Ln(NO ₃) ₄ (H ₂ O) ₃] триклін.; P1̄	Pr–Sm	62 90 162	162	–	355	Температурні дані для координаційної сполуки Nd
Cs[Ln(NO ₃) ₄ (H ₂ O) ₂]·H ₂ O моноклін.; P2 ₁ /n	Y, Gd–Lu	155 175 252	155	–	–	Температурні дані для координаційної сполуки Yb

Встановлено, що координаційні числа Ln³⁺ церієвої підгрупи – 12, виявлені у низькотемпературних асоційованих формах, зберігаються незмінними аж до утворення стійких високотемпературних багатокомпонентних оксидних фаз MeLnO₂, Me₂Ln₂Ti₃O₁₀. Це вказує на доцільність використання лужних координаційних нітратних РЗЕ-вмісних прекурсорів у подібних технологічних інноваційних рішеннях. Логічним практичним застосуванням й апробацією

одержаного комплексу емпіричних знань про сумісну поведінку і властивості складових компонентів в досліджуваних системах було розроблення одного із можливих способів нанесення і формування композиційних фотокаталітично активних покриттів на структуровані металеві носії (для мінімізації внеску адсорбційної складової у загальний ефект комплексного перетворення) із дисперсій субмікронних розмірів анатазнаї фази TiO₂ і

подібних частинок з модифікованими властивостями їхнього поверхневого шару та послідовного проведення їхнього порівняльного тестового випробування на активність у процесах фотодеструкції парів органічних речовин (на прикладі етанолу) у повітряному середовищі під впливом УФ опромінення. Модифікування активності центрів окислення об'єктів здійснювали двох стадійним термообробленням зразків з роздільним етапуванням процесів: нанесенням і фіксуванням покриттів на основу із водно-суспензійних систем TiO_2 і розчинних нітратних прекурсорів Ln та лужних металів, взятих у заданих співвідношеннях; формуванням дифузійних потоків на міжфазних границях складових гетерогенних композиційних систем та регулюванням складу й умов перероблення розплавів нітратних прекурсорів (з метою створення сприятливих умов для асоціації титанатів з шаруватою структурою $\text{Me}_2\text{Ln}_2\text{Ti}_3\text{O}_{10}$). Такі композиційні системи являються кристалізаційно-конденсованими структурами твердіння.

(Вимоги до структурованих функціонально активних подібних покриттів, процедури їх підготовки та характеристики обговорюються в декількох оглядових статтях [26, 27]).

На підготовчих стадіях термооброблення вихідних водно-суспензійних систем із заправкою компонентного складу розчинних нітратних попередників у співвідношенні, що відповідає утворенню тришарового перовскітоподібного титанату $\text{Me}_2\text{Nd}_2\text{Ti}_3\text{O}_{10}$ ($\text{Me}_2\text{O} \cdot \text{Nd}_2\text{O}_3 \cdot 3\text{TiO}_2$, Me – Li, Na, K; Ln – La-Nd) відбувається видалення розчинника із пересиченого дисперсійного середовища, і згідно з вивченою політермічною діаграмою розчинності [1, 2] (перетворення в області концентраційної конгруентності) поступове виділення зародків кристалів лужних координаційних нітратів Ln. Подальше нагрівання

цих проміжних новоутворених фаз веде до їх плавлення і послідовного розкладання з виділенням оксидів азоту і кисню.

В [7] встановлено, що перехід субмікронного порошку діоксиду титану з анатаза в рутил відбувається при 750-850 °C, а роль первинних елементів структури в таких дисперсних об'єктах грають частки порошку.

TiO_2 (анатаз) у досліджуваній багатокомпонентній гетерогенній системі веде себе хімічно індиферентно по відношенню до складових структурних компонентів до моменту зародження у продуктах термолізу розплаву нітратних координаційних прекурсорів слабо закристалізованих хімічно активних частинок подвійних оксидів MeLnO_2 ($\text{Me}_2\text{O} \cdot \text{Ln}_2\text{O}_3$) [4]. При збільшенні енергії активації системи ($t > 520$ °C) і, відповідно, енергії теплового руху структурних елементів виникає можливість їх зближення на більш короткі відстані, проявляється зміцнення координаційного хімічного зв'язку між складовими катіонами відповідних металів і атомами кисневих аніонів з утворенням перовскітоподібних, в тій чи іншій мірі, катіоновпорядкованих тришарових оксидних структур $\text{Me}_2\text{Ln}_2\text{Ti}_3\text{O}_{10}$ з набором притаманних їм властивостей.

В роботі дослідження фотокаталітичної активності синтезованих матеріалів проводили на прикладі тестової реакції окислення парів етанолу в статичному реакторі. Процес відбувається з утворенням в газовій фазі проміжного продукту - ацетальдегіду, який в кінцевому рахунку повністю окислюється до CO_2 . Для сформованих композиційних фотокаталізаторів спостерігалось збільшення швидкості окислення субстрату в порівнянні з чистим TiO_2 (рис. 2, табл. 4). В результаті це призводило до зменшення часу видалення субстрату і інтермедиату з газової фази та зниження максимальної концентрації ацетальдегіду в останній.

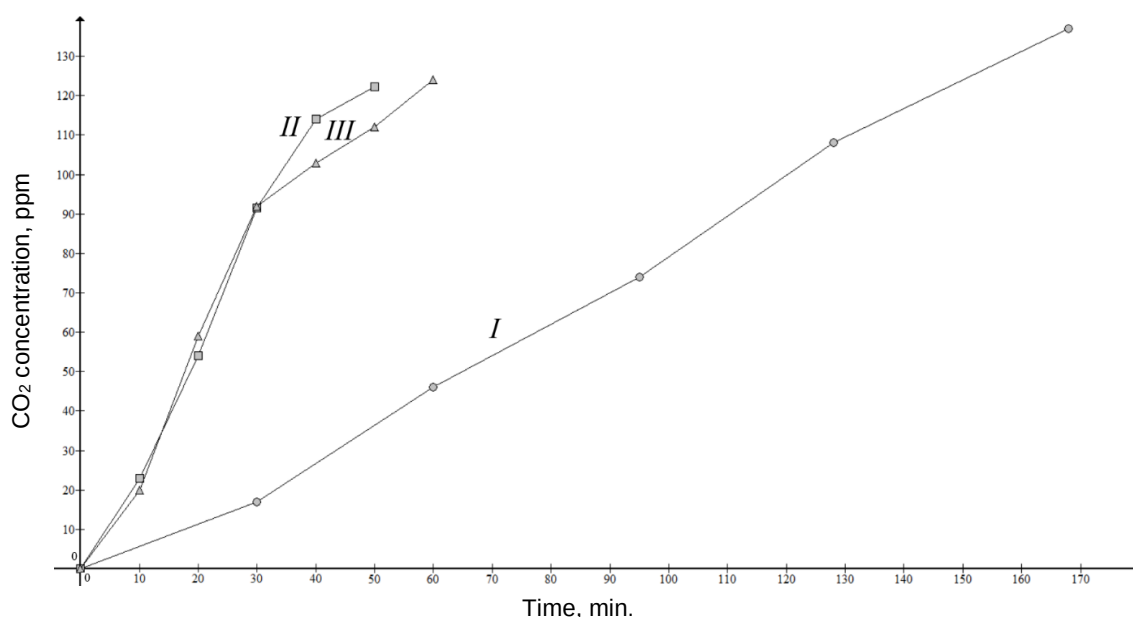


Рис. 1. Кінетика процесів окислення парів етанолу при ультрафіолетовому опроміненні фотокаталізаторів діоксиду титану (анатаз) (I) і композицій (II, III), модифікованих структурою шаруватих перовскітоподібних оксидів $\text{K}_2\text{Nd}_2\text{Ti}_3\text{O}_4$

Таблиця 4 – Оцінювання умовної активності зразків фотокаталізаторів

Катіоновпорядкований тришаровий $K_2Nd_2Ti_3O_{10}$	TiO_2 (модифікація анатазу)
$S=218 \text{ см}^2$	$S=395 \text{ см}^2$
$t = 18,8^\circ\text{C}$	$t = 19,0^\circ\text{C}$
$\tau_{\Sigma} = 40 \text{ хв.}$	$\tau_{\Sigma} = 128 \text{ хв.}$
$\Delta CO_2 = 114 \text{ ppm}$	$\Delta CO_2 = 108 \text{ ppm}$
$v_{\text{сер.}} = 2,85 \text{ ppm / хв.}$	$v_{\text{сер.}} = 0,84 \text{ ppm / хв.}$
$v_{\text{сер. акт.}} = 13,1 \cdot 10^{-3} \text{ ppm / см}^2 \cdot \text{хв.}$	$v_{\text{сер. акт.}} = 2,1 \cdot 10^{-3} \text{ ppm / см}^2 \cdot \text{хв.}$

Примітка: I зразок – на основі TiO_2 (анатаз) і II зразок – на основі композиції тришарового оксиду $K_2Nd_2Ti_3O_{10}$ у процесах розкладання етанолу (рис. 1, для ділянок з пропорційними трендами)

Швидкість окислення субстрату залежала від способу виготовлення покриття, природи застосованих модифікаторів властивостей, послідовності стадій і режимів оброблення, сумісності покриття і матеріала основи та характеру його підготовки, застосованого компонентного складу і вмісту просочувальних систем, способу введення в реактор оброблюваних субстратів.

Для можливості порівняння активності розроблених фотокаталізаторів наважки маси TiO_2 у вихідних заправках водно-суспензійних систем зразків з чистим діоксидом титану і зразків з застосованими модифікованими композиціями брали однакові.

Катіоновпорядковані тришарові композиційні матеріали $K_2Nd_2Ti_3O_{10}$ за відповідних умов можуть виступати альтернативами TiO_2 анатазної модифікації, умовний питомий показник активності яких у 6,2 рази вищий у порівнянні з такою характеристикою діоксиду титану.

Одержані відомості дозволяють оптимізувати умови формування катіоновпорядкованих шаруватих титанатів; з'ясувати умови і виявити температурний діапазон застосування даного класу сполук. Отримані власні і літературні фізико-хімічні, термохімічні та структурні дані, а також результати їх інтерпретації є етапом розвитку експериментальної і теоретичної наукової бази даних про шаруваті сполуки і процеси з їх участю.

Досліджуваний клас шаруватих сполук є перспективною основою для створення функціональних матеріалів з унікальними властивостями, які визначаються двовимірним характером побудови міжшарового простору, спотворенням структури титан-кисневих октаедрів перовскітового шару і високою рухливістю катіонів лужних металів. Такі їх властивості можуть знайти

своє застосування в інноваційних областях науки, техніки, енергетики, електроніки, екології.

Висновки

1. Результати дослідження свідчать, що процеси одержання оксидних РЗЕ-вмісних конструкційних і функціональних матеріалів різного призначення із використанням нітратів елементів різної електронної структури хімічним змішуванням вихідних компонентів при спільному виділенні продуктів із рідкої фази послідовним чи сумісним осадженням з наступним термообробленням відбуваються стадійно, через утворення низки проміжних фаз. Дані про їхній склад, вміст і поведінку в кожному конкретному випадку потребують попередніх системних емпіричних знань в повних концентраційних співвідношеннях у заданому температурному інтервалі.

2. Виявлено відмінності в поведінці структурних компонентів у системах лантанодів церієвої й ітрієвої підгруп, в їхньому характері взаємодії, стадійності, особливостях і закономірностях перебігу.

3. Одержані нові дані становлять основу для:

– пошуку способів збільшення активності Ln-форм;

– з'ясування природи послідовних термічних перетворень у нітратних РЗЕ-вмісних багатокомпонентних системах різних агрегатних станів у ході їхнього термооброблення; умов утворення й існування, властивостей проміжних фаз; впливаючих факторів; можливих способів керування одержання цільового продукту;

– створення сучасних досконалих низькозатратних технологій формування функціональних матеріалів різного призначення із відтворюваними властивостями.

Конфлікт інтересів

Автори декларують, що не мають конфлікту інтересів стосовно даного дослідження, в тому числі фінансового, особистісного характеру, авторства чи іншого характеру, що міг би вплинути на дослідження та його результати, представлені в даній статті.

Фінансування

Дослідження проводилося без фінансової підтримки.

Доступність даних

Рукопис не має пов'язаних даних.

Використання засобів штучного інтелекту

Автори підтверджують, що не використовували технології штучного інтелекту при створенні представленої роботи.

REFERENCE

1. Vandana, Ya., Sachin, P., and Sitharaman, U. (2025), "Dion–Jacobson ($n = 2$) Layered Perovskites ($A'SmTa_2O_7$; $A' = Rb, Cs$) As Potential Ferroelectric Materials", *The Journal of Physical Chemistry C*, vol. 129, no. 11, pp. 5666–5673, doi: <https://doi.org/10.1021/acs.jpcc.4c08181>

2. Sachin P., and Sitharaman U. (2024), "Realization of new $n = 3$ and $n = 4$ Dion-Jacobson layered perovskite series with interlayer K^+ ions and perovskite blocks stabilized by Pr^{3+} ions", *Ceramics International*, vol. 50, no. 17, pp. 31540–31547, doi: <https://doi.org/10.1016/j.ceramint.2024.05.461>
3. Дрючко О.Г., Стороженко Д.О., Бунякіна Н.В. та др. Фізико-хімічне охарактеризування координаційних нітратів РЗЕ і лужних металів – прекурсорів оксидних поліфункціональних матеріалів. *Вісник НТУ «ХПІ»*, 2018, № 39 (1315). С. 3–13, doi: <https://doi.org/10.20998/2079-0821.2018.39.01>
4. Dryuchko, O., Storozhenko, D., Vigdorichik, A. et al. (2018), "Features of transformations in REE-containing systems of nitrate precursors in preparatory processes of formation of multifunctional oxide materials", *Molecular Crystals and Liquid Crystals*, vol. 72(1), pp. 199–214, doi: <https://doi.org/10.1080/15421406.2018.1542066>
5. Dryuchko, O.G., Soloviev, V.V., Bunyakina, N.V. et al. (2024), "Formation of multifunctional perovskite-like layered oxide materials using precursors based on coordinated nitrates of alkali and rare-earth metals", *Funct. Mater.*, vol. 31 (1), pp. 128–139, doi: <https://doi.org/10.15407/fm31.01.128>
6. Varma, A., Mukasyan, A.S., Rogachev, A.S. et al. (2016), "Solution Combustion Synthesis of Nanoscale Materials", *American Chemical Society. Chem. Rev.*, vol. 116, pp. 14493–14586, doi: <https://doi.org/10.1021/acs.chemrev.6b00279>
7. Murray, J. and Wriedt, H. (1987), "The O–Ti (oxygen-titanium) system", *Journal of Phase Equilibria*, vol. 8(2), pp. 148–165, doi: <https://doi.org/10.1007/BF02873201>
8. Fujishima, A., Zhang, X. and Trykc, D.A. ("2008), TiO₂ photocatalysis and related surface phenomena", *Surface Science Reports*, vol. 63(12), pp. 515–582, doi: <https://doi.org/10.1016/j.surfrep.2008.10.001>
9. Dryuchko, O.G., Soloviev, V.V., Shefer, O.V. et al. (2026), "Cooperative transformations in REE-containing nitrate precursor systems during preliminary processes of perovskite-like oxide material property modification", *Funct. Mater.*, 2026, vol. 33(2), pp. 160–167, doi: <https://doi.org/10.15407/fm33.02.160>
10. Lagaly, G. (1986), "Interaction of alkylamines with different types of layered compounds", *Solid State Ionics.*, vol. 22, no. 1, pp. 43–51, doi: [10.1016/0167-2738\(86\)90057-3](https://doi.org/10.1016/0167-2738(86)90057-3)
11. Machida, M., Miyazaki, K., Matsushima, Sh. and Arai M, (2003), "Photocatalytic properties of layered perovskite tantalates, $MLnTa_2O_7$ ($M = Cs, Rb, Na$, and H ; $Ln = La, Pr, Nd$, and Sm)", *J. Mater. Chem.*, vol. 13, pp. 1433–1437, doi: <https://doi.org/10.1039/b301938c>
12. Gopalakrishnan, J., Sivakumar, T., Ramesha, K. et al. (2000), "Transformations of Ruddlesden-Popper oxides to new layered perovskite oxides by metathesis reactions", *Chem. Phys.*, vol. 9, pp. 6237–6241, doi: <https://doi.org/10.1021/ja9914644>
13. Schaak R. E., and Mallouk T.E. (2001), "KLnTiO₄ ($Ln=La, Nd, Sm, Eu, Gd, Dy$): A New Series of Ruddlesden–Popper Phases Synthesized by Ion-Exchange of HLnTiO₄", *J. Solid State Chem.*, vol. 161(2), pp. 225–232, doi: <https://doi.org/10.1006/jssc.2001.9303>
14. Zhu, W. J., Feng, H.H., and Hor, P.H. (1996), "Synthesis and characterization of layered titanium oxides $NaRTiO_4$ ($R = La, Nd$ and Gd)", *Mater. Res., Bull.* 31(1), pp.107–111, doi: [https://doi.org/10.1016/0025-5408\(95\)00167-0](https://doi.org/10.1016/0025-5408(95)00167-0)
15. Cui, W, Guo, D, Liu, L, Hu, J, Rana, D, and Liang, Y. (2014), "Preparation of $ZnIn_2S_4/K_2La_2Ti_3O_{10}$ composites and their photocatalytic H₂ evolution from aqueous Na_2S/Na_2SO_3 under visible light irradiation", *Catal Commun.*, vol. 48, pp. 55–59, doi: <https://doi.org/10.1016/j.catcom.2014.01.026>
16. Kanji S., Masashi, M., Tomohiko, O. et al. (2024), "Designed functions of oxide/hydroxide nanosheets via elemental replacement/doping", *Chem. Soc. Rev.*, vol. 53 (21), pp. 10523–10574, doi: <https://doi.org/10.1039/d4cs00339j>
17. Kato, M., Kajita, T., Hanakago, R. et al. (2006), "Search for new superconductors by the Li-intercalation into layered perovskites", *Phys. C Supercond.*, vol. 445, pp. 26–30, doi: <https://doi.org/10.1016/j.physc.2006.03.070>
18. Thangadurai, V., Subbanna, G.N., and Gopalakrishnan, J. (1998), " $Ln_2Ti_2O_7$ ($Ln = La, Nd, Sm, Gd$): a novel series of defective Ruddlesden-Popper phases formed by topotactic dehydration of $HLnTiO_4$ ", *J. Chem. Commun.*, vol. 7, pp. 1299–1300, doi: <https://doi.org/10.1039/a802644k>
19. Ranmohotti, K.G.S., Josepha, E., Choi, J. et al. (2011), "Topochemical manipulation of perovskites: low-temperature reaction strategies for directing structure and properties", *Adv. Mater.*, vol. 23(4), pp. 442–460, doi: <https://doi.org/10.1002/adma.201002274>
20. Pogliani, L., Ameta, S. C., and Haghi, A. K. (2020), *Chemistry and Industrial Techniques for Chemical Engineers*, Francis and Taylor, 388 p., doi: <https://doi.org/10.1201/9780429286674>
21. Руднева С.І., Сахненко М.Д., А.В. Дженюк. Гетерогенні рівноваги в хімічній технології. Харків: НТУ «ХПІ», 2020. 116 с. URL: <https://web.kpi.kharkov.ua/fchem/wp-content/uploads/sites/30/2022/09/Geterogenni-rivnovagi-SIR.pdf>
22. Thompson, L. (2014), *Fundamentals of analytical chemistry and quantitative chemical analysis*, Basi6 International, Irving, TX, U.S.A., available at: <https://www.abebooks.com/9781781543313/Fundamentals-Analytical-Chemistry-Quantitative-Chemical-1781543313/plp>
23. Dryuchko, O.G., Soloviev, V.V., Bunyakina, N.V. et al. (2025), "Some aspects of the production of multifunctional oxide REE-containing materials using nitrate systems", *Functional Materials*, vol. 32, no. 4, pp. 532–543, doi: <https://doi.org/10.15407/fm32.04.532>
24. Дрючко О.Г., Шефер О.В., Галай В.М., Боряк Б.Р., Захарченко Р.В. Модифікування характеристик фотокаталітично-активних функціональних вузлів адаптивних систем очищення повітря з використанням координаційних нітратів РЗЕ та лужних металів. *Системи управління, навігації та зв'язку*, 2022. № 3. С. 22–32, doi: <https://doi.org/10.26906/SUNZ.2022.3.022>
25. Eriksson, B., Larsson, L.O., Niinisto L. et al. (1980), "Crystal and molecular structure of potassium diaquapentanitratolanthanate (III), *Acta Chem. Scand.*, A 34, No. 8, pp. 567–572, doi: <https://doi.org/10.3891/acta.chem.scand.34a-0567>
26. Meille, V. (2006), "Review on methods to deposit catalysts on structured sur-faces", *Appl. Catal.*, vol. 315, pp. 1–17, doi: <https://doi.org/10.1016/j.apcata.2006.08.031>

27. Avila, P., Montes, M., and Miró E.E. (2005), "Monolithic reactors for environmental applications: A review on preparation technologies", *Chem. Eng. J.*, vol. 109, pp. 11–36, doi: <https://doi.org/10.1016/j.cej.2005.02.025>

Received (Надійшла) 13.06.2026

Accepted for publication (Прийнята до друку) 26.08.2026

Publication date (Дата публікації) 25.09.2026

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Дрючко Олександр Григорович – кандидат хімічних наук, доцент, доцент кафедри автоматики, електроніки та телекомунікацій, Національний університет «Полтавська політехніка імені Юрія Кондратюка», Полтава, Україна;
Oleksandr Dryuchko – Candidate of Chemical Sciences, Associate Professor, Associate Professor of the Department of Automation, Electronics and Telecommunications, Yuriy Kondratyuk Poltava Polytechnic National University, Poltava, Ukraine;
e-mail – itm.dryuchko@nupp.edu.ua; ORCID Author ID: <https://orcid.org/0000-0002-2157-0526>;
Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=6504199218>.

Бунякіна Наталія Володимирівна – кандидат хімічних наук, доцент, доцент кафедри хімії та фізики, Національний університет «Полтавська політехніка імені Юрія Кондратюка», Полтава, Україна;
Natalia Bunyakina – Candidate of Chemical Sciences, Associate Professor, Associate Professor of the Department of Chemistry and Physics, Yuriy Kondratyuk Poltava Polytechnic National University, Poltava, Ukraine;
e-mail – nning.bunyakina@nupp.edu.ua; ORCID Author ID: <https://orcid.org/0000-0003-4241-5127>;
Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=6506551562>.

Шефер Олександр Віталійович – доктор технічних наук, професор, завідувач кафедри автоматики, електроніки та телекомунікацій, Національний університет «Полтавська політехніка імені Юрія Кондратюка», Полтава, Україна;
Oleksandr Shefer – Doctor of Technical Sciences, Professor, Head of the Department of Automation, Electronics and Telecommunications, Yuriy Kondratyuk Poltava Polytechnic National University, Poltava, Ukraine;
e-mail: itm.ovshefer@nupp.edu.ua; ORCID Author ID: <https://orcid.org/0000-0002-3415-349X>;
Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57210203269>.

Галай Василь Миколайович – кандидат технічних наук, доцент, доцент кафедри автоматики, електроніки та телекомунікацій, Національний університет «Полтавська політехніка імені Юрія Кондратюка», Полтава, Україна;
Vasyl Halai – Candidate of Technical Sciences, Associate Professor, Associate Professor of the Department of Automation, Electronics and Telecommunications, Yuriy Kondratyuk Poltava Polytechnic National University, Poltava, Ukraine;
e-mail – itm.galayvm@nupp.edu.ua; ORCID Author ID: <https://orcid.org/0000-0002-1205-7923>;
Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57221716022>.

Захарченко Руслан Володимирович – кандидат технічних наук, доцент, доцент кафедри автоматики, електроніки та телекомунікацій, Національний університет «Полтавська політехніка імені Юрія Кондратюка», Полтава, Україна;
Ruslan Zakharchenko – Candidate of Technical Sciences, Associate Professor, Associate Professor of the Department of Automation, Electronics and Telecommunications, Yuriy Kondratyuk Poltava Polytechnic National University, Poltava, Ukraine; e-mail – itm.rvzakharchenko@nupp.edu.ua; ORCID Author ID: <https://orcid.org/0000-0002-4651-0159>;
Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57221734157>.

Synthesis features and characteristic properties of perovskite-like layered oxide materials of rare-earth and transition elements as fundamental aspects for the practice of innovative photocatalytic solutions

О. Г. Дрючко, Н. В. Бунякіна, О. В. Шефер, В. М. Галай, Р. В. Захарченко

Abstract. Relevance. The use of perovskite-like layered oxide phases and their precursors of complex lanthanide nitrates opens an innovative approach for the controlled synthesis of highly active nanocomposites of transition and rare-earth elements. **Research objects:** Precursors - alkaline coordination nitrates of rare-earth elements. Target systems - single-layer and layered nanostructured oxide highly ordered systems of lanthanides and transition elements (in particular, coatings based on highly dispersed anatase titanium dioxide). **Purpose of the article:** Generalization of information on the conditions of formation, structure and properties of coordination lanthanide nitrates and development on their basis of a methodology for physicochemical modification of the oxidation centers of the diffuse layer of nanostructured TiO₂ oxide systems to increase the photocatalytic activity of the latter. **Research results.** Fundamental data on alkaline coordination nitrate precursors have been systematized – 1) the conditions of their existence, the nature of the chemical bond, the composition, the features of the coordination polyhedra Ln, the types of ligand coordination have been determined, the existence of isotype series in terms of the stoichiometry of the composition, structure, and the characteristic properties detected has been established; 2) a method of chemical modification of oxidation centers in the surface layer of anatase TiO₂ by heat treatment in contact with the products of thermolysis of melts of alkaline coordination nitrates of lanthanides has been proposed; 3) a high photocatalytic destruction of vapors of organic substrates (for example, ethanol) on the obtained modified coating samples has been experimentally revealed. **Conclusions.** The obtained data (as primary information) are the basis for the detection, identification, and control of the phase state of processing objects in the preparatory stages, the selection of criteria for the compatibility of components in the formation of single-layer and layered nanostructured oxide composite systems of lanthanides and transition elements of wide application, with catalytic and photocatalytic activity, a coating capable of self-cleaning with hydrophilic properties; the development of various combined methods of their activation and the establishment of technological and functional dependencies; controlled modification of the properties of the resulting target products.

Keywords: alkaline coordination nitrates of lanthanides; conditions of formation; crystal structure of compounds; characteristic properties; transformations with physical activation; highly ordered oxide systems of lanthanides and transition elements.

Dmytro Diachenko, Maksym Shostak

Kharkiv National University of Radio Electronics, Kharkiv, Ukraine

METHOD FOR ASSESSING THE RISK OF DEFECT MANIFESTATION IN .NET MICROSERVICES IN DISTRIBUTED COMPUTER SYSTEMS BASED ON OPERATIONAL METRICS

Abstract. Background. Software defects undetected during functional testing may manifest under load, making risk assessment based on operational metrics relevant. **Subject of the study:** methods and tools for assessing the risk of defects in .NET microservices within distributed computer systems, taking operational metrics into account. **Objective:** to develop a method for assessing the risk of defect manifestation in .NET microservices within distributed computer systems based on a comprehensive analysis of operational metrics. **Research tasks.** To achieve the purpose, a set of operational metrics was defined, and a model for forming an integral risk indicator based on their normalized values was developed. **Methods.** Methods of system analysis, normalization and integral assessment, load testing, and statistical processing were used. **Results.** A method for the integral assessment of the risk of defect manifestation in .NET microservices based on normalized operational metrics was proposed. **Conclusions.** The proposed method makes it possible to assess the risk of defect manifestation in software components of distributed computer systems without modifying their program code, based on operational data obtained during load testing.

Keywords: defect manifestation risk; .NET microservices; distributed computer systems; operational metrics; integral risk indicator; load testing; throughput; Kubernetes; Docker; software reliability.

1. Introduction

1.1 Relevance of the Topic

At the present stage of software system development, there is an active transition to cloud deployment models, within which microservice architecture has become one of the leading approaches to building scalable, fault-tolerant, and flexible applications. This approach involves decomposing a software system into a set of relatively independent services, each of which performs a separate functional role, is deployed autonomously, and can be updated or scaled independently of other components. Combined with containerization, the microservice model facilitates application portability across different execution environments, provides a high degree of execution environment isolation, and simplifies the automation of deployment processes. In this regard, container orchestration systems, particularly Kubernetes, have become important tools for managing service lifecycles, load balancing, failure recovery, and horizontal autoscaling.

Under these conditions, ensuring the effective horizontal scaling of microservice applications becomes particularly important. Unlike vertical scaling, which involves increasing the resources of an individual node, horizontal scaling is based on increasing the number of service instances to distribute the load among them. This approach is more natural for cloud and distributed environments; however, its effectiveness is determined not only by infrastructure parameters but also by the implementation characteristics of the service itself. The practical use of container platforms demonstrates that services with identical functional purposes may differ significantly in performance, response time stability, throughput, and resource consumption patterns, which directly affects the operation of autoscaling mechanisms.

This issue is particularly relevant for .NET microservices because the .NET platform provides a wide range of tools for building high-performance web applications, including an asynchronous request-

processing model, memory optimization mechanisms, minimal API approaches, caching, and various methods of organizing parallel execution. However, the availability of these tools alone does not guarantee high efficiency in real-world container deployments. Service performance in Kubernetes is influenced by a combination of factors, including the load-processing model, thread-blocking patterns, container startup time, request-processing latency, CPU and memory utilization profiles, and service behavior when the horizontal autoscaling mechanism is triggered. Consequently, there is a need not only to create a microservice application on the .NET platform but also to purposefully investigate its behavior in an orchestrated environment to identify operational indicators that may suggest an increased risk of software implementation defects manifesting.

The traditional approach to deploying microservices in a container environment is often limited to verifying the functional operability of a service and performing basic configuration of the horizontal autoscaling mechanism. However, under real-world conditions, this is insufficient for a substantiated assessment of the reliability of a software component and the risk of defects in its implementation manifesting. With formally identical scaling parameters, two functionally similar services may demonstrate different sensitivities to load, use computing resources differently, and differ significantly in average response time, tail latencies, throughput, and result stability. Such differences may be associated with thread blocking, inefficient organization of parallel execution, excessive memory utilization, lack of caching, or other software implementation characteristics that are not always detected during traditional functional testing.

Under intensive load, latent deficiencies in a software component may manifest as violations of acceptable response time, an increased percentage of failed requests, unstable resource utilization, reduced throughput, and an unpredictable response of the autoscaling mechanism. At the same time, the individual use of each of these metrics

does not provide a sufficiently comprehensive characterization of the state of a software component. For example, an increase in the number of replicas may be accompanied by persistently high latencies, while an acceptable average response time may conceal a significant increase in its 95th percentile. Therefore, the risk of defect manifestation should be assessed based on the comprehensive consideration of operational indicators characterizing the performance, stability, resource efficiency, and scalability of the component.

A promising approach to addressing this task is a comparative study of basic and optimized .NET microservice implementations in a container orchestration environment using load testing. This approach makes it possible to determine how software implementation characteristics affect average response time, the 95th percentile of response time, result variability, throughput, the percentage of failed requests, CPU and memory consumption, and changes in the number of service instances resulting from the activation of the autoscaling mechanism. The combination of these characteristics can be used to form an integral indicator of the risk of defect manifestation in a software component under varying load conditions.

The relevance of the study is further reinforced by the fact that, in cloud service development practice, performance and scaling indicators are primarily used for infrastructure configuration, whereas the possibility of applying them to assess the risk of defect manifestation in software components remains insufficiently investigated. Autoscaling is often regarded as a universal means of resolving performance problems, although the actual behavior of a distributed system is determined by the interaction among the orchestrator, load characteristics, and the internal implementation of the service. Under these conditions, a comprehensive analysis of operational metrics makes it possible not only to assess scaling efficiency but also to identify software components whose behavior is characterized by increased instability and a risk of defect manifestation under load.

Therefore, it is advisable to develop a method for assessing the risk of defect manifestation in .NET microservices within distributed computer systems that provide for the formation of an integral assessment based on timing characteristics, throughput, the percentage of failed requests, computing resource utilization, and component behavior during horizontal scaling. An experimental comparison of the basic and optimized implementations of the same functional service under identical conditions makes it possible to establish the relationship between software implementation characteristics and the level of risk associated with its unstable operation.

1.2 Literature Review and Problem Definition

Assessing the risk of defect manifestation in software components of distributed computer systems requires consideration not only of functional testing results but also of operational indicators characterizing component behavior under varying load conditions. For microservice applications, such indicators include request-processing timing characteristics, throughput, the percentage of failed requests, computing resource consumption, and the

response of autoscaling mechanisms. In this regard, it is advisable to analyze current studies devoted to microservice scaling, the selection of operational metrics, the prediction of resource requirements, and the verification of component operational stability in container environments.

In article [1], the authors consider the problem of microservice autoscaling under the need to ensure optimal resource allocation while meeting quality-of-service requirements and service-level agreements. The paper summarizes current approaches to microservice autoscaling, with an emphasis on verifying the correctness of the proposed algorithms. The authors systematize the publication results according to three interrelated areas: microservices, verification, and autoscaling. Studies on scaling verification in cloud systems are also analyzed. The article demonstrates that formal methods have long been used to verify autoscaling policies in cloud environments; however, such approaches are only beginning to become widespread in microservice systems. The paper also identifies the main factors affecting microservice scaling and the corresponding assessment metrics and establishes that probabilistic model checking is the most widely used formal method in this subject area. Open problems are separately outlined, including the verification of existing microservice scaling mechanisms and the validation of approaches based on machine learning.

Article [2] considers the problem of predictive scaling of microservice applications on Kubernetes platforms using machine learning methods. The paper proposes an approach to estimating the future resource requirements of microservices based on data obtained during the operation of the standard horizontal autoscaling mechanism, followed by using machine learning models to predict the resources required as the number of users increases. The article describes an experimental setup that includes Kubernetes, Google Online Boutique as a microservice application, Locust for load simulation, and Prometheus for metric collection. It also presents a procedure for selecting horizontal scaling parameters, forming datasets, and training several models, including linear regression, SVM, and MLP. The obtained results demonstrated that machine learning models can predict future computing resource requirements with high accuracy for both individual pods and the cluster, with the MLP model showing the best results in generalized prediction. At the same time, the paper emphasizes limitations associated with extrapolating predictions beyond the available data, the dependence of results on the size of the collected observation dataset, and the need for further model improvement before their use in production environments.

Article [3] considers the research and applied problem of microservice autoscaling while accounting for quality-of-service requirements and efficient resource utilization in a Kubernetes environment. The paper proposes the MS-RA approach, which is based on service-level objectives and implemented using the MAPE-K loop. The proposed solution makes it possible to make real-time scaling decisions while considering not only the need to improve performance but also the need to prevent excessive use of computing resources. The article describes the architecture of the approach and its input parameters, including service

objectives, server types, and scaling parameters, and presents the results of an experimental comparison with horizontal scaling in Kubernetes. The obtained results demonstrated that the proposed approach ensures compliance with the established SLOs while using significantly less CPU, memory, and fewer replicas compared with standard horizontal scaling. The paper also emphasizes the importance of aligning scaling with service requirements, while future research directions include extending the approach to other types of microservice systems and combining reactive mechanisms with machine learning methods.

Article [4] is devoted to improving the efficiency of automatic resource scaling in cloud environments using machine learning methods. The paper proposes an approach based on the statistical selection of the metrics most relevant to a particular application to improve the accuracy of scaling decisions. The authors proceed from the assumption that different applications require different indicators for the correct estimation of required resources; therefore, universal metric sets do not provide sufficient efficiency. The article demonstrates that using a metric-selection mechanism makes it possible to improve the system response speed to load changes, reduce resource utilization, and stabilize quality-of-service indicators. Attention is paid to the limitations of existing approaches, including their dependence on fixed metrics, the difficulty of adapting to dynamic load scenarios, and the need for more flexible resource orchestration tools.

Article [5] considers the problem of automatic resource scaling in cloud computing environments to improve the efficiency of computing infrastructure utilization, ensure quality of service, and reduce costs. The paper systematizes current autoscaling methods, including reactive and proactive approaches such as threshold-based rules, queuing theory, reinforcement learning, fuzzy logic, machine learning, and time-series analysis. The authors analyze the practical application areas of these approaches in cloud services, telecommunications infrastructure, big data processing systems, healthcare services, and smart city platforms. The article demonstrates that autoscaling plays a key role in ensuring the elasticity of cloud systems; however, its effectiveness is limited by the difficulty of accurate load prediction, environmental dynamics, and issues related to energy efficiency, cost, security, and compatibility with existing infrastructure.

In article [6], the authors analyze the impact of microservice autoscaling policies on host energy consumption and the risk of violating energy constraints in cloud environments. The paper proposes an approach based on probabilistic model checking that makes it possible to formally assess the correctness of scaling decisions as early as the design stage. The authors constructed four variants of an MDP model that account for different scaling constraints based on Kubernetes horizontal scaling and used them to analyze autoscaling policies from an energy-efficiency perspective. The obtained results demonstrated that policies simultaneously accounting for latency and energy consumption are the most appropriate, as this approach provides a better balance between quality of service and energy costs. The paper also establishes that the number of pods, the scale of replica-increase operations, and the

selection of scaling thresholds have a significant impact on energy consumption.

Article [7] considers the scalability of microservice architecture. The paper summarizes the results of 55 studies devoted to microservice scaling approaches, the metrics used, and autoscaling tools and mechanisms. The authors demonstrate that existing studies focus primarily on autoscaling, particularly on machine-learning-based approaches, while the main challenges include collecting valid metrics, making dynamic scaling decisions, balancing performance, cost, and security, and addressing the risk of cascading degradation in quality of service. It is separately emphasized that, despite significant progress in this field, substantial research gaps remain due to the lack of standardized autoscaling metrics and the need to develop more adaptive and resilient solutions for real-world dynamic environments.

Article [8] considers the problem of improving the scalability and performance of .NET microservices deployed in an Azure Kubernetes Service environment. The paper analyzes the impact of various configurations and optimization approaches, including the configuration of node pools, the Kubernetes scheduler, horizontal autoscaling mechanisms, VPA, and the use of a service mesh, on the key operational indicators of a microservice application. The authors conducted an experimental evaluation using metrics such as CPU utilization, memory utilization, response time, throughput, and pod-scaling efficiency. The obtained results demonstrated that optimized configurations make it possible to reduce resource consumption, shorten response time, and increase throughput compared with the standard Azure Kubernetes Service deployment configuration. The authors note that the greatest effect is achieved by combining optimized resource allocation, specialized scheduling, and autoscaling mechanisms, although implementing such solutions increases configuration complexity and the requirements for system monitoring and administration.

Article [9] considers software defect prediction based on optimized machine learning models using NASA MDP datasets. The authors compared the k-nearest neighbors method, decision trees, logistic regression, linear discriminant analysis, a multilayer perceptron, and the support vector machine method by applying PCA, SMOTE, and random hyperparameter search; the k-nearest neighbors method demonstrated the best results for most datasets. At the same time, the research focuses on defect prediction using software project data and does not provide an integral assessment of the risk of their manifestation based on operational metrics of components underload.

Article [10] compares traditional and ensemble machine learning algorithms for defect prediction using four datasets. The Gradient Boosting model with boosting demonstrated the highest accuracy of 94.5%; however, the assessment of the risk of defect manifestation based on operational metrics of components under load was not considered in the research. Article [11] presents a systematic review of 63 studies and a meta-analysis of the application of deep learning to software defect prediction. Deep learning was found to outperform traditional models in terms of accuracy, F-measure, and AUC; however, operational metrics of components under load were not considered.

Article [12] compares machine learning methods for test case prioritization in the continuous integration process. The effectiveness of the models was found to depend on the size of the test history and the available execution time; however, the integral assessment of the risk of defect manifestation based on operational metrics was not considered. Article [13] compares online and offline learning models for the just-in-time prediction of defect-inducing changes in program code. The proposed BORB approach enables the use of offline models in an online mode but does not consider operational metrics of components under load. Article [14] proposes an ensemble method for predicting defects in software modules based on PROMISE and GitHub data. At the same time, the integral assessment of the risk of defect manifestation based on operational indicators was not considered.

Study [15] proposes a Top-k classification method with label ranking, the effectiveness of which was confirmed using text and image datasets. However, the method is not directly intended for defect prediction or the analysis of operational metrics of software components. Article [16] proposes the APFD-Net model for test case prioritization based on variational learning to rank and a differentiable loss function. The approach is aimed at the early detection of defects during testing but does not assess the risk of their manifestation based on operational metrics. Article [17] presents a systematic analysis of the application of machine learning to defect prediction from the perspective of its practical use in commercial development. At the same time, a comprehensive assessment of the operational risk of software components was not considered. Study [18] presents SQuaD, a multidimensional temporal dataset of quality indicators from 450 software projects, covering more than 700 metrics, version histories, defects, and vulnerabilities. The dataset is intended for software quality research but does not contain operational metrics obtained through microservice load testing.

Article [19] presents the Defectors dataset for defect prediction at the change and code-line levels, containing approximately 213,000 files from 24 Python projects. It supports the training of defect prediction models but does not consider the behavior of components under load. Article [20] systematizes datasets, data validation methods, artificial intelligence models, and defect prediction tools. At the same time, the assessment of the risk of defect manifestation based on operational metrics of components under load was not considered. Article [21] compares online and offline learning models for the just-in-time prediction of defect-inducing changes in program code. The proposed BORB approach enables the use of offline models in an online mode but does not consider operational metrics of components under load. Article [22] proposes a general process for predicting defects in large-scale complex information systems using machine learning and static software characteristics. However, the behavior of components under load and the integral operational risk are not assessed. Article [23] presents a systematic review of the application of deep learning models to software defect prediction. The predominant use of supervised learning, software metrics, and convolutional neural networks was identified; however, the operational indicators of microservices under load were not considered.

The analysis of current scientific publications has demonstrated that ensuring the stable operation of microservices in container environments is one of the key problems in cloud computing and distributed software systems. Existing studies investigate formal methods for verifying autoscaling policies, predicting resource requirements, meeting service-level objectives, selecting informative metrics, ensuring energy efficiency, and improving the performance of microservice applications. The results obtained by the authors confirm that response time, throughput, CPU and memory utilization, the number of replicas, and the stability of indicators under load are important characteristics of the state of microservice components.

At the same time, most of the analyzed studies focus primarily on the infrastructure level of the problem, namely scaling policies, load prediction, decision-making algorithms, and resource orchestration mechanisms. Operational metrics in such studies are used primarily for scaling control and performance assessment but are not considered as a set of indicators that can be used to determine the risk of defect manifestation in a software component. The relationship between the internal implementation characteristics of a microservice and the emergence of such adverse indicators under load as increased average response time and its 95th percentile, significant result variability, reduced throughput, an increased percentage of failed requests, unstable resource utilization, and a disproportionate response of the autoscaling mechanism remain insufficiently investigated.

Studies analyzing the impact of .NET platform tools, asynchronous request processing, thread blocking, caching, and the organization of parallel execution on the risk of unstable operation of a software component in a Kubernetes environment are particularly limited. Existing approaches also do not provide a unified integral assessment that comprehensively accounts for timing characteristics, throughput, the percentage of failed requests, resource consumption, and component behavior during horizontal scaling. In this regard, it is advisable to develop a method for assessing the risk of defect manifestation in .NET microservices based on a comprehensive analysis of operational metrics obtained during load testing in a container orchestration system. Comparing the basic and optimized implementations of a functionally similar microservice under identical conditions makes it possible to establish the relationship among software implementation characteristics, the stability of operational indicators, and the level of risk of defect manifestation in a distributed computer system.

1.3 Purpose and Objectives of the Study

The purpose of the study is to develop a method for assessing the risk of defect manifestation in .NET microservices within distributed computer systems based on a comprehensive analysis of operational metrics. This will make it possible to identify software components with potentially unstable behavior under varying load conditions by considering request-processing timing characteristics, throughput, the percentage of failed requests, computing resource utilization, and the response of the horizontal scaling mechanism.

To achieve this purpose, the following objectives were set:

- to identify a set of operational metrics that characterize the operational stability of .NET microservices and can be used as indicators of the risk of software implementation defects manifesting;
- to develop a mathematical model for forming an integral assessment of the risk of defect manifestation in a software component;
- to define the criterion and risk assessment indicators, considering timing characteristics, throughput, request-processing errors, CPU and memory utilization, and component behavior during horizontal scaling;
- to develop a method for assessing the risk of defect manifestation in .NET microservices based on the normalization and comprehensive processing of operational metrics;
- to conduct an experimental study of the basic and optimized .NET microservice implementations under identical container orchestration conditions and compare the risk levels formed for them.

2. Materials and Research Methods

The object of the study is the process of assessing the risk of defect manifestation in software components of distributed computer systems during their operation under varying load conditions. The paper considers .NET microservices containerized using Docker and deployed in a Kubernetes environment. The software components process streams of HTTP requests, use the computing resources of cluster nodes, and interact with the horizontal autoscaling mechanism.

Within this study, defect manifestation is understood as the occurrence of an undesirable operational state of a software component that may be associated with the characteristics of its internal implementation and is characterized by exceeding the acceptable response time, significant tail latencies, failed requests, reduced throughput, unstable resource consumption, or inefficient behavior during scaling. The proposed approach is not intended to directly detect defects in the source code but provides an assessment of the risk of their manifestation based on observations of component behavior.

The main research hypothesis is that the comprehensive consideration of timing, resource, and scaling characteristics provides a more substantiated assessment of the risk of defect manifestation in a software component compared with the use of individual performance indicators. A software implementation characterized by a shorter average response time, a lower 95th percentile value, lower result variability, higher throughput, and more stable resource utilization is expected to have a lower integral risk level. The following conditions were used in developing the risk assessment method.

Condition 1. The basic and optimized microservice implementations have the same functional purpose, support identical HTTP interfaces, and process requests with the same structure.

Condition 2. Both implementations are deployed in the same Kubernetes cluster with identical resource request and limit values, ensuring the comparability of the experimental results.

Condition 3. The horizontal autoscaling mechanism is configured identically for both services: the minimum number of replicas is one, the maximum is five, and the target CPU utilization is 50%.

Condition 4. The load is generated at three intensity levels corresponding to 10, 20, and 50 virtual users. Each test lasts 60 seconds, and each experimental scenario is repeated three times.

Condition 5. The operational metrics of the basic and optimized implementations are recorded according to the same rules and at the same points during the experiment.

Condition 6. The risk assessment is comparative and reflects the relative level of risk of defect manifestation within the investigated load scenarios. The obtained value is not considered the probability of a defect being present in the source code.

During the development of the defect manifestation risk assessment method, a decomposition method was applied, according to which the assessment process is divided into three interrelated components: performance assessment, resource utilization stability assessment, and component behavior assessment during horizontal scaling. The relationship among the components is maintained through a common set of operational metrics and an integral risk indicator.

3. Results of the Development and Study of the Method for Assessing the Risk of Defect Manifestation in .NET Microservices

3.1 Mathematical Model for Assessing the Risk of Defect Manifestation in .NET Microservices

To formally represent the results of each experimental run, the following vector of operational characteristics is used:

$$X_{s,l,j} = [t_{avg}, t_{95}, q, e, c, m, n]^T, \quad (1)$$

where s is the software component variant, with $s=B$ corresponding to the basic implementation and $s=O$ to the optimized implementation; l is the load level specified by the number of virtual users; j is the repeated run number; t_{avg} is the average response time, ms; t_{95} is the 95th percentile of response time, ms; q is the throughput, requests per second; e is the percentage of failed requests; c is the CPU utilization; m is the memory utilization; n is the number of service replicas after completion of the load test; T is the vector transposition symbol.

For each software component variant and each load level, the mean value of the metric under study is determined as follows:

$$\bar{x}_{s,l} = \frac{1}{J} \sum_{j=1}^J x_{s,l,j}, \quad (2)$$

where $\bar{x}_{s,l}$ is the value of the metric under study during the j -th run; J is the total number of repeated runs. In the conducted experiments, $J=3$ for each scenario.

To assess the dispersion of the obtained results, the standard deviation is calculated as follows:

$$\sigma_{s,l} = \sqrt{\frac{1}{J-1} \sum_{j=1}^J (x_{s,l,j} - \bar{x}_{s,l})^2}, \quad (3)$$

where $\sigma_{s,l}$ is the standard deviation of the metric values for the corresponding service variant and load level.

To compare the variability of indicators with different units of measurement and value scales, the coefficient of variation is used:

$$v_{s,l} = \sigma_{s,l} / \bar{x}_{s,l}. \quad (4)$$

An increase in the coefficient of variation indicates reduced predictability of software component behavior. Therefore, increased variability in response time, its 95th percentile, and throughput is considered one of the indicators of the risk of defect manifestation under load.

3. 2 Criterion and Indicators for Assessing the Risk of Defect Manifestation in .NET Microservices

Since operational metrics have different units of measurement, they are normalized before being combined. For indicators whose increase signifies an increase in risk, the following formula is used:

$$r_i = \frac{x_i - x_i^{\min}}{x_i^{\max} - x_i^{\min}}, \quad (5)$$

where x_i is the current value of the indicator; $x_i^{\min}$ and $x_i^{\max}$ are its minimum and maximum values in the experimental dataset, respectively; r_i is the normalized risk component.

According to formula (5), the average response time, the 95th percentile of response time, the percentage of failed requests, CPU utilization, memory utilization, and variability indicators are normalized. Throughput has the opposite direction of influence: its decrease indicates deterioration in the component state. Therefore, the normalized risk component associated with insufficient throughput is determined using the following formula:

$$r_q = (q^{\max} - q) / (q^{\max} - q^{\min}). \quad (6)$$

As a result of normalization, all partial assessments fall within the interval from 0 to 1. A value of 0 corresponds to the lowest recorded adverse impact, while a value of 1 corresponds to the highest. If the maximum and minimum values of a particular metric are equal, the corresponding component is set to zero because this metric does not make it possible to distinguish between the investigated variants. The normalized assessment of the risk associated with timing characteristics is formed based on the average response time and its 95th percentile:

$$r_t = \alpha r_{avg} + (1 - \alpha) r_{95}, \quad (7)$$

where r_{avg} is the normalized assessment of the average response time; r_{95} is the normalized assessment of the 95th percentile; α is the coefficient determining the ratio of their influence, $\alpha \in [0; 1]$.

Within the experimental study, it is advisable to set $\alpha=0.4$, resulting in a weight of 0.4 for the average response time and 0.6 for the 95th percentile. The greater weight of the 95th percentile is explained by the fact that it better characterizes adverse request-processing conditions. The risk component associated with computing resource utilization is determined based on the normalized CPU and memory utilization indicators:

$$r_r = \beta r_c + (1 - \beta) r_m, \quad (8)$$

where r_c – is the normalized assessment of CPU utilization; r_m – is the normalized assessment of memory utilization; β – is the coefficient determining the ratio of their influence. If CPU and memory are considered equally significant, $\beta=0.5$ is used.

The risk component characterizing the instability of software component behavior is determined based on the normalized coefficients of variation:

$$r_v = \frac{r_{v,avg} + r_{v,95} + r_{v,q}}{3}, \quad (9)$$

where $r_{v,avg}$ is the normalized variability of the average response time; $r_{v,95}$ is the normalized variability of the 95th percentile; $r_{v,q}$ is the normalized variability of throughput.

To account for component behavior during horizontal scaling, a scaling inefficiency indicator is used:

$$r_s = \frac{n}{n_{max}} \cdot \frac{r_t + r_q}{2}, \quad (10)$$

where n is the actual number of replicas after completion of the test; n_{max} is the maximum number of replicas specified in the autoscaling settings; r_t is the risk component associated with timing characteristics; r_q is the insufficient-throughput component. The r_s indicator increases when an increase in the number of replicas is not accompanied by a sufficient reduction in response time and an increase in throughput. Therefore, the number of replicas itself is not considered an indicator of a defect but is considered only in conjunction with scaling effectiveness. To reduce assessment subjectivity, it is advisable to determine the weighting coefficients of the partial criteria using the entropy method. First, the relative normalized value is calculated for each experimental scenario and each criterion:

$$p_{ak} = r_{ak} / \sum_{a=1}^M r_{ak}, \quad (11)$$

where a is the experimental scenario number; k is the partial criterion number; M is the number of compared scenarios; r_{ak} is the normalized value of the k -th criterion for the a -th scenario. For two service implementations and three load levels, $M=6$. The entropy of an individual criterion is determined using the following formula:

$$E_k = -\frac{1}{\ln M} \sum_{a=1}^M p_{ak} \ln p_{ak}, \quad (12)$$

where E_k characterizes the homogeneity of the values of the corresponding criterion. The more strongly the indicator differs among the investigated scenarios, the greater its informational significance. The weighting coefficient of the criterion is determined using the formula:

$$w_k = (1 - E_k) / \sum_{h=1}^K (1 - E_h), \quad (13)$$

where K is the total number of partial criteria. The proposed method uses six criteria: timing characteristics, failed requests, insufficient throughput, resource utilization, result variability, and scaling inefficiency; therefore, $K=6$.

If the value of a particular criterion is zero in all scenarios, as may occur in the absence of failed requests, its weighting coefficient is set to zero. The weights of the remaining criteria are then renormalized so that their sum equals one. The integral indicator of the risk of defect manifestation for an individual experimental scenario is determined as the weighted sum of the partial components:

$$R_a = \sum_{k=1}^K w_k r_{ak}, \quad (14)$$

In expanded form, the formula is as follows:

$$R = w_t r_t + w_e r_e + w_q r_q + w_r r_r + w_v r_v + w_s r_s. \quad (15)$$

The weighting coefficients must satisfy the conditions:

$$w_t + w_e + w_q + w_r + w_v + w_s = 1, \quad (16)$$

The value of the integral indicator belongs to the interval $R \in [0; 1]$. When $R \in [0; 0,33]$ the risk level is considered low; when $R \in [0,33; 0,67]$ it is considered medium, and when $R \in [0,67; 1]$, it is considered high. This assessment is comparative and reflects the risk of unstable operation of the software component under the investigated conditions rather than the probability of a defect being present in its source code.

3. 3 Method for Assessing the Risk of Defect Manifestation in .NET Microservices and Its Experimental Study

The proposed method for assessing the risk of defect manifestation in .NET microservices involves sequential acquisition, statistical processing, and comprehensive assessment of the operational metrics of a software component. The application of the method includes the following stages:

- determining the software component variant and the parameters of the load scenario;
- conducting a series of repeated experimental runs;
- collecting timing characteristics, throughput indicators, the percentage of failed requests, CPU and memory utilization, and the number of service replicas;
- calculating mean values, standard deviations, and coefficients of variation;
- normalizing the obtained operational indicators;
- forming partial risk components based on timing, resource, and scaling characteristics;
- determining the weighting coefficients of the partial criteria;
- calculating the integral risk indicator and determining the corresponding level of risk of defect manifestation.

The experimental study was conducted to verify the applicability of the proposed method for assessing the impact of a .NET microservice software implementation on the risk of defect manifestation during its operation in the Kubernetes container orchestration system. As part of the experiment, a testbed was created to reproduce a typical deployment environment for containerized microservice applications and compare the behavior of two functionally similar services under identical load conditions. The study is based on a comparison of the basic and optimized service implementations. This makes it possible to determine how changes in the internal request-

processing logic affect the average response time, the 95th percentile of response time, throughput, the percentage of failed requests, result variability, CPU and memory utilization, and the response of the horizontal autoscaling mechanism.

The experimental testbed was built on the .NET 10 platform. Two web services in the minimal API format, ServiceBasic and ServiceOptimized, were created for the study. Both services implement the same set of HTTP interfaces, including a root route, the /health health-check endpoint, and a group of test endpoints for simulating different types of loads. The /api/heavy endpoint was selected as the primary target of the experimental load because it simulates computationally intensive request processing and creates conditions under which the differences between the two software implementations are most evident.

The basic implementation performs computations with additional blocking delays and represents a typical request-processing arrangement without specialized optimization tools. The optimized implementation uses an asynchronous execution model, reduces the number of blocking operations during computations, applies simplified caching mechanisms, and improves the load-processing structure. Thus, both services are functionally similar but differ in their internal execution organization, making it possible to investigate the impact of the software implementation on the operational risk profile.

A separate Dockerfile was created for each service using a multi-stage build with the `mcr.microsoft.com/dotnet/sdk:10.0` image at the compilation stage and the `mcr.microsoft.com/dotnet/aspnet:10.0` image at the runtime stage. Both services were configured to listen on port 8080, ensuring their correct operation in the container environment. The build process produced the local Docker images `service-basic:1.0` and `service-optimized:1.0`, which were used for subsequent deployment in Kubernetes.

The services were deployed in a local Kubernetes cluster operating in the Docker Desktop environment. Separate Deployment and Service YAML configurations were created for each service. At the Deployment level, one initial replica, identical `imagePullPolicy` values, container port 8080, and identical `resources.requests` and `resources.limits` parameters for CPU and memory were specified. Defining the resource parameters is necessary because the horizontal autoscaling mechanism determines the relative CPU utilization based on `requests.cpu`. The `NodePort` type was used at the Service level, while the `kubectl port-forward` command was applied for local load testing, ensuring that test traffic was consistently directed to the corresponding service in the cluster.

A separate horizontal autoscaling object was configured for each deployment. The corresponding Deployment was specified as the scaling target. Identical scaling limits of one to five replicas and a target CPU utilization of 50% were set for the basic and optimized services. This configuration ensures identical infrastructure conditions and makes it possible to associate the identified differences in operational indicators primarily with the software implementation characteristics of the services.

To support the operation of the horizontal autoscaling mechanism, metrics-server was additionally installed in the cluster to collect and provide information on computing resource consumption. During its configuration, the `--kubelet-insecure-tls` parameter was added, which is required for the stable operation of the metrics service in the local Docker Desktop cluster.

After the infrastructure had been prepared, load testing was performed using k6. Separate `basic-test.js` and `optimized-test.js` scenarios were created to send requests to the `/api/heavy` endpoint through local port forwarding for the corresponding service. The load was simulated at three intensity levels: 10, 20, and 50 virtual users. Each test lasted 60 seconds. Each experimental scenario was repeated three times, which made it possible to reduce the influence of random fluctuations, assess the repeatability of the results, and calculate mean values, standard deviations, and coefficients of variation. The B10_1–B10_3, B20_1–B20_3, and B50_1–B50_3 run series were formed for the basic service, while the O10_1–O10_3, O20_1–O20_3, and O50_1–O50_3 series were formed for the optimized service.

During each run, k6 recorded the average response time, the 90th and 95th percentiles of response time, the maximum response time, throughput, and the percentage of failed requests. At the same time, Kubernetes tools were used to determine the number of service replicas after the test, CPU utilization, and memory utilization. The obtained values were used to calculate the partial components and the integral indicator of the risk of defect manifestation in accordance with the mathematical model presented in Subsections 3.1 and 3.2. Fig. 1 shows an example of the structure of a CSV file containing performance metrics, load parameters, and horizontal scaling characteristics of the services.

Service	Load Level	avg_ms	p90_ms	p95_ms	max_ms	rps	failed_percent	CPU	Mem	Replicas
ServiceBasic	10	1050	1450	1550	1650	10	0	10%	20%	1
ServiceBasic	20	650	850	950	1050	20	0	20%	40%	2
ServiceBasic	50	1445	2500	2600	2700	50	0	50%	80%	5
ServiceOptimized	10	164	176	180	185	10	0	10%	20%	1
ServiceOptimized	20	164	176	180	185	20	0	20%	40%	2
ServiceOptimized	50	164	176	180	185	50	0	20%	40%	2

Fig. 1. Fragment of the Experimental Dataset Generated from the Results of .NET Microservice Load Testing

During each test run, the system records a set of indicators characterizing both service performance and scaling behavior. From the k6 output data for each run, the average response time `avg_ms`, the p90 and p95 response-time percentiles `p90_ms` and `p95_ms`, the maximum response time `max_ms`, the number of requests per second `rps`, and the percentage of failed requests `failed_percent` are extracted. At the same time, horizontal scaling and resource consumption metrics are collected using Kubernetes tools. After each run, the `kubectl get hpa` and `kubectl top pods` commands are used to record the number of service replicas before and after the test, the current CPU load per pod, and memory utilization. Since the number of replicas may change when the horizontal autoscaling mechanism is triggered, CPU and memory values are averaged across the available pods of the corresponding service at the time of

measurement. All these data are entered into a single `experiment_results.csv` file, where each row corresponds to a separate experimental run.

After the series of runs has been completed, the results collected are processed in Google Colab using Python, the pandas library for tabular analysis, and matplotlib for plotting graphs. At this stage, the data are grouped according to two attributes: service type and load level. For each group, the mean values and standard deviations of the average response time, the 95th percentile of response time, throughput, the number of replicas after the test, CPU utilization, and memory utilization are calculated. Fig. 2 shows the dependence of the average response time of the two .NET microservice implementations on the load level specified by the number of virtual users. The results obtained demonstrate that the optimized implementation consistently provides lower average response times throughout the investigated load range. While the average response time of the basic service varies from approximately 659 to 1,445 ms and increases significantly under a load of 50 virtual users, this indicator remains almost constant for the optimized service at approximately 164–176 ms. This indicates the higher efficiency of the optimized service implementation under container deployment conditions.

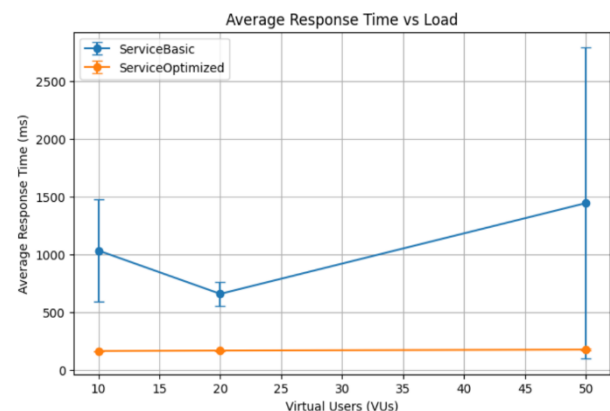


Fig. 2. Dependence of the Average Response Time of the Services on the Load Level

An important feature of the graph presented is also the difference in the variability of the results. The basic implementation exhibits significantly greater deviations, especially under high load, indicating lower predictability of its behavior. In contrast, the optimized service demonstrates minimal dispersion of values, confirming its stability across repeated experimental runs.

Fig. 3 shows the dependence of the p95 response-time percentile on the load level for the basic and optimized .NET microservice implementations. Unlike the average response time, this indicator characterizes service behavior at high latency values and makes it possible to assess the stability of request processing under less favorable operating conditions.

The obtained results indicate that, for the basic implementation, the 95th-percentile response-time values remain substantially higher across all scenarios studied and increase particularly sharply under a load of 50 virtual users. At the same time, the corresponding indicator for the optimized service remains at a significantly lower level and changes only slightly as the load increases.

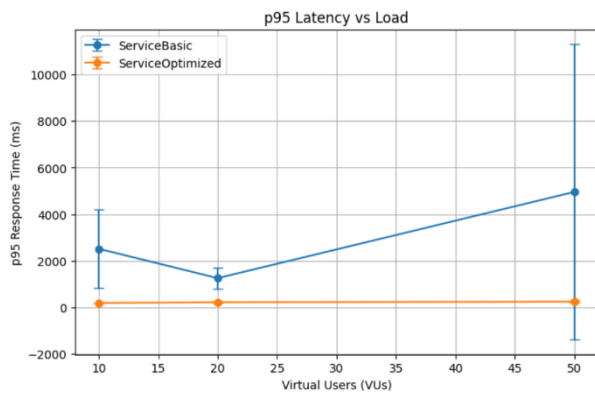


Fig. 3. Dependence of the p95 Response-Time Percentile of the Services on the Load Level

This relationship between the results confirms that the optimized implementation not only reduces the average request processing time but also ensures greater predictability of system operation under peak loads.

The considerable dispersion of values for the basic service indicates its increased sensitivity to rising request intensity and less stable operation under scaling conditions. In contrast, the low p95 response-time percentile values and minor variability between repeated experiments for the optimized service indicate greater resilience to load. Fig. 4 shows the dependence of service throughput on the load level specified by the number of virtual users. The obtained results indicate that, as the load increases, both implementations increase the number of processed requests per second; however, the optimized service demonstrates higher throughput in all scenarios. This difference becomes particularly noticeable at 50 virtual users, where the advantage of the optimized implementation is most pronounced.

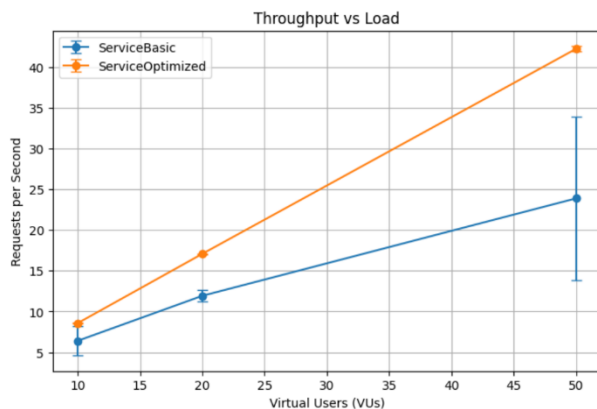


Fig. 4. Dependence of the Throughput of the Services on the Load Level

The presented graph confirms that the applied .NET platform tools and software improvements ensure more efficient resource utilization during request processing. The basic implementation also demonstrates greater variability in the results, whereas the optimized service operates more stably. In the context of this research, this confirms the feasibility of the proposed approach to implementing microservices in container orchestration systems. Fig. 5 shows the change in the number of service replicas after load testing at different request intensity levels.

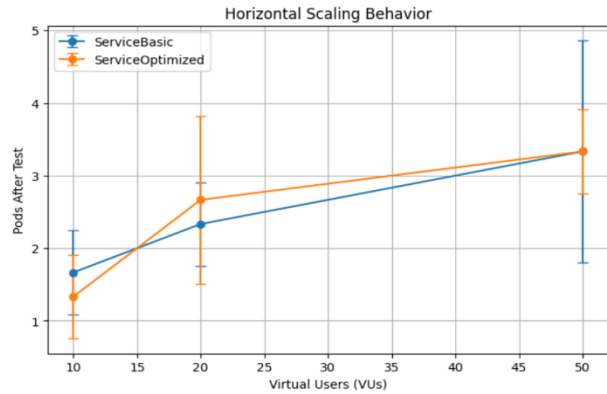


Fig. 5. Change in the Number of Service Replicas During Horizontal Scaling

The obtained results demonstrate that both implementations respond to increased load by increasing the number of replicas, which means that the horizontal autoscaling mechanism operates correctly for both the basic and optimized services. At the same time, the optimized implementation demonstrates a slightly higher average scaling level even under medium load, whereas at 50 virtual users, the values for both services become nearly identical.

The presented graph confirms that the proposed approach makes it possible to study not only the timing characteristics and throughput of .NET microservices but also the specific features of their behavior during horizontal autoscaling in the Kubernetes environment.

The weights were determined using the entropy method based on six aggregated scenarios and are presented in Fig. 6. The weight w_e is equal to zero because the percentage of failed requests was 0% in all 18 runs. This does not mean that the indicator was excluded from the mathematical model – in this experiment, it simply had no discriminative capability. The obtained operational metrics were used to calculate the partial and integral indicators of the risk of defect manifestation. They are presented on Fig. 7. Before performing the calculations, the results of the three repeated runs were averaged separately for each service implementation and load level. For the timing characteristics and throughput, the mean values, standard deviations, and coefficients of variation were determined.

Component	Designation	Weights
Timing characteristics	w_t	0,233
Failed requests	w_e	0
Insufficient throughput	w_q	0,182
Resource utilization	w_r	0,129
Result variability	w_v	0,206
Scaling inefficiency	w_s	0,251

Fig. 6. Weights

Realization	Users	r_t	r_e	r_q	r_r	r_v	r_s	R	Level
ServiceBasic	10	0,556	0	0,67	0,68	0,48	0,205	0,489	medium
ServiceBasic	20	0,285	0	0,56	0	0,2	0,198	0,26	low
ServiceBasic	50	1	0	1	0,74	1	0,8	0,917	high
ServiceOptimized	10	0	0	0,02	0,36	0	0,003	0,052	low
ServiceOptimized	20	0,006	0	0	0,15	0,06	0,001	0,033	low
ServiceOptimized	50	0,011	0	0,05	0,08	0,04	0,022	0,036	low

Fig. 7. Partial and Integral Indicators of the Risk of Defect Manifestation

Since the number of requests processed per second depends on the load intensity, relative throughput, defined as the ratio of the average number of requests per second to the number of virtual users, was used to ensure comparability between the scenarios. The operational indicators were then normalized, and the timing risk component r_t , failed-request component r_e , insufficient-throughput component r_q , resource component r_r , variability component r_v , and scaling-inefficiency component r_s were calculated. The weighting coefficients of the partial indicators were determined using the entropy method. The highest weights were assigned to the scaling-inefficiency component, $w_s = 0.251$, the timing component, $w_t = 0.233$, and the result-variability component, $w_v = 0.206$. The weights of insufficient throughput and resource utilization were $w_q = 0.182$ and $w_r = 0.129$, respectively. Since the percentage of failed requests was zero in all experimental runs, the r_e indicator had no discriminative capability, and its weight within the conducted experiment was set to zero.

The calculation results showed that, for the basic implementation with 10 virtual users, the integral indicator was $R = 0.489$, corresponding to a medium risk level. This value is explained by the considerable variability in the average response time and its 95th percentile between repeated runs. With 20 virtual users, the indicator decreased to $R = 0.260$, corresponding to a low risk level. This decrease is associated with the more stable results of runs B20_1–B20_3 and lower average resource utilization. Under a load of 50 virtual users, the integral indicator of the basic implementation increased to $R = 0.917$, corresponding to a high-risk level. The obtained value was primarily influenced by increases in the average response time and its 95th percentile, a decrease in relative throughput, considerable dispersion of the results, and an increase in the number of replicas without sufficient stabilization of the timing characteristics. For the optimized implementation, the integral indicator values were $R = 0.052$, $R = 0.033$, and $R = 0.036$ for 10, 20, and 50 virtual users, respectively. All the obtained values correspond to a low risk level. The minor change in the indicator as the load increases confirms the stability of the optimized software implementation and its ability to maintain predictable timing characteristics and throughput throughout the investigated load range. Fig. 8 shows the dependence of the integral indicator of the risk of defect manifestation for the basic and optimized .NET microservice implementations on the number of virtual users.

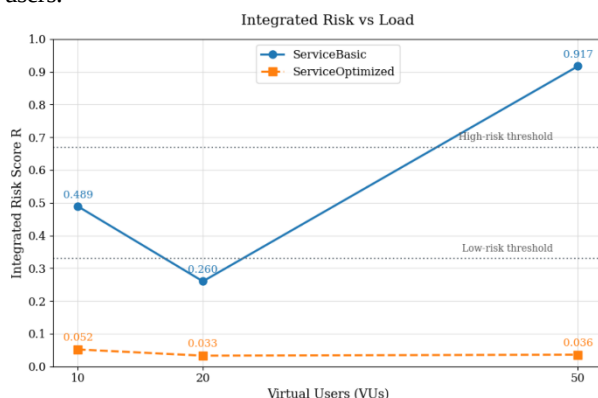


Fig. 8. Dependence of the Integral Indicator of the Risk of Defect Manifestation in .NET Microservices on the Load Level

A comparison of the obtained values shows that the optimized implementation reduced the integral risk indicator by 89.5% with 10 virtual users, by 87.3% with 20 virtual users, and by 96.1% with 50 virtual users compared with the basic implementation. The results obtained confirm the operability of the proposed method and the possibility of its application for the comparative assessment of the risk of defect manifestation in functionally similar software components based on operational metrics.

4. Discussion of the Results of the Study of the Method for Assessing the Risk of Defect Manifestation in .NET Microservices

The proposed mathematical model for assessing the risk of defect manifestation in .NET microservices is based on the comprehensive consideration of timing characteristics, throughput, the percentage of failed requests, computational resource utilization, result variability, and service behavior during horizontal scaling. A distinctive feature of the model is the integration of heterogeneous operational metrics into a single integral indicator R . The experimental study revealed substantial differences between the risk profiles of the basic and optimized implementations. For Service Basic with 10 virtual users, the value $R = 0.489$ was obtained, corresponding to a medium risk level. With 20 users, the indicator decreased to $R = 0.260$ and corresponded to a low risk level. This decrease is explained by the lower variability of the B20 series results compared with the B10 series, which exhibited considerable dispersion in the average response time and its 95th percentile. With 50 virtual users, the integral indicator of the basic implementation increased to $R = 0.917$, corresponding to a high-risk level. The main reasons were a sharp increase in the timing characteristics, a decrease in relative throughput, increased result variability, and an increase in the number of replicas without sufficient performance stabilization. For Service Optimized, the integral indicator values were $R = 0.052$, $R = 0.033$, and $R = 0.036$ with 10, 20, and 50 virtual users, respectively. All values correspond to a low risk level and change only slightly as the load increases. Compared with the basic implementation, optimization of the software logic reduced the integral risk indicator by 89.5% with 10 users, by 87.3% with 20 users, and by 96.1% with 50 users. The results obtained are explained using asynchronous request processing, reduced blocking, caching, and more efficient organization of computations. The absence of failed requests in all experimental scenarios resulted in zero values for the partial component r_e and its weighting coefficient. This does not indicate the absence of defects in the software implementation, since their manifestation may be associated not only with failed request completion but also with excessive delays, instability of timing characteristics, reduced throughput, and inefficient resource utilization. Therefore, the proposed method provides an assessment of a broader class of adverse states of a software component.

Unlike the approaches presented in [1, 6], which focus on the formal verification of autoscaling policies, the proposed method assesses the risk associated with the behavior of the software implementation itself. In [2–5, 7], operational metrics are used primarily for predicting resource requirements, managing scaling, and ensuring

quality of service, whereas the proposed method considers them as interrelated indicators of the risk of defect manifestation. Unlike [8], the proposed approach involves forming an integral assessment that enables the quantitative comparison of the operational profiles of different software implementations. The research results should be applied with consideration of the fact that the risk assessment is comparative and does not represent the probability that a defect is present in the program code. The experiments were conducted in a local Kubernetes cluster for two implementations of one microservice, three load levels, and a fixed test duration. Normalization was performed within the generated dataset; therefore, when applying the method to other systems, the threshold values of the metrics and the weighting coefficients must be determined again.

Further research should focus on validating the method for a larger number of software components, extending the set of load scenarios, and conducting experiments in production clusters. The use of data on confirmed defects, the simulation of individual component failures, and the consideration of network delays. This will make it possible to refine the boundaries of the risk levels and move from a comparative integral assessment to predicting the probability of defect manifestation.

5. Conclusions

1. A mathematical model for assessing the risk of defect manifestation in .NET microservices within distributed computer systems was proposed. A distinctive feature of the model is the comprehensive consideration of the average response time, its 95th percentile, throughput, the percentage of failed requests, CPU and memory utilization, the number of replicas, and the variability of the experimental results. The statistical processing of repeated runs made it possible to consider not only the mean metric values but also the stability of the software component's behavior. To ensure the correct comparison of scenarios with different load intensities, relative throughput, defined as the number of requests per second per virtual user, was used.

2. Partial indicators and an integral criterion for assessing the risk of defects in a software component were defined. The timing risk component, failed-request component, insufficient-throughput component, resource component, result-variability component, and horizontal-scaling-inefficiency component were formed. To reduce the subjectivity of the assessment, the weighting

coefficients were determined using the entropy method. The integral indicator (R) takes values from 0 to 1 and enables the operational state of the component to be classified as having a low, medium, or high-risk level. In the experiments conducted, the percentage of failed requests was zero; therefore, the corresponding component did not affect the integral assessment but was retained in the general mathematical model.

3. A method for assessing the risk of defect manifestation in .NET microservices based on the collection, statistical processing, normalization, and comprehensive analysis of operational metrics was developed. The experimental study was conducted for basic and optimized implementations of a functionally similar service under identical Kubernetes container orchestration conditions. For ServiceBasic, the integral indicator was $R = 0.489$ with 10 virtual users, $R = 0.260$ with 20 users, and $R = 0.917$ with 50 users, corresponding to medium, low, and high-risk levels, respectively. For ServiceOptimized, the value obtained were $R = 0.052$, $R = 0.033$, and $R = 0.036$, all of which corresponded to a low risk level. The optimized implementation reduced the integral risk indicator by 89.5–96.1% compared with the basic implementation. The obtained results confirmed the operability of the proposed method and the possibility of using it for the comparative assessment of the risk of defect manifestation in software components based on operational data without directly analyzing their program code.

Conflict of Interest

The authors declare that they have no conflicts of interest related to this research, including those of a financial, personal, authorship-related, or other nature that could have influenced the research and its results presented in this article.

Funding

The research was conducted without financial support.

Data Availability

The manuscript has no associated data.

Use of Artificial Intelligence Tools

The authors confirm that they did not use artificial intelligence technologies in the preparation of the presented work.

REFERENCES

1. Jawaddi S.N.A., Johari M.H., and Ismail, A. (2022), "A review of microservices autoscaling with formal verification perspective", *Softw Pract Exper.*, vol. 52, is.11, pp. 2476–2495, doi: <https://doi.org/10.1002/spe.3135>
2. Taheri, J., Gördén, A., and Al-Dulaimy, A. (2024), "Using Machine Learning to Predict the Exact Resource Usage of Microservice Chains", *UCC '23*, article 25, pp. 1–9, doi: <https://doi.org/10.1145/3603166.3632166>
3. Nunes, J. P. K. S., Nejati, S., Sabetzadeh, M., and Nakagawa E. Y. (2024), "Self-adaptive, Requirements-driven Autoscaling of Microservices", *SEAMS '24*, pp. 168–174, doi: <https://doi.org/10.1145/3643915.3644094>
4. Pintye, I., Kovács, J., and Lovas, R. (2024), "Enhancing Machine Learning-Based Autoscaling for Cloud Resource Orchestration", *J Grid Computing*, vol. 22(68), doi: <https://doi.org/10.1007/s10723-024-09783-1>
5. Alharthi, S., Alshamsi, A., Alseiyari, A., and Alwarafy, A. (2024), "Auto-Scaling Techniques in Cloud Computing: Issues and Research Directions", *Sensors*, vol. 24(17), art. 5551, doi: <https://doi.org/10.3390/s24175551>
6. Jawaddi, S.N.A., Ismail, A. et al. (2025), "Analyzing Energy-Efficient and Kubernetes-Based Autoscaling of Microservices Using Probabilistic Model Checking", *J Grid Computing*, vol. 23(3), doi: <https://doi.org/10.1007/s10723-024-09789-9>
7. Gurung, N., Shrestha, S., and Chulyadyo, R. (2025), "Scalability in Microservices: A systematic literature review", *Journal of Computer Science and Technology*, vol. 25(2), e.11, doi: <https://doi.org/10.24215/16666038.25.e11>
8. Medavarapu, S.V., and Medavarapu, S.S. (2023), "Enhancing Scalability and Performance of .NET Microservices on Azure Kubernetes Service (AKS)", *J. Sc. and Eng. Res.*, vol. 10(6), pp. 133–139, doi: <https://doi.org/10.5281/zenodo.13597405>

9. Siswanto, M., and Yuhana, U.L. (2023), "Software Defect Prediction Based on Optimized Machine Learning Models: A Comparative Study", *Teknika*, vol. 12(2), 1pp. 66–172, doi: <https://doi.org/10.34148/teknika.v12i2.634>
10. Kabir, S.M.H., Rahman, Md. T., and Mridul, A.H. (2025), "Software Defect Prediction Using Traditional Machine Learning and Ensemble Learning Algorithms", *Smart Wearable Tech.*, vol. 1, A9, doi: <https://doi.org/10.47852/bonviewSWT52025645>
11. Zain, Z.M., Sakri, S., and Ismail, N.H.A. (2023), "Application of Deep Learning in Software Defect Prediction: Systematic Literature Review and Meta-analysis", *Inf. Softw. Technol.*, vol. 158, doi: <https://doi.org/10.1016/j.infsof.2023.107175>
12. Marijan, D. (2023), "Comparative study of machine learning test case prioritization for continuous integration testing", *Software Qual J.*, vol. 31, pp. 1415–1438, doi: <https://doi.org/10.1007/s11219-023-09646-0>
13. Cabral, G.G., Minku, L.L. et. al. (2023), "An investigation of online and offline learning models for online Just-in-Time Software Defect Prediction", *Journal of Systems and Software*, vol. 28, doi: <https://doi.org/10.1007/s10664-023-10335-6>
14. Bayramova, T., (2023), "Software defect prediction using the machine learning methods", *Problems of Information Technology*, vol. 14(2), doi: <http://doi.org/10.25045/jpit.v14.i2.03>
15. Chen, Z., Wang, K., Yang, C., Cao, B., and Fan, J., (2025), "Learning Top-k Classification with Label Ranking", *Communications in Computer and Information Science*, vol. 2755, doi: https://doi.org/10.1007/978-981-95-4094-5_9
16. Tang, P., Wang, J., and Liu, M. (2025), "Variational learning to rank for Test Case Prioritization via prioritizing metric inspired differentiable loss", *Engineering Appl. of Artificial Intelligence*, vol. 141, doi: <https://doi.org/10.1016/j.engappai.2024.109776>
17. Stradowski, S., and Madeyski L. (2023), "Machine learning in software defect prediction: A business-driven systematic mapping study", *Information and Software Technology*, vol. 155, doi: <https://doi.org/10.1016/j.infsof.2022.107128>
18. Robredo, M., Esposito, M., Taibi, D., Peñaloza, R., and Lenarduzzi, V., (2025), *SQuAD: The Software Quality Dataset - Dataset* [Data set], doi: Zenodo. <https://doi.org/10.5281/zenodo.17566691>
19. Mahbub, P., Shuvo, O., and Rahman, M.M. (2023), "Defectors: A Large, Diverse Python Dataset for Defect Prediction", 2023 IEEE/ACM 20th Int. Conf. on MSR, pp. 393-397, doi: <https://doi.org/10.1109/MSR59073.2023.00085>
20. Pachouly, J., Ahirrao, S., Kotecha, K., Selvachandran, G., and Abraham A. (2022), "A systematic literature re-view on software defect prediction using artificial intelligence: Datasets, Data Validation Methods, Approaches, and Tools", *Engineering Applications of Artificial Intelligence*, vol. 111, pp. 1–33, doi: <https://doi.org/10.1016/j.engappai.2022.104773>
21. Cabral, G.G., Minku, L.L., Oliveira, A.L., Pessoa, D.A., and Tabassum, S. (2023), "An investigation of online and offline learning models for online Just-in-Time Software Defect Prediction", *Journal of Systems and Software*, vol. 28, 121, doi: <https://doi.org/10.1007/s10664-023-10335-6>
22. Guo, Z., Sun, Y., Fu, B., and Cheng, X., (2022), "Research on software defect prediction method based on machine learning oriented to large-scale complex information system", *ICIEIS2021*, doi: <https://doi.org/10.1117/12.2627233>
23. Giray, G., Bennin, K. E., Köksal, Ö., Babur, Ö., and Tekinerdogan, B. (2023), "On the use of deep learning in software defect prediction", *Journal of Systems and Software*, vol. 195, pp. 1–26, doi: <https://doi.org/10.1016/j.jss.2022.111537>

Received (Надійшла) 06.06.2026

Accepted for publication (Прийнята до друку) 19.08.2026

Publication date (Дата публікації) 25.09.2026

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Дяченко Дмитро Олександрович – аспірант кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Dmytro Diachenko – PhD student, Department of Electronic Computers, Kharkiv National University of Radio Electronics Kharkiv, Ukraine;

e-mail: dmytro.diachenko2@nure.ua; ORCID Author ID: <http://orcid.org/0009-0006-5751-3511>.

Шостак Максим Віталійович – аспірант кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Maksym Shostak – PhD student, Department of Electronic Computers, Kharkiv National University of Radio Electronics Kharkiv, Ukraine;

e-mail: maksym.shostak@nure.ua; ORCID Author ID: <https://orcid.org/0009-0003-1510-7425>.

Метод оцінки ризику прояву дефектів .NET-мікросервісів у розподілених комп'ютерних системах на основі експлуатаційних метрик

Д. О. Дяченко, М. В. Шостак

Анотація. Актуальність. Дефекти програмного забезпечення, не виявлені під час функціонального тестування, можуть проявлятися під навантаженням, що зумовлює актуальність оцінювання ризику на основі експлуатаційних метрик.

Предмет дослідження: методи та засоби оцінки ризику прояву дефектів .NET-мікросервісів у розподілених комп'ютерних системах з урахуванням експлуатаційних метрик. **Мета дослідження:** розробка методу оцінки ризику прояву дефектів .NET-мікросервісів у розподілених комп'ютерних системах на основі комплексного аналізу експлуатаційних метрик. **Завдання дослідження.** Для досягнення мети визначено сукупність експлуатаційних метрик і розроблено модель формування інтегрального показника ризику на основі їх нормованих значень. **Методи дослідження.** Використано методи системного аналізу, нормування та інтегрального оцінювання, навантажувального тестування, а також статистичної обробки. **Результати дослідження.** Запропоновано метод інтегральної оцінки ризику прояву дефектів .NET-мікросервісів за нормованими експлуатаційними метриками. **Висновки.** Запропонований метод дає змогу оцінювати ризик прояву дефектів програмних компонентів розподілених комп'ютерних систем без втручання в їхній програмний код на основі експлуатаційних даних, отриманих під час навантажувального тестування.

Ключові слова: ризик прояву дефектів; .NET-мікросервіси; розподілені комп'ютерні системи; експлуатаційні метрики, інтегральний індикатор ризику; навантажувальне тестування; пропускна здатність; Kubernetes; Docker; надійність програмного забезпечення.

A. Kapiton¹, T. Franchuk², D. Tyshchenko², R. Zakharov²

¹National University «Yuri Kondratyuk Poltava Polytechnic», Poltava, Ukraine

²State University of Trade and Economics, Kyiv, Ukraine

DEVELOPMENT OF DECENTRALIZED FORECASTING SYSTEMS BASED ON MULTI-CLOUD INFRASTRUCTURE

Abstract. Relevance. The study substantiates the need for a cloud environment that ensures a high level of system scalability, enabling the integration of new production facilities, workshops, and equipment without the need for capital investment in proprietary server infrastructure. **The subject of the study** is the processes of developing and operating decentralized forecasting systems based on multi-cloud infrastructure. **The aim of the study** is to analyze modern methods for developing decentralized forecasting systems based on multi-cloud infrastructure and to enhance their efficiency, reliability, and fault tolerance. **The research objectives** are to investigate the architectural aspects of creating decentralized forecasting systems in a cloud environment; to examine the effectiveness of using distributed computing for technical diagnostics; and to determine a software solution for time-series forecasting using recurrent neural networks. to analyze cloud platforms and machine learning libraries for distributed time-series modeling; to identify a key challenge regarding the end-to-end analysis of various neural network architectures and formulate technical requirements to overcome this limitation.

Research methods. The methodology encompasses systems analysis, mathematical modeling, blockchain technologies, and machine learning methods. **Research results.** The feasibility of employing cloud and distributed architectures to address diagnostic tasks is substantiated, driven by the specific requirements for processing large volumes of data continuously generated by primary sensors. It is determined that a necessary prerequisite is the development and investigation of a time-series forecasting software system based neural network models utilizing cloud computing; such a system must implement the complete data processing cycle from acquisition to forecast generation and quality assessment thereby enabling its subsequent application in the field of computer engineering. A structured plan for implementing the system's end-to-end operational cycle is presented. **Conclusions.** The implementation and deployment of such forecasting systems within a cloud infrastructure, utilizing containerization and orchestration via managed services align with modern engineering standards that ensure high fault tolerance and horizontal scalability for analytical platforms.

Keywords: forecasting systems; cloud infrastructure; distributed systems; data synchronization; network latency; data orchestration; machine learning; computer engineering; information technology.

1. Introduction

1.1 Relevance of the topic

A cloud-based, distributed time-series forecasting system for vibration diagnostics is essential for preventing industrial accidents and transitioning from scheduled repairs to condition-based maintenance. It analyzes the oscillations and vibrations of machine tools, turbines, or engines to detect hidden defects well before the equipment fails. The system's primary purpose is to predict equipment failures (Predictive Maintenance) and calculate remaining useful life (RUL) for precise maintenance planning. Continuous, automated real-time sensor monitoring eliminates the need for manual measurements.

The system ensures early defect detection by identifying micro-cracks, imbalance, bearing wear, and shaft misalignment. This approach minimizes financial losses by eliminating the risk of sudden conveyor shutdowns and completely preventing prolonged industrial production downtime.

1.2 Analysis of literature sources and problem definition

A thorough exploration of the topic requires examining methodological approaches to evaluating decentralized forecasting systems in a cloud environment. Since this area has been overlooked in recent publications, the design of such systems currently represents one of the most pressing directions in contemporary research.

C. Chatfield addresses general issues of time series forecasting [1].

D. Popovic investigates computational intelligence in time series forecasting [2].

G. James, D. Witten, T. Hastie, and R. Tibshirani examine the theoretical and practical aspects of statistical learning [3].

In other the works are explore the applied aspects of encoder-decoder architectures based on recurrent neural networks [4].

S. Amari discusses the backpropagation and stochastic gradient descent methods [5].

Researchers are specialists in wavelet analysis and financial time series forecasting [6, 7].

The authors describe the characteristics and issues associated with the problems of investigate applied aspects, specifically forecasting on the NSE stock market using deep learning models [8].

Other articles open the topic of examine the learning stages involved in recognizing long-term dependencies using gradient descent, noting it as a complex task [9].

A number of works are devoted to the problems to analyze this issue through the lens of Long Short-Term Memory and investigate and empirically evaluate recurrent neural networks with gating mechanisms in sequence modeling tasks, and also conduct a comparative analysis of financial time series forecasting [10–13].

A number of researchers examine modern development trends and architectural features of

computers, as well as the fundamentals of modeling decentralized forecasting systems based on cloud platforms [14–16]. An objective study of the problem requires an analysis of existing solutions in the field of cloud-based distributed forecasting systems. Previous studies overlooked this aspect, making the development methodology for such decentralized platforms a central topic of current scientific discussions.

1.3 Aim and objectives of the study

The aim of the study is to analyze methods for developing decentralized forecasting modern systems, which based on multi-cloud infrastructure, enhance their efficiency, reliability, fault tolerance.

To achieve the stated objective, the following scientific and practical tasks have been formulated:

- investigating architectural solutions for constructing decentralized forecasting systems in a cloud environment;
- evaluating the effectiveness of using distributed computing in technical diagnostics processes based on the continuous analysis of large datasets from primary sensors;
- identify a software solution for time-series forecasting using recurrent neural networks with GRU and LSTM architectures;
- analyze cloud platforms and machine learning libraries for distributed time-series modeling;
- identify a key challenge for the end-to-end analysis of various neural network architectures and formulate technical requirements to address this limitation.

2. Research materials and methods

The object of the study is the process of forecasting distributed time series (large volumes of diagnostic data) in decentralized multi-cloud infrastructures. Research hypothesis: The implementation and deployment of forecasting systems within a cloud infrastructure, utilizing Docker containerization and orchestration via managed services, enable high fault tolerance and horizontal scalability for analytical platforms in accordance with modern engineering standards.

The study formalizes the stages of the development lifecycle for predictive systems at the intersection of Data Science methodology, cloud engineering, and the diagnostic domain, enabling the formulation of a structured architecture for an end-to-end operational cycle. A comparative analysis of cloud technologies relevant to forecasting distributed time series in computer engineering was conducted. Neural network modeling tools based on the TensorFlow, PyTorch, and Prophet libraries were investigated.

A critical scientific and practical barrier was identified: the lack of an integrated environment with a graphical user interface for the comparative analysis of heterogeneous architectures within a unified computational pipeline. Based on the identified limitations, a set of technical requirements was established for the proposed software system, designed to minimize these shortcomings and ensure the practical value of the results obtained.

3. Results of development and research

3.1 Design and practical implementation of decentralized intelligent forecasting systems

The design and practical implementation of decentralized intelligent forecasting systems based on heterogeneous multi-cloud infrastructure entail a complex set of critical architectural and technological challenges. Key among these are the systemic complexity of orchestrating distributed data arrays, heightened risks regarding the confidentiality and integrity of the information perimeter, and deterministic latencies in inter-node data exchange. The primary factors negatively affecting the efficiency of such systems are the processes of ensuring real-time data consistency and synchronization, the structural and functional heterogeneity of the cloud platforms employed, and the high computational and network overhead associated with maintaining communication between distributed components. The need to employ cloud and distributed architectures for vibration diagnostics tasks stems from the specific requirements of processing large datasets (Big Data) continuously generated by primary sensors. Since modern vibration sensors produce high-frequency signals - yielding thousands of measurements per second - parallel processing of these data streams is implemented using distributed computing frameworks. Utilizing a cloud environment ensures a high level of system scalability, enabling the integration of new production facilities, workshops, and machinery without the need for capital investment in on-premises server infrastructure. Furthermore, mathematical modeling and the training of predictive AI neural network structures - specifically recurrent network and transformer architectures - require substantial computational resources, which are efficiently provided by cloud-based graphics processing units (GPUs).

Above all, this architecture implements the concept of data centralization, aggregating diagnostic information from all the enterprise's remote sites into a single, consolidated monitoring interface to facilitate management decision-making [1, 4, 8].

Time series forecasting is a fundamental applied task that is critical to the financial, energy, telecommunications, medical, and computer science sectors. Traditional statistical approaches - such as ARIMA and exponential smoothing - demonstrate satisfactory performance only when dealing with stationary data and stable variance. However, their predictive capability is significantly limited in the presence of heteroscedasticity, or structural shifts. To adequately model this processes in such series, modern researchers employ networks.

In particular, Hochreiter and Schmidhuber's LSTM architecture overcame tasks, enabling some effective detection of long-range dependencies in data sequences. A simplified variant, the Gated Recurrent Unit, used fewer parameters while maintaining comparable accuracy. A comparative study by Siami-Namini et al. confirmed that LSTM and GRU reduce RMSE by an average of 84–87% compared to ARIMA when applied to financial time series [1, 5, 6, 13].

Despite the powerful toolsets for building neural network models available in libraries such as TensorFlow, PyTorch, and Prophet, they lack an integrated environment with a graphical user interface for the comparative analysis of architectures within a unified computational pipeline. Consequently, the development of a custom software suite offering these capabilities constitutes a practical contribution of this research. Implementing and deploying such predictive systems within a cloud infrastructure leveraging Docker containerization and orchestration via managed services (such as AWS Fargate, Google Cloud Run, or Azure Container Apps) reflects modern engineering standards that ensure high fault tolerance and horizontal scalability for analytical platforms.

3.2 A structured plan for implementing the end-to-end system work cycle

A necessary requirement is to develop and investigate a time-series forecasting software system based on GRU and LSTM neural network models using cloud computing; this system must implement the complete data processing cycle - from data loading to generating a forecast and evaluating its quality - for subsequent application in computer engineering. The development and research of such a system are broken down into distinct stages that combine Data Science methodology, cloud engineering, and the specifics of vibration diagnostics.

A structured plan for implementing the end-to-end system work cycle is presented in Fig.1.

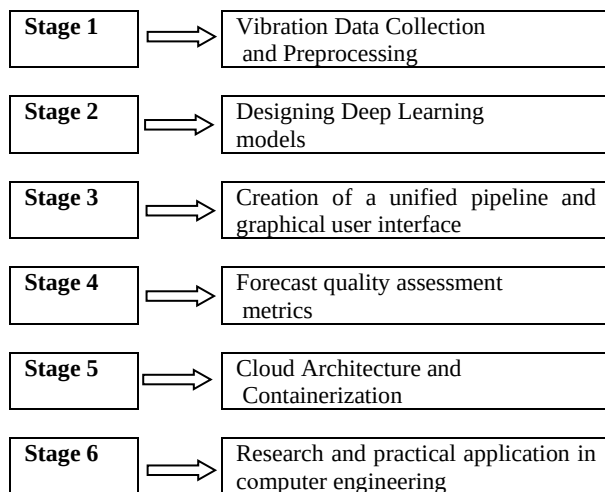


Fig. 1. A structured plan for implementing the end-to-end system work cycle

Stage 1. Vibration Data Collection and Preprocessing (Data Engineering). Vibration signals are high-frequency time series (typically ranging from 2 to 50 kHz). The full cycle begins with data ingestion and adaptation. Loading: Organizing streaming or batch data ingestion via APIs (e.g., from industrial IoT gateways such as those from SKF or Siemens). Cleaning and Resampling: Removing anomalies, filling in missing values, downsampling (if long-term trend analysis is required), or applying sliding windows to generate data samples. Feature Engineering:

The effectiveness of GRU/LSTM architectures depends directly on the amplitude-frequency properties of the data. Temporal parameters (RMS, crest factor, kurtosis) and spectral features calculated via FFT or DWT are evaluated. Normalization: Scaling data to the $[0, 1]$ or $[-1, 1]$ range using MinMaxScaler or StandardScaler (critical for the convergence of recurrent networks) [1, 6-14].

Stage 2. Designing Deep Learning models. Given the time-dependent nature of vibration processes, where the current state of structural elements (such as rotor shafts or rolling-element bearings) is directly determined by their prior history.

In particular, the Long Short-Term Memory (LSTM) architecture utilizing a mechanism of such functional gates enables the effective approximation of long-term nonlinear wear trends and failure prediction. As a more computationally optimized alternative, the Gated Recurrent Unit (GRU) modification is employed; it implements a simplified structure based on two gates that control data flow.

It is computationally "lighter," trains faster in the cloud, and often demonstrates similar or superior accuracy on short time series.

Hybrid approaches (optional): A combination of 1D-CNN (for automatic feature extraction from vibration signals) with subsequent LSTM/GRU layers for trend forecasting [1, 8, 14].

Stage 3. Creation of a unified pipeline and graphical user interface. To overcome the limitations of standard libraries, a monolithic or microservice-based application is being developed (e.g., using Python/FastAPI/Streamlit or Dash): Configuration module: Via the GUI, the user selects the model (LSTM or GRU) and specifies hyperparameters (number of layers, batch size, epochs, forecasting window).

Unified pipeline: Single-button process execution: Data loading → Preprocessing → Training/Validation → Result visualization [1, 4, 8].

Stage 4. Forecast quality assessment metrics: For comparative analysis, the interface includes a mathematical module that evaluates the deviation of the predicted vibration from the actual values: MAE (Mean Absolute Error): Indicates the average absolute error in measurement units (e.g., mm/s for vibration velocity). RMSE (Root Mean Squared Error): More sensitive to large outliers and anomalous vibration spikes. MAPE (Mean Absolute Percentage Error): Relative error expressed as a percentage, useful for business reports [1, 4, 8].

Stage 5. Cloud Architecture and Containerization (Cloud Deployment) To ensure horizontal scalability and availability, the system is deployed to the cloud: Containerization: The application (FastAPI backend and Streamlit frontend) is packaged into Docker containers. Data Storage: Raw vibration data is stored in cloud storage services (AWS S3, Google Cloud Storage), while prediction results and metrics are stored in time-series databases (InfluxDB, TimescaleDB).

Deployment: Serverless or managed container services are used. They automatically allocate resources: as the load increases, the cloud spins up new container

instances. GPU-based training: Cloud instances with GPU support are utilized for periodic model retraining [1, 4, 8].

Stage 6. Research and practical application in computer engineering. The final stage of the research involves testing the system against real-world benchmarks (such as the well-known NASA IMS Bearing Dataset or data from a specific industrial facility): Comparative analysis: Running LSTM and GRU models on identical vibration data using the same hyperparameters. Performance evaluation: Comparison based on accuracy metrics (MAE/RMSE) and training/inference speed.

Predictive diagnostics: Based on the generated forecast, the system identifies the point at which vibration levels will exceed critical thresholds (in accordance with ISO 10816 standards), enabling the calculation of the machine's Remaining Useful Life (RUL) and the timely issuance of maintenance alerts [1, 4, 8].

The developed software suite operates as a cross-platform desktop application based on a two-tier architecture.

The presentation layer (interface) facilitates interactive user engagement and the visualization of analytical data, while the business logic layer (computational module) handles data import, neural network training, and the generation of forecasts. This clear separation of responsibilities isolates the interface components from the algorithmic logic, significantly simplifying future system modernization and scaling.

3.3 Improving the sequential data processing procedure in the system

The sequential information processing pipeline within the system comprises the following stages: automatic retrieval of time series via the Yahoo Finance API structuring data into a pandas.

To facilitate inter-level interaction, an asynchronous queue mechanism has been integrated: upon completion of the model training procedure, the computational component encapsulates the output data into a `ModelResult` object and delegates it to the `result_queue`, while the main GUI thread retrieves this object via cyclic polling using the `root.after()` method. This architectural approach isolates the time-consuming neural network training processes, preventing the blocking of the Tkinter main event loop. Consequently, the user interface remains continuously responsive and stable while resource-intensive background operations are being executed.

List of the software system's main components and their functional purpose:

- `VibrationForecastApp` (Class, `main.py`) - the application's main class responsible for window initialization, GUI component construction, launching the background training thread, and visualizing vibration diagnostics results and remaining useful life (RUL) forecasts;

- `ModelResult` (Dataclass, `forecasting.py`) - a data structure designed to store the training results of a single model, including actual and predicted vibration

velocity/acceleration values, accuracy metrics, future forecasts, and the loss curve;

- `fetch_vibration_data` (Function, `forecasting.py`) - a function for loading raw time-series data from vibration sensors (via industrial CSV files, time-series databases, or IoT gateways) and performing initial filtering;

- `extract_vibration_features` (Function, `processing.py`) - a module for calculating vibration signal diagnostic features in the time and frequency domains (RMS, crest factor, kurtosis, FFT);

- `train_and_evaluate` (Function, `forecasting.py`) - a function implementing the full workflow for LSTM/GRU models, covering feature normalization (MinMaxScaler), sliding window sequence generation, neural network training, quality metric calculation, and predictive wear trend generation;

- `result_queue` (`queue.Queue`, `main.py`) - an asynchronous queue designed to safely transfer analysis and training results from the background computational thread to the main GUI thread, preventing the interface from freezing.

The system consists of two Python modules: `main.py` and `forecasting.py`. This division adheres to the single responsibility principle: the first module handles only the GUI, while the second manages the computational logic. Separating the interface from the calculation logic simplifies future code maintenance and testing. An analysis of the BTC/USD time series for the last five years revealed no missing values in the target `'Close'` variable. Nevertheless, this validation procedure remains an integral part of the pipeline, as data may arrive in a fragmented state due to network disruptions or technical errors on the market data provider's API side.

4. Discussion of research results

To enable multi-user access to the forecasting system and eliminate the need for local deployment of the Python interpreter on client workstations, a shift to a client-server architecture involving the migration of the computational core to cloud infrastructure has been justified. This architectural paradigm allows the algorithmic logic developed for the `'forecasting.py'` module to remain entirely intact, while transforming the graphical user interface into a web-based frontend or a "thin client" implementation.

The cloud deployment of the analytics platform is carried out through a sequential three-stage process, beginning with the build phase, which involves compiling the Docker software image and subsequently publishing it to a centralized container registry. During the operational phase, the integrated load balancer evenly distributes client requests among active system instances, while a monitoring subsystem based on Prometheus or Cloud Monitoring continuously tracks response latency and error rates.

The creation of decentralized forecasting systems in a multi-cloud environment is complicated by data management issues, security vulnerabilities, and network latency.

Key challenges stem from information reconciliation, platform heterogeneity, and high loads on communication channels.

5. Conclusions

1. A comprehensive analysis entails a detailed examination of recommendations regarding the operation of distributed forecasting models based on cloud platforms. The limited attention paid to this aspect in recent literature has made the development of decentralized forecasting tools a key trend in contemporary research.

2. Based on the research results, a distributed software system for predictive time-series analytics utilizing GRU and LSTM recurrent neural network models and featuring integrated support for cloud deployment was developed and comprehensively analyzed for industrial vibration diagnostics applications. The set objective has been fully achieved: an end-to-end pipeline for processing diagnostic data was successfully implemented spanning from the import of raw high-frequency vibration sensor signals to the generation of a predictive wear trend and the calculation of statistical accuracy metrics.

3. As part of the project, a comparative study of two deep learning architectures was conducted using real-world vibration monitoring data from industrial equipment, and a conceptual framework was established for scaling the local prototype into a containerized cloud service.

4. An analytical review has demonstrated that classical statistical methods (specifically ARIMA and exponential smoothing) are effective only for stationary processes, yet they rapidly lose accuracy when analyzing nonlinear, highly dynamic, and noisy fluctuations in vibration velocity and acceleration. In contrast, modern recurrent neural networks (LSTM, GRU) ensure minimal prediction error for non-

stationary vibration time series associated with the degradation of mechanical components. A critical analysis of existing software confirmed a lack of available tools for the comparative rapid analysis of neural network architectures within a unified computational pipeline. The developed software suite offers practical value and can serve both as a methodological foundation for specialized courses in machine learning and cloud computing and as a functional prototype of an analytical platform for monitoring the actual condition of industrial equipment.

5. Future system modernization will focus on transitioning to a microservices architecture that isolates modules for sensor telemetry collection, background model training, and real-time forecast generation, thereby enabling independent horizontal scaling.

Conflict of interest

The authors declare that they have no conflicts of interest regarding this study whether financial, personal, related to authorship, or of any other nature, that could influence the research and the results presented in this article.

Funding

The study was conducted without financial support.

Data availability

The manuscript has no linked data.

Use of artificial intelligence tools

The authors confirm that they did not use artificial intelligence technologies in the creation of the presented work. The authors confirm that they did not use artificial intelligence technologies in the creation of the presented work.

REFERENCES

1. Hamilton, J. (1994), *Time Series Analysis*, Princeton University Press, Princeton, doi: <https://doi.org/10.2307/j.ctv14jx6sm>
2. Chatfield, C. (2005), "Time-series forecasting", *Significance*, vol. 2, no. 3, pp. 131–133, doi: <https://doi.org/10.1111/j.1740-9713.2005.00117.x>
3. Popovic, D. (2026), *Computational Intelligence in Time Series Forecasting*, Springer-Verlag, London, doi: <https://doi.org/10.1007/1-84628-184-9>
4. Cho, K., Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., and Bengio, Y. (2014), "Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation", *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, Doha, Qatar, pp. 1724–1734, doi: <https://doi.org/10.3115/v1/D14-1179>
5. Rashid, T. (2016), *Make Your Own Neural Network* Scotts Valley: CreateSpace, available at: https://app-01-us-dallas-tx.cashcowfarmer.com/virtual-library/YsnnFO/8AD266/make-your-own-neural-network_tariq-rashid.pdf
6. James, G., Witten, D., Hastie, T., and Tibshirani R. (2013), *An Introduction to Statistical Learning: with Applications in R*, Springer, New York, available at: https://www.stat.berkeley.edu/~rabbee/s154/ISLR_First_Printing.pdf
7. Amari, S. (1993), "Backpropagation and stochastic gradient descent method", *Neurocomputing*, vol. 5, no. 4-5, pp. 185–196, doi: [https://doi.org/10.1016/0925-2312\(93\)90006-O](https://doi.org/10.1016/0925-2312(93)90006-O)
8. Moroz, V., and Milovska, K. (2022), "Wavelet analysis and forecasting of financial time series", *Bulletin of the National Technical University «KhPI». Series: Information and Modeling*, vol. 1, no/ 1-2 (7-8), doi: <https://doi.org/10.20998/2411-0558.2022.02.11>
9. Hiransha, M., Gopalakrishnan, E., Menon, V. and Soman, K. (2018), "NSE Stock Market Prediction Using Deep-Learning Models Procedia", *Computer Science*, vol. 132, pp. 1351–1362, doi: <https://doi.org/10.1016/j.procs.2018.05.050>
10. Bengio, Y., Simard, P., and Frasconi, P. (1994), "Learning long-term dependencies with gradient descent is difficult", *IEEE Transactions on Neural Networks*, vol. 5, no. 2, pp. 157–166, doi: <https://doi.org/10.1109/72.279181>
11. Hochreiter, S., and Schmidhuber, J. (1997), "Long Short-Term Memory", *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, doi: <https://doi.org/10.1162/neco.1997.9.8.1735>
12. Chung, J., Gulcehre, C., Cho, K. H., and Bengio, Y. (2014), "Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling", *arXiv*, arXiv:1412.3555, doi: <https://doi.org/10.48550/arXiv.1412.3555>
13. Chung, J., Gulcehre, C., Cho K. H., and Bengio, Y. (2019), "Comparative Analysis of Forecasting Financial Time Series Using ARIMA, LSTM, and BiLSTM", *arXiv*, arXiv:1911.09512, doi: <https://doi.org/10.48550/arXiv.1911.09512>

14. Franchuk, T., Tyshchenko, D., Desiatko, A., and Karpunin, I. (2025), "Features of accounting digitalization processes", *Galician economic journal*, vol. 95, no 1, pp. 61–66, doi: https://doi.org/10.33108/galicianvisnyk_tntu2025.01.061
15. Kapiton, A., Tyshchenko, D., Franchuk, T., and Desiatko, A. (2025), "Modern Trends in Development and Architectural Features of Computers", *Herald of Khmelnytskyi National University. Technical Sciences*, vol. 359(6.1), pp. 230–234, doi: <https://doi.org/10.31891/2307-5732-2025-359-32>
16. Kapiton, A., Tyshchenko, D., Franchuk, T., Desiatko, A., and Svystun M. (2025), "Advantages and disadvantages of iOS 26", *Information technology and society*, vol. 4 (19), pp. 74–77, doi: <https://doi.org/10.32689/maup.it.2025.4.12>

Received (Надійшла) 15.07.2026

Accepted for publication (Прийнята до друку) 02.09.2026

Publication date (Дата публікації) 25.09.2026

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Капітон Алла Мирославівна – доктор педагогічних наук, професор, професор кафедри комп'ютерних та інформаційних технологій і систем, Національний університет «Полтавська політехніка імені Юрія Кондратюка», Полтава, Україна;
Alla Kapiton – Doctor of Pedagogical Sciences, Professor, Professor of the Department of Computer and Information Technologies and Systems, Yuriy Kondratyuk Poltava Polytechnic National University, Poltava, Ukraine;
e-mail: kits_seminar@ukr.net; ORCID Author ID: <https://orcid.org/0000-0002-7845-0883>;
Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=60159322400>.

Франчук Тамара Михайлівна – кандидат економічних наук, доцент кафедри інженерії програмного забезпечення та кібербезпеки; Державний торговельно-економічний університет, Київ, Україна;
Tamara Franchuk – Candidate of Economic Sciences, Associate Professor of the Department of Software Engineering and Cybersecurity; State University of Trade and Economics, Kyiv, Ukraine;
e-mail: Tamara.Franchuk@ukr.net; ORCID Author ID: <https://orcid.org/0000-0001-7615-1276>;
Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57871557200>.

Тищенко Дмитро Олександрович – доцент, кандидат економічних наук, доцент кафедри інженерії програмного забезпечення та кібербезпеки; Державний торговельно-економічний університет, Київ, Україна;
Dmytro Tyshchenko – Associate Professor, Candidate of Economic Sciences, Associate Professor of the Department of Software Engineering and Cybersecurity; State University of Trade and Economics, Kyiv, Ukraine;
e-mail: tyshchenko_d@knute.edu.ua; ORCID Author ID: <https://orcid.org/0000-0002-2193-9012>;
Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57223189094>.

Захаров Ростислав Геннадійович – доктор філософії, доцент кафедри інженерії програмного забезпечення та кібербезпеки; Державний торговельно-економічний університет, Київ, Україна;
Rostyslav Zakharov – PhD, Associate Professor of the Department of Software Engineering and Cybersecurity; State University of Trade and Economics, Kyiv, Ukraine;
e-mail: r.zakharov@knute.edu.ua; ORCID Author ID: <https://orcid.org/0009-0003-1433-7755>;
Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=58614274000>.

Розробка децентралізованих систем прогнозування на базі мультихмарної інфраструктури

А. М. Капітон, Т. М. Франчук, Д. О. Тищенко, Р. Г. Захаров

Анотація. Актуальність. У дослідженні обґрунтовано потребу в хмарному середовищі, яке забезпечує високий рівень масштабованості системи, уможливаючи інтеграцію нових виробничих потужностей, цехів та обладнання без необхідності капітальних вкладень у власну серверну інфраструктуру. **Предмет дослідження:** процеси розробки та функціонування децентралізованих систем прогнозування на базі мультихмарної інфраструктури. **Мета дослідження:** аналіз сучасних методів розробки децентралізованих систем прогнозування на базі мультихмарної інфраструктури, підвищення їх ефективності, надійності та відмовостійкості. **Завдання дослідження:** дослідити архітектурні аспекти створення децентралізованих систем прогнозування в хмарному середовищі; розглянути ефективність використання розподілених обчислень для технічної діагностики; визначити програмне рішення для прогнозування часових рядів за допомогою рекурентних нейронних мереж; виконати аналіз хмарних платформ і бібліотек машинного навчання для моделювання розподілених часових рядів; виявити ключову проблему для наскрізного аналізу різних нейромережевих архітектур, та сформульовано технічні вимоги для усунення цього обмеження. **Методи дослідження.** Методологія охоплює системний аналіз, математичне моделювання, блокчейн-технології та методи машинного навчання. **Результати дослідження.** Обґрунтовано доцільність застосування хмарних і розподілених архітектур для вирішення діагностичних завдань, що зумовлено специфічними вимогами до обробки великих обсягів даних, які безперервно генеруються первинними датчиками. Визначено, що необхідною умовою є розробка та дослідження програмної системи прогнозування часових рядів на основі обраних моделей нейронних мереж із використанням хмарних обчислень; така система повинна реалізовувати повний цикл обробки даних від їх отримання до формування прогнозу та оцінювання його якості, а отже і для подальшого застосування в галузі комп'ютерної інженерії. Представлено структурований план реалізації наскрізного операційного циклу системи. **Висновки.** Впровадження та розгортання таких систем прогнозування в хмарній інфраструктурі з використанням контейнеризації та оркестрації через керувані сервіси відповідає сучасним інженерним стандартам, що гарантують високу відмовостійкість і горизонтальну масштабованість аналітичних платформ.

Ключові слова: системи прогнозування; хмарна інфраструктура; розподілені системи; синхронізація даних; мережеві затримки; оркестрація даних; машинне навчання; комп'ютерна інженерія; інформаційні технології.

А. А. Коваленко, Г. А. Кучук

Харківський національний університет радіоелектроніки, Харків, Україна

МЕТОД АДАПТИВНОЇ КЛАСТЕРИЗАЦІЇ МОБІЛЬНИХ ВУЗЛІВ ТУМАННОГО ШАРУ ІНТЕРНЕТУ РЕЧЕЙ

Анотація. Актуальність. Основним критерієм кластеризації туманних пристроїв є близькість розташування туманних вузлів, що входять до кластера. Але при наявності мобільних вузлів структура кластера може порушитися внаслідок виходу будь-якого мобільного вузла за межі кластера. Це призведе до непередбачених затримок при обробці даних Інтернету речей. Тому для таких систем додається критерій стійкості структури кластера. **Предметом дослідження** є методи кластеризації туманного шару Інтернету речей з високою щільністю. **Метою дослідження** є підвищення стійкості мобільних компонентів туманного шару за рахунок модифікації методу кластеризації за рахунок розробки методу формування стійкого мобільного кластеру. **Завдання дослідження:** розробка підходу до формування архітектури мобільної компоненти, моделі та методу забезпечення стійкості мобільного кластера туманного шару Інтернету речей з високою щільністю. **Результати дослідження.** Для організації взаємодії мобільних пристроїв розроблено архітектуру мобільної компоненти туманного шару, побудовану з урахуванням його децентралізованої структури та особливостей функціонування мобільних пристроїв Інтернету речей. Запропонована архітектура має чотири рівні та передбачає виділення на нижньому рівні окремих мобільних кластерів туманних пристроїв. Для дослідження процесів формування мобільних кластерів розроблено відповідну модель, у якій враховано випадковий характер переміщення пристроїв Інтернету речей. Моделювання просторового розподілу мобільних компонентів здійснюється на основі точкового процесу Томаса. На основі запропонованої моделі виконано модифікацію алгоритму кластеризації FOREL, у якій визначення положення центра кластера здійснюється з використанням вагових коефіцієнтів, що характеризують стійкість його компонентів. **Висновки.** Результати моделювання засвідчили підвищення стійкості сформованих мобільних кластерів у середовищах із високою та надвисокою щільністю пристроїв. Встановлено, що зі збільшенням середнього відносного відхилення фактичних траєкторій руху пристроїв від запланованих стійкість кластерної структури зростає. Така поведінка пояснюється зміщенням центра кластера у напрямку найбільш нестійких його компонентів, що забезпечує адаптацію кластерної структури до змін просторового розташування мобільних пристроїв. Запропонований метод доцільно застосовувати для організації кластеризації мобільних компонентів туманного шару Інтернету речей за умов, коли середнє відхилення руху пристроїв від заданих траєкторій не перевищує 20 % радіуса кластера.

Ключові слова: мобільний пристрій; інтернет речей, кластеризація, туманний шар, точковий процес Томаса.

1. Вступ

1.1 Актуальність дослідження

Сучасний Інтернет речей (IoT) базується на технології хмарних обчислень. Однак, в останні роки виникли труднощі, які є наслідком таких факторів [1]: наявність географічного розподілу компонент IoT; збільшення мережних затримок; висока вартість каналів зв'язку; наявність мобільності пристроїв IoT. Більшість з цих труднощів були вирішені введенням додаткового шару, що був названий туманними обчисленнями [2].

Туманні обчислення є децентралізованою підмережею екосистеми IoT, що знаходиться в найбільш ефективній точці між джерелами даних і хмарою. Туманні обчислення є перспективною концепцією, яка припускає наближення обробки даних до кінцевих пристроїв мереж і є подальшим розвитком хмарної концепції [3, 4]. Пристрої туманного шару використовуються там, де дані потрібно аналізувати дуже швидко. Туманні вузли забезпечують часткову обробку та тимчасове зберігання даних перед їх надсиланням у хмару. Середовище туманних обчислень може використовувати багато пристроїв і цілей залежно від підтримуваної системи IoT. Також туманні обчислення забезпечують аналіз даних у режимі, наближеному до реального часу, що є першочерговою потребою для мінімізації затримки.

Крім того, будь-який туманний вузол може бути реалізованим за допомогою мобільного пристрою IoT. Зазвичай, мобільні пристрої IoT підпорядковані якійсь низці правил, але у більшості випадків траєкторія їхнього руху є випадковою величиною [5].

У мережах Інтернету речей з високою щільністю туманний шар зазвичай розбивається на низку кластерів [6, 7]. Зазвичай, основним критерієм кластеризації є близькість розташування туманних вузлів, що входять до кластера. Але при наявності в інфраструктурі туману мобільних вузлів структура кластера може порушитися внаслідок виходу якогось мобільного вузла за межі кластера [8]. Це призведе до непередбачених затримок при обробці даних IoT. Тому для систем IoT з високою щільністю при проведенні кластеризації необхідно враховувати такий критерій як стійкість структури кластера.

Отже, для усунення непередбачених часових затримок при обробці даних IoT з високою щільністю необхідно формувати у туманному шарі стійкі мобільні кластери.

1.2 Аналіз літературних джерел і визначення проблеми

Проблеми кластеризації у середовищах, що мають високу щільність елементів, доволі часто розглядаються у сучасних наукових дослідженнях. При розгляді цієї проблеми треба враховувати, що туманний шар мережі підтримки Інтернету речей

практично завжди має розподілену децентралізовану структуру, але завдання кластеризації його обчислювальних компонент суттєво ускладнюється у випадку наявності мобільних туманних компонент. Розглянемо наукові роботи за даною тематикою, які можна застосувати для туманного шару.

При реалізації методики кластеризації у туманному шарі, розглянутої в статті [9], враховується динаміка мобільних вузлів туманного шару. Крім того, використовується апарат сигмоїдальних нейронних мереж. Але група мобільних пристроїв не розглядається окремо від статичних вузлів, відповідно питання стійкості сформованих кластерів не розглядаються. Також у статті [10] питання кластеризації не орієнтовані на розгляд стійкості віртуальних кластерів, що містять мобільні пристрої IoT. У дослідженні, наведеному у статті [11], пропонується багаторівнева архітектура децентралізованої системи. У даній архітектурі мобільні та статичні кластери пристроїв туманного середовища IoT можуть розглядатися окремо. Це сприяє зменшенню затримок при обробці даних. Але сформовані кластери не враховують ймовірне відхилення мобільних елементів від запланованого руху. Такі відхилення враховуються у дослідженні, наведеному у статті [12]. Дослідження спрямоване на управління ресурсами у розподілених системах, але орієнтоване тільки на нескладні топології системи. Тому його не можна застосовувати для систем IoT з високою та надвисокою щільністю, як і алгоритм, запропонований у [13]. Даний алгоритм враховує особливості мобільних пристроїв IoT, оптимізує їх продуктивність та енергоспоживання, розрахований тільки на нескладні топології розглядаємої інфраструктури.

У статті [14] наведено результати досліджень щодо задіяння туманних мобільних компонент для виконання інтелектуальних застосунків. Але залишилися невирішеними питання, пов'язані із стійкістю мобільних компонент. Причиною цього є орієнтація дослідження на захищеність задіяних мобільних елементів. Один із можливих варіантів урахування питань стійкості мобільних компонент запропонований у статті [15]. В даному дослідженні розподілена мобільна компонента туманного шару використовується для вирішення оперативних завдань промислового IoT (IIoT). При цьому одним із головних показників ефективності запропонованого підходу є стійкість розподілених складових мобільної компоненти, але не враховується ймовірне відхилення задіяних мобільних пристроїв IIoT від запланованого руху. Це може призвести до руйнування структури мобільної компоненти. Такі ж проблеми можливі і при реалізації системи управління доступом до мобільних компонент туманного шару, запропонованої в [16]. Крім того, у розглянутих роботах проведені дослідження базуються на статичній кластеризації, що виконана за територіальною ознакою. При цьому сформовані кластери попередньо не аналізувались на стійкість.

Зазначимо, що запропонований у статті [17] декомпозиційний підхід можна застосовувати при

проведенні кластеризації туманного шару систем IoT високої щільності. у. Але даний підхід реалізований тільки для статичної інфраструктури туманного середовища. Для врахування наявних мобільних компонент можна використовувати алгоритм перенесення обчислювального навантаження, наведений у [18]. Даний алгоритм орієнтований на зменшення затримок під час виконання завдань у розподілених системах, але не враховує специфічні характеристики задіяної інфраструктури туманного шару систем IoT високої або надвисокої щільності.

Отже, розглянуті наукові роботи [9–18] при кластеризації туманного шару Інтернету речей недостатньо враховують характерні особливості IoT з високою щільністю. Крім того, у більшості робіт для мобільних компонент не враховується такий критерій як стійкість туманного середовища. Це може призвести до руйнування структури мережі туманних пристроїв. Тому доцільним є проведення дослідження, присвяченого підвищенню стійкості мобільних компонент туманного шару за рахунок модифікації методу кластеризації.

1.3 Мета і завдання дослідження

Метою дослідження є підвищення стійкості мобільних компонент туманного шару за рахунок модифікації методу кластеризації за рахунок розробки методу формування стійкого мобільного кластеру. Це дасть можливість підвищити стійкість групи мобільних пристроїв IoT, задіяних у туманному шарі, за рахунок наближення нестійких елементів до центрів кластерів.

Для досягнення поставленої мети дослідження треба вирішити такі часткові завдання:

- розробити підхід до формування архітектури мобільної компоненти туманного шару Інтернету речей з високою щільністю;
- розробити модель мобільного кластера туманного шару Інтернету речей з високою щільністю;
- розробити та дослідити алгоритм забезпечення стійкості мобільного кластера туманного шару Інтернету речей з високою щільністю.

2. Матеріали і методи дослідження

В роботі розглядаються мобільні пристрої IoT, які є складовими туманного шару. Дані пристрої зазвичай мають суттєво обмежені обчислювальні ресурси [19, 20]. При проведенні дослідження були враховані такі об'єктивні умови: туманний шар містить територіально рознесені мобільні компоненти, які складаються з пристроїв IoT; мобільні пристрої IoT в процесі руху прагнуть зберегти його параметри; практично всі мобільні пристрої IoT, що використовуються у туманному шарі, мають обмежені обчислювальні ресурси, а система пристроїв IoT має високу щільність у тривимірному просторі, що розглядається.

Для реалізації запропонованого методу кластеризації туманного шару Інтернету речей була застосована низка різноманітних методів. При розробці моделі туманного шару для опису

взаємного положення елементів мережі використовувалась модель точкового поля Томаса з ефектами випадкових скупчень [21]. В даній моделі передбачається, що в деякій області простору D випадково розподілені точкові об'єкти $c_1, c_2, \dots, c_n$. Кожний з цих об'єктів може розглядатися як деякий центр скупчення. Відповідно, кожному центру скупчення c_i ставиться у відповідність деяка випадкова величина y_i , яка характеризує кількість об'єктів, що входять до нього. Випадкові величини $y_1, y_2, \dots, y_n$ зазвичай вважаються взаємно незалежними і однаково розподіленими. Ефекти угруповання точкових об'єктів зазвичай спостерігаються у околі центрів скупчення. Положення кожного елемента у скупченні характеризується тривимірним нормальним розподілом в області простору D .

При розробці метода забезпечення стійкості кластера туманного шару були розглянуті центроїдні алгоритми кластеризації [22]. При формуванні кластера дані алгоритми зазвичай як критерій використовують мінімізацію величини Θ – сумарної відстані від центра кластера до його елементів:

$$\Theta \xrightarrow{\text{pos}(C)} \min; \Theta = \sum_{i=1}^k d_{i0}, \text{ pos}(C) \in D, \quad (1)$$

$$k \in \overline{1, n}, i \in \overline{1, k},$$

де C – центр кластера, d_{i0} – відстань від i -го елемента до центру кластера, k – кількість елементів в кластері, n – кількість елементів множини, що підлягає кластеризації.

Мінімізація функції (1) проводиться за рахунок вибору центрів кластерів, котрі визначаються як координати центру мас:

$$C = (x_c, y_c, z_c),$$

$$x_c = \frac{1}{k} \sum_{i=1}^k x_{i0}; \quad y_c = \frac{1}{k} \sum_{i=1}^k y_{i0}; \quad z_c = \frac{1}{k} \sum_{i=1}^k z_{i0}, \quad (2)$$

де (x_{i0}, y_{i0}, z_{i0}) – просторові координати i -го елемента кластера.

В якості базового алгоритму кластеризації розглядався центроїдний алгоритм FOREL, який використовується у багатьох системах підтримки IoT [23]. Алгоритм надає можливість операцій над центрами кластерів, які відомі у процесі роботи алгоритму. Після виконання алгоритму над множиною отриманих кластерів можна робити деякі дії. Так, можна обрати найбільш репрезентативні об'єкти із кожного кластера. Можна вибирати центри кластерів або кілька об'єктів із кожного кластера, враховуючи апіорні знання про необхідну репрезентативність вибірки. Отже, з готової кластеризації маємо можливість будувати найбільш репрезентативну вибірку. Також можливо переформування отриманої множини кластерів у багаторівневу структуру з використанням методу найкоротшого незамкненого шляху [24].

Головна властивість алгоритму – кількість кластерів заздалегідь не визначена. Кожен кластер є шаром, радіус якого визначений заздалегідь. В

результаті виконання алгоритму кожний j -й із m отриманих кластерів містить k_j елементів. Мінімізації належить такий функціонал:

$$Func = \sum_{j=1}^m \sum_{i=1}^{k_j} \rho(e_i, C_j) \xrightarrow{\forall(e_i, C_j)} \min,$$

де $\rho(e_i, C_j)$ – значення відстані елемента e_i від центра кластера C_j , до якого він належить при поточному розподілі. Порядок дій алгоритму описаний нижче.

1. Задаємо вхідне значення радіусу R .

2. Обираємо довільну точку простору та додаємо новий об'єкт F .

3. Формуємо поточну множину K , що складається із всіх об'єктів, до яких відстань від F є меншою ніж R .

4. Знаходимо центр мас всіх об'єктів із множини K ($\text{card } K=k$), котрий буде співпадати для об'єктів IoT з центром мас, тобто:

$$x_c = \sum_{e \in K} x_e / k, \quad y_c = \sum_{e \in K} y_e / k, \quad z_c = \sum_{e \in K} z_e / k.$$

Далі до отриманого нового центру мас переноситься об'єкт F . Коректуємо значення радіусу R за допомогою коефіцієнту β ($0,5 < \beta \leq 1$): $R = \beta \cdot R$ та переходимо до кроку 3. Цикл із кроків 3 та 4 повторюється, поки множина K не стабілізується.

5. Коли множина K стабілізується, вона оголошується новим кластером. Об'єкти, що потрапили до K , видаляються для розглядання із вибірки, що кластеризується.

6. У випадку, якщо розглядаєма множина об'єктів IoT не порожня, повертаємося до кроку 1, відновлюючи значення радіусу R . У іншому випадку даний алгоритм закінчує роботу.

Для оцінки ефективності кластеризації використовувався показник стійкості системи m мобільних кластерів [25]:

$$\eta = \max \left(P_j^- \mid j \in \overline{1, m} \right), \quad (3)$$

де P_j^- – ймовірність того, що за фіксований проміжок часу структура j -го кластера буде зруйнована, тобто, чим менше даний показник, тим система стійкіша.

3. Результати розробки і дослідження методу адаптивної кластеризації мобільних вузлів туманного шару

3.1 Формування архітектури мобільної компоненти туманного шару

Стрімке поширення технологій Інтернету речей зумовило подальший розвиток його архітектури та появу туманного шару [26]. За своїм функціональним призначенням він розташований між хмарною інфраструктурою, представленою центрами обробки даних та кінцевими пристроями і сенсорами Інтернету речей. Основні архітектурні принципи хмарних обчислень, застосовувані для забезпечення функціонування Інтернету речей, можуть бути поширені й на туманну модель обчислень [27, 28]. Водночас практична реалізація такої архітектури є

складнішою, вважаючи як на розподілений характер туманного шару, так й на обмежені ресурси значної частини його вузлів. Взаємодія з хмарними центрами обробки даних здійснюється з використанням розподіленої децентралізованої інфраструктури, яка виконує функції ядра туманного шару [29, 30]. У межах такої інфраструктури формується єдине середовище доступу та агрегування, до якого підключаються вузли Інтернету речей, що володіють певними обчислювальними можливостями. Використання цих ресурсів дає змогу перенести частину обчислювальних операцій із хмарного рівня ближче до джерел формування даних.

Функціональну структуру апаратної платформи туманного шару можна представити сукупністю таких основних модулів:

- мережних модулів апаратної платформи, призначених для підтримання надійного з'єднання з використанням провідних та безпроводних технологій;

- апаратних прискорювачів, які інтегруються у платформу та призначені для підвищення продуктивності виконання окремих операцій;

- обчислювальних компонентів, що формують необхідний для функціонування туманного вузла обсяг обчислювальних ресурсів;

- модулів пам'яті, які забезпечують зберігання даних і є невід'ємною складовою обчислювальної платформи;

- засобів інформаційної безпеки забезпечення захисту даних та ресурсів туманного вузла;

- модулів керування, за допомогою яких здійснюється координація роботи складових апаратної платформи;

- засобів віртуалізації та контейнеризації, призначених для логічного розподілу апаратних ресурсів і створення середовищ виконання віртуальних машин або контейнерів для розгортання туманних сервісів.

Під час формування архітектури туманного шару необхідно враховувати характеристики конкретних класів систем Інтернету речей. Однією з актуальних тенденцій є залучення пристроїв з обмеженими обчислювальними та комунікаційними можливостями до виконання функцій туманних обчислювальних вузлів. Додатково ускладнює організацію таких систем зростання кількості мереж Інтернету речей з високою та надвисокою щільністю розміщення пристроїв. Такі мережі, як правило, мають розподілену просторову структуру, у зв'язку з чим виникає необхідність раціонального групування їх компонентів. Застосування кластеризації туманних пристроїв дає змогу підвищити ефективність використання доступних мережних та обчислювальних ресурсів системи Інтернету речей.

Окремим фактором, який необхідно враховувати при проектуванні перспективних мереж Інтернету речей, є просторове переміщення окремих пристроїв або цілих сегментів. Мобільність таких компонентів передбачає зміну їхнього положення у просторі протягом функціонування системи. Одним із напрямів розвитку туманних обчислень є

розміщення туманних вузлів на мобільних платформах, здатних змінювати своє просторове положення, відповідні рішення можуть передбачати використання безпілотних літальних апаратів як кінцевих пристроїв Інтернету речей або як мобільних вузлів, що входять до складу туманних кластерів.

Таким чином, дослідження туманних обчислень з урахуванням мобільності Fog-пристроїв є окремим актуальним напрямом. Функції Fog-пристроїв можуть виконувати різні елементи мережної інфраструктури: обладнання сегмента «останньої милі», користувацькі пристрої та безпосередньо пристрої Інтернету речей. Використання мобільних компонентів розширює можливості організації мережного слайсингу. При цьому об'єднання мобільних Fog-пристроїв у кластери може здійснюватися відповідно до певних характеристик, що визначають їх функціональну сумісність. Такими характеристиками можуть бути параметри переміщення пристроїв, зокрема швидкість і напрямок руху, а також часові характеристики виконання обчислювальних завдань на вузлах відповідного кластера. З урахуванням зазначених особливостей для побудови архітектури туманного шару Інтернету речей із надвисокою щільністю пристроїв та наявністю мобільних компонентів доцільно виділити чотири функціональні рівні:

- хмарний рівень, призначений для централізованого зберігання та оброблення даних IoT із використанням центрів обробки даних;

- рівень взаємодії з хмарною інфраструктурою, функції якого забезпечуються серверами, що формують ядро туманного шару та організують обмін даними між туманним і хмарним шарами;

- рівень туманних пристроїв, на якому вузли об'єднуються у функціональні кластери та взаємодіють із серверами ядра туманного шару;

- рівень кінцевих пристроїв Інтернету речей, що охоплює сенсори, камери, мобільні та інші пристрої, які безпосередньо отримують інформацію від фізичних об'єктів і процесів.

Також слід враховувати, що ефективність функціонування системи IoT з високою щільністю суттєво залежить від якості проведеної кластеризації. Розбиття туманних пристроїв на кластери впливає на показники QoS, особливо при виконанні транзакцій, для яких основними є часові критерії. В свою чергу, час обробки транзакції IoT для статичних кластерів в основному залежить від характеристик внутрішньо-кластерних каналів зв'язку. Але для мобільних кластерів на перше місце виходять стійкісні показники кластерної структури, для дослідження яких треба розробити відповідну модель.

3. 2 Математична модель функціонування мобільного кластеру туманного шару

Розміщення пристроїв туманного шару мережі підтримки Інтернету речей з високою щільністю характеризується обмеженим об'ємом. При цьому туманні пристрої розміщуються на відносно невеликій відстані один від одного. Величину такої відстані, наприклад, можна визначити дальністю

зв'язку, яка забезпечується стандартом мережі доступу. Крім того, на структуру мережі суттєво впливає взаємне розташування туманних пристроїв. Також треба враховувати мобільність туманних пристроїв, внаслідок чого архітектура та взаємне розташування пристроїв можуть змінюватися. Крім того, вихід мобільного пристрою туманному шару із зони зв'язку, призводить до часткового руйнування структури туману, тобто до незапланованих затримок у обробці інформації, що надходить з датчиків IoT, та зниження показників QoS.

В більшості випадків для динамічних туманних обчислень, що базуються на мобільних пристроях Інтернету речей, оперують групами пристроїв – роями. У роях пристрої прагнуть зберегти загальні параметри руху. Наприклад, параметри руху автомобілів визначаються прийнятими правилами та особливостями дороги. При формуванні динамічних туманних кластерів в системах Інтернету речей великої щільності обираємо групи рухливих щільно розміщених пристроїв. Такі пристрої намагаються зберегти своє положення щодо інших членів групи.

Для описи взаємного положення елементів мережі динамічних туманних обчислень підтримки Інтернету речей пропонується використовувати модель точкового процесу Томаса. Для цього розглянемо розташування групи мобільних пристроїв IoT у тривимірному просторі на початку руху. Нехай окремі елементи групи мобільних пристроїв IoT $e_1, \dots, e_k, \dots$ входять до кластера K_j та розташовані в поточних позиціях. Ці позиції задані правилами організації для пристроїв IoT та на даний момент фіксовані щодо центру даного кластера, позначеного на схемі як C_j . Усі елементи кластера разом можуть переміщатися з рівними швидкостями в одному і тому ж напрямку. Отже, відстань між пристроєм IoT і центром кластера K_j в тривимірному просторі можна визначати за таким виразом:

$$d_{ji0} = \sqrt{(x_{ji} - x_{jc})^2 + (y_{ji} - y_{jc})^2 + (z_{ji} - z_{jc})^2}, \quad (4)$$

де $x_{ji}, x_{jc}, y_{ji}, y_{jc}, z_{ji}, z_{jc}$ – поточні координати пристроїв IoT та координатора кластера K_j – точки C_j .

При цьому, якщо центр кластера C_j прийняти за початок координат, то елементи кластера будуть відносно нерухомими щодо його центру.

Але в процесі реального руху швидкості і напрямки руху окремих мобільних елементів можуть дещо відрізнитися. Це призводить до їх відхилення від заданих позицій в кластері. Отже, координати x_{ji}, y_{ji}, z_{ji} , є випадковими величинами. Зауважимо, що мобільні пристрої IoT у більшості випадків зберігають власне становище щодо інших членів групи. Зробимо припущення, що випадкові координати динамічних пристроїв IoT мають нормальний розподіл. Нехай у системі координат кластера K_j початкове положення i -го пристрою – точка $A_{ji0}(x_{ji0}, y_{ji0}, z_{ji0})$, а поточне положення – точка $A_{ji}(x_{ji}, y_{ji}, z_{ji})$. Тоді відхилення пристрою від початкового положення розраховується як:

$$\xi_{ji} = \sqrt{(x_{ji} - x_{ji0})^2 + (y_{ji} - y_{ji0})^2 + (z_{ji} - z_{ji0})^2}. \quad (5)$$

Зауважимо, що величина ξ_{ji} не залежить від вибору систем координат, тобто $\xi_{ji} = \xi_i \forall i$. Відхилення ξ_i є випадковою величиною, розподіленою за нормальним законом із такою функцією щільності ймовірності:

$$f(\xi_i) = \frac{1}{\sigma_i \sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2}\left(\frac{\xi_i}{\sigma_i}\right)^2\right), \quad (6)$$

де σ_i – середньоквадратичне відхилення випадкової величини ξ_i від початкового положення пристрою i .

Розглянемо пристрій IoT e_i , який належить кластеру K_j з радіусом зони досяжності R_j , а d_{ji0} – це відстань між початковим положенням e_i та центром кластера K_j . В процесі руху пристрій e_i практично може відхилитися від початкового положення не більше, ніж на $3\sigma_i$. Отже всі поточні положення пристрою повинні знаходитися у сфері радіусом R_j з центром у початковому положенні. Вихід будь-якого мобільного пристрою за межі даної сфери призведе до руйнування структури кластера K_j .

Розрахуємо ймовірність того, що пристрій IoT e_i буде знаходитися у зоні досяжності кластера K_j :

$$P_{ji}^+ \rightarrow P_{ji}^{(theor+)} = \begin{cases} 1, & \text{якщо } 3\sigma_i + d_{ji0} \leq R_j; \\ 0.5 + \Phi(3\sigma_i + d_{ji0} - R_j), & \text{інакше,} \end{cases} \quad (7)$$

де $\Phi(\xi_i) = \frac{1}{\sqrt{2\pi}} \int_0^{\xi_i} \exp\left(-\frac{t^2}{2}\right) dt$ – значення функції

Лапласа для розподілу (8).

Тоді ймовірність того, що структура кластера K_j буде зруйнована, тобто хоча б один із пристроїв вийде за межі зони досяжності координатора є такою:

$$P_j^- = 1 - \sum_{i=1}^{k_j} P_{ji}^+, \quad (8)$$

де k_j – кількість елементів кластера K_j .

В якості критерію формування мобільного кластера приймемо максимальну стійкість кластера туманних обчислень. Під стійкістю кластера будемо розуміти ймовірність збереження його структури, тобто ймовірність наявності зв'язку між центральним пристроєм та всіма елементами кластера. Тоді цільовою функцією при формуванні кластерної структури буде такий вираз:

$$\max_j \left(\min_{pos(C_j)} \left(1 - \sum_{i=1}^{k_j} P_{ji}^+ \right) \right),$$

$$pos(C_j) \in D, \quad k_j \in \overline{1, n}, \quad (9)$$

де D – зона обслуговування IoT; $pos(C_j)$ – позиція центру кластера K_j в тривимірному просторі; n – розмір групи мобільних пристроїв системи IoT, що кластеризується.

3. 3 Розробка та дослідження ефективності алгоритму забезпечення стійкості мобільного кластера туманного шару

Стійкість кластера, що складається із мобільних пристроїв IoT, забезпечується за рахунок збереження його структури. При розгортанні туманних обчислень на базі наземних мобільних систем головним координуючим пристроєм може бути мережний пристрій повітряного сегменту IoT. Він повинен мати таку позицію в кластері, яка максимально забезпечувала б стійкість кластера при русі групи мобільних пристроїв, треба мінімізувати ймовірність виходу будь-якого пристрою із кластера при випадкових переміщеннях пристроїв IoT.

При кластеризації пристроїв туманного шару Інтернету речей з високою і надвисокою щільністю зазвичай використовуються центроїдні алгоритми кластеризації. Дані алгоритми мінімізують сумарну відстань від туманних пристроїв до центрів кластерів відповідно з виразом (1). Для кластеризації групи мобільних пристроїв D був обраний центроїдний алгоритм FOREL, який надає можливість варіювати параметрами кластерів та їх центрів. Алгоритм FOREL базується на принципі формування кластерів у місцях згущення об'єктів IoT. Згідно з даним алгоритмом та виразом (1) для групи D при кластеризації мінімізується функціонал Θ :

$$\Theta \rightarrow \min_{pos(C_j)}; \quad \Theta = \sum_{j=1}^m \sum_{i=1}^{k_j} d_{ji0},$$

$$pos(C_j) \in D, \quad k_j \in \overline{1, n}, \quad j \in \overline{1, m}, \quad (10)$$

де m – кількість кластерів у групі пристроїв IoT, отриманих після виконання алгоритму FOREL.

Зауважимо, що стійкість мобільного кластера збільшується при наближенні центру мобільного туманного пристрою до центру кластера. Тому бажано центр кластера змістити ближче до скупчення пристроїв. Це можна зробити за рахунок введення вагових коефіцієнтів, які впливають на позицію центру кластера. Якщо ваговий коефіцієнт якогось елемента набагато більший ніж вагові коефіцієнти інших елементів, то центр кластера буде зміщуватися до цього елемента. Введемо вагові коефіцієнти у формулу (2), за якою розраховуються координати центра мас кластера:

$$C_j = (x_{jc}, y_{jc}, z_{jc}); \quad x_{jc} = \frac{1}{k_j} \sum_{i=1}^{k_j} \gamma_i x_{ji0};$$

$$y_{jc} = \frac{1}{k_j} \sum_{i=1}^{k_j} \gamma_i y_{ji0}; \quad z_{jc} = \frac{1}{k_j} \sum_{i=1}^{k_j} \gamma_i z_{ji0}, \quad (11)$$

де γ_i – ваговий коефіцієнт мобільного пристрою e_i .

Врахуємо критерій стійкості кластера за допомогою вагового коефіцієнта мобільного пристрою IoT e_i :

$$w_i = \sigma_i^s, \quad 0 \leq s \leq 1, \quad (12)$$

де s – коефіцієнт, що визначає чутливість методу до випадкових відхилень елементів кластеру від початкової позиції.

Вибір значення коефіцієнта визначається значимістю критерію стійкості кластера для завдання. При $s=0$ стійкість кластера як критерій не враховується. При $s=1$ критерій стійкості кластера максимально впливає на формування кластерів.

Також модифікується третій крок алгоритму FOREL. На цьому кроці до поточного кластеру відбираються тільки ті мобільні об'єкти IoT, для яких виконана умова $P_{ji}^{(theor+)} = 1$. Це робить ймовірність

виходу будь-якого мобільного об'єкту із кластера мінімальною. Звісно, кількість кластерів може збільшитися, але стійкість отриманих кластерів суттєво підвищиться.

Зауважимо, що для завдання кластеризації із застосуванням модифікованого алгоритму FOREL потрібні додаткові вхідні дані:

– початкові координати елементів групи мобільних пристроїв IoT, що є математичними сподіваннями координат елементів;

– середньоквадратичні відхилення положень мобільних пристроїв IoT.

Такі дані можна отримати в результаті збору відповідної статистики у перебігу достатнього часового інтервалу.

Для оцінки ефективності модифікованого методу була використана імітаційна модель. Елементи групи мобільних пристроїв у моделі утворювали гетерогенні скупчення, тобто різні за складом та характером розміщення. Координати мобільних пристроїв розглядалися як значення випадкових величин, розподілених за нормальним законом. Параметри відповідних нормальних розподілів для кожного елемента e_i задавалися нульовим математичним сподіванням $\mu_i=0$ і середньоквадратичним відхиленням (СКВ) σ_i . Розкид відхилень елементів від початкових позицій був різним для різних елементів. На СКВ накладено обмеження $R/60 < \sigma_i < R/12$, яке продиктовано особливостями мобільних пристроїв IoT.

Оцінювалась ефективність модифікованого методу у порівнянні з методом формування кластера, що базується тільки на критерії відстані між елементами. Оцінка проводилася за показником стійкості системи, розрахованим за виразом (6) та переведеним у відсотки. При моделюванні коефіцієнт s , що визначає чутливість методу до випадкових відхилень, змінювався дискретно з кроком 0,2. Були змодельовані різні варіанти розміру групи мобільних пристроїв: $n \in \{50, 150, 300\}$. Також у моделі використовувався показник середньостатистичної границі можливих відхилень мобільних пристроїв від запланованих траєкторій:

$$\sigma = 3 \sum_{i=1}^n \sigma_i / n. \quad (15)$$

Узагальнені результати моделювання наведені у табл. 1.

Таблиця 1 – Результати імітаційного моделювання (відсоток незруйнованих кластерів)

σ/R	n	m, s	m_0	$s=0$	$s=0.2$	$s=0.4$	$s=0.6$	$s=0.8$	$s=1.0$	m_1
0.05	50		4	0.4 %	0.4 %	0.3 %	0.3 %	0.3 %	0.2 %	4
	150		9	2.2 %	1.7 %	1.2 %	1.0 %	0.7 %	0.5 %	10
	300		14	3.4 %	2.7 %	2.1 %	1.7 %	1.3 %	1.0 %	19
0.10	50		4	3.6 %	2.8 %	2.1 %	1.6 %	1.1 %	0.5 %	6
	150		9	10.5 %	3.9 %	5.3 %	3.1 %	1.9 %	0.8 %	14
	300		15	21.7 %	2.9 %	7.0 %	5.3 %	2.8 %	1.3 %	37
0.25	50		4	27.6 %	14.1 %	7.9 %	6.8 %	2.7 %	0.9 %	11
	150		10	32.0 %	16.8 %	8.8 %	5.2 %	3.3 %	1.1 %	34
	300		16	44.5 %	29.9 %	11.8 %	9.1 %	7.0 %	1.9 %	87

У табл. 1 стовпець « $s=0$ » відповідає кластеризації, при якій відхилення мобільних елементів туманного шару не враховуються. Кількість отриманих кластерів при такій кластеризації наведена у стовпці « m_0 ». Стовпець « $s=1$ » відповідає кластеризації, яка максимальна орієнтована на стійкість мобільних кластерів. При цьому кількість отриманих кластерів наведена у стовпці « m_1 ». Результати моделювання показали, що ефективність запропонованого методу підвищується в таких випадках: зростання значень СКВ мобільних пристроїв від запланованих траєкторій руху; збільшення кількості мобільних туманних пристроїв. Це пояснюється тим, що алгоритм кластеризації обирає положення центру кластера, за рахунок наближення центру до найбільш нестійких елементів.

Але при відносно великих значеннях СКВ кількість кластерів у порівнянні із стандартним методом починає збільшуватися. Це ускладнює керування компонентами туманного шару.

4. Обговорення результатів дослідження методу адаптивної кластеризації

Розроблена архітектура мобільної компоненти туманного шару Інтернету речей високої або надвисокої щільності. Наявність окремих мобільних компонент є суттєвою відмінністю від існуючих архітектурних рішень побудови туманної інфраструктури. Запропонована модель окремого мобільного кластера туманного шару. Головна відмінність даної моделі – орієнтованість не тільки на просторові показники, але й на стійкісні характеристики кластера. Для реалізації моделі використана модель точкового процесу Томаса Дана модель дозволила розробити метод забезпечення стійкості мобільного кластера туманного шару. Доведена ефективність проведення кластеризації з використанням модифікованого центроїдного алгоритму FOREL. Даний алгоритм модифікований за рахунок введення вагових коефіцієнтів при знаходженні положення центру мобільного кластера, що надало можливість наблизити нестабільні компоненти до центру кластера.

Досліджена ефективність запропонованого метода забезпечення стійкості мобільного кластера туманного шару Інтернету речей високої або надвисокої щільності. Порівняльне тестування стандартного та модифікованого методів (табл. 1) показало такі результати: при зростанні значень СКВ

мобільних пристроїв від запланованих траєкторій руху стійкість мобільного кластера підвищується; при значенні відношення σ/R від 0.05 до 0.2 використання модифікованого методу дозволяє підвищити стійкість мобільної структури до 20 %; при $\sigma/R > 0.2$ використання модифікованого методу недоцільно внаслідок ускладнення керування компонентами туманного шару; при $\sigma/R < 0.05$ стійкість мобільної структури підвищується не більше, ніж на 5 %. Отримані результати дослідження метода забезпечення стійкості мобільного кластера можна пояснити наближенням центрів мобільних кластерів до нестійких компонент.

Як розвиток даного дослідження можна відзначити наступне. При підвищенні нестабільності групи мобільних пристроїв IoT деякі пристрої стають ізольованими, тобто не входять до сформованих кластерів, тобто треба проведення додаткового дослідження щодо ізольованих компонент у групі мобільних кластерів.

5. Висновки

Розроблений підхід до формування архітектури мобільної компоненти туманного шару Інтернету речей високої або надвисокої щільності. При розробці враховувались специфічні особливості мобільних пристроїв IoT, які використовуються в якості динамічних вузлів туманного шару. Це дозволило запропонувати трирівневу архітектуру туманної інфраструктури, в якій мобільні кластери є окремими компонентами.

Розроблена модель окремого мобільного кластера туманного шару, яка враховує випадковість руху мобільних пристроїв IoT та базується на точковому процесі Томаса. Головною відмінністю даної моделі від існуючих є орієнтованість не тільки на просторові показники, але й на стійкісні характеристики компонент кластера. Дана модель дозволила розробити метод забезпечення стійкості мобільного кластера туманного шару на базі запропонованої архітектури мобільної компоненти туманного шару Інтернету речей.

3. Розроблений та досліджений метод забезпечення стійкості мобільного кластера туманного шару підтримки Інтернету речей високої або надвисокої щільності. Підвищення стійкості досягається за рахунок ітеративного наближення центру кластера до найбільш нестійких його компонент. Порівняна стійкість мобільних кластерів, побудованих стандартним та модифікованим

методами. Якщо середнє відносне відхилення елементів кластера знаходиться в межах від 0.05 до 0.2, то стійкість підвищується до 20 %. Доведено, що при значенні середнього відносного відхилення більше 0.2, використання модифікованого методу недоцільно внаслідок ускладнення керування компонентами туманного шару.

Доступність даних

Рукопис не має пов'язаних даних.

Конфлікт інтересів

Автори декларують, що не мають конфлікту інтересів стосовно даного дослідження, в тому числі фінансового, особистісного характеру, авторства чи

іншого характеру, що міг би вплинути на дослідження та його результати, представлені в даній статті.

Використання засобів штучного інтелекту

Автори підтверджують, що не використовували технології штучного інтелекту при створенні представленої роботи.

Фінансування

Дослідження профінансовано Міністерством освіти та науки України в рамках дослідницького проекту 0125U001544 на тему “Методологія забезпечення процесів моніторингу та контролю виконання портфелів проектів та програм для проектних офісів в умовах відбудови України”.

СПИСОК ЛІТЕРАТУРИ

1. Hu, N. (2024). Internet of things edge data mining technology based on cloud computing model. *International Journal of Innovative Computing, Information and Control*, 20(6), 1749–1763. <http://dx.doi.org/10.24507/ijicic.20.06.1749>
2. Hunko, M., Tkachov, V., Kovalenko, A., Kuchuk, H. (2023). Advantages of Fog Computing: A Comparative Analysis with Cloud Computing for Enhanced Edge Computing Capabilities. *KhPI Week 2023 - Conference Proceedings*, 1–5. <https://doi.org/10.1109/KhPIWeek61412.2023.10312948>
3. Kuchuk, H., Malokhvii, E. (2024). Integration of iot with Cloud, Fog and Edge computing: a review. *Advanced Information Systems*. 8(2), 65–78. <https://doi.org/10.20998/2522-9052.2024.2.08>
4. Fatlawi, A., Al Dujaili, M.J. (2023). Integrating the Internet of Things (IoT) and Cloud Computing Challenges and Solutions: A Review. *AIP Conference Proceedings*, 2977(1), 020067. <http://dx.doi.org/10.1063/5.0181842>
5. Singh, C., Khilari, S., Taware, R. (2024). Active Machine-to-Machine (M2M) and IoT Communication Architecture for Mobile Devices and Sensor Nodes. *Lecture Notes in Networks and Systems*, pp. 25–38. https://doi.org/10.1007/978-981-97-5786-2_3
6. Kuchuk, N., Kovalenko, A., Ruban, I., Shyshatskyi, A., Zakovorotnyi, O., Sheviakov, I. (2023). Traffic Modeling for the Industrial Internet of NanoThings. 2023 IEEE 4th KhPI Week on Advanced Technology, KhPI Week 2023 - Conference Proc., 194480. <http://dx.doi.org/10.1109/KhPIWeek61412.2023.10312856>
7. Alsadie, D. (2024). Advancements in heuristic task scheduling for IoT applications in fog-cloud computing: challenges and prospects. *PeerJ Computer Science*, 10, e2128. <https://doi.org/10.7717/PEERJ-CS.2128>
8. Sobchuk, V., Pykhivskyi, R., Barabash, O., Korotin, S., Omarov, S. (2024). Sequential intrusion detection system for zero-trust cyber defense of IoT/IIoT networks. *Advanced Information Systems*, 8(3), 92–99. <https://doi.org/10.20998/2522-9052.2024.3.11>
9. Saurabh, Dhanaraj, R.K. (2024). Enhance QoS with fog computing based on sigmoid NN clustering and entropy-based scheduling. *Multimed Tools Appl*, 83, 305–326. <https://doi.org/10.1007/s11042-023-15685-3>
10. Kuchuk, H., Mozhaiev, O., Kuchuk, N., Tiulieniev, S., Mozhaiev, M., Gnusov, Y., Tsuranov, M., Bykova, T., Klivets, S., Kuleshov, A. (2024). Devising a method for the virtual clustering of the Internet of Things edge environment. *Eastern-European Journal of Enterprise Technologies*, 1(9)(127), 60–71. <https://doi.org/10.15587/1729-4061.2024.298431>
11. Sharma, Sh. Saini H. (2019). A novel four-tier architecture for delay aware scheduling and load balancing in fog environment. *Sustainable Computing: Informatics and Systems*, 24. <https://doi.org/10.1016/j.suscom.2019.100355>
12. Li, G., Liu, Y., Wu, J., Lin, D., Zhao, Sh. (2019). Methods of Resource Scheduling Based on Optimized Fuzzy Clustering in Fog Computing. *Sensors*, 19(9). <https://doi.org/10.3390/s19092122>
13. Jamil, B. Shojafar, M., Ahmed, I., Ullah, A., Munir, K., Ijaz, H. (2020). A job scheduling algorithm for delay and performance optimization in fog computing. *Concurrency and Computation: Practice and Experience*, 32(7). <https://doi.org/10.1002/cpe.5581>
14. Ding, H., Ding, X., Xia, F., Zhou, F. (2023). An Efficient Method for Implementing Applications of Smart Devices Based on Mobile Fog Processing in a Secure Environment. *International Journal of Advanced Computer Science and Applications*, 14(10), 93–105. <https://doi.org/10.14569/IJACSA.2023.0141011>
15. Qayyum, T., Trabelsi, Z., Waqar Malik, A., Hayawi, K. (2022). Mobility-aware hierarchical fog computing framework for Industrial Internet of Things. *Journal of Cloud Computing*, 11(1), 72. <https://doi.org/10.1186/s13677-022-00345-y>
16. Routray, K., Bera, P. (2024). Fog-Assisted Dynamic IoT Device Access Management Using Attribute-Based Encryption. *ACM International Conference Proceeding Series*, 346–352. <https://doi.org/10.1145/3631461.3631466>
17. Kuchuk, H., Kalinin, Y., Dotsenko, N., Chumachenko, I., Pakhomov, Y. (2024). Decomposition of integrated high-density IoT data flow. *Advanced Information Systems*, 8(3), 77–84. <https://doi.org/10.20998/2522-9052.2024.3.09>
18. Proietti Mattia, G., Beraldi, R. (2023). P2PFaaS: A framework for FaaS peer-to-peer scheduling and load balancing in Fog and Edge computing. *SoftwareX*, 21, 101290. <https://www.sciencedirect.com/science/article/pii/S2352711022002084>
19. Lu, S., Wu, J., Wang, N., Duan, Y., Liu, H., Zhang, J., Fang, J. (2023). Resource provisioning in collaborative fog computing for multiple delay-sensitive users. *Software – Practice and Experience*, 53(2), 243–262. <https://doi.org/10.1002/spe.3000>
20. Pardo, C., Wei, R., Ivens, B.S. (2022). Integrating the business networks and internet of things perspectives: A system of systems (SoS) approach for industrial markets. *Industrial Marketing Management*, 104, 258–275. <https://doi.org/10.1016/j.indmarman.2022.04.012>
21. Drabech, Z., Douimi, M., Zemmouri, E. (2024). A Markov random field model for change points detection. *Journal of Computational Science*, 83, 102429. <https://doi.org/10.1016/j.jocs.2024.102429>

22. Zhu, Q., Hu, L., Wang, R. (2022). Image Clustering Algorithm Based on Predefined Evenly-Distributed Class Centroids and Composite Cosine Distance. *Entropy*, 24(11), 1533. <https://doi.org/10.3390/e24111533>
23. Mutambik, I. (2024). An Entropy-Based Clustering Algorithm for Real-Time High-Dimensional IoT Data Streams. *Sensors*, 24(22), 7412. <https://doi.org/10.3390/s24227412>
24. Petrovska, I., Kuchuk, N., Pochebut, M., Kuchuk, H., Mozhaiev, O., Onishchenko Y. (2023). Sequential Series-Based Prediction Model in Adaptive Cloud Resource Allocation for Data Processing and Security. The 13th IEEE Int. Conf. on Dependable Systems, Services and Technologies, DESSERT'2023, Athens, Greece, pp. 1–6. <https://doi.org/10.1109/DESSERT61349.2023.10416496>
25. Filosi, M., Visintainer, R., Riccadonna, S., Jurman, G., Furlanello, C. (2014). Stability indicators in network reconstruction. *PLoS ONE*, 9(2), e89815. <https://doi.org/10.1371/journal.pone.0089815>
26. Petrovska I., Kuchuk, H. Mozhaiev M. (2022). Features of the distribution of computing resources in cloud systems. 2022 IEEE 3rd KhPI Week on Advanced Technology, KhPI Week 2022 - Conference Proceedings, 03-07 October 2022. <https://doi.org/10.1109/KhPIWeek57572.2022.9916459>
27. Kuchuk, N., Kashkevich, S., Radchenko, V., Andrusenko, Y., Kuchuk, H. (2024). Applying edge computing in the execution IoT operative transactions. *Advanced Information Systems*, 8(4), 49–59. <https://doi.org/10.20998/2522-9052.2024.4.07>
28. Thomas, P., Jose, D.V. (2023). Towards Computation Offloading Approaches in IoT-Fog-Cloud Environment: Survey on Concepts, Architectures, Tools and Methodologies. *Lecture Notes in Networks and Systems*, 613 LNNS, 37–52. https://doi.org/10.1007/978-981-19-9379-4_4
29. Mani Kiran, C.V.N.S., Jagadeesh Babu, B., Singh, M.K. (2023). Study of Different Types of Smart Sensors for IoT Application Sensors. *Smart Innovation, Systems and Technologies*, 290, 101–107. https://doi.org/10.1007/978-981-19-0108-9_11
30. Emami Khansari, M., Sharifian, S. (2024). A scalable modified deep reinforcement learning algorithm for serverless IoT microservice composition infrastructure in fog layer. *Future Generation Computer Systems*, 153, 206–221. <https://doi.org/10.1016/j.future.2023.11.022>

Received (Надійшла) 14.06.2026

Accepted for publication (Прийнята до друку) 02.09.2026

Publication date (Дата публікації) 25.09.2026

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Коваленко Андрій Анатолійович – доктор технічних наук, професор, завідувач кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Andriy Kovalenko – Doctor of Technical Sciences, Professor, Head of the Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: andriy.kovalenko@nure.ua; ORCID Author ID: <https://orcid.org/0000-0002-2817-9036>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=56423229200>.

Кучук Георгій Анатолійович – доктор технічних наук, професор, професор кафедри комп'ютерної інженерії та програмування, Національний технічний університет "Харківський політехнічний інститут", Харків, Україна;

Heorhii Kuchuk – Doctor of Technical Sciences, Professor, Professor of Computer Engineering and Programming Department, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;

e-mail: kuchuk56@ukr.net; ORCID Author ID: <http://orcid.org/0000-0002-2862-438X>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57057781300>.

A Method for Adaptive Clustering of Mobile Nodes in the Fog Layer of the Internet of Things

Andriy Kovalenko, Heorhii Kuchuk

Abstract. Relevance. The primary criterion for clustering fog devices is the proximity of the fog nodes belonging to a cluster. However, when mobile nodes are present, the cluster structure may be disrupted if a mobile node moves beyond the cluster boundaries. This may result in unexpected delays in IoT data processing. Therefore, for such systems, an additional criterion related to the stability of the cluster structure should be considered. **The subject of the study** is clustering methods for high-density fog-layer Internet of Things networks. **The aim of the study** is to improve the stability of mobile components of the fog layer by modifying the clustering approach through the development of a method for forming a stable mobile cluster. **The objectives of the study** are to develop an approach to the architecture of the mobile component, as well as a model and a method for ensuring the stability of a mobile cluster in a high-density fog-layer Internet of Things network. **Research results.** To organize the interaction of mobile devices, an architecture of the mobile component of the fog layer has been developed, taking into account its decentralized structure and the specific operational characteristics of mobile Internet of Things devices. The proposed architecture consists of four levels and provides for the formation of separate mobile clusters of fog devices at the lower level. To investigate the formation of mobile clusters, a corresponding model has been developed that accounts for the random nature of the movement of Internet of Things devices. The spatial distribution of mobile components is modeled using the Thomas point process. Based on the proposed model, the FOREL clustering algorithm has been modified. In the modified algorithm, the position of the cluster center is determined using weighting coefficients that characterize the stability of its components. **Conclusions.** Simulation results demonstrated increased stability of the formed mobile clusters in environments with high and ultra-high device density. It was found that an increase in the mean relative deviation of the actual device trajectories from their planned trajectories leads to an increase in the stability of the cluster structure. This behavior is explained by the displacement of the cluster center toward its least stable components, which enables the cluster structure to adapt to changes in the spatial distribution of mobile devices. The proposed method can be used for clustering mobile components of the fog layer of the Internet of Things when the mean deviation of device movement from the specified trajectories does not exceed 20% of the cluster radius.

Keywords: mobile device; Internet of Things; clustering; fog layer; Thomas point process.

Yaroslav Lashyn, Oleksandr Voitko, Vitalii Fedoriienko

National Defence University of Ukraine, Kyiv, Ukraine

HYBRID NEURO-MATHEMATICAL MODEL FOR DETECTING INFORMATION THREATS: INTEGRATING LARGE LANGUAGE MODELS AND MODIFIED DIFFUSION PROCESSES

Abstract. Background. In the context of the rapid evolution of cyber threats, where destructive narratives spread with unprecedented speed, traditional isolated monitoring systems lose their effectiveness. Subject of research: analytical systems for detecting coordinated information operations in modern cognitive security. **Purpose of research:** to develop a unified neuro-mathematical architecture that maps non-linear semantic features onto differential equations of spatio-temporal diffusion. **Research objectives.** To construct a hybrid model based on a large language model, transform semantic outputs into continuous kinematic parameters, and empirically validate the model's efficacy during real information attacks. **Research methods.** A mixed-methods empirical approach was employed, integrating quantitative mathematical modeling with computational linguistics, the Llama-3.1-8B-Instruct model utilizing the Adaptive Self-Consistency algorithm, and modifying the Bass diffusion model alongside the Neural Hawkes process. **Research results.** The implementation of thermodynamic logit scaling reduced semantic classification entropy to 0.160 bits. Integrating continuous kinematic parameters into differential equations allowed the system to identify the phase transition of an attack 3 to 4 hours prior to its statistical peak. It was revealed that granular emotional subtype classification without fine-tuning leads to mathematical collapse, entropy surging to 1.349 bits. **Conclusions.** The research expands the Bass diffusion theory and Hawkes processes by replacing static constants with dynamic, AI-derived forcing functions. It mathematically proves the necessity of macro-polarization in feature spaces for stable differential modeling without network fine-tuning.

Keywords: information security; information operations; large language models; natural language processing; social network analysis; hybrid neuro-mathematical model; diffusion processes; stochastic models.

1. Introduction

1.1 Research Background

The contemporary digital landscape is characterized by an unprecedented degree of complexity, wherein sociotechnical information dissemination networks are continuously subjected to coordinated manipulation and cognitive warfare [1]. The transition from scripted, deterministic automated agents to autonomous, cognitively capable social bots has fundamentally altered the parameters of online social interactions. Modern artificial intelligence algorithms, specifically generative multimodal architectures and Large Language Models (LLMs), are increasingly deployed to execute highly sophisticated information operations that exhibit emergent properties such as long-term strategic planning, emotional mimicry, and hyper-personalized persuasion [2]. Furthermore, the democratization of these advanced generative models has enabled malicious actors to turbocharge disinformation operations, creating synthetic content farms that disseminate propaganda across multiple languages with near-zero marginal cost [3].

Under these modern conditions, it is necessary to conduct scientific research on hybrid threat detection models because existing analytical approaches are systematically inadequate. Analyzing text semantics using natural language processing [4, 5] without understanding their kinematic propagation vectors [6, 7] fails to distinguish between organic, user-driven virality and algorithmic, state-sponsored amplification [8]. Conversely, classical mathematical diffusion models [9, 10] rely on static coefficients that mathematically cannot account for the rapid, dynamic interventions characteristic of modern botnets.

In practice, the results of this research can provide security practitioners and threat analysts with a unified

framework to filter out organic social media noise and enable the early detection of artificially amplified cognitive threats. By mathematically transforming discrete linguistic classifications into continuous temporal indicators, the results will allow these cognitive variables to function as dynamic forcing functions within stochastic mass-servicing systems, facilitating the neutralization of bot-driven panic injections prior to mass audience infection [11, 12]. Therefore, research dedicated to the development of integrated hybrid neuro-mathematical models for detecting information threats is highly relevant.

1.2 Literature review and problem statement

In works [4, 8], deep learning and Natural Language Processing (NLP) are utilized to analyze the semantic payload of individual messages. These studies successfully identify discrete elements of state-sponsored propaganda and synthetic text. However, the evaluation of the temporal dimension of information dissemination remains an unresolved issue in these works. This question remains unresolved because the applied semantic transformer architectures inherently analyze texts in a vacuum and lack mechanisms to track network dynamics or propagation vectors over time [13–15].

To mitigate the information entropy of LLMs, works [16] and [17] propose Adaptive Self-Consistency (ASC) decoding algorithms. While this approach improves reasoning on complex tasks by dynamically adjusting the sampling budget, the stable integration of these stochastic classifications into rigid mathematical diffusion models remains unresolved. This issue remains unresolved because standard ASC implementations lack deliberate thermodynamic temperature scaling, which is fundamentally necessary to mathematically bound the variance of the neural network's outputs before feeding them into sensitive differential equations [18, 19].

Works [9, 10], and [20] apply the Bass diffusion model to evaluate the macroscopic kinematic spread of information based on external influence and internal imitation. However, adapting this model to the hyperdynamic environment of modern botnets remains an unresolved problem. This limitation persists because the proposed classical models strictly rely on static coefficients, which mathematically cannot account for the rapid, dynamic interventions and artificial volume injections characteristic of modern cyber-warfare [21, 22].

The microscopic dynamics of information cascades are considered in works [23–25], which propose Neural Hawkes point processes to dynamically learn non-linear intensity functions from temporal data. Nevertheless, evaluating the underlying cognitive payload that drives the organic virality is an unresolved issue in these studies. This question remains unresolved because the employed point processes are mathematically blind to semantic features; they rely exclusively on the metadata of temporal event clustering and ignore the real-time emotional volatility of the content [26].

Systematizing the identified local problems reveals a fundamental methodological gap: semantic models lack a temporal dimension, kinematic diffusion models lack cognitive awareness, and advanced generative algorithms lack the rigid mathematical bounding required for continuous differential systems. Therefore, the general unresolved problem is the systematic inability of existing, isolated analytical frameworks to reliably identify anomalous structural shifts within the state vector of a sociotechnical network during coordinated information operations. Resolving this problem requires the formalization of a non-linear mathematical mapping that systematically projects a highly dimensional space of linguistic features into a continuous state vector of spatio-temporal diffusion.

1.3 Research Aim and Objectives

The aim of this research is to construct and empirically validate a unified hybrid neuro-mathematical architecture that seamlessly integrates tensor-based semantic inference with analytical spatio-temporal diffusion models to detect anomalous structural shifts in sociotechnical networks during coordinated information operations. To achieve this aim, the following objectives were formulated:

- to develop a non-linear semantic contour based on a quantized large language model Llama-3.1-8B-Instruct utilizing a modified Adaptive Self-Consistency algorithm with thermodynamic logit scaling to minimize classification entropy;
- to synthesize continuous early warning indicators, specifically the Sentiment Volatility Coefficient and Bot-Injection Index, by mathematically transforming discrete semantic classifications into dynamic time-series scalars;
- to construct a dynamic differential contour by modifying the classical Bass diffusion model and the Neural Hawkes point process, substituting their static adoption constants with the AI-derived kinematic functions;
- to empirically validate the predictive sensitivity of the proposed hybrid model on real-world data from a live

information operation and determine the mathematical boundary limits of zero-shot generative models in granular emotional classification.

2. Materials and Methods

The object of the research is the sociotechnical information dissemination networks that are continuously subjected to coordinated manipulation and cognitive warfare during information operations. The paper considers the non-linear semantic features and continuous kinematic parameters, formed by generative large language models and diffusion equations, for early detection of anomalous structural shifts within the state vector of a sociotechnical network. The mathematical layer transposes the discrete semantic classifications into continuous temporal sequences to govern stochastic mass-servicing models. The main hypothesis of the research: the implementation of a unified hybrid neuro-mathematical architecture that maps non-linear semantic features onto differential equations of spatial-temporal diffusion will allow for reducing the semantic classification entropy and identifying the phase transition of coordinated cognitive attacks hours before traditional volume-based anomaly detectors. The method is based on the foundational theoretical postulate that modern information warfare possesses an inextricable dual nature: it is simultaneously semantic, comprising cognitive narratives, emotional manipulation, and rhetorical framing, and kinematic (comprising propagation velocity, dissemination channels, and network topology) [7]. This will ensure the early detection of artificially amplified cognitive threats prior to mass audience infection. When developing the hybrid neuro-mathematical method, the following conditions were used:

Condition 1. Organic information diffusion generally traces a smooth cumulative trajectory, whereas coordinated bot injections manifest mathematically as unit impulses, closely approaching the Dirac delta function or Heaviside step function [27].

Condition 2. The rate of adoption in classical diffusion linearly depends on the fraction of the population that has already adopted the information [21].

Condition 3. To maintain mathematical convergence without fine-tuning, the neural feature space is strictly confined to macro-level polarization, simplifying the highly granular emotional subsets that typically cause entropy surges.

Condition 4. The semantic outputs are transformed into continuous kinematic parameters, which dynamically govern the phase-space of a modified Bass diffusion model and a Neural Hawkes process. In the process of developing the method for detecting information threats, a number of various methods and algorithms were applied.

When developing the mathematical model of the Semantic Non-linear Contour, the principle of the Adaptive Self-Consistency decoding algorithm was used [16]. The principle consists in applying thermodynamic temperature scaling to the model's output logits to generate a stochastic ensemble of independent conclusions, diverging from standard greedy decoding. The probability distribution for the generation of the i -th

token (class) is computed utilizing a thermodynamically modified Softmax function:

$$p_i = \frac{\exp(z_i/\tau)}{\sum_j \exp(z_j/\tau)}, \quad (1)$$

where $p_i \in [0, 1]$ represents the normalized probability of selecting the i -th class from the total vocabulary of features, ensuring the sum of all p_i equals 1; $z_i, z_j \in \mathbb{R}$ represent the unscaled logits extracted from the final fully connected layer of the neural network, indicating the raw algorithmic “confidence” in class i or j ; and $\tau \in (0, \infty)$ denotes the generative temperature hyperparameter.

The application of this mathematical principle when building the dynamic process model provides the following advantages:

A detailed mathematical analysis of its boundary conditions reveals that as $\lim_{\tau \rightarrow 0} p_i$, the distribution degenerates into a discrete Dirac delta function (equivalent to $\arg \max$), losing its capacity to navigate semantically ambiguous texts. Conversely, as $\lim_{\tau \rightarrow \infty} p_i = \frac{1}{|V|}$, all exponential terms approach unity, transforming the distribution into absolute uniformity, whereby the system loses classification capability. Therefore, extensive multi-factor experimentation dictated the strict structural bifurcation of the parameter τ . For the classification of narrative and thematic topics, an elevated temperature of $\tau_{topic} = 0.85$ deliberately amplifies the sampling variance, compelling the neural network to traverse peripheral semantic association graphs, which is critical for defeating adversarial attempts to bypass monitoring systems via synonymic masking. For the classification of sentiment, the temperature was strictly compressed to $\tau_{sent} = 0.40$, which forces the probability distribution to cluster tightly around the dominant emotional vector, ensuring absolute convergence among the ensemble's independent inferences.

When quantifying algorithmic certainty in the Semantic Non-linear Contour, the principle of the mathematical confidence coefficient (Confidence margin) was used. The principle consists in calculating the differential between the volume of votes for the triumphant majority class and the closest competing alternative class within a stochastic ensemble. The mathematical margin is formalized as follows:

$$\gamma_i = \frac{\sum_{j=1}^N \mathbb{I}(y_j = c^*) - \max_{k \neq c^*} \sum_{j=1}^N \mathbb{I}(y_j = k)}{N}, \quad (2)$$

where N is the total ensemble size; $\mathbb{I}(\cdot)$ is the discrete indicator function returning 1 if the condition is true and 0 if false; c^* is the triumphant majority class; and k is the closest competing alternative class.

The application of this principle when evaluating algorithmic certainty provides the following advantages: it acts as a vital algorithmic entropy penalty. By purging classification outputs where $\gamma_i < \theta_{conf}$ was empirically optimized to 0.75, it creates an impenetrable safety barrier, preventing the infiltration of semantic noise into the highly sensitive differential equations of the dynamic contour. When computing the constituent elements of the system's macro-state vector S_t , the principle of five distinct mathematical Early Warning Indicators (EWIs),

specifically the hybridized impact function, was used. The principle consists in weighting each source's emission to counteract the obfuscation of social graph topologies and mitigate the impact of automated bot inflation, commonly referred to as Sybil attacks. The hybridized impact function is calculated as:

$$EIS_{it} = \alpha \frac{v_{it}}{\max(v_t)} + \beta \frac{s_{it}}{\max(s_t)}, \quad (3)$$

where v_{it} represents the median number of unique content views for the i -th source; s_{it} denotes the subscriber count of the emission source; $\max(v_t)$ and $\max(s_t)$ are denominators enforcing strict Euclidean normalization within the interval; $\alpha = 0.75$ and $\beta = 0.25$ are specific influence constants.

The application of this principle when computing the impact of sources provides the following advantages:

It enforces the mathematical dominance of kinematic interaction (views) over static topological metrics (subscribers), which is a fundamental prerequisite for gradient stability in downstream neural network layers. This effectively nullifies the efficacy of bot networks where the subscriber count approaches the maximum $s_i \rightarrow \max(s)$ but genuine human engagement approaches zero $v_i \rightarrow 0$. When isolating kinematic anomalies, the principle of the Early Amplification Ratio (EAR) was used [27]. The principle consists in computing the area under the intensity curve strictly during the initial hours of an attack's deployment, distinguishing coordinated bot injections, which manifest mathematically as unit impulses, closely approaching the Dirac delta function or Heaviside step function from organic information diffusion, which generally traces a smooth cumulative trajectory.

$$EAR(t_0, \Delta t) = \frac{\int_{t_0}^{t_0 + \Delta t} V_{topic}(t) dt}{\mu_V \cdot \Delta t}, \quad (4)$$

where $V_{topic}(t)$ represents the definite integral of content generation velocity, messages per hour, over a highly restricted analytical window $\Delta t \leq 2$ hours, computed via numerical integration methods; $\mu_V \cdot \Delta t$ represents the historical mathematical expectation of background intensity for the given narrative under non-anomalous conditions.

The application of this principle when isolating kinematic anomalies provides the following advantages:

An $EAR \geq 3.0$ algorithmic threshold signifies a 300% deviation from expected statistical behavior, immediately flagging a critical kinetic anomaly. When formalizing the sociological phenomenon of emotional manipulation – specifically the deliberate injection of panic, fear, or unbridled aggression into an otherwise neutral information stream – the principle of the Sentiment Volatility Coefficient (SVC) was used. The principle consists in measuring the standardized emotional dispersion:

$$SVC_t = \sqrt{\frac{1}{|\mathcal{M}_t|} \sum_{i=1}^{|\mathcal{M}_t|} (\sigma_i - \bar{\sigma}_t)^2}, \quad (5)$$

where $|\mathcal{M}_t|$ is the cardinality of the message set within the temporal window t ; $\sigma_i \in \{-1, 0, 1\}$ is the discrete emotional polarity derived from the semantic contour;

$\bar{\sigma}_t$ is the moving mathematical expectation of the sentiment for the set. The application of this principle when measuring emotional dispersion provides the following advantages:

The utilization of standard deviation over absolute deviation is mathematically vital; squaring the differentials exponentially amplifies the systemic weight of extreme emotional outliers, rendering the metric highly sensitive to artificial emotional injections. When inextricably linking the kinematic anomaly EAR with the cognitive anomaly SVC into a singular, unified threat marker, the principle of the Bot-Injection Index (BII) was used. The principle consists in calculating a synthetic, non-linear, multi-dimensional scalar utilizing the hyperbolic tangent function:

$$BII_t = \tanh\left(w_{ear} \cdot EAR_t + w_{svc} \cdot \frac{SVC_t}{SVC_{norm}}\right), \quad (6)$$

where $w_{ear} = 0.15$ and $w_{svc} = 0.35$ are weighting coefficients empirically optimized via gradient descent methodologies on historical datasets of prior information operations; SVC_{norm} represents a constant of normal background volatility, required to bring disparate multidimensional values to a common denominator. The application of this principle when linking anomalies provides the following advantages:

The integration of the hyperbolic tangent function constitutes a critically important architectural safeguard that explicitly eliminates the pervasive problem of gradient explosion when these metrics are subsequently fed into differential equations. It guarantees the compression of potentially infinite inputs, massive bot spam causing $EAR \rightarrow 100$, into a highly stable, bounded computational space of $[0, 1)$, ensuring the computational resilience of the overall system. When concentrating the model's computational resources exclusively on the destructive vector of the data stream, the principle of the Thematic Risk Factor (TRF) was used. The principle consists in acting as a complex spatial filter by aggregating the aforementioned scalars:

$$TRF_t = \sum_{m_i \in \mathcal{M}_t} EIS_{it} \cdot \mathbb{I}(\theta_i \in \Theta_{destruct}) \cdot \gamma_i, \quad (7)$$

where $\mathbb{I}(\theta_i \in \Theta_{destruct})$ is an indicator function that strictly zeroes out any message where the identified topic θ_i does not belong to the set of predefined destructive narratives $\Theta_{destruct}$; γ_i ensures that classifications with low neural certainty contribute proportionally less to the total accumulated risk. The application of this principle when filtering the data stream provides the following advantages:

It completely eliminates organic, neutral noise from the risk calculation, accelerating system response time. When evaluating the persistence of an operation, the principle of the geometric Campaign Longevity metric (NTP) was used. The principle consists in calculating the definite integral of the area where the risk exceeds the background norm:

$$NTP = \int_{T_{obs}} \max(0, TRF_t - \mu_{TRF}) dt, \quad (8)$$

where: TRF_t is the Thematic Risk Factor; μ_{TRF} is the background norm; and $\max(0, \dots)$ is a non-linear

operation functioning identically to a Rectified Linear Unit (ReLU) activation within neural networks, truncating all sub-threshold risk fluctuations to zero. The application of this principle when evaluating operation persistence provides the following advantages:

It ensures the integral accumulates area exclusively above the baseline anomaly threshold, allowing analysts to mathematically differentiate transient, organic information noise from sustained, highly resourced cognitive campaigns executed by state-sponsored actors. When achieving the mathematical transformation of classical, isotropic diffusion models into dynamic, predictive systems, the principle of non-linear dynamic modification of the Bass diffusion model was used [20]. The principle consists in replacing static constants with dynamic control functions derived directly from the cognitive-kinematic EWIs. The coefficient of external injection p_t^* is calculated as:

$$p_t^* = p_0 \cdot (1 + \lambda_p \cdot BII_t \cdot \overline{EIS}_t), \quad (9)$$

where p_0 is the base innovation rate; λ_p is a hyperparameter dictating sensitivity to bot attacks; $\overline{EIS}_t$ represents high-impact channels; and BII_t represents automated spam.

The application of this principle when modifying external injection provides the following advantages:

It mathematically simulates the real-world operational tactics of bot-farms, which utilize high-impact channels and automated spam to artificially inflate the external visibility of a narrative, driving an exponential surge in p_t^* . Concurrently, when calculating the coefficient of social imitation, the principle of emotional resonance function was used. The principle consists in mathematically modeling the proven sociological phenomenon that high-volatility content, inducing fear or anger, spreads horizontally at exponential rates [28]:

$$q_t^* = q_0 \cdot \left(1 + \lambda_q \cdot \max\left(0, \frac{SVC_t}{SVC_{norm}} - 1\right)\right), \quad (10)$$

where q_0 is the base organic imitation rate; λ_q reflects the audience's sensitivity to emotional fluctuations.

The application of this principle when modifying social imitation provides the following advantages:

If the emotional dispersion SVC_t spikes relative to the norm due to the injection of panic-inducing narratives, the multiplier becomes positive, mathematically accelerating the network's inherent horizontal virality. When mapping the microscopic topology of information cascades, the exact sequence of individual retweets over time, the principle of self-exciting Hawkes processes was used [23]. The principle consists in exponentially amplifying the baseline intensity of the point process by the Thematic Risk Factor derived from the LLM contour:

$$\lambda(t) = \mu_0 \cdot \exp(\alpha \cdot TRF_t) + \sum_{t_i < t} \phi(t - t_i), \quad (11)$$

where $\lambda(t)$ is the conditional intensity of the Poisson point process, the probability of a new publication in the next microsecond; $\phi(\tau) = \kappa \cdot e^{-\omega\tau}$ is the decay kernel representing the viral echo effect of all previous events t_i

[25]. The application of this principle when mapping information cascades provides the following advantages:

By embedding the TRF_t scalar directly into the exponent of the base intensity, the system generates a critical mathematical alert indicating an impending structural phase shift, the transition to mass “Attack” phase 2 to 4 hours prior to full operational deployment, significantly outperforming traditional delayed-response models.

3. Results of the Development and Empirical Validation of the Hybrid Neuro-Mathematical Model

3.1 Development of the Semantic Non-linear Contour and Entropy Minimization

To evaluate the efficacy of the proposed method and achieve complete scientific validation of the developed neuro-mathematical apparatus, an extensive computational experiment was executed utilizing real-world data. The empirical dataset comprised $N = 999$ lemmatized messages extracted from $K = 50$ high-density Telegram channels over an observation window of $\Delta t = 24$ hours, meticulously capturing the live deployment and dissemination of a verified information operation. Comprehensive data preprocessing protocols were applied, including the removal of stop words and strict tokenization constrained to a maximum of 512 tokens to

ensure complete compatibility with the LLM’s context window. Following the execution of the semantic contour utilizing the Adaptive Self-Consistency algorithm on the quantized Llama-3.1 model, the system successfully categorized 711 messages as neutral, 20 as positive, and 268 as destructive narratives $P_{neg} = 26.8\%$. A parallel unsupervised topic modeling process (utilizing BERTopic) independently corroborated these findings, isolating four robust topical clusters with an exceptionally high Adjusted Rand Index $ARI = 0.925$. The implementation of the thermodynamic logit scaling mechanism yielded an unprecedented reduction in classification entropy, which plummeted to $H_{topic} = 0.160$ bits, accompanied by a high mean algorithmic confidence margin of $\gamma = 0.878$. This definitively established the efficacy of the proposed thermodynamic technique in neutralizing the semantic uncertainty that typically plagues zero-shot LLM inference.

3.2 Synthesis of Continuous Early Warning Indicators

The continuous aggregation of these validated semantic metrics into the macro-state vector S_t , with an hourly discretization step $\Delta t = 1$ hr, provided empirical proof of the model’s extreme predictive sensitivity. The results of the modeling for these indicators are presented in Table 1.

Table 1 – Fragment of the calculation matrix of IO macro-indicators

Temporal Slice (t_i)	Intensity (V)	Sentiment ($\bar{\sigma}$)	Volatility SVC_t	EAR_{t_0}	Bot Index BII_t	Risk Factor TRF_t	Z-score TRF
Background t_{-1}	11	-0.15	0.34	1.0	0.05	1.45	0.12
Trigger t_0	48	-0.42	0.68	4.36	0.874	4.8	2.95
Attack t_1	95	-0.65	0.82	8.63	0.96	11.2	4.8

An exhaustive analysis of the mathematical behaviors observed during the critical phase transition at t_0 , the preparatory “Trigger” phase, illustrates the core superiority of the hybrid framework. The absolute volume metric $V = 48$ messages/hour remained insufficiently high to breach the static thresholds of conventional monitoring systems, which would have inevitably resulted in a false negative. However, the advanced kinematic apparatus detected severe localized deviations: the EAR surged to 4.36, signifying an immediate, non-organic injection of volume that was over four times the expected background rate. Concurrently, the emotional dispersion SVC experienced a 100% amplification jumping from a baseline of 0.34 to 0.68, mathematically confirming the artificial induction of panic and aggression into the information flow. Inserting these discrete scalars into the normalized BII formula yielded irrefutable mathematical evidence of algorithmic manipulation:

$$BII_{t_0} = \tanh\left(0.15 \cdot 4.36 + 0.35 \cdot \frac{0.68}{0.34}\right) \approx 0.874. \quad (12)$$

3.3 Construction of the Dynamic Differential Contour

The BII value of 0.874, an asymptotic push toward the theoretical limit of 1.0, alongside a standardized Z-score of 2.95 for the TRF, comfortably exceeding the critical 2.0σ anomaly threshold, triggered the

immediate, non-linear exponential scaling of the innovation parameter p_t^* in the modified Bass equation. This explicit neural-to-differential linkage enabled the system’s detection module to mathematically declare a phase transition to the “Attack” state 3 to 4 hours prior to the actual empirical peak volume occurring at t_1 .

3.4 Empirical Validation and Mathematical Limits of Generative Models

A secondary, yet equally profound result of the experiment was the mathematical documentation of the fundamental boundary limits inherent in contemporary artificial intelligence paradigms. During an extended experimental phase, the semantic feature space Σ was expanded beyond basic macro-polarization: positive, negative and neutral to encompass highly granular emotional subsets: anger, fear, anxiety, sorrow. The ensuing statistical analysis recorded an absolute collapse in semantic inference quality. The mean information entropy H for granular sentiment spiked exponentially to a critical 1.349 bits – dangerously approaching the maximum theoretical entropy limit for a 3 to 4 class distribution – while the ensemble confidence margin collapsed to a marginal $\gamma = 0.292$. The semantic space experienced severe divergence; the LLM ensemble began outputting near-uniform probability distributions, entirely incapable of distinguishing between adjacent

psycho-emotional states, confusing raw geopolitical anger with general hate speech. This empirically proved that zero-shot generative architectures are structurally incapable of executing stable, highly granular emotional classifications on complex texts without extensive, parametric fine-tuning, thereby defining the absolute operational limits of the hybrid model.

4. Discussion of the Results of the Hybrid Neuro-Mathematical Model Empirical Validation

The proposed hybrid neuro-mathematical model of the information threat detection process. The specific feature of this model is its dual-nature architecture, integrating semantic non-linear filtering with dynamic differential diffusion equations. The criteria and quality indicators of the threat detection process were determined. The specific feature of the proposed Early Warning Indicators (EWIs) is the application of continuous kinematic scalar aggregation, formulas for EAR, SVC, and BII. This allowed for evaluating the absolute mathematical limits of generative models and formulating the dynamic control parameters for the modified Bass diffusion model and Neural Hawkes process.

The evaluation of the proposed model's efficacy on Table 1, showed the following results: the implementation of thermodynamic logit scaling yielded an unprecedented reduction in classification entropy, which plummeted to $H_{topic} = 0.160$ bits, accompanied by a high mean algorithmic confidence margin of $\gamma = 0.878$. During the critical phase transition, the kinematic apparatus detected an EAR surge to 4.36 and a 100% amplification in emotional dispersion SVC, which mathematically confirmed the artificial induction of panic. The explicit neural-to-differential linkage enabled the system's detection module to mathematically declare a phase transition to the "Attack" state 3 to 4 hours prior to the actual empirical peak volume. Expanding the semantic feature space beyond basic macro-polarization to encompass highly granular emotional subsets without fine-tuning resulted in an absolute collapse in semantic inference quality, with entropy spiking to a critical 1.349 bits. The obtained results of the proposed model research can be explained by the application of thermodynamic stabilization and the integration of the hyperbolic tangent function within the BII metric, which explicitly eliminates the pervasive problem of gradient explosion when these metrics are fed into differential equations. In contrast to [4, 5], where deep learning sequence models or standard classification algorithms are proposed to parse complex disinformation, the proposed model incorporates the temporal dimension necessary to forecast the velocity and trajectory of information propagation. In addition, the proposed model, unlike [6], transforms rigorous semantic assessments into continuous kinetic variables, resolving the core deficiency of pure mathematical epidemiology when applied to highly volatile social networks.

In contrast to [27], where traditional systems remain perpetually reactive by isolating the cognitive payload but ignoring the dissemination physics, the proposed model guarantees system stability under duress by

mathematically bounding the parameters. Also, unlike, where it is assumed that massive, pre-trained models possess generalized psychological profiling capabilities out-of-the-box, the proposed model mathematically proves that zero-shot generative architectures are structurally incapable of executing stable, highly granular emotional classifications on complex texts without extensive parametric fine-tuning. Therefore, the developed hybrid neuro-mathematical model allowed for the early detection and mitigation of information anomalies long before they reach critical mass within the target population. This made it possible to filter out organic social media noise and mathematically differentiate transient information noise from sustained cognitive campaigns. But the proposed results should be applied under the following limitations:

- the application of the model is strictly confined to macro-level polarization of the neural feature space;
- the presence of a mathematically sound thermodynamic temperature scaling setup tailored for specific classification tasks: topic and sentiment.

In addition, the proposed model is expedient to use if the text corpus is sufficiently large and the temporal window provides a reliable historical background intensity baseline. As a drawback of this study, it should be noted that the paper lacks an analysis of the model's performance when integrated with specialized Low-Rank Adaptation (LoRA) or extensive supervised fine-tuning protocols. To eliminate this drawback, the hybrid model needs to be expanded with an additional procedure to align the model's latent space with the specific nuances of granular emotional classification in the operational environment. As a development of this research, the following can be noted.

First, it is necessary to investigate the integration of advanced LoRA techniques to expand the emotional granularity of the semantic contour while maintaining the convergence of the dynamic differential equations.

Second, it is necessary to conduct research regarding the dynamic optimization of the weighting coefficients w_{ear} and w_{svc} using automated machine learning (AutoML) frameworks across diverse types of sociotechnical platforms beyond Telegram.

5. Conclusions

1. The proposed mathematical model of the information threat detection process. The specific feature of this model is its dual-nature architecture, integrating semantic non-linear filtering with dynamic differential diffusion equations. The application of Adaptive Self-Consistency decoding with thermodynamic logit scaling yielded an unprecedented reduction in classification entropy. This allowed for formulating strict boundary conditions to prevent the collapse of coupled differential systems and minimize semantic uncertainty.

2. The criteria and quality indicators of the threat detection process were determined. The specific feature of the proposed Early Warning Indicators is the application of continuous kinematic scalar aggregation. This allowed for assessing the temporal dimension of information dissemination and formulating continuous dynamic control functions for the Bass diffusion equation and Neural Hawkes point process.

3. The developed unified hybrid neuro-mathematical architecture seamlessly integrates tensor-based semantic inference with analytical spatio-temporal diffusion models. In the first stage, the Llama-3.1-8B-Instruct model resolves semantic ambiguity via the ASC algorithm; in the second stage, linguistic outputs are transformed into dynamic early warning indicators. The results allowed for evaluating the efficacy of the developed method. During empirical testing, the method significantly outperformed traditional detectors by accurately identifying the phase shift of cognitive attacks hours before the empirical peak volume. But, expanding the feature space to highly granular emotional subsets resulted in an absolute collapse in inference quality, establishing the mathematical limitations of zero-shot models without fine-tuning.

Conflicts of interest

The authors declare that they have no conflicts of interest regarding this research, including financial, personal, authorship, or any other nature that could influence the research and the results presented in this article.

Funding

The research was conducted without financial support.

Data Availability

The manuscript has no associated data.

Use of artificial intelligence

The authors confirm that they did not use artificial intelligence technologies in the creation of the presented work.

REFERENCES

1. Parmar, M. (2026), "Safety, security, and cognitive risks in world models", *arXiv*, arXiv.2604.01346, doi: <https://doi.org/10.48550/arXiv.2604.01346>
2. Li, Y., Yang, Y., and Huang, H. (2025), "Advances in social bots [Editorial]", *Electronics*, vol. 14, no. 24, 4885, doi: <https://doi.org/10.3390/electronics14244885>
3. Sadeghi, M., Dimitriadis, D., Padovese, V., Pozzi, G., Badilini, S., Vercellone, C., Huet, N., Fishman, Z., Pfaller, L., and Adams, N. (2026), "Tracking AI-enabled misinformation: 3,749 AI content farm sites (and counting), plus the top false claims generated by artificial intelligence tools", *NewsGuard*, available at: <https://www.newsguardtech.com/special-reports/ai-tracking-center>
4. Sabri, M. I. F., Hettiarachchi, H., and Ranasinghe, T. (2025), "Detecting fake news in the era of language models", *Proceedings of the 15th International Conference on Recent Advances in Natural Language Processing*, pp. 1036–1043, available at: <https://aclanthology.org/2025.ranlp-1.119/>
5. Burghardt, K., Chen, K., and Lerman, K. (2024), "Large language models reveal information operation goals, tactics, and narrative frames", *Proceedings of the ICWSM Workshops*, doi: <https://doi.org/10.48550/arXiv.2405.03688>
6. Stanovov, V., Grabljevec, S., Akhmedova, S., Semenkin, E., Stojanović, R., Rozman, Č., and Škraba, A. (2022), "Identification of COVID-19 spread mechanisms based on first-wave data, simulation models, and evolutionary algorithms", *PLOS One*, doi: <https://doi.org/10.1371/journal.pone.0279427>
7. Whiteside, R. D. (2025), "Narrative manipulation, malinfluence operations, and cognitive warfare through large language model poisoning with adversarial noise", *Military Intelligence Professional Bulletin*, available at: <https://mipb.ikn.army.mil/>
8. Ezzeddine, F., Ayoub, O., Giordano, S., Nogara, G., Sbeity, I., Ferrara, E., and Luceri, L. (2023), "Exposing influence campaigns in the age of LLMs: A behavioral-based AI approach to detecting state-sponsored trolls", *EPJ Data Science*, vol. 12, no. 1, 46, doi: <https://doi.org/10.1140/epids/s13688-023-00423-4>
9. Cosguner, K., and Seetharaman, P. B. (2022), "Dynamic pricing for new products using a utility-based generalization of the Bass diffusion model", *Management Science*, vol. 68, no. 3, pp. 1904–1922, doi: <https://doi.org/10.1287/mnsc.2021.4257>
10. Liang, L. (2021), "Quantitative analysis of diffusion of public opinion by the modified Bass model for the dyeing steamed bread incident in China", *Saudi Journal of Business and Management Studies*, vol. 6, no. 1, pp. 24–27, doi: <https://doi.org/10.36348/sjbms.2021.v06i01.004>
11. Ask, T. F., Sütterlin, S., Müller, L., Lugo, R., Østby, L., Jensen, M., Nordheim, K., Gjøvik, J., and Lervik, E. (2025), "Cognitive security: The study and practice of protecting the human mind and other cognitive assets from cognitive threats", *OSF*, doi: https://doi.org/10.31234/osf.io/2ftqc_v1
12. Stella, M., Ferrara, E., and De Domenico, M. (2018), "Bots increase exposure to negative and inflammatory content in online social systems", *Proceedings of the National Academy of Sciences*, vol. 115, no. 49, pp. 12435–12440, doi: <https://doi.org/10.1073/pnas.1803470115>
13. Munir, B., Abbasi, M. Z., Wilson, W. B., and Colombo, A., Jr. (2025), "Evaluating AI in legal operations: A comparative analysis of accuracy, completeness, and hallucinations in ChatGPT-4, Copilot, DeepSeek, Lexis+ AI, and Llama 3", *International Journal of Legal Information*, vol. 53, no. 2, pp. 103–114, doi: <https://doi.org/10.1017/iji.2025.10052>
14. Barbera, I. (2025), *AI privacy risks & mitigations - large language models (LLMs). Practical guide for conducting risk assessments to mitigate privacy and data protection risks in systems using LLMs*, European Data Protection Board Support Pool of Experts, available at: <https://www.edpb.europa.eu/system/files/2025-04/ai-privacy-risks-and-mitigations-in-llms.pdf>
15. Vassilev, A., Oprea, A., Fordyce, A., Anderson, H., Davies, X., and Hamin, M. (2025), *Adversarial machine learning: A taxonomy and terminology of attacks and mitigations* (NIST AI 100-2e2025), National Institute of Standards and Technology, doi: <https://doi.org/10.6028/NIST.AI.100-2e2025>
16. Jeong, E., and Choi, Y. S. (2025), "ACR: Adaptive confidence re-scoring for reliable answer selection among multiple candidates", *Applied Sciences*, vol. 15, no. 17, 9587, doi: <https://doi.org/10.3390/app15179587>
17. Alcaraz, A. (2023), *Achieving greater self-consistency in large language models*, Towards Data Science, available at: <https://towardsdatascience.com/achieving-greater-self-consistency-in-large-language-models-6e6cb5f3c5b7>
18. Xie, Z., Guo, J., Yu, T., and Li, S. (2024), *Calibrating reasoning in language models with internal consistency*, NeurIPS 2024, available at: <https://neurips.cc/virtual/2024/poster/93260>
19. Taherinezhad, F., Momeni Nezhad, M. J., Karimi, S., et. al (2026), "Large language model adaptation strategies in speech-based cognitive screening: Systematic evaluation", *JMIR AI*, vol. 5, e82608, doi: <https://doi.org/10.2196/82608>

20. Bass, F. M., Krishnan, T. V., and Jain, D. C. (1994), "Why the Bass model fits without decision variables", *Marketing Science*, vol. 13, no. 3, pp. 203–223, doi: <https://doi.org/10.1287/mksc.13.3.203>
21. Manshadi, V., Johari, R., and Misra, S. (2015), *A modified Bass model for product growth in social networks*, available at: https://www-2.rotman.utoronto.ca/MSOM-2015/Documents/MSOM2015_0325.pdf
22. Baek, J., Farias, V. F., Georgescu, A., Levi, R., Peng, T., Sinha, D., Wilde, J. T., and Zheng, A. T. (2025), "The limits to learning a diffusion model", *Management Science*, doi: <https://doi.org/10.1287/mnsc.2022.02953>
23. Laub, P. J., Lee, Y., Pollett, P. K., and Taimre, T. (2025), "Hawkes models and their applications", *Annual Review of Statistics and Its Application*, vol. 12, pp. 233–258, doi: <https://doi.org/10.1146/annurev-statistics-112723-034304>
24. Li, J., Yang, Y., Zhang, Y., Hu, Q., Zhao, A., and Gao, H. (2025), "Public opinion field effect and Hawkes process join hands for information popularity prediction", *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 39, no. 11, pp. 12076–12083, doi: <https://doi.org/10.1609/aaai.v39i11.33315>
25. Okawa, M., Iwata, T., Tanaka, Y., Toda, H., Kurashima, T., and Kashima, H. (2021), "Dynamic Hawkes processes for discovering time-evolving communities' states behind diffusion processes", *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, pp. 1276–1286, doi: <https://doi.org/10.1145/3447548.3467248>
26. Zeng, F., and Gao, W. (2022), "Early rumor detection using neural Hawkes process with a new benchmark dataset", *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics*, pp. 4105–4117, doi: <https://doi.org/10.18653/v1/2022.naacl-main.302>
27. Kuntur, S., Wróblewska, A., Paprzycki, M., and Ganzha, M. (2025), "Under the influence: A survey of large language models in fake news detection", *IEEE Transactions on AI*, vol. 6, no. 2, pp. 458–476, doi: <https://doi.org/10.1109/TAI.2024.3471735>
28. Eryarsoy, E., Delen, D., Davazdahemami, B., and Topuz, K. (2021), "A novel diffusion-based model for estimating cases, and fatalities in epidemics: The case of COVID-19", *Journal of Business Research*, vol. 124, pp. 163–178, doi: <https://doi.org/10.1016/j.jbusres.2020.11.054>

Received (Надійшла) 21.06.2026

Accepted for publication (Прийнята до друку) 02.09.2026

Publication date (Дата публікації) 25.09.2026

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Лашин Ярослав Олександрович – ад'юнкт кафедри соціальної комунікації та публічної дипломатії, Інститут стратегічних комунікацій, Національний університет оборони України, Київ, Україна;

Yaroslav Lashyn – Ph. D. Student at the Department of Social Communication and Public Diplomacy, Institute of Strategic Communications, National Defence University of Ukraine, Kyiv, Ukraine;

e-mail: y.lashyn@edu.nuou.org.ua; ORCID Author ID: <https://orcid.org/0009-0000-5533-9210>;

Войтко Олександр Володимирович – доктор військових наук, доцент, начальник інституту стратегічних комунікацій, Національний університет оборони України, Київ, Україна;

Oleksandr Voitko – Doctor of Military Sciences, Associate Professor/Head of the Institute of Strategic Communications, National Defence University of Ukraine, Kyiv, Ukraine;

e-mail: o.voitko@edu.nuou.org.ua; ORCID Author ID: <https://orcid.org/0000-0002-4610-4476>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57210362201>.

Федорієнко Віталій Анатолійович – кандидат технічних наук, професор кафедри соціальної комунікації та публічної дипломатії, Інститут стратегічних комунікацій, Національний університет оборони України, Київ, Україна;

Vitalii Fedorienko – PhD, Professor at the Department of Social Communication and Public Diplomacy, Institute of Strategic Communications, National Defence University of Ukraine, Kyiv, Ukraine;

e-mail: v.fedorienko@edu.nuou.org.ua; ORCID Author ID: <https://orcid.org/0000-0002-0921-3390>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57221718554>.

**Гібридна нейроматематична модель виявлення інформаційних загроз:
інтеграція великих мовних моделей та модифікованих дифузійних процесів**

Я. О. Лашин, О. В. Войтко, В. А. Федорієнко

Анотація. Актуальність. У контексті стрімкої еволюції кіберзагроз, коли деструктивні нарративи поширюються з безпрецедентною швидкістю, традиційні ізольовані системи моніторингу втрачають ефективність. **Предмет дослідження:** аналітичні системи виявлення скоординованих інформаційних операцій у сучасному середовищі когнітивної безпеки. **Мета дослідження:** розробити єдину нейроматематичну архітектуру, яка відображає нелінійні семантичні ознаки за допомогою диференціальних рівнянь просторово-часової дифузії. **Завдання дослідження:** побудувати гібридну модель на основі великої мовної моделі, перетворити семантичні вихідні дані на неперервні кінематичні параметри та емпірично оцінити ефективність моделі під час реальних інформаційних атак. **Методи дослідження:** застосовано змішаний емпіричний підхід, що поєднує кількісне математичне моделювання, комп'ютерну лінгвістику, модель Llama-3.1-8B-Instruct з використанням алгоритму Adaptive Self-Consistency, а також модифіковану модель дифузії Басса та нейронний процес Гокса. **Результати дослідження:** застосування термодинамічного логіт-масштабування знизило ентропію семантичної класифікації до 0,160 біта. Інтеграція неперервних кінематичних параметрів у диференціальні рівняння дала змогу системі виявити фазовий перехід інформаційної атаки за 3–4 години до досягнення її статистичного піку. Встановлено, що детальна класифікація емоційних підтипів без тонкого налаштування моделі призводить до математичної нестабільності, унаслідок чого ентропія зростає до 1,349 біта. **Висновки:** дослідження розширює теорію дифузії Басса та процеси Гокса шляхом заміни статичних констант динамічними функціями впливу, отриманими за допомогою штучного інтелекту. Математично обґрунтовано необхідність макрополяризації у просторі ознак для забезпечення стабільного диференціального моделювання без тонкого налаштування нейронної мережі.

Ключові слова: інформаційна безпека; інформаційні операції; великі мовні моделі; оброблення природної мови; аналіз соціальних мереж; гібридна нейроматематична модель; дифузійні процеси; стохастичні моделі.

Oleksandr Skorlupin, Andrii Podorozhniak

National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine

YOLO FAMILY ARCHITECTURES FOR EXPLOSIVE OBJECT DETECTION IN EMBEDDED SYSTEMS OF AUTONOMOUS UNMANNED VEHICLES: A COMPARATIVE ANALYSIS

Abstract. Relevance. The need to choose an effective YOLO architecture for explosive object detection in unmanned aerial vehicle embedded systems is driven by limited computational resources, real-time constraints, and energy efficiency requirements. **The object of research** is explosive object detection in embedded autonomous unmanned vehicle systems using YOLO family architectures. **The subject of the research** is the comparative efficiency of YOLO models in terms of detection accuracy, throughput, model size, and energy consumption on embedded hardware platforms. **The purpose of this paper** is to develop and apply an integral efficiency criterion based on neural network methods for selecting the most suitable YOLO architecture for explosive object detection and humanitarian demining tasks. **Research results.** The proposed criterion was applied to six YOLO models (v5n, v8n, v8s, v10n, v11n, v11s) deployed on Raspberry Pi 3 Model B+ and Raspberry Pi 4 Model B with 4 GB of RAM. Comparative evaluation under three scenarios showed substantial differences in detection quality, throughput, model size, and energy consumption. The ranking confirms the practical value of the methodology for model selection in embedded systems. **Conclusions.** The methodology provides a transparent and reproducible basis for comparing YOLO architectures in resource-constrained environments. Real-time operation remains challenging on low-power platforms, which justifies further optimization of deployment and inference.

Keywords: YOLO; explosive object detection; unmanned aerial vehicle; embedded systems; integral efficiency criterion; neural network; humanitarian demining; Raspberry Pi.

Introduction

Relevance. The task of automated detection of explosive objects using an embedded computer system onboard an autonomous unmanned aerial vehicle is an important research direction in the field of humanitarian demining [1–3]. The scale of contamination of Ukraine's territory by explosive remnants of war, which international organizations estimate to cover tens of thousands of square kilometers, combined with the productivity of manual demining at only 15–25 m² per deminer per day, creates a strong need for automated systems for preliminary terrain survey [4, 5].

Therefore, one of the most promising ways to address this problem is the deployment of a variety of robotic systems, including unmanned aerial vehicles (UAVs) and unmanned ground vehicles (UGVs) [6].

An overview of scientific works. In its most general form, the architecture of such a system includes image acquisition from an optical camera, resizing the image to the fixed input resolution of the neural network detector (typically 640×640 pixels), forward propagation through a YOLO-family convolutional neural network, and generation of georeferenced detection results in near real time [7].

In recent publications, methods are proposed for quantitatively assessing the effectiveness of systems of this type using classification quality metrics (Precision, Recall, F1), a spatial localization metric (IoU), and performance-oriented indicators (FPS, decomposition of the time budget for processing a single frame) [8, 9]. These methods assume that the target model has been predefined in advance and make it possible to evaluate its compliance with the functional requirements of the task.

At the system design stage, the opposite task arises: selecting, on a well-founded basis, the most appropriate architecture from several potentially suitable candidates

using the same criteria, but without conducting a full cycle of field tests for each of them.

The technical specificity of the task is complicated by several circumstances that must be taken into account in the selection methodology. The current YOLO family includes at least six architectural variants whose computational complexity is compatible with the resources of single-board computers such as the Raspberry Pi 3 Model B+ (quad-core Cortex-A53, 1 GB LPDDR2) and Raspberry Pi 4 Model B (quad-core Cortex-A72, 4 GB LPDDR4). This list includes the v5n, v8n, v8s, v10n, v11n, and v11s versions, ranging from the 2021 release to the 2025 architectures. Each of the named models can be run in several software inference formats – PyTorch, ONNX Runtime, TensorFlow Lite, and NCNN – each of which has its own trade-off between speed and deployment convenience on ARM processors [10]. Special attention should be given to the fact that the latest-generation models, especially YOLOv10, implement a training strategy based on dual label assignment; this eliminates the need for Non-Maximum Suppression (NMS) during inference and provides a deterministic latency gain, which is especially noticeable on platforms with limited computing power [11].

The combination of these factors makes model selection non-trivial in nature, and an arbitrary choice of a default architecture is methodologically unjustified. In the context of research focused on embedded detector deployment, the issue of selecting a well-founded architectural variant acquires independent scientific significance.

The aim of the article is to formalize and practically apply a methodology for the justified selection of a YOLO architecture for the task of detecting explosive objects by an embedded computer system onboard an unmanned aerial vehicle. The proposed methodology is developed as a logical continuation of the

criterion framework presented by the authors in, with the analytical apparatus extended to an integral efficiency criterion that accounts for the specific features of both target platforms and the operational conditions of the system in support of humanitarian demining.

Problem statement and characteristics of the onboard computing platforms

The task considered is the automated detection of explosive objects across six classes – anti-personnel and anti-tank mines, mortar and artillery shells, hand grenades, and small arms ammunition – in images received from an optical camera onboard an unmanned aerial vehicle, using the onboard computing system. The class nomenclature is inherited from earlier works [2, 6, 13] and corresponds to the most common types of explosive remnants of war in Ukraine [4, 14, 15].

Detection should be performed directly onboard, without transmitting the video stream to a ground station, because analyses of edge-computing efficiency in highly mobile computing networks show that local processing enables decision-making that is on average 15% faster than a centralized architecture and preserves node operability under communication loss, which is of fundamental importance for post-war applications [16]. Raspberry Pi 3 Model B+ and Raspberry Pi 4 Model B, each in the 4 GB RAM configuration, were selected as the onboard computing platforms. The choice is motivated by three considerations: both platforms are widely available, have a well-established software ecosystem, and exhibit documented behavior under sustained computational load, which improves result reproducibility; prior experience with deploying computer vision systems on Raspberry Pi for license plate recognition is described in [17].

The computational-performance gap between the Pi 3B+ and Pi 4, documented at about 2.5 – 4 times [18], covers the typical range of onboard computers used in small UAVs, so the obtained estimates can be interpolated to intermediate configurations. The power consumption of both platforms, measured in watts, is compatible with the energy budget of a small multirotor aircraft: at a flight power of 200 – 400 W, onboard computing accounts for only a few percent.

Raspberry Pi 3 Model B+ is built around the Broadcom BCM2837B0 system-on-chip (SoC) with four ARM Cortex-A53 cores running at 1.4 GHz and 1 GB of LPDDR2 RAM. The Cortex-A53 microarchitecture is an energy-efficient, in-order design and provides about 2.24 DMIPS/MHz per core; the limited memory places an upper bound on the model size that can be loaded together with the operating system and inference software stack, while the platform's power use is about 1.9 – 2.1 W at idle and up to 3.6 W under full computational load.

Raspberry Pi 4 Model B uses the BCM2711 SoC with four Cortex-A72 cores clocked at 1.5 GHz and 4 GB of LPDDR4-3200 RAM. Cortex-A72 is a superscalar out-of-order microarchitecture that delivers about 4.72 DMIPS/MHz, roughly twice that of Cortex-A53 at the same frequency; combined with the faster memory subsystem, this provides a 2.5 – 4× increase in convolutional-network inference performance compared with the Pi 3B+. Power consumption is 3.4 – 4.0 W at idle and 6.0 – 7.6 W under load; during prolonged full-load operation, the platform requires a passive heatsink to avoid thermal throttling, which above 80°C reduces clock frequency and distorts speed measurements. The consolidated platform characteristics are presented in Table 1.

Table 1 – Characteristics of onboard computing platforms

Technical specification	Pi 3B+	Pi 4 (4 Gb)	Pi 5 (8 Gb)	Jetson Orin Nano Super
SoC	BCM2837B0	BCM2711	BCM2712	6×Cortex-A78AE + GPU 1024 CUDA
Processor	4×Cortex-A53	4×Cortex-A72	4×Cortex-A76	1,7
Clock frequency, GHz	1,4	1,5	2,4	8 ГГц LPDDR5
Memory	1 Гб LPDDR2	4 Гб LPDDR4	8 Гб LPDDR4X	7,0
Performance, DMIPS/MHz	2,24	4,72	6,4	7
Power (idle), W	1,9-2,1	3,4-4,0	2,7–3,0	15–25
Power (load), W	3,6	6,0-7,6	7,5–9,0	249
Approximate cost, USD	35	55	80	6×Cortex-A78AE + GPU 1024 CUDA

To complete Table 1 for a fuller comparison, next-generation platforms, Raspberry Pi 5 and NVIDIA Jetson Orin Nano Super, are also included, although they are not part of the study's comparative set. Raspberry Pi 5, based on the Cortex-A76 processor, delivers roughly four times higher performance than Raspberry Pi 4, but its power consumption under load rises to 7.5–9.0 W, while the total cost of the platform together with mandatory active cooling and a higher-power supply is about 1.5 times that of Raspberry Pi 4.

The NVIDIA Jetson Orin Nano Super platform, equipped with an Ampere-architecture GPU accelerator,

outperforms the considered single-board computers by an order of magnitude in computational performance; however, its power consumption in maximum-performance mode reaches 25 W, and its cost is about 4.5 times higher than that of Raspberry Pi 4.

For a small multirotor UAV, both factors are highly significant: higher power consumption shortens flight time for a fixed battery capacity, and the cost of the computing module becomes decisive in serial deployment, where loss of the aircraft in a demining zone is an expected operational risk. These platforms are therefore considered a direction for further expansion of

the study, provided that the energy and cost constraints are relaxed.

An essential part of the problem formulation is the choice of the software format for model execution. For ARM processors without hardware accelerators, several inference stacks are available: native PyTorch, ONNX Runtime, TensorFlow Lite, and the NCNN framework, which is optimized for mobile ARM platforms using vectorized NEON instructions. According to official measurements [18, 19], for the YOLOv8n model on Raspberry Pi 4, the NCNN format provides an inference time of 414.7 ms versus 560.0 ms in ONNX Runtime and 950.3 ms in TensorFlow Lite, corresponding to advantages of 1.35 and 2.29 times, respectively. In view of this, the subsequent analysis is carried out for the NCNN format as the primary one, with data for the other formats given for comparison in the comparative analysis.

The system requirements are formulated as follows. The target video-processing rate is set at 25 frames per second, which corresponds to the standard PAL frame rate and typical onboard camera settings. The detector input resolution is 640×640 pixels, inherited from the training configuration of the YOLO family of models. The flight altitude in the calculation scenario is 20 m, which, at a typical camera field of view [20], yields a ground sampling distance on the order of centimeters per pixel and is sufficient for detecting anti-personnel mine-class objects. It is expected in advance that the target rate will not be achieved by any model on general-purpose processors without accelerators; quantifying this gap and identifying ways to reduce it are among the results of the work.

Overview of the compared YOLO family architectures

Six models from four YOLO generations are included in the comparison: YOLOv5n, YOLOv8n, YOLOv8s, YOLOv10n, YOLOv11n, and YOLOv11s. The selection criterion was the basic compatibility of model computational complexity with the resources of the target platforms: the smallest nano configurations of each generation were considered, along with two small-class configurations to assess whether moving to larger models on Raspberry Pi 4 is justified. Medium-class configurations and above were excluded because their computational complexity, exceeding 60 GFLOPs, leads to frame-processing times of several seconds even on Pi 4, which makes such a comparison practically meaningless. The 2025 architectures are included in the summary table of characteristics as reference points, but they are not used in the computational part for the reasons given below. Two-stage detectors such as Mask R-CNN, whose application to recognition tasks is discussed [21], are excluded from the comparison because their computational requirements are fundamentally higher and incompatible with the resources of the target platforms.

YOLOv5n (2021) is the smallest configuration of the fifth generation and contains 1.9 million parameters with a computational complexity of 4.5 GFLOPs. The

architecture uses a CSPDarknet backbone with C3 modules and an anchor-based detection scheme. The model achieves 28.0% mAP50-95 on the COCO validation set, which is the lowest among the compared models, but its minimal computational complexity makes its speed a benchmark for resource-constrained platforms.

YOLOv8n and YOLOv8s (2023) belong to the eighth generation, which introduced an anchor-free detection scheme with a decoupled head and C2f modules in the backbone. YOLOv8n has 3.2 million parameters at 8.7 GFLOPs and achieves 37.3% mAP50-95; YOLOv8s has 11.2 million parameters at 28.6 GFLOPs and 44.9%, respectively. The eighth generation is treated as the baseline for explosive-object detection because earlier work [2, 6] and the evaluation methodology were built on YOLOv8n, and the transfer of this generation to applied domains is documented, including in remote sensing tasks [22].

YOLOv10n (2024) uses a fundamentally different training scheme with a dual label assignment mechanism: during training, a one-to-many branch operates in parallel to provide a strong gradient signal, while a one-to-one branch produces a consistent output without duplicated detections. As a result, the non-maximum suppression stage is removed from inference [11]. For embedded platforms, this matters in practice because NMS is executed serially on the CPU, and on ARM architectures without a vectorized implementation its cost can reach tens of milliseconds per frame. YOLOv10n has 2.3 million parameters at 6.7 GFLOPs and achieves 38.5% mAP50-95, outperforming YOLOv8n in both accuracy and speed.

YOLOv11n and YOLOv11s (2024) represent the current generation, in which the C2f modules have been replaced by more compact C3k2 blocks, a C2PSA spatial attention module has been added, and the number of parameters has been reduced while preserving accuracy. With 2.6 million parameters (6.5 GFLOPs), YOLOv11n achieves 39.5% mAP50-95, the highest result among the nano-class models; YOLOv11s, with 9.4 million parameters (21.5 GFLOPs), reaches 47.0%, which is the maximum among all compared models. The consolidated characteristics are presented in Table 2.

Table 2 – Characteristics of the compared models (input 640×640, COCO val)

Model	Year	Param., million	GFLOPs	mAP50-95, %
YOLOv5n	2021	1,9	4,5	28,0
YOLOv8n	2023	3,2	8,7	37,3
YOLOv8s	2023	11,2	28,6	44,9
YOLOv10n	2024	2,3	6,7	38,5
YOLOv11n	2024	2,6	6,5	39,5
YOLOv11s	2024	9,4	21,5	47,0
YOLOv12n	2025	2,6	6,5	40,6
YOLOv26n	2025	2,4	5,4	40,1

Two newer architectures are included separately in Table 2, but they are excluded from the computational part of the study. YOLOv12n implements a spatial attention mechanism in place of part of the convolutional blocks and, according to formal metrics, outperforms YOLOv11n by 1.1 percentage points in mAP50-95 at the same computational complexity.

However, achieving the claimed speed requires optimized attention implementations designed for graphics accelerators; on general-purpose processors, attention operations are less efficiently vectorized than convolutions, so the actual inference time on ARM platforms is higher than that of YOLOv11n despite identical GFLOPs. An additional factor is lower stability during export: support for attention mechanisms in inference stacks for embedded platforms is incomplete, and conversion results vary across tool versions. YOLO26n is built on an anchor-free scheme without non-maximum suppression and, according to published measurements, outperforms YOLOv11n in both accuracy and speed. The limitation of this architecture lies in the requirements for the software environment: the model needs up-to-date versions of the training and export toolchain, support for which in stable builds of inference frameworks for ARM platforms was not available at the time of the study, and the full training cycle imposes higher computational demands than architectures of earlier generations. In view of this, both models are considered a direction for further research after the tool support has stabilized.

The comparison across generations reveals a clear trend: each successive nano-class generation delivers higher accuracy at lower or comparable computational complexity.

YOLOv11n outperforms YOLOv5n by 11.5 percentage points in mAP50-95 with only a 44% increase in complexity, while YOLOv10n achieves accuracy comparable to YOLOv11n with no NMS stage at all. At the same time, moving from the nano class to the small class within the same generation increases computational complexity by 3.3 times, whereas the accuracy gain is only 7.5–7.6 percentage points; the quantitative impact of this asymmetry for resource-constrained platforms is assessed below.

Comparison criteria and integral efficiency indicator

Model comparison is performed using four partial indicators: detection quality, effective speed, model size, and the energy consumed to process a single frame. These partial indicators are reduced to a single integral criterion, which formalizes model selection under a given requirement profile and eliminates the subjectivity of comparing models by only one characteristic.

Detection quality is characterized by the mAP50-95 metric on the COCO validation set, according to the official data provided by the developers [11, 19, 23]. Although the target domain of the task – images of explosive objects captured from flight altitude – differs from the COCO distribution, the relative ranking of models by accuracy after fine-tuning on a specialized dataset, as shown by the practice of transferring YOLO-

family models to applied domains [17, 21, 22], is preserved, which makes this indicator suitable for comparative analysis.

Speed is evaluated through a frame-processing time model. Inference time is assumed to be proportional to the model's computational complexity and inversely proportional to the platform's effective performance:

$$t_{\text{inf}} \approx C / P_{\text{eff}}, \quad (1)$$

where C – model's computational complexity, GFLOPs; P_{eff} – platform's effective performance for convolutional inference tasks, GFLOPS. The value of P_{eff} is determined by calibration against the measured inference time of a reference model: according to [19, 20], for YOLOv8n ($C=8.7$ GFLOPs) on Raspberry Pi 4 in NCNN format, the inference time is 414.7 ms, from which $P_{\text{eff}} \approx 21.0$ GFLOPS. For models for which direct measurements are unavailable, inference time is obtained by extrapolation using equation (1), calibrated with the reference YOLOv8n model for the nano class and YOLOv8s for the small class.

The total frame processing time t_{full} consists of preprocessing, inference, and postprocessing:

$$t_{\text{full}} = t_{\text{pre}} + t_{\text{inf}} + t_{\text{nms}}, \quad (2)$$

where t_{pre} – time spent on resizing the image resolution and normalization; t_{nms} – non-maximum suppression time, which for models with a classical detection head on ARM processors is estimated at 20 ms for Pi 4 and 30 ms for Pi 3B+, while for YOLOv10n it is zero by architectural design [11]. The effective system throughput FPS_{eff} is defined as:

$$FPS_{\text{eff}} = 1 / t_{\text{full}}. \quad (3)$$

The energy characteristic is expressed as the energy consumed to process a single frame:

$$E_{\text{fr}} = P_{\text{load}} \cdot t_{\text{full}}, \quad (4)$$

where P_{load} – the platform's average power consumption under computational load, in watts. Based on measurements [20], 6.0 W is assumed for Raspberry Pi 4 and 3.6 W for Raspberry Pi 3B+. Equation (4) reflects the dual effect of slow inference on a weaker platform: despite lower absolute power consumption, the energy per frame for Pi 3B+ may exceed that of Pi 4 because of the longer processing time, which, for a fixed onboard battery capacity, directly reduces the number of frames processed during a flight. The model size in single-precision format is estimated by the number of parameters:

$$M = 4 \cdot N_{\text{par}}, \quad (5)$$

where N_{par} – the number of model parameters; the coefficient 4 corresponds to the size of a single-precision number in bytes. For the Pi 3B+ platform with 1 GB of RAM, this metric limits the set of feasible models once the memory used by the operating system, the inference stack, and intermediate network activations is taken into account.

Because the partial metrics have different units and opposite optimization directions, they are normalized to the range before aggregation. For metrics to be maximized (mAP, FPS_{eff}), direct min-max normalization is used:

$$\tilde{X}_i = (X_i - X_{\min}) / (X_{\max} - X_{\min}), \quad (6)$$

for metrics to be minimized (λ), inverse normalization is used:

$$\tilde{X}_i = (X_{\max} - X_i) / (X_{\max} - X_{\min}), \quad (7)$$

where X_i – the value of the indicator for the i -th model; $X_{\min}$ and $X_{\max}$ are the minimum and maximum values of the indicator among the compared models, respectively.

The integral efficiency criterion of the i -th model E_i is defined as a weighted sum of the normalized partial indicators:

$$E_i = (w_1 \cdot mAP_i + w_2 \cdot FPS_i + w_3 \cdot M_i + w_4 \cdot E_i), \quad (8)$$

where $w_1 \dots w_4$ – the weighting coefficients reflect task priorities and satisfy the normalization condition that their sum equals one.

The weights are justified on substantive grounds. For the task of detecting explosive objects, missing an object has the highest cost, so the weight of the detection-quality indicator cannot be lower than the weights of the resource-cost indicators. Three scenarios are considered. The balanced scenario ($w = 0.40; 0.30; 0.15; 0.15$) corresponds to general-purpose requirements without a pronounced priority. The detection-completeness-priority scenario ($w = 0.55; 0.20; 0.125; 0.125$) models a careful survey of a limited area, where flight speed is reduced and the cost of a miss is maximal.

The speed-priority scenario ($w = 0.20; 0.55; 0.125; 0.125$) corresponds to a first-pass survey of large areas, where detections are subject to further verification and the decisive factor is the survey rate.

Table 4 – Performance and energy per frame rating (NCNN, 640×640)

Model	t, ms (Pi4)	FPS (Pi4)	E, J (Pi4)	t, ms (3B+)	FPS (3B+)	E, J (3B+)
YOLOv5n	235	4,26	1,41	674	1,48	2,43
YOLOv8n	435	2,30	2,61	1275	0,78	4,59
YOLOv10n	320	3,13	1,92	959	1,04	3,45
YOLOv11n	330	3,03	1,98	960	1,04	3,46
YOLOv8s	1062	0,94	6,37	3156	0,32	11,36
YOLOv11s	803	1,24	4,82	2380	0,42	8,57

Analysis of the data in Table 4 allows us to state the following. On the Raspberry Pi 4, nano-class models provide an effective speed of 2.3 - 4.3 frames per second; the fastest is YOLOv5n, the most accurate among the class is YOLOv11n at a speed of 3.0 frames per second. YOLOv10n, due to the absence of the NMS stage, reaches 3.1 frames per second - the second fastest result with an accuracy inferior to YOLOv11n by only one percentage point. The small class models on the Pi 4 operate at the limit of one frame per second, which is only suitable for frame-by-frame analysis of still scenes. On the Raspberry Pi 3B+, no model exceeds 1.5 frames per second, and small class

Comparative analysis based on benchmark data

The comparison is based on official performance measurements on the Raspberry Pi 4 platform and model characteristics from the developers' documentation [18, 19]. Table 3 presents the measured inference time of the YOLOv8n model on the Raspberry Pi 4 across the main software execution formats.

Table 3 – YOLOv8n inference time on Raspberry Pi 4 by execution format (640×640, FP32)

Format	Time, ms	Relative to NCNN
NCNN	414,7	1,00
OpenVINO	534,9	1,29
ONNX Runtime	560,0	1,35
TF SavedModel	816,5	1,97
TensorFlow Lite	950,3	2,29
PyTorch	1068,4	2,58

Data in Table 3 confirm NCNN as the primary execution format: its advantage over ONNX Runtime is 35%, and over TensorFlow Lite it is 129%. This is due to the optimization of NCNN's computational kernels for ARM NEON vector instructions and the more efficient distribution of workload across four cores without the overhead of general-purpose runtimes.

Table 4 summarizes the estimated total frame processing time, effective throughput, and energy per frame for six models on both platforms, obtained using formulas (1) – (4) calibrated against the measured values for YOLOv8n and YOLOv8s. For Raspberry Pi 3B+, slowdown coefficient of 3.0 relative to Pi 4 was applied, corresponding to the midpoint of the documented range 2.5 – 4.0 [18]; postprocessing time was increased in proportion to processor performance.

models require 2.4-3.2 s per frame; additionally, the deployment of YOLOv8s and YOLOv11s on a platform with 1 GB of memory is complicated by competition for memory with the operating system and inference stack. Energy estimates show that despite the lower power consumption, the Pi 3B+ consumes 4.59 J of 4.59 J on frame processing with the YOLOv8n model versus 2.61 J of the Pi 4, i.e. 1.76 times more. With a fixed energy budget of the onboard battery, this directly reduces the number of frames that can be processed in a single flight and makes the Pi 4 platform more energy efficient despite the higher absolute consumption.

The issue of external hardware inference accelerators, which are potentially capable of bringing performance closer to the target values, requires separate consideration. For Raspberry Pi platforms, two main solutions are available - Google Coral USB Accelerator with 4 TOPS performance in integer format and Hailo-8L modules with 13 TOPS performance. The use of both solutions as part of the on-board system based on the platforms in question was evaluated negatively for the following reasons.

The Hailo-8L module requires a PCIe interface, which is physically absent on the Raspberry Pi 3B+ and Pi 4, and is compatible only with the Raspberry Pi 5 platform through a specialized adapter, that is, it is not suitable for the target platforms of the study. Coral USB Accelerator is compatible with the USB interface, but is limited to the execution of only fully quantized integer models in TensorFlow Lite format with compilation for Edge TPU; the conversion of modern YOLO models to this format is accompanied by a degradation of accuracy in operations that are not supported by the accelerator and are rolled back to the CPU, and software support for the product by the developer is actually terminated.

Additional factors are an increase in the power consumption of the system by 2.0-2.5 W, complication of the power circuit due to peak currents of more than 900 mA, additional payload mass and the appearance of an additional failure point in the flight configuration. Given the totality of the above limitations, further analysis was performed for the platforms' own computing resources.

Calculation example of the application of the integral criterion

The application of criterion (8) is demonstrated for the calculation scenario adopted in Section 1: flight altitude 20 m, target video stream frequency 25 frames per second, Raspberry Pi 4 main platform. Normalization of partial indicators is performed according to formulas (6) and (7) on a set of six models; the initial data are the values of mAP50-95 from Table. 2, effective speed and energy per frame from Table. 4, size of models in ONNX format according to the documentation [18]. Normalized values of partial indicators and final values of the integral criterion for three scenarios of weighting factors are presented in Table. 5.

Table 5 – Normalized indicators and integral criterion according to three scenarios (Raspberry Pi 4)

Model	mAP	FPS	M	E	E (Balance)	E (Completeness)	E (Speed)
YOLOv5n	0,000	1,000	1,000	1,000	0,600	0,450	0,800
YOLOv8n	0,489	0,409	0,869	0,758	0,563	0,554	0,526
YOLOv10n	0,553	0,659	0,963	0,897	0,698	0,668	0,705
YOLOv11n	0,605	0,629	0,926	0,885	0,702	0,685	0,693
YOLOv8s	0,889	0,000	0,000	0,000	0,356	0,489	0,178
YOLOv11s	1,000	0,091	0,185	0,313	0,502	0,630	0,312

In a balanced scenario, the first two positions are occupied by YOLOv11n with an integral score of 0.702 and YOLOv10n with a score of 0.698; the difference of 0.004 is within the margin of error of the initial data, so the models are recognized as almost equivalent. In the scenario of the priority of completeness of detection, the leadership of YOLOv11n becomes more pronounced - 0.685 versus 0.668 for YOLOv10n - due to the highest accuracy among the nano class; the third position is occupied by YOLOv11s with a score of 0.630, for which the maximum accuracy partially compensates for the low performance. In the performance priority scenario, the first place is occupied by YOLOv5n with a score of 0.800, the minimum computational complexity of which outweighs the lowest accuracy; YOLOv10n with a score of 0.705 is ahead of YOLOv11n due to the absence of a non-maximal suppression stage. Small class models in all three scenarios are located at the bottom of the ranking: on the platforms in question, the increase in accuracy does not compensate for the loss of performance and energy efficiency.

The results obtained allow us to formulate practical recommendations. For further experimental research, the YOLOv11n model was chosen as the main one, which ranks first in two of the three scenarios, in particular, in the critical scenario for the problem of the priority of

completeness of detection. The YOLOv10n model is identified as an equivalent alternative, the feasibility of which increases with more stringent latency requirements; comparison of both models on a specialized dataset of images of explosive objects is the subject of the next stage of research. The YOLOv5n model retains a niche for the Raspberry Pi 3B+ platform, where its speed of 1.5 frames per second is the most achievable among the configurations considered.

A comparison of the achieved indicators with the target frequency of 25 frames per second reveals a gap of 5.8 times for the fastest configuration (YOLOv5n on the Pi 4) and 8.3 times for the recommended configuration (YOLOv11n on the Pi 4). Bridging the gap is possible in three ways, which are not mutually exclusive. Reducing the input resolution from 640×640 to 320×320 reduces the computational complexity by about four times, which is estimated (1) to increase the performance of the YOLOv11n to 10-12 frames per second at the cost of degrading detection of small objects. Thinning of the frame flow with the processing of each k-th frame allows you to match the detection rate with the flight speed: at a speed of 5 m/s and a viewing bandwidth of more than 15 m, the processing of three frames per second provides multiple overlap of the terrain without gaps. An additional reserve is the quantization of the model to an

integer format and other optimization methods that can provide an increase in performance by 1.5-2 times; the quantification of this reserve on target platforms is the subject of a separate work.

Conclusions

The problem of a reasonable choice of the architecture of the neural network detector of the YOLO family for the built-in system for detecting explosive objects on the computing platforms Raspberry Pi 3 Model B+ and Raspberry Pi 4 Model B is formulated.

A comparative analysis of six models of four generations of the YOLO family was performed according to official benchmarks with extrapolation to the research platforms. It was found that on the Raspberry Pi 4, nano-class models provide an effective speed of 2.3-4.3 frames per second in NCNN format, while small class models do not exceed 1.25 frames per second, and on the Raspberry Pi 3B+, only nano-class configurations retain practical suitability. The inexpediency of using external hardware accelerators for target platforms is substantiated: the Hailo-8L module is instructively incompatible due to the lack of a PCIe interface, and the Coral USB Accelerator is limited by an outdated software stack and creates additional energy and mass costs in the flight configuration.

Based on the results of the calculation of the integral criterion, the YOLOv11n model is recommended as the main one for further experimental research, since it ranks first in the balanced scenario and the completeness of detection priority scenario, and the YOLOv10n model as an almost equivalent alternative with the architectural advantage of the absence of the non-maximal suppression stage. The gap between the achievable speed and the target frequency of 25 frames per second by 5.8-

8.3 times is stated and the ways to overcome it are determined: reducing the resolution of the detector input, thinning the frame flow taking into account the flight speed and quantizing the model. Quantitative study of methods for optimizing models on target platforms is the direction of further work.

Author Contributions

Conceptualization, A.P. and O.S.; methodology, A.P.; validation, A.P. and O.S.; formal analysis, A.P.; investigation, A.P.; resources, O.S.; data curation, O.S.; writing-review and editing, A.P. and O.S. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Conflicts of interest

The authors declare that they have no conflicts of interest in relation to the current study, including financial, personal, authorship, or any other, that could affect the study, as well as the results reported in this paper.

Data Availability Statement

No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Use of artificial intelligence

During the preparation of this manuscript, the authors used ChatGPT (GPT-5) for the purposes of grammar checking and language refinement. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

REFERENCE

1. Crowther, E., & Shepel, N. (2024), "Phase II Testing of UAV/Sensor Systems for Humanitarian Mine Action in Ukraine", *The Journal of Conventional Weapons Destruction*, 28(3), Article 8. Available at: <https://commons.lib.jmu.edu/cisr-journal/vol28/iss3/8/>.
2. Podorozhniak, A., Liubchenko, N., Skorpupin, O., Korolenko, S., & Stas, A. (2025), "Mobile explosive object detection system for humanitarian demining needs", *2025 IEEE 6th KhPI Week on Advanced Technology (KhPIWeek 2025)*, Kharkiv, Ukraine, pp. 1–6. doi: <https://doi.org/10.1109/KHPIWEEK61436.2025.11288620>
3. Levchenko, D., Podorozhniak, A., & Liubchenko, N. (2025), "Tools and methods for explosive objects detection using artificial intelligence and computer vision", *Control, Navigation and Communication Systems*, 3(81), pp. 117–121. doi: <https://doi.org/10.26906/SUNZ.2025.3.117>
4. United Nations Mine Action Service (2026), *Ukraine*. Available at: <https://unmas.org/en/where-we-work/ukraine>
5. Humanity & Inclusion (2026), *Landmines: use, contamination and civilian harm in Ukraine*. Available at: https://www.hi.org/sn_uploads/document/HI-Factsheet-Landmine-use-in-Ukraine-2026-final-ENG-V2_1.pdf
6. Podorozhniak, A., & Skorpupin, O. (2026), "Landmine detection using robotic systems based on multispectral video imaging", *Control, Navigation and Communication Systems*, 2026, 2(84), pp. 160–164, doi: <https://doi.org/10.26906/SUNZ.2026.2.160>.
7. Vivoli, E., Bertini, M., & Capineri, L. (2024), "Deep Learning-Based Real-Time Detection of Surface Landmines Using Optical Imaging", *Remote Sensing*, 16(4), 677. doi: <https://doi.org/10.3390/rs16040677>
8. Al-Slemani, A.S.A., & Abubakr, G. (2024), "Adaptive Landmine Detection and Recognition in Complex Environments using YOLOv8 Architectures", *Journal of Smart Systems Research*, 5(2), pp. 110–120. doi: <https://doi.org/10.58769/joinssr.1542886>
9. Dodić, D., Vujović, V., Jovković, S., Milutinović, N., & Trpkoski, M. (2025), "SAHI-Tuned YOLOv5 for UAV Detection of TM-62 Anti-Tank Landmines: Small-Object, Occlusion-Robust, Real-Time Pipeline", *Computers*, 14(10), 448. doi: <https://doi.org/10.3390/computers14100448>
10. Sapkota, R., & Karkee, M. (2025), "Ultralytics YOLO Evolution: An Overview of YOLO26, YOLO11, YOLOv8, and YOLOv5 Object Detectors for Computer Vision and Pattern Recognition", *arXiv:2510.09653 [cs.CV]*. doi: <https://doi.org/10.48550/arXiv.2510.09653>
11. Wang, A., Chen, H., Liu, L., Chen, K., Lin, Z., Han, J., & Ding, G. (2024), "YOLOv10: Real-Time End-to-End Object Detection", *arXiv:2405.14458 [cs.CV]*. doi: <https://doi.org/10.48550/arXiv.2405.14458>
12. Podorozhniak, A., Liubchenko, N., Skorpupin, O., & Oleksiienko, D. (2026), "Unmanned mine detection system using an optimized YOLO26 model for edge computing on STM32", *Proceedings of 16-th International Scientific and Technical*

- Conference "Current directions in the development of information and communication technologies and control tools", 2026, vol. 5, pp. 50–52, available at: <https://repository.kpi.kharkov.ua/handle/KhPI-Press/103564>.
13. Skorlupin, O., & Podorozhniak, A. (2024), "Optical methods for detecting explosive objects using autonomous unmanned systems", *Proceedings of 12-th International Scientific and Technical Conference "Problems of informatization"*, 2024, vol. 2, p. 132, available at: <https://repository.kpi.kharkov.ua/handle/KhPI-Press/87663>.
 14. Popov, M.O. (2022), "Remote mine detection technology based on the analysis of survey materials from unmanned aerial vehicles: status and prospects", *Visnyk of the National Academy of Sciences of Ukraine*, no. 5, pp. 25–36. doi: <https://doi.org/10.15407/vsn2022.05.056>
 15. Mishchuk, V., & Podorozhniak, A. (2024), "Analysis of Trade-Offs Between Accuracy and Speed of Real-Time Object Detectors for the Tasks of Explosive Ordnance Detection", *2024 IEEE 5th KhPI Week on Advanced Technology, KhPIWeek 2024*, Kharkiv, Ukraine, pp. 1–5. doi: <https://doi.org/10.1109/KHPIWEEK61434.2024.10878035>
 16. Kuchuk, H. and Malokhvii, E. (2024), "Integration of IoT with cloud, fog, and edge computing: a review", *Advanced Information Systems*, 8(2), pp. 65–78. doi: <https://doi.org/10.20998/2522-9052.2024.2.08>
 17. Podorozhniak, A., Liubchenko, N., Petrukovich, D., & Onishchenko, D. (2023), "Implementation of automatic license plate recognition system with Raspberry Pi", *Control, Navigation and Communication Systems*, 1(71), pp. 134–137. doi: <https://doi.org/10.26906/SUNZ.2023.1.134>
 18. Ultralytics Inc. (2024), "Raspberry Pi with Ultralytics YOLO: Benchmarks and Setup Guide", available at: <https://docs.ultralytics.com/guides/raspberry-pi/>
 19. Ultralytics Inc. (2026), "Ultralytics YOLO Documentation", available at: <https://docs.ultralytics.com/>
 20. Hashimov, E., Pashayev, A., & Khaligov, G. (2025), "Camera control algorithm and image quality assessment method to obtain a quality image", *Advanced Information Systems*, 9(3), pp. 50–56, 2025, doi: <https://doi.org/10.20998/2522-9052.2025.3.06>
 21. Podorozhniak, A., Liubchenko, N., Sobol, M., & Onishchenko, D. (2023), "Usage of Mask R-CNN for automatic license plate recognition", *Advanced Information Systems*, 7(1), pp. 54–58. doi: <https://doi.org/10.20998/2522-9052.2023.1.09>
 22. Bajić, M., & Potočnik, B. (2023), "UAV Thermal Imaging for Unexploded Ordnance Detection by Using Deep Learning", *Remote Sensing*, 15(4), 967. doi: <https://doi.org/10.3390/rs15040967>
 23. Khanam, R., & Hussain, M. (2024), "YOLOv11: An Overview of the Key Architectural Enhancements", *arXiv:2410.17725 [cs.CV]*. doi: <https://doi.org/10.48550/arXiv.2410.17725>

Received (Надійшла) 22.06.2026

Accepted for publication (Прийнята до друку) 09.09.2026

Publication date (Дата публікації) 25.09.2026

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Скорлупін Олександр Володимирович – аспірант кафедри комп'ютерної інженерії та програмування, Національний технічний університет "Харківський політехнічний інститут", Харків, Україна;

Oleksandr Skorlupin – postgraduate student, Computer Engineering and Programming Department, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;

e-mail: oleksandr.skorlupin@cit.khpi.edu.ua; ORCID Author ID: <https://orcid.org/0009-0001-4295-4115>;

Scopus ID: : <https://www.scopus.com/authid/detail.uri?authorId=60435005700>.

Подорожняк Андрій Олексійович – кандидат технічних наук, доцент, професор кафедри комп'ютерної інженерії та програмування, Національний технічний університет "Харківський політехнічний інститут", Харків, Україна;

Andrii Podorozhniak – Candidate of Technical Sciences, Associate Professor, Professor of Computer Engineering and Programming Department, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;

e-mail: andrii.podorozhniak@khpi.edu.ua; ORCID Author ID: <https://orcid.org/0000-0002-6688-8407>;

Scopus ID: <https://www.scopus.com/authid/detail.uri?authorId=57202229410>.

Архітектури сімейства YOLO для виявлення вибухонебезпечних об'єктів у вбудованих системах автономних безпілотних засобів: порівняльний аналіз

О. В. Скорлупін, А. О. Подорожняк

Анотація. Актуальність. Необхідність вибору ефективної архітектури YOLO для виявлення вибухонебезпечних об'єктів у вбудованих системах безпілотних літальних апаратів зумовлена обмеженими обчислювальними ресурсами, вимогами до роботи в реальному часі та енергоефективності. **Об'єктом дослідження** є виявлення вибухонебезпечних об'єктів у вбудованих автономних безпілотних системах із використанням архітектур сімейства YOLO. **Предметом дослідження** є порівняльна ефективність моделей YOLO за точністю виявлення, пропускну здатністю, розміром моделі та енергоспоживанням на вбудованих апаратних платформах. **Метою цієї статті** є розроблення та застосування інтегрального критерію ефективності, заснованого на неймережевих методах, для вибору найпридатнішої архітектури YOLO для виявлення вибухонебезпечних об'єктів і завдань гуманітарного розмінування. **Результати дослідження.** Запропонований критерій було застосовано до шести моделей YOLO (v5n, v8n, v8s, v10n, v11n, v11s), розгорнутих на Raspberry Pi 3 Model B+ та Raspberry Pi 4 Model B з 4 ГБ оперативної пам'яті. Порівняльне оцінювання за трьома сценаріями показало суттєві відмінності в якості виявлення, пропускій здатності, розмірі моделі та енергоспоживанні. Отримане ранжування підтверджує практичну цінність методики для вибору моделі у вбудованих системах. **Висновки.** Методика забезпечує прозору й відтворювану основу для порівняння архітектур YOLO в ресурсно обмежених середовищах. Робота в реальному часі залишається складним завданням на малопотужних платформах, що зумовлює подальшу оптимізацію розгортання та інференсу.

Ключові слова: YOLO; виявлення вибухонебезпечних об'єктів; безпілотний літальний апарат; вбудовані системи; інтегральний критерій ефективності; нейронна мережа; гуманітарне розмінування; Raspberry Pi.

А. В. Шостак

Національний аерокосмічний університет «Харківський авіаційний інститут», Харків, Україна

ЕКСПЕРИМЕНТАЛЬНЕ ДОСЛІДЖЕННЯ МЕТОДІВ ПОШУКУ В АВЛ-ДЕРЕВІ З УРАХУВАННЯМ ІЄРАРХІЇ ПАМ'ЯТІ

Анотація. Актуальність. Дослідження впливу просторової локальності даних, передбачення розгалужень та програмного попереднього завантаження даних на продуктивність пошуку становить значний інтерес при розробці високопродуктивних структур даних. **Предмет дослідження:** методи пошуку в АВЛ-дереві, представленому у вигляді масиву з використанням розміщення Ейтцингера, та вплив ієрархії пам'яті процесора на їхню швидкість. **Мета дослідження:** експериментально оцінити ефективність різних методів пошуку в АВЛ-дереві з поданням на масивах залежно від розміру масиву та положення робочого набору даних щодо рівнів ієрархії пам'яті процесора. **Завдання дослідження.** Реалізувати та експериментально дослідити алгоритми пошуку Binary, AVL, Eytzinger, EytzingerB та EytzingerBP; визначити діапазони розмірів масивів, які відповідають різним рівням ієрархії пам'яті; порівняти швидкість досліджуваних алгоритмів; встановити вплив просторової локальності даних, помилок передбачення розгалужень та програмного попереднього завантаження даних на ефективність пошуку. **Методи дослідження.** Використані методи алгоритмічного аналізу, експериментального дослідження продуктивності програмного забезпечення та статистичної обробки результатів вимірів. Експериментальна оцінка виконана засобами BenchmarkDotNet на платформі .NET 8 з використанням апаратних лічильників продуктивності процесора. **Результати дослідження.** Встановлено, що ефективність алгоритмів пошуку визначається як їх обчислювальна складність, а й особливостями взаємодії з ієрархією пам'яті процесора. Показано, що при розміщенні робочих даних у кешах L1/L2 найменший час пошуку забезпечує алгоритм EytzingerB, тоді як при переході до даних, розміщених у кешах L3 та оперативної пам'яті, найкращі результати демонструє алгоритм EytzingerBP, який використовує програмне попереднє завантаження даних. Отримано експериментальні залежності часу пошуку від розміру масиву, визначено галузі ефективного застосування досліджуваних алгоритмів та кількісно оцінено вираз найефективніших методів пошуку. Експериментально встановлено зміна найефективнішого алгоритму пошуку при переході робочих даних між рівнями ієрархії пам'яті процесора: у діапазоні L1/L2 мінімальний час пошуку забезпечує алгоритм EytzingerB, тоді як у діапазоні L3 та оперативної пам'яті лідирує алгоритм EytzingerBP. **Висновки.** Подання АВЛ-дерева в масиві Ейтцингера забезпечує істотне підвищення швидкості пошуку в порівнянні з класичною реалізацією дерева на покажчиках. Експериментально підтверджено, що вибір оптимального методу пошуку визначається рівнем ієрархії пам'яті, в якому знаходиться робочий набір даних. Отримані результати можуть бути використані для проектування високопродуктивних пошукових структур даних та оптимізації алгоритмів пошуку для сучасних багаторівневих підсистем пам'яті.

Ключові слова: АВЛ-дерево, розміщення Ейтцингера, алгоритми пошуку, програмне попереднє завантаження даних, кеш-пам'ять, локальність даних, BenchmarkDotNet, експериментальна оцінка продуктивності.

1. Вступ

1.1 Актуальність тематики

Пошук елементів у впорядкованих структурах даних є однією з фундаментальних операцій сучасної інформатики та широко використовується при побудові систем керування базами даних, інформаційно-пошукових систем, компіляторів, файлових систем, систем підтримки прийняття рішень та інших програмних комплексів [1–3]. Ефективність виконання операцій пошуку багато в чому визначає загальну продуктивність подібних систем, що зумовлює постійний інтерес до розробки та вдосконалення алгоритмів пошуку.

Класичні методи аналізу алгоритмів оцінюють ефективність пошуку переважно за асимптотичною обчислювальною складністю. Для бінарного пошуку, пошуку в АВЛ-дереві та інших збалансованих структур даних складність пошуку становить $O(\log n)$ [3, 4]. Проте розвиток архітектури сучасних процесорів призвело до того, що однакова асимптотична складність алгоритмів не гарантує однакового швидкості за її практичної реалізації.

Сучасні обчислювальні системи використовують багаторівневу ієрархію пам'яті, що включає кеші першого (L1), другого (L2) та третього (L3) рівнів, а

також оперативну пам'ять. Час звернення до різних рівнів пам'яті відрізняється на порядки, внаслідок чого ефективність алгоритму все більшою мірою визначається просторовою та тимчасовою локальністю доступу до даних, кількістю промахів кешу, ефективністю апаратного прогнозування розгалужень та можливістю попереднього завантаження даних. За таких умов дослідження взаємодії алгоритмів пошуку з підсистемою пам'яті стає не менш важливим, ніж аналіз їх обчислювальної складності.

Одним із перспективних напрямів підвищення продуктивності пошуку є представлення дерев пошуку у вигляді масиву з використанням розміщення Ейтцингера. Таке уявлення забезпечує послідовне розміщення вузлів дерева в пам'яті, покращує просторову локальність доступу до даних і дозволяє ефективніше використовувати механізми кешування та апаратної попередньої вибірки процесора. Додаткове підвищення швидкості може бути досягнуто за рахунок застосування алгоритмів пошуку без умовних переходів та програмного попереднього завантаження даних.

У зв'язку з цим експериментальне дослідження впливу різних способів організації пошуку у масивному уявленні АВЛ-дерева на сучасних процесорах є актуальним науково-практичним завданням.

1.2 Аналіз літературних джерел і визначення проблеми

Теоретичні основи алгоритмів пошуку та збалансованих дерев докладно викладені у фундаментальних роботах [1-4], де розглядаються властивості бінарного пошуку, бінарних дерев пошуку та AVL-дерев, аналізується їх обчислювальна складність та доводяться основні оцінки ефективності.

Сучасні підходи до реалізації AVL-дерев та особливостей їх практичного застосування представлені також у роботі [5]. Більшість класичних досліджень розглядає ефективність алгоритмів переважно з позицій кількості операцій, що виконуються, припускаючи однакову вартість звернення до пам'яті. Проте це припущення не відповідає архітектурі сучасних обчислювальних систем, у яких час доступу до різних рівнів пам'яті істотно відрізняється. Вплив архітектури пам'яті на продуктивність програм детально розглядається у працях [6, 7], де показано, що локальність доступу до даних та організація кеш-пам'яті мають визначальний вплив на практичну продуктивність алгоритмів. Аналогічні питання оптимізації програм для сучасних процесорів детально висвітлено в посібниках Intel [8–10], а також у працях [11–14], присвячених особливостям мікроархітектур сучасних процесорів, механізмів передбачення розгалужень та оптимізації доступу до пам'яті. Одним із найважливіших досліджень, присвячених впливу розміщення даних на ефективність пошуку, є робота [15]. Автори показали, що представлення бінарного дерева пошуку в масиві з використанням розміщення Ейтцингера дозволяє значно підвищити продуктивність пошуку порівняно з традиційним бінарним пошуком завдяки поліпшенню просторової локальності даних та використанню програмного попереднього завантаження. Однак отримані результати стосуються переважно ідеального бінарного дерева пошуку і не враховують особливості застосування цих підходів до AVL-дерев. У попередніх дослідженнях автора проведено порівняльний аналіз класичної реалізації AVL-дерев на показниках та масивного представлення AVL-дерев [16]. Отримані результати підтвердили істотну перевагу масивного представлення за швидкістю пошуку.

Незважаючи на значну кількість робіт у цій галузі, питання порівняльної ефективності різних методів пошуку в масивному представленні AVL-дерев при зміні положення робочого набору даних відносно рівнів ієрархії пам'яті процесора залишається недостатньо дослідженим. Практично відсутні експериментальні дослідження, що дозволяють визначити діапазони ефективного застосування класичного пошуку Eytzinger, його branchless-модифікації та алгоритмів із програмним попереднім завантаженням даних. Розв'язання цієї задачі представляє інтерес як з теоретичної точки зору, так і для практичного проектування високопродуктивних структур даних.

1.3 Мета і завдання дослідження

Метою дослідження є експериментальна оцінка ефективності методів пошуку в AVL-дереві,

представленому у вигляді масиву з використанням розміщення Ейтцингера, з урахуванням впливу ієрархії пам'яті сучасних процесорів. Для досягнення поставленої мети було сформульовано такі завдання дослідження:

1. Реалізувати методи пошуку Binary, AVL, Eytzinger, Eytzinger Branchless (EytzingerB) та Eytzinger Prefetch (EytzingerBP) у єдиному програмному середовищі.

2. Виконати експериментальне дослідження швидкодії зазначених методів для різних розмірів масивів, що відповідають розміщенню робочих даних у кешах L1, L2, L3 та оперативної пам'яті.

3. Провести порівняльний аналіз часу пошуку у сценаріях успішного (Hit) та неуспішного (Miss) пошуку.

4. Дослідити вплив просторової локальності даних, кількості промахів кешу та помилок передбачення розгалужень на продуктивність аналізованих алгоритмів пошуку з використанням апаратних лічильників продуктивності процесора.

5. Визначити сфери ефективного застосування досліджуваних методів пошуку в залежності від положення робочого набору даних щодо рівнів ієрархії пам'яті процесора.

2. Матеріали і методи дослідження

Об'єктом дослідження було AVL-дерево, представлене у вигляді одновимірного масиву з використанням розміщення Ейтцингера. На відміну від класичної реалізації AVL-дерев на основі динамічно пов'язаних вузлів, при такому представленні вершини дерева розташовуються послідовно в масиві у порядку обходу за рівнями. Для вершини з індексом i лівий нащадок розташовується за індексом $(2i+1)$, а правий — за індексом $(2i+2)$. Таке розміщення забезпечує високу просторову локальність даних і дозволяє ефективно використовувати механізми кешування сучасних процесорів.

Для формування досліджуваних структур використовувалося канонічне AVL-дерево, побудоване з попередньо відсортованого масиву унікальних цілих чисел. Побудова здійснювалася шляхом рекурсивного вибору медіанного елемента кожного підмасиву, що забезпечувало мінімально можливу висоту дерева та відповідало класичній структурі збалансованого бінарного дерева пошуку. Після побудови дерева виконувалося його перетворення в масив Ейтцингера шляхом обходу по рівнях.

У роботі досліджено п'ять методів пошуку:

- Binary – класичний бінарний пошук у відсортованому масиві [1, 2];

- AVL – пошук у класичному AVL-дереві, реалізованому на показниках [3–5];

- Eytzinger – пошук у масивному представленні AVL-дерев з розміщенням Ейтцингера [15];

- EytzingerB – модифікація пошуку Eytzinger, що мінімізує кількість умовних переходів за рахунок використання арифметичних операцій під час обчислення напрямку переходу [17–19];

- EytzingerBP – модифікація методу EytzingerB, що використовує програмне попереднє завантаження

даних для зменшення латентності звернень до пам'яті [20, 21].

Усі алгоритми було реалізовано мовою C# із використанням платформи .NET 8. Експериментальні дослідження проводилися за допомогою бібліотеки BenchmarkDotNet v0.15.8, яка забезпечує автоматичну організацію вимірювань, попереднє прогрівання середовища виконання, багаторазове повторення експериментів та статистичну обробку результатів.

Вимірювання проводилися на персональному комп'ютері з такими характеристиками: процесор Intel Core i5-12400; оперативна пам'ять – 32 ГБ; операційна система Microsoft Windows 10 x64.

Для оцінки впливу ієрархії пам'яті використовувалися масиви 32-бітних ключів різних розмірів, що дозволяли послідовно переносити робочий набір даних із кеш-пам'яті першого рівня (L1) у кеші L2 та L3, а потім в оперативну пам'ять. Розміри досліджуваних масивів обиралися з урахуванням характеристик кеш-пам'яті використовуваного процесора та наведені в табл. 1 [7, 8, 13]. Для кожного розміру масиву виконувалися два сценарії пошуку: - Hit – пошук елемента, який є в дереві; - Miss – пошук елемента, якого немає.

Таблиця 1 – Відповідність розмірів досліджуваних масивів рівням ієрархії пам'яті

№	Розмірність масиву	Куди поміщається
1	500	L1
2	16000	L1/L2
3	500000	L1/L2/L3
4	2000000	L1/L2/L3
5	8000000	L1/L2/L3/RAM
6	16000000	L1/L2/L3/RAM
7	32000000	L1/L2/L3/RAM
8	64000000	L1/L2/L3/RAM

Для підвищення статистичної достовірності використовували 10 ітерацій прогрівання середовища виконання та 30 вимірювальних ітерацій BenchmarkDotNet. Підсумкові результати подавалися у вигляді середнього часу виконання операції пошуку, стандартного відхилення та довірчого інтервалу. Окрім часу пошуку, аналізувалися апаратні характеристики виконання алгоритмів із використанням апаратних лічильників продуктивності процесора. Для кожного алгоритму реєструвалися такі показники: – загальна кількість процесорних циклів на одну операцію пошуку; – кількість промахів кеш-пам'яті на одну операцію (CacheMisses/Op); – кількість помилок передбачення розгалужень на одну операцію (BranchMispredictions/Op). Використання апаратних лічильників дозволило зіставити експериментально вимірний час пошуку з особливостями роботи підсистеми пам'яті та механізму передбачення розгалужень процесора, що дало змогу пояснити спостережувані відмінності в продуктивності досліджуваних методів.

Порівняння методів здійснювалося за середнім часом виконання операції пошуку, а також за

відносним прискоренням порівняно з класичним бінарним пошуком і класичною реалізацією АВЛ-дерева на покажчиках. Додатково аналізувалася зміна ефективності алгоритмів при збільшенні розміру масиву та перенесенні робочих даних між рівнями ієрархії пам'яті процесора.

У роботі перевірялося припущення про те, що різні методи пошуку в масивному представленні АВЛ-дерева мають різну чутливість до характеристик ієрархії пам'яті процесора. Внаслідок цього при зміні обсягу робочого набору даних найбільш ефективний метод пошуку може змінюватися. Визначення областей ефективного застосування кожного з досліджуваних методів було одним із основних завдань даного дослідження.

3. Результати дослідження

Для оцінки ефективності досліджуваних методів пошуку було проведено серію обчислювальних експериментів для масивів різного обсягу. Для кожного розміру масиву вимірювався середній час пошуку існуючого елемента (сценарій Hit) та відсутнього елемента (сценарій Miss).

На рис. 1, 2 представлено залежність середнього часу пошуку від розміру масиву в логарифмічному масштабі для досліджуваних алгоритмів. У табл. 2, 3 наведено чисельні результати експериментів.

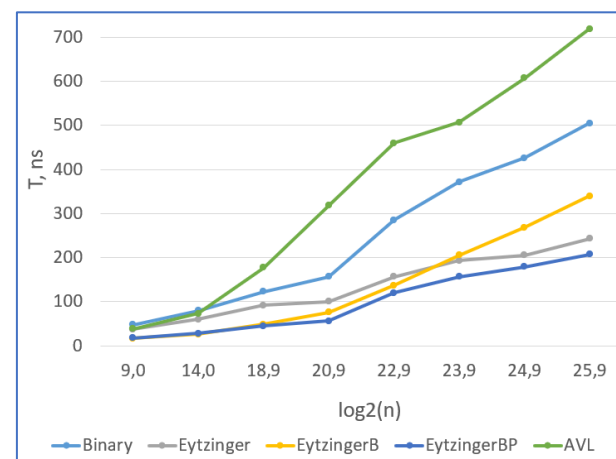


Рис. 1. Залежність часу пошуку T від розміру масиву n (сценарій Hit)

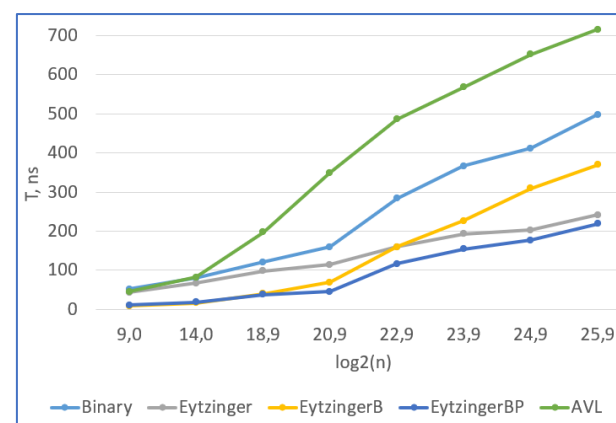


Рис. 2. Залежність часу пошуку T від розміру масиву n (сценарій Miss)

На рис. 3 наведено залежність часу пошуку досліджуваних алгоритмів пошуку відносно часу пошуку алгоритму Binary від розмірності масивів у логарифмічному масштабі для сценарію Hit.

У табл. 2 наведено мінімальний час пошуку для сценаріїв Hit і Miss серед усіх розглянутих алгоритмів пошуку.

Для масивів, що повністю розміщуються в кешах L1 і L2, мінімальний час пошуку забезпечує алгоритм EytzingerB (рис. 1–3, табл. 2). Перевага алгоритму EytzingerB для розмірів $n=500$, 16000 пов'язана з тим, що в цьому випадку масив з даними вже повністю міститься в кешах L1 або L2, вартість доступу до пам'яті є незначною, тому програмне попереднє завантаження не встигає принести користь і лише збільшує кількість виконуваних інструкцій.

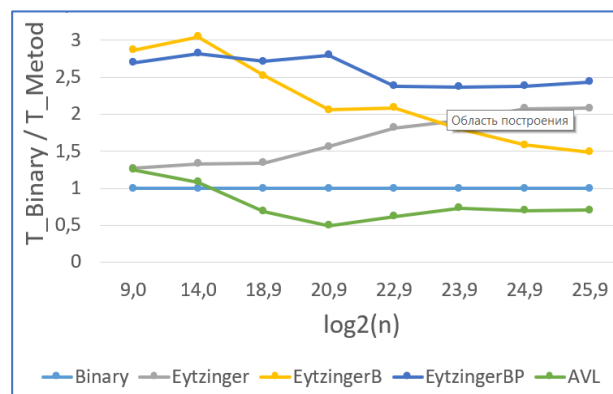


Рис. 3. Відносна залежність часу пошуку (Binary=1) від розміру масиву n

Таблиця 2 – Мінімальний час пошуку T залежно від розміру масивів n

n	Метод	Сценарій	T, ns
500	EytzingerB	Hit	16,684
		Miss	8,367
16000	EytzingerB	Hit	25,998
		Miss	16,024
500000	EytzingerBP	Hit	45,14
		Miss	37,25
2000000	EytzingerBP	Hit	55,97
		Miss	45,38
8000000	EytzingerBP	Hit	119,70
		Miss	116,70
16000000	EytzingerBP	Hit	156,80
		Miss	154,40
32000000	EytzingerBP	Hit	178,80
		Miss	177,10
64000000	EytzingerBP	Hit	207,20
		Miss	219,20

У табл. 3 наведено фрагмент результатів дослідження алгоритмів пошуку. Стовець 1 таблиці 3 відповідає алгоритму пошуку, стовець 2 – розміру масиву n, стовець 3 – усередненому часу пошуку T (ns), стовець 4 – значенню апаратного лічильника кількості помилок передбачення розгалуження на операцію, стовець 5 – значенню апаратного лічильника промахів кешу на операцію.

Таблиця 3 – Витяг із результатів дослідження алгоритмів пошуку

1	2	3	4	5
Binary_Hit	2000000	156,56	11	0
Binary_Miss	2000000	159,93	12	0
AVL_Hit	2000000	317,82	10	17
AVL_Miss	2000000	349,01	11	19
Eytzinger_Hit	2000000	100,23	10	0
Eytzinger_Miss	2000000	113,63	11	0
EytzingerB_Hit	2000000	75,99	1	0
EytzingerB_Miss	2000000	69,29	0	0
EytzingerBP_Hit	2000000	55,97	1	0
EytzingerBP_Miss	2000000	45,38	0	0
Binary_Hit	8000000	285,08	12	33
Binary_Miss	8000000	283,58	12	33
AVL_Hit	8000000	460,07	11	26
AVL_Miss	8000000	486,22	12	28
Eytzinger_Hit	8000000	156,94	11	5
Eytzinger_Miss	8000000	159,94	12	5
EytzingerB_Hit	8000000	136,36	1	1
EytzingerB_Miss	8000000	159,62	0	2
EytzingerBP_Hit	8000000	119,70	5	1
EytzingerBP_Miss	8000000	116,70	6	0
Binary_Hit	16000000	371,13	12	65
Binary_Miss	16000000	365,93	12	65
AVL_Hit	16000000	507,00	12	30
AVL_Miss	16000000	568,13	13	33
Eytzinger_Hit	16000000	193,52	11	11
Eytzinger_Miss	16000000	192,70	12	11
EytzingerB_Hit	16000000	205,53	1	2
EytzingerB_Miss	16000000	225,79	0	3
EytzingerBP_Hit	16000000	156,80	7	1
EytzingerBP_Miss	16000000	154,40	10	0

Для масивів, що розміщуються в кешах L1/L2/L3 та в L1/L2/L3/RAM, мінімальний час пошуку забезпечує алгоритм EytzingerBP (рис. 1–3, табл. 2, 3). При переході до масивів, розміри яких перевищують обсяг кеш-пам'яті L2 та L3, вартість звернення до пам'яті суттєво зростає. У цих умовах алгоритм EytzingerBP дозволяє заздалегідь ініціювати завантаження необхідних кеш-ліній, завдяки чому затримка доступу частково перекривається обчисленнями, що призводить до зменшення загального часу пошуку.

Значне скорочення часу пошуку в діапазонах пам'яті L1/L2/L3 та L1/L2/L3/RAM алгоритму EytzingerBP порівняно з іншими алгоритмами забезпечується програмним попереднім завантаженням даних на 4–5 рівнів дерева вперед відносно поточного вузла. Таке попереднє завантаження даних у зазначеному діапазоні пам'яті мінімізує час доступу до пам'яті. Ефективність такого завантаження при реалізації алгоритму EytzingerBP підтверджується практично повною відсутністю промахів кешу (стовпець 5 табл. 3) та відсутністю або незначною кількістю помилок передбачення розгалужень (стовпець 4 табл. 3).

У всіх досліджених діапазонах розмірів масивів класичні алгоритми AVL і Binary демонструють найбільший час виконання серед розглянутих методів.

Для виявлення причин відмінностей у продуктивності досліджуваних методів поряд із часом виконання пошуку були використані апаратні лічильники продуктивності процесора.

На рис. 4 представлено залежність середньої кількості промахів кеш-пам'яті від розміру масиву для сценарію Hit для досліджуваних методів пошуку.

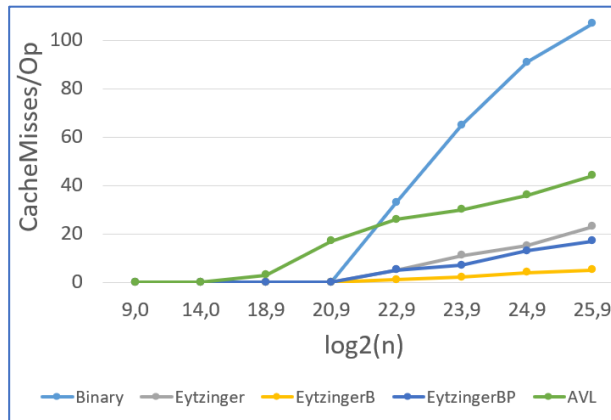


Рис. 4. Залежність середньої кількості промахів кеш-пам'яті від розміру масиву n

Аналіз отриманих результатів показує, що зі збільшенням обсягу масиву кількість промахів кеш-пам'яті закономірно зростає для всіх досліджуваних методів. Найбільш інтенсивне зростання спостерігається після виходу робочого набору даних за межі кеш-пам'яті верхніх рівнів, що відповідає переходу до обслуговування частини звернень із кеш-пам'яті L3 та оперативної пам'яті.

Для алгоритмів AVL і Binary характерна найбільша кількість промахів кеш-пам'яті практично в усьому дослідженому діапазоні розмірів масивів. Для алгоритму Binary це пояснюється особливостями послідовності звернень до елементів масиву, за якої сусідні за часом звернення відповідають значно віддаленим адресам пам'яті. У результаті просторова локальність використовується недостатньо ефективно, а ймовірність потрапляння потрібних даних у кеш виявляється нижчою.

Розміщення вузлів дерева за схемою Ейтцингера забезпечує більш компактне розташування елементів, що належать до верхніх рівнів дерева, завдяки чому зменшується кількість промахів кеш-пам'яті (рис. 4). Цей ефект найбільш помітний під час обробки масивів великого обсягу, коли вартість звернення до пам'яті починає суттєво перевищувати вартість виконання арифметичних операцій та умовних переходів.

На рис. 5 представлені залежності середньої кількості помилок передбачення розгалужень від розміру масиву для сценарію Hit для досліджуваних методів пошуку (графіки для алгоритмів Eytzinger і EytzingerB збігаються).

Для алгоритмів EytzingerB та EytzingerBP кількість помилкових прогнозів розгалужень зменшується практично на порядок порівняно з іншими реалізаціями пошуку (рис. 5). Отримані результати підтверджують, що усунення умовних

переходів дозволяє істотно знизити вплив механізму прогнозування розгалужень процесора. Разом з тим зменшення кількості помилкових прогнозів не завжди супроводжується пропорційним скороченням часу пошуку, оскільки під час роботи з великими масивами визначальним фактором стає латентність доступу до пам'яті.

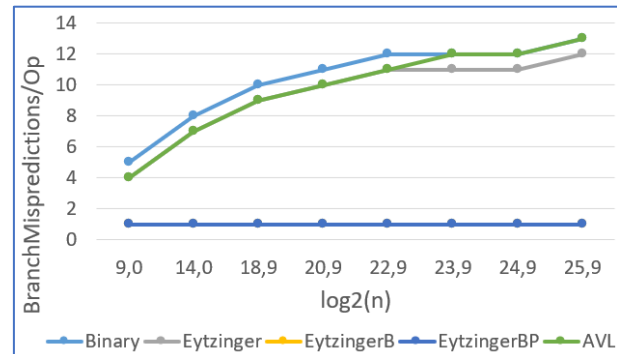


Рис. 5. Залежність середньої кількості помилок прогнозування розгалужень від розміру масиву n

Дослідження методу EytzingerBP із програмним попереднім завантаженням показує, що збільшення кількості звернень до пам'яті, спричинене виконанням інструкцій попереднього завантаження, не призводить до погіршення продуктивності. Навпаки, під час обробки великих масивів спостерігається скорочення загального часу пошуку. Це свідчить про те, що частина затримок звернень до пам'яті успішно перекривається виконанням обчислень, а необхідні дані виявляються завантаженими в кеш до моменту їх використання.

Таким чином, аналіз апаратних лічильників підтверджує, що відмінності в продуктивності досліджуваних методів зумовлені не лише їхньою алгоритмічною організацією, а й особливостями взаємодії з ієрархією пам'яті сучасного процесора. Отримані результати слугують експериментальним підтвердженням того, що просторова локальність розміщення даних, зменшення кількості умовних переходів та використання програмного попереднього завантаження по-різному впливають на продуктивність пошуку залежно від розміру робочого набору даних.

4. Обговорення результатів

Отримані результати підтверджують висунуту в роботі гіпотезу про те, що ступінь впливу різних мікроархітектурних факторів змінюється під час переходу між рівнями ієрархії пам'яті процесора. Відповідно, ефективність конкретного алгоритму визначається не лише його структурою, а й обсягом робочого набору даних відносно розмірів кеш-пам'яті.

Першим чинником, що визначає відмінності в продуктивності, є просторова локальність розміщення даних. У бінарному пошуку послідовні звернення здійснюються до елементів масиву, розташованих на значній відстані один від одного. У міру збільшення розміру масиву це призводить до

зростання кількості промахів кеш-пам'яті та збільшення часу доступу до даних.

Розміщення елементів за схемою Ейтцингера суттєво змінює характер звернень до пам'яті. Вузли верхніх рівнів дерева розташовуються в пам'яті послідовно, завдяки чому елементи, до яких найчастіше звертаються, ефективно розміщуються в кеш-пам'яті процесора. Результати експериментів показують, що саме цей ефект забезпечує стійку перевагу алгоритму Eytzinger порівняно з бінарним пошуком практично на всьому досліджуваному діапазоні розмірів масивів (рис. 1–3).

Другим фактором є вплив механізму передбачення розгалужень. Використання branchless-реалізації практично усуває помилки передбачувача переходів, що підтверджується вимірами апаратного лічильника BranchMispredictions/Op. Якщо для класичного пошуку кількість помилкових передбачень сягає 11–13 на одну операцію пошуку, то для branchless-варіантів цей показник зменшується до одного або практично до нуля (рис. 5).

Однак проведені експерименти показують, що зменшення кількості помилкових прогнозів розгалужень саме по собі не гарантує максимальної продуктивності. При роботі з масивами невеликого обсягу виграш від виключення умовних переходів виявляється обмеженим, оскільки вартість звернень до пам'яті невелика, а основний час визначається виконанням обчислювальних операцій. При подальшому збільшенні обсягу масиву вплив латентності пам'яті стає значно суттєвішим, ніж вартість помилкового передбачення переходу.

Третім фактором є можливість перекриття латентності пам'яті за допомогою програмного попереднього завантаження даних. Експериментальні результати показують, що використання інструкцій попереднього завантаження не забезпечує переваг при розміщенні даних у кеш-пам'яті верхніх рівнів, а в ряді випадків призводить до невеликого збільшення часу пошуку внаслідок появи додаткових інструкцій (табл. 2).

Зовсім інша ситуація спостерігається під час обробки масивів, розміри яких значно перевищують обсяг кеш-пам'яті процесора. У цьому випадку програмне попереднє завантаження дозволяє розпочати перенесення даних у кеш до моменту їх безпосереднього використання. Завдяки перекриттю обчислень з операціями звернення до пам'яті зменшується ефективна латентність завантаження даних, що призводить до скорочення загального часу виконання пошуку.

Для алгоритмів без програмного попереднього завантаження відмінності між сценаріями Hit і Miss виявляються відносно невеликими й визначаються головним чином довжиною пройденого шляху пошуку. Однак для алгоритмів, що використовують програмне попереднє завантаження даних, відмінності стають більш вираженими. Ефективність попереднього завантаження залежить від того, чи будуть попередньо завантажені дані дійсно використані під час наступних звернень. У разі

успішного пошуку частина заздалегідь завантажених вузлів може виявитися непотрібною через дострокове завершення алгоритму, тоді як у разі невдалого пошуку практично весь шлях пошуку проходить до листового рівня, що дозволяє ефективніше використовувати результати попереднього завантаження (рис. 1–3).

Отримані в ході роботи експериментальні дані підтверджують цю закономірність. Для невеликих масивів відмінності між сценаріями Hit і Miss є незначними, оскільки затримки звернення до пам'яті невеликі й практично повністю компенсуються кеш-пам'яттю процесора. У міру збільшення обсягу даних вплив латентності пам'яті зростає, внаслідок чого ефективність різних стратегій організації пошуку починає істотно залежати від характеру пошукового запиту.

Цю обставину необхідно враховувати під час вибору алгоритму для прикладних систем, у яких співвідношення успішних і неуспішних пошукових запитів може істотно відрізнятися.

Проведені експерименти показують, що кількісна величина виграшу визначається характеристиками конкретної апаратної платформи, зокрема розмірами кеш-пам'яті, латентністю різних рівнів пам'яті та особливостями реалізації апаратного попереднього відбору процесора.

Таким чином, результати дослідження свідчать про те, що вибір найбільш ефективного методу пошуку повинен здійснюватися з урахуванням не тільки алгоритмічних характеристик, але й особливостей архітектури пам'яті обчислювальної системи.

Для сучасних процесорів визначальними факторами продуктивності стають локальність розміщення даних, ефективність використання кеш-пам'яті, зменшення кількості помилкових передбачень розгалужень та можливість приховування латентності пам'яті за допомогою програмного попереднього завантаження даних.

5. Висновки

У роботі проведено експериментальне дослідження методів пошуку в AVL-дереві, представленому у масиві з використанням розміщення Ейтцингера, з урахуванням особливостей ієрархії пам'яті процесора. Отримані результати дають змогу зробити такі висновки.

1. Експериментально підтверджено, що при однаковій асимптотичній складності $O(\log n)$ продуктивність алгоритмів пошуку суттєво визначається особливостями організації пам'яті процесора. У міру збільшення обсягу масиву вплив латентності пам'яті стає визначальним фактором швидкодії, тоді як вплив обчислювальної складності алгоритму поступово зменшується.

2. Показано, що розміщення вузлів AVL-дерева в масиві за схемою Ейтцингера забезпечує більш ефективне використання просторової локальності даних порівняно з бінарним пошуком і представленням дерева на покажчиках, що призводить до стійкого скорочення часу пошуку

практично в усьому дослідженому діапазоні розмірів масивів.

3. Встановлено, що використання алгоритму пошуку без розгалужень EytzingerB практично усуває помилки прогнозування розгалужень, що підтверджується вимірами апаратного лічильника BranchMispredictions/Op. Однак експеримент показав, що зменшення кількості помилкових прогнозів саме по собі не гарантує максимальної швидкодії. При збільшенні обсягу даних і виходу робочого набору за межі кеш-пам'яті визначальним фактором стає латентність доступу до пам'яті.

4. Експериментально підтверджено, що застосування програмного попереднього завантаження даних дозволяє додатково зменшити час пошуку під час обробки великих масивів. Найбільш помітний ефект досягається при розміщенні робочого набору даних у кеш-пам'яті третього рівня та оперативній пам'яті, де програмне попереднє завантаження дозволяє частково приховати затримки звернення до пам'яті шляхом перекриття обчислень із завантаженням даних.

5. Аналіз апаратних лічильників CacheMisses/Op та BranchMispredictions/Op показав високий ступінь узгодженості з вимірними часовими характеристиками алгоритмів. Отримані результати підтверджують, що відмінності у швидкодії зумовлені насамперед особливостями взаємодії алгоритмів із підсистемою пам'яті процесора, а не лише відмінностями у кількості виконуваних операцій.

6. При розміщенні даних у кеш-пам'яті першого та другого рівнів найкращі результати демонструє алгоритм EytzingerB, який поєднує високу просторову локальність із практично повним усуненням помилок передбачення розгалужень. При збільшенні обсягу масиву та переході робочого набору даних в область кеш-пам'яті третього рівня, а потім оперативної пам'яті перевага переходить до алгоритму EytzingerBP, що забезпечує зменшення латентності доступу до пам'яті за рахунок програмного попереднього завантаження даних.

7. У порівнянні з бінарним пошуком розроблені методи пошуку в розміщенні Ейтцингера забезпечують скорочення часу до 3,05 разів (рис. 3, табл. 3). У порівнянні з традиційною реалізацією АВЛ-дерева на показниках виграш виявляється ще більш значним (до 5,67 разів) завдяки усуненню

великої кількості нерегулярних звернень до пам'яті та поліпшенню просторової локальності розміщення даних.

Таким чином, основним результатом дослідження є експериментальне підтвердження того, що ефективність пошуку в деревах визначається не лише алгоритмічною складністю $O(\log n)$, а й характером взаємодії алгоритму з ієрархією пам'яті процесора.

Отримані результати показують, що під час перенесення робочого набору даних між рівнями ієрархії пам'яті відбувається зміна домінуючих факторів продуктивності: якщо для невеликих масивів істотну роль відіграє вартість умовних переходів, то під час роботи з великими масивами основним обмежувальним фактором стає латентність доступу до пам'яті, яка ефективно компенсується програмним попереднім завантаженням даних. Отримані результати можуть бути використані при проектуванні високо-продуктивних структур даних та алгоритмів пошуку для систем, що працюють із великими обсягами інформації. Перспективним напрямом подальших досліджень є детальний аналіз стратегій програмного попереднього завантаження даних для розміщення Ейтцингера, включаючи вибір оптимальної глибини попереднього завантаження, дослідження різних схем попереднього відбору та оцінку їхньої ефективності на процесорах різних архітектур.

Конфлікт інтересів

Автори декларують, що не мають конфлікту інтересів стосовно даного дослідження, в тому числі фінансового, особистісного характеру, авторства чи іншого характеру, що міг би вплинути на дослідження та його результати, представлені в даній статті.

Фінансування

Дослідження проводилося без фінансової підтримки.

Доступність даних

Рукопис не має пов'язаних даних.

Використання засобів штучного інтелекту

Автори підтверджують, що не використовували технології штучного інтелекту при створенні представленої роботи.

СПИСОК ЛІТЕРАТУРИ

1. Knuth, D. (1997), *The Art of Computer Programming. Sorting and Searching*, vol. 3, Addison-Wesley, 2nd edition, 780 p., available at: <https://www.scribd.com/document/438657664/The-Art-Of-Computer-Programming-Sorting-and-Searching-2nd-edition-Volume-3-pdf>
2. Knuth, D. (2010), *Selected papers on design of algorithms*, vol. 7, Stanford, California, 453 pp., available at: <https://press.uchicago.edu/ucp/books/book/distributed/S/bo8930293.html>
3. Sedgewick R., and Wayne K., (2024), *Algorithms*, 4nd edition, Princeton University, Addison-Wesley, 955 p. available at: <https://algs4.cs.princeton.edu/home/>
4. Cormen, H. Thomas, Leiserson, E. Charles, and Rivest, L. Ronald (2022), Clifford Stein. Introduction to algorithms, MIT Press, 1312 p., available at: <https://mitpress.mit.edu/9780262046305/introduction-to-algorithms/>
5. Wiener, R. (2022), "AVL Trees", *Generic Data Structures and Algorithms in Go: An Applied Approach Using Concurrency, Genericity and Heuristics*, Apress, Berkeley, CA, pp. 315–347, doi: https://doi.org/10.1007/978-1-4842-8191-8_10
6. Drepper, U. (2007), *What every programmer should know about memory*, 114 p., available at: <https://akkadia.org/drepper/cpumemory.pdf>

7. Pikus, F. G. (2021), *The Art of Writing Efficient Programs*, Packt Publishing, Birmingham, 445 p., available at: https://ivms.com.safespacenj.org/Tech_Books/writ%20eff%20pgms.pdf
8. (2024), *Intel® 64 and IA-32 Architectures Optimization Reference Manual*, vol. 1, Order Number 248966-049, available at: <https://cdrdv2-public.intel.com/814198/248966-Optimization-Reference-Manual-V1-049.pdf>
9. (2026), *Intel® 64 and IA-32 Architectures Software Developer's Manual Combined*, vol. 2, Instruction Set Reference Order Number 325383-092US, available at: <https://cdrdv2.intel.com/v1/dl/getContent/671110>
10. (2026), *Intel® 64 and IA-32 Architectures Software Developer's Manual Combined*, vol. 3, System Programming Guide, 182 p., available at: <https://cdrdv2.intel.com/v1/dl/getContent/671447>
11. Fog, A. (2026), *Optimizing software in C++: An optimization guide for Windows, Linux and Mac platforms*, available at: https://www.agner.org/optimize/optimizing_cpp.pdf
12. Fog, A. (2025), *Optimizing subroutines in assembly language An optimization guide for x86 platforms*, available at: https://www.agner.org/optimize/optimizing_assembly.pdf
13. Fog, A. (2026), *The microarchitecture of Intel, AMD, and VIA CPUs*, available at: <https://www.agner.org/optimize/microarchitecture.pdf>
14. Fog, A. (2025), *Instruction tables: Lists of instruction latencies, throughputs and micro-operation breakdowns for Intel, AMD and VIA CPUs*, available at: https://www.agner.org/optimize/instruction_tables.pdf
15. Khuong P.-V., and Morin, P. (2017), "Array Layouts for Comparison-Based Searching", *Journal of Experimental Algorithmics*, vol. 22, pp. 1–39, doi: <https://doi.org/10.1145/3053370>
16. Shostak, A. (2026), "Comparative experimental analysis of methods for implementing the AVL tree", *Control, Navigation and Communication Systems*, no. 1, pp. 131–134, doi: <https://doi.org/10.26906/SUNZ.2026.1.131>
17. Anderson, S.E. (2026), *Bit Twiddling Hacks*, available at: <https://graphics.stanford.edu/~seander/bithacks.html>
18. Warren H. S. (2013), *Hacker's Delight*, Addison-Wesley, 2nd ed., 512 p., available at: <https://dl.acm.org/doi/10.5555/2462741>
19. (2026), *Branchless Programming*, available at: <https://en.algorithmica.org/hpc/pipelining/branchless/>
20. Mowry, T.C., Lam, M. S., and Gupta A. (1992), "Design and Evaluation of a Compiler Algorithm for Prefetching", *Proceedings of ASPLOS-V*, pp. 62–73, available at: <https://suif.stanford.edu/papers/mowry92.pdf>
21. Luk C.-K., and Mowry, T.C. (1996), "Compiler-Based Prefetching for Recursive Data Structures", *Proceedings of ASPLOS-VII*, pp. 222–233, available at: <https://www.cs.tufts.edu/comp/150IPL/papers/luk96prefetch.pdf>

Received (Надійшла) 22.05.2026

Accepted for publication (Прийнята до друку) 14.08.2026

Publication date (Дата публікації) 25.09.2026

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Шостак Анатолій Васильович – кандидат технічних наук, доцент, доцент кафедри кібербезпеки та інтелектуальних інформаційних технологій Національного аерокосмічного університету України «ХАІ», Харків, Україна;
Anatolii Shostak – Candidate of Technical Sciences, Associate Professor, Associate Professor of a Department of Cybersecurity and Intelligent Information Technologies of the National Aerospace University of Ukraine "KhAI", Kharkiv, Ukraine;
 e-mail: a.shostak@csn.khai.edu; ORCID Author ID: <https://orcid.org/0009-0003-7687-3641>

Experimental Study of Search Methods in an AVL Tree Considering the Memory Hierarchy

Anatolii Shostak

Abstract. Background. Investigating the effects of data spatial locality, branch prediction, and software prefetching on search performance is of considerable interest in the development of high-performance data structures. **Subject of the study:** search methods in an AVL tree represented as an array using the Eytzinger layout, and the impact of the processor memory hierarchy on their performance. **Purpose of the study:** to experimentally evaluate the efficiency of various search methods in an array-based AVL tree depending on the array size and the location of the working data set relative to the levels of the processor memory hierarchy. **Research objectives.** To implement and experimentally investigate the Binary, AVL, Eytzinger, EytzingerB, and EytzingerBP search algorithms; to determine the ranges of array sizes corresponding to different levels of the memory hierarchy; to compare the performance of the investigated algorithms; and to determine the effects of data spatial locality, branch mispredictions, and software prefetching on search efficiency. **Research methods.** The study employs algorithmic analysis, experimental software performance evaluation, and statistical processing of measurement results. The experimental evaluation was performed using BenchmarkDotNet on the .NET 8 platform with processor hardware performance counters. **Research results.** It was established that the efficiency of search algorithms is determined not only by their computational complexity but also by the characteristics of their interaction with the processor memory hierarchy. The results show that when the working data are located in the L1/L2 caches, the EytzingerB algorithm provides the shortest search time, whereas when the data are located in the L3 cache and main memory, the EytzingerBP algorithm, which uses software prefetching, demonstrates the best performance. Experimental relationships between search time and array size were obtained, the ranges of effective application of the investigated algorithms were determined, and the performance gains achieved by the most efficient search methods were quantitatively evaluated. The experiments demonstrated a change in the most efficient search algorithm as the working data move between levels of the processor memory hierarchy: within the L1/L2 range, the EytzingerB algorithm provides the shortest search time, whereas within the L3 and main-memory range, the EytzingerBP algorithm is the most efficient. **Conclusions.** Representing an AVL tree using the Eytzinger array layout provides a substantial increase in search performance compared with the conventional pointer-based tree implementation. Experimental results confirm that the choice of the optimal search method is determined by the level of the memory hierarchy in which the working data set resides. The obtained results can be used in the design of high-performance search data structures and in the optimization of search algorithms for modern multi-level memory subsystems.

Keywords: AVL tree; Eytzinger layout; search algorithms; software prefetching; cache memory; data locality; BenchmarkDotNet; experimental performance evaluation.

Elshan Hashimov^{1,2}, Ramil Akhundov², Aziz Talibov¹, Marko Radovanović³

¹Azerbaijan Technical University, Baku, Azerbaijan

²National Defense University, Baku, Azerbaijan

³Military Academy, University of Defence in Belgrade, Belgrade, Serbia

A RISK- AND TIME-BASED MODEL FOR PROCESSING INFORMATION FLOWS IN THE PHYSICAL PROTECTION OF SPECIAL-PURPOSE AND CRITICAL INFRASTRUCTURE FACILITIES

Abstract. The effectiveness of a physical protection system depends not only on sensor sensitivity, barrier delay, and response-force readiness, but also on the timely, complete, and undistorted delivery of heterogeneous security data to the decision-making level. This study proposes an integrated risk- and time-based model for processing information flows in the physical protection of special-purpose and critical infrastructure facilities. The information environment is represented as a directed graph. Transmission, processing, and queuing delays are combined with completeness, accuracy, consistency, reliability, and cyber-resilience indicators. A partially non-compensatory information-quality index is developed by combining a weighted geometric mean with the weakest component and a time-decay factor. Information latency is explicitly embedded in the detection, delay, and response time balance, and analytical expressions are formulated for the probability of timely intervention and residual risk. A 10,000-run Monte Carlo experiment is performed for a hypothetical three-zone facility. The distributed backup route achieved an integral information-quality score of 0.849, compared with 0.782 for the centralized primary route and 0.590 for the congested and partially compromised route. In the insider-bypass scenario, the distributed route increased the probability of satisfying the time condition from 0.442 to 0.562 and the mean integrated interruption probability from 0.246 to 0.326. The proposed approach enables weaknesses in information channels to be evaluated jointly with physical barriers and allows information routes to be selected dynamically according to event risk and the time available.

Keywords: physical protection system; information flow; information latency; data fusion, risk; time sufficiency; critical infrastructure; partially non-compensatory model.

1. Introduction

Protecting special-purpose and critical infrastructure facilities is essential to the uninterrupted operation of energy, transport, communications, water supply, industry, defense, and public administration systems. At such facilities, a physical protection system integrates the detection of an adversary, the delay of adversary movement, and the timely intervention of response forces within a single organizational and technical framework. Under the classical approach, system effectiveness is assessed using the probability of detection, the delay created by physical barriers, and the time required for response forces to reach the scene [1, 2]. The widespread deployment of digital sensors, networked cameras, biometric access-control devices, unmanned surveillance platforms, and analytical software has shown that the interaction among these three functions depends heavily on the information infrastructure. If a sensor signal is not verified in time, an event is misclassified, or a command reaches the response force too late, the time reserve created by physical barriers may lose part or all of its operational value.

Within the information environment of a physical protection system, data about the same event arrive from heterogeneous sources. A perimeter radar provides the target coordinates and speed, a thermal camera supplies a heat signature, a video surveillance system provides visual features, an access-control system reports an identification result, and a security officer contributes contextual observations. These data differ in frequency, format, accuracy, and reliability. Processing them separately

increases the operator's information load, while combining them without proper correlation may produce contradictory decisions. More data do not automatically result in better situational awareness. On the contrary, a high false-alarm rate, repeated notifications, and multiple interfaces can delay the appearance of a critical signal in the overall information stream [3, 4].

The role of information flow in physical protection is defined by two interdependent properties. The first is information quality. Data used for decision-making must be complete, authentic, consistent, traceable, and resilient to cyber effects. The second is the time value of information. Correct information may be tactically useless if it arrives after the decision window has closed. An information channel therefore cannot be evaluated solely in terms of packet loss or availability. The entire route from source data to a decision and an actionable command must be considered, together with the route's latency and reliability and the time remaining in the event scenario.

Current regulatory and methodological documents require defense in depth and a balanced combination of detection, delay, and response functions in physical protection [2, 5]. The NIST approach to operational technology security emphasizes information-security requirements such as availability, integrity, monitoring, event logging, and network segmentation [6, 7]. Nevertheless, an important methodological gap remains at the intersection of physical protection and information processing. Information quality is often studied as a separate cyber-resilience problem, whereas time sufficiency is addressed as a separate physical protection

problem. Consequently, there is no unified indicator that measures how information latency reduces the time reserve created by physical barriers and influences the choice of an optimal route.

The purpose of this study is to develop a methodological approach for modeling the generation, transmission, processing, multisource fusion, assessment, and conversion of data into management decisions in the physical protection of special-purpose and critical infrastructure facilities using risk and time criteria. To achieve this purpose, the study formalizes the information environment as a directed graph; determines route latency and reliability; proposes a partially non-compensatory information-quality indicator; incorporates information latency into the time balance of physical protection; develops a criterion for event prioritization and alternative-route selection; and verifies the approach through scenario analysis and Monte Carlo simulation.

The scientific contribution of the study consists of three main results. First, the quality of an information route is assessed not only through an average indicator, but also by considering its weakest component and the loss of information value over time. Second, information-processing time is treated not as an external parameter of the detection, delay, and response model, but as an internal variable that directly reduces the available time reserve. Third, route selection is formulated as a multicriteria optimization problem involving information quality, residual risk, violation of the time condition, and resource consumption. This approach provides a quantitative basis for demonstrating that a bottleneck in an information channel can be as dangerous as a weak link in a physical barrier.

2. Literature review

The theoretical foundation of physical protection systems lies in the functional compatibility of detection, delay, and response. Garcia relates the effectiveness of a physical protection system to the probability that an adversary will be detected and stopped by response forces before completing a critical action [1]. The IAEA recommendations identify defense in depth, a graded approach proportional to the threat, assessment of detection alarms, and the integration of response measures as core principles [2, 5]. Wadoud et al. linked detection, delay, and response functions to probabilistic and temporal indicators when evaluating the physical protection of a nuclear facility [8]. Although these studies establish the time logic of physical protection, signal transmission through a communications network, queuing at processing nodes, and conversion into decision-relevant information are generally subsumed within an aggregated detection time.

In digital physical protection systems, sensors, software, communications, and actuating elements create cyber-physical interdependencies. IAEA Nuclear Security Series No. 42-G requires the integrity and availability of computer-based systems that perform nuclear security functions to be protected [5]. NIST SP 800-82 Rev. 3 emphasizes real-time requirements, high availability, deterministic response, continuous monitoring, and risk-informed network architecture in

operational technology environments [6]. NIST SP 800-160 Vol. 2 Rev. 1 presents cyber resilience at the systems-engineering level as the ability to continue, recover, and adapt essential functions under adverse conditions [9]. These sources show that information infrastructure is inseparable from security functions, but they do not provide a calculation mechanism that directly links the current state of an information route to the time remaining in a physical attack scenario.

In the field of multisensor data fusion, Hall and Llinas systematized the basic principles of integration at the data, feature, and decision levels [10]. The refined JDL model proposed by Steinberg, Bowman, and White distinguishes among object assessment, situation assessment, impact assessment, and process refinement [11]. Khaleghi et al. provided an extensive review of uncertainty, conflict, correlation, and synchronization problems in heterogeneous data fusion [12]. Castanedo grouped data-fusion methods into data association, state estimation, and decision-fusion approaches [13]. Tsanousa et al. compared architectures for storing, indexing, and fusing sensor data in smart environments [14]. This body of research demonstrates that combining multiple sources can improve accuracy and event recognition. In many models, however, latency is treated only as a constraint or a computational cost, while its effect on the probability of timely intervention by physical response forces is not evaluated separately.

Under dynamic security conditions, situational awareness is a major determinant of operator decision quality. Endsley describes situational awareness through three levels: perception of environmental elements, comprehension of their meaning, and projection of their future status [3]. In physical protection, this sequence corresponds to receiving a sensor signal, classifying the event, assessing the adversary's intent, and selecting a response option. As the volume of information presented to the operator increases, selection and prioritization become more important. Situational awareness is improved not by the number of data items displayed on a screen, but by information that is relevant to the decision and delivered in time.

Studies of security operations centers demonstrate the practical consequences of a high notification load. Agyepong et al. established the need for systematic performance indicators to measure analyst work [15]. Forsberg and Frantti proposed a framework of technical metrics for detection, investigation, and response processes [16]. Tariq et al. identified false alarms, disconnected dashboards, staff shortages, and excessive notification volumes as key causes of alert fatigue [4]. The probabilistic cyber-alert management model developed by Chen et al. demonstrates the value of prioritization based on risk and uncertainty [17]. These findings are equally relevant to the control center of a physical protection system. The key difference is that, in a physical intervention scenario, the consequences of a delayed decision may emerge immediately in the real facility and affect people, rather than remaining confined to the information system.

Recent research has shown growing interest in the joint assessment of physical and cyber effects. Sandia

studies indicate that cyber effects on digital detection and delay measures can weaken physical protection functions, while alternative detection and response architectures can improve overall system effectiveness [18, 19]. The scenario-based assessment by Wadoud et al. further strengthens the route and time logic [8]. Nevertheless, the partially non-compensatory aggregation of information completeness, authenticity, consistency, channel reliability, and cyber resilience, together with the direct subtraction of information latency from the available time reserve, has not yet been sufficiently developed within a unified model.

A comparative review of the literature reveals three gaps. First, data-fusion models focus primarily on information content and accuracy but generally do not account for the decision deadline imposed by physical protection. Second, physical protection models assess time sufficiency but do not explicitly describe the internal delay structure of the information route. Third, cyber-resilience documents protect availability and integrity but do not formalize dynamic route selection according to the risk level of the current physical attack. The proposed methodology brings these three directions together within a unified risk-time framework.

Several English-language studies included in the reference list examine conceptual models of multilevel physical protection, the limitations of normative design under dynamic threats, justification of requirements, systems analysis of the design process, and the integration of machine learning with physical-process models [20–22]. Findings on the probabilistic formalization of risk, individual, collective, and societal risk indicators, time sufficiency, and sensitivity indicators strengthen the theoretical basis of the risk-time framework [23, 24].

The role of information processes in conceptual design, the derivation of measurable requirements, and information-to-action requirements in high-tempo operations have also been investigated [25]. Research on selecting cyber-defense architectures for critical information infrastructure, the cyber resilience of communications and command systems under hybrid attacks, the efficiency of MIMO-based multiservice networks, the probabilistic-temporal limitations of software-defined network controllers, and network indicators under dynamic threats provides a basis for defining reliability criteria for information routes [26–28].

At the sensor level, findings on advanced sensing technologies, intelligent monitoring, deep learning,

multisensor detection of ground targets, multimodal fusion, and multisensor architectures for unmanned aerial vehicles support the joint assessment of information completeness, authenticity, and consistency [29, 30].

Research on accelerating decision-making through C4ISR architecture, visual analytics in security systems, and the use of C4ISR systems in operations provides a methodological foundation for processing information flows at the decision level [31–33]. A systems approach to protecting critical infrastructure in urban environments, monitoring special-purpose information networks, analyzing information flows in technical support, and assessing the efficiency of critical communications infrastructure extends the practical scope of the model [34, 35].

3. Methodology

3.1. Research design and basic assumptions

The study combines conceptual modeling, graph theory, probabilistic assessment, multicriteria aggregation, scenario analysis, and Monte Carlo simulation. The physical protection facility is represented by three levels: an external perimeter, an internal protected zone, and a critical zone. Each zone contains detection sources, preliminary-processing devices, communication nodes, and decision-making elements. The endpoint of the information flow is not limited to the central control room, but also includes the operational command delivered to the response force. The route therefore covers every stage from signal generation to receipt of an actionable instruction.

The model is based on the following assumptions: event scenarios are assigned to predefined threat classes; the processing time and reliability of each information node can be estimated; at least one functional backup route exists; quality indicators are normalized within the interval from 0 to 1; event probability and impact severity are obtained from the risk register; and the response-force travel time is represented as a scenario-dependent random variable. In an operational system, the model can be calibrated using logs, network telemetry, sensor tests, and exercise results.

Fig. 1 traces the sequence from data generation to response action and feedback. At the sensor level, the initial signal undergoes filtering and feature extraction before being correlated with other sources (Table 1) at the multisource fusion node.

Table 1 – Information sources and assessment characteristics in physical protection

Information source	Primary data	Typical risk	Characteristic assessed in the model
Perimeter radar and vibration sensor	Coordinates, speed, boundary violation	Atmospheric effects, false alarms	Authenticity, reliability, latency
Daylight and thermal cameras	Image, object features, trajectory	Degraded visibility, loss of video stream	Completeness, consistency, cyber resilience
Access-control system	Identification, authorization, time of passage	Credential forgery, insider threat	Authenticity, traceability, integrity
Security officer	Observation, context, confirmation	Subjectivity, delayed reporting	Authenticity, processing time
OT and building automation	Status of doors, locks, power, and equipment	Cyber interference, synchronization error	Availability, consistency, cyber resilience

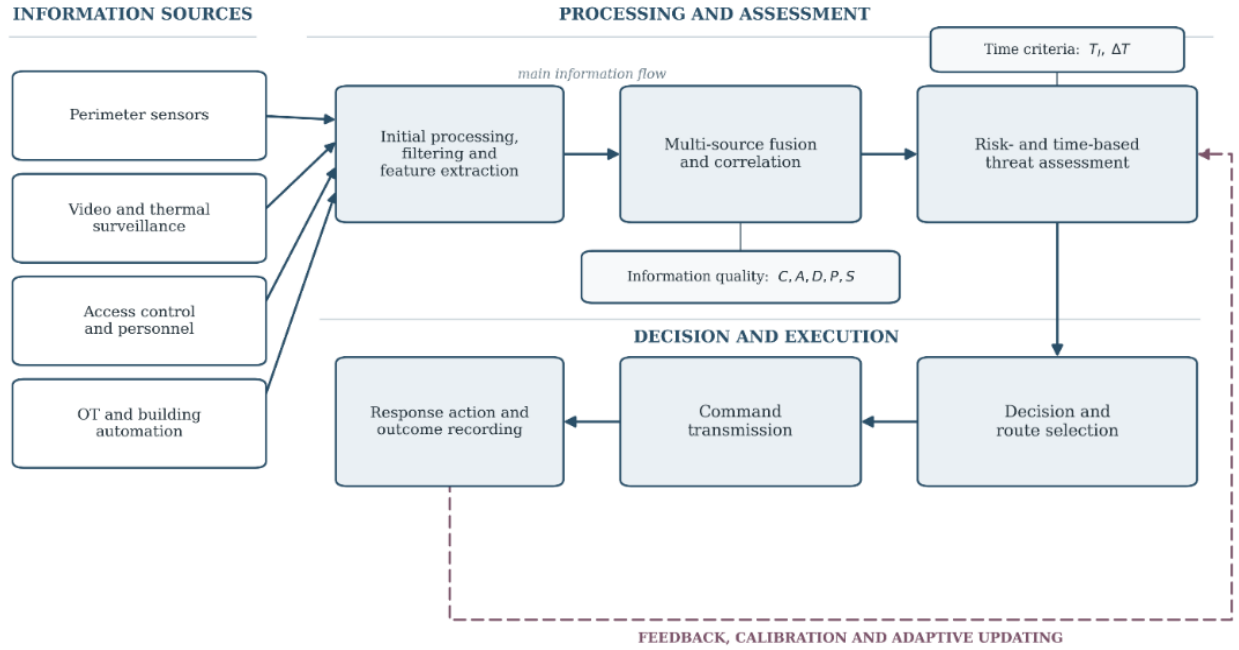


Fig. 1. Structural and functional model of information flow in a physical protection system

The threat assessment determines the event probability, target criticality, and remaining time reserve. The decision module assigns a priority and selects a route, the command is transmitted to the response force, and the outcome of the action generates feedback for updating the model parameters.

3.2. Graph model of the information environment

The information environment is represented as a directed, attributed graph:

$$G_I = (V_I, E_I, \theta_V, \theta_E), \quad r \in \mathcal{R}_s. \quad (1)$$

Here, V_I is the set of sensor, processing, fusion, decision, and execution nodes; E_I is the set of communication links between nodes; θ_V contains the node parameters; θ_E contains the channel parameters; and $\mathcal{R}_s$ is the set of feasible information routes for scenario s . Each arc (i, j) of the graph is assigned a transmission time t_{ij}^{tr} , a queuing time t_{ij}^{que} , a successful-transmission probability p_{ij} , and the current channel load u_{ij} . Each node j is characterized by a processing time t_j^{pr} , a correct-processing probability p_j^{pr} , and a computational reserve.

The total information latency along route r for event s is defined as follows:

$$T_{I,s}(r) = \sum_{(i,j) \in r} (t_{ij,s}^{tr} + t_{ij,s}^{que}) + \sum_{j \in r} t_{j,s}^{pr} + t_s^{dec} + t_s^{cmd}. \quad (2)$$

Here, t_s^{dec} is the time required to formulate the decision, while t_s^{cmd} is the time required to transmit the operational command to the response force. This decomposition makes it possible to identify the source of the delay. For example, a high $t_{ij,s}^{que}$ may indicate network or server congestion, whereas a high t_s^{dec} may indicate inefficiency in the operator interface, procedure, or event-classification process.

Assuming independent components, the information reliability of a route is estimated using the following product:

$$P_{I,s}(r) = \prod_{(i,j) \in r} p_{ij,s} \prod_{j \in r} p_{j,s}^{pr}. \quad (3)$$

When common-cause failures exist among components, the product estimate may be optimistic. In operational applications, dependencies such as a shared power source, synchronization service, network switch, or software platform should be represented by separate group coefficients. For this purpose, a correction coefficient $\chi_g \in [0,1]$ for common-cause failure can be applied to the resulting value of $P_{I,s}(r)$.

3.3. Partially non-compensatory assessment of information quality

Information quality is described by five basic components: completeness $C_{s,r}$, authenticity $A_{s,r}$, consistency $D_{s,r}$, route reliability $P_{s,r}$, and cyber resilience $S_{s,r}$. Completeness represents the extent to which the attributes required to describe an event have been obtained. Authenticity expresses how accurately the information reflects the actual situation. Consistency indicates the absence of conflicts among sources. Reliability is the probability that information will be processed and transmitted without loss. Cyber resilience is the ability to preserve the critical information function under cyberattack or failure conditions.

Initial indicators measured on different scales are converted into benefit-type criteria. A benefit indicator is normalized using the following expression:

$$q_{s,r,k} = \frac{x_{s,r,k} - x_k^-}{x_k^+ - x_k^-}, \quad 0 \leq q_{s,r,k} \leq 1. \quad (4)$$

Here, x_k^+ and x_k^- are the best and worst reference values, respectively, for criterion k . Cost-type indicators, such as latency and packet loss, are normalized in the

opposite direction. The reference values should preferably be derived from acceptance-test protocols, service-level requirements, and threat scenarios rather than solely from expert judgment.

A simple weighted sum allows a weak indicator to be fully offset by other indicators with high values. Such compensation may produce an unacceptable result in physical protection. For example, high completeness should not compensate for low reliability if otherwise complete and accurate information has been altered by a cyberattack or arrives after the decision deadline. The integral information-quality score is therefore calculated using the following partially non-compensatory function:

$$Q_{I,s}(r) = \left[\prod_{k=1}^m q_{s,r,k}^{w_k} \right]^{1-\eta} [\min_{1 \leq k \leq m} q_{s,r,k}]^{\eta} \exp[-\lambda_s T_{I,s}(r)]. \quad (5)$$

Here, $w_k \geq 0$ and $\sum_{k=1}^m w_k = 1$ are the criterion weights; $\eta \in [0,1]$ is the sensitivity coefficient for the weakest link; and $\lambda_s > 0$ is the rate at which the time value of information decays. When $\eta = 0$, the model reduces to weighted geometric aggregation. As η increases, the model responds more strongly to the weakest component. The exponential factor gives preference to the route with lower latency when two routes have the same underlying quality. The value of λ_s varies with event dynamics. It should be small for a slowly developing event and larger for a rapid armed intrusion or an adversary already close to the critical zone. The criterion weights can be determined through expert pairwise comparison, the entropy method, or a combination of both. In the present calculation, the weight vector for completeness, authenticity, consistency, reliability, and cyber resilience is $\mathbf{w} = (0.20, 0.25, 0.15, 0.20, 0.20)$. Authenticity receives a relatively high weight because it directly affects decision correctness, while the minimum-component factor creates a mandatory protective threshold for all indicators. The weakest-link sensitivity is set to $\eta = 0.25$, and the information-aging parameter is set to $\lambda_s = 0.025 \text{ s}^{-1}$.

3.4. Time sufficiency, intervention probability, and residual risk

The principal decision constraint in physical protection is the time available before an adversary reaches the target and completes the critical action. Let $T_{A,s}^0$ denote the time required for the adversary's unimpeded movement and action, $T_{G,s}$ the additional delay created by physical barriers, $T_{D,s}$ the interval between the first detectable adversary action and the generation of a reliable alarm, and $T_{R,s}$ the time required for the response force to reach the intervention point after receiving the command. The time margin for scenario s and information route r is then defined as:

$$\Delta T_{s,r} = T_{A,s}^0 + T_{G,s} - T_{D,s} - [T_{I,s}(r) + T_{R,s}]. \quad (6)$$

The condition $\Delta T_{s,r} \geq 0$ indicates that the response force can intervene before the adversary completes the critical action. A negative value means that the information and response chain exceeds the time reserve created by the physical delay measures. Equation (6)

makes information latency an explicit variable in the physical-protection time balance. Under this formulation, every additional second of network or decision latency reduces the time margin by the same amount.

Because the parameters are probabilistic, a strict threshold of zero alone is insufficient. The probability of timely intervention is evaluated smoothly using a logistic function:

$$P_{T,s}(r) = \frac{1}{1 + \exp[-\kappa_s \Delta T_{s,r}]}. \quad (7)$$

Here, κ_s determines the sharpness of the time threshold. When $\Delta T_{s,r}$ is strongly positive, $P_{T,s}(r)$ approaches one; when it is strongly negative, the probability approaches zero. Near the threshold, the logistic function provides a more realistic representation of measurement and execution uncertainty.

Timely and correct event detection, information-route quality, satisfaction of the time condition, and the neutralization capability of the response force jointly determine the integrated intervention probability:

$$P_{int,s}(r) = P_{D,s} Q_{I,s}(r) P_{T,s}(r) P_{N,s}. \quad (8)$$

Here, $P_{D,s}$ is the probability of detecting the event, and $P_{N,s}$ is the probability of neutralizing the adversary once a timely engagement has occurred. Multiplying the factors preserves the direct influence of any weakened function on the final result.

If π_s is the probability that scenario s will occur during the period under consideration and L_s is the impact severity of its successful realization, the residual risk is calculated as:

$$R_{res,s}(r) = \pi_s L_s [1 - P_{int,s}(r)]. \quad (9)$$

This expression shows that route selection affects not only technical quality but also the criticality of the protected asset.

The same improvement in route quality produces a greater reduction in risk for a scenario with more severe consequences.

3.5. Event prioritization and route selection

When several alerts are generated at the same time, their processing order should not be determined solely by arrival time. The following risk-time priority index is proposed for event h :

$$\Pi_h = \frac{\hat{p}_h \hat{L}_h U_h K_h}{\max(\hat{T}_{remaining,h}, \varepsilon)}. \quad (10)$$

Here, $\hat{p}_h$ is the current probability that the event represents a genuine threat; $\hat{L}_h$ is the normalized severity of the potential damage; U_h is the information-uncertainty coefficient; K_h is the criticality of the target; $\hat{T}_{remaining,h}$ is the time remaining for a decision; and ε is a small constant that prevents division by zero. A high threat probability, severe potential damage, a critical target, and little remaining time all increase the priority. Including the uncertainty coefficient prevents an unconfirmed but potentially high-consequence event from being assigned an unduly low priority.

Route selection is formulated as a multicriteria optimization problem:

$$r_s^* = \operatorname{argmin}_{r \in \mathcal{R}_s} \left\{ \alpha [1 - Q_{I,s}(r)] + \beta \frac{R_{res,s}(r)}{R_s^{\max}} + \gamma \frac{c_r}{c^{\max}} + \delta \frac{\max(0, -\Delta T_{s,r})}{T_{ref}} \right\}. \quad (11)$$

Here, c_r is the route's consumption of communication, computing, and energy resources; α , β , γ , and δ are the weights of the objective function. For critical events, β and δ may be assigned relatively high values, whereas γ may be larger during routine operation. The selection is performed subject to the following mandatory thresholds:

$$Q_{I,s}(r) \geq Q_s^{\min}, P_{I,s}(r) \geq P_s^{\min}, T_{I,s}(r) \leq T_{I,s}^{\max}. \quad (12)$$

If a route fails to meet any mandatory threshold, a favorable average value in the objective function is not sufficient for its selection. This property retains the partially non-compensatory logic at the optimization stage.

Figure 2 presents the logical sequence of the algorithm. After receiving an event, the system evaluates source reliability and information quality, then calculates latency and the time margin for the feasible routes. Routes that fail to meet the mandatory thresholds are excluded. The remaining alternatives are compared using risk, quality, time, and resource criteria. If the current route falls below a threshold, the system switches to a backup route and reassesses the outcome through feedback.

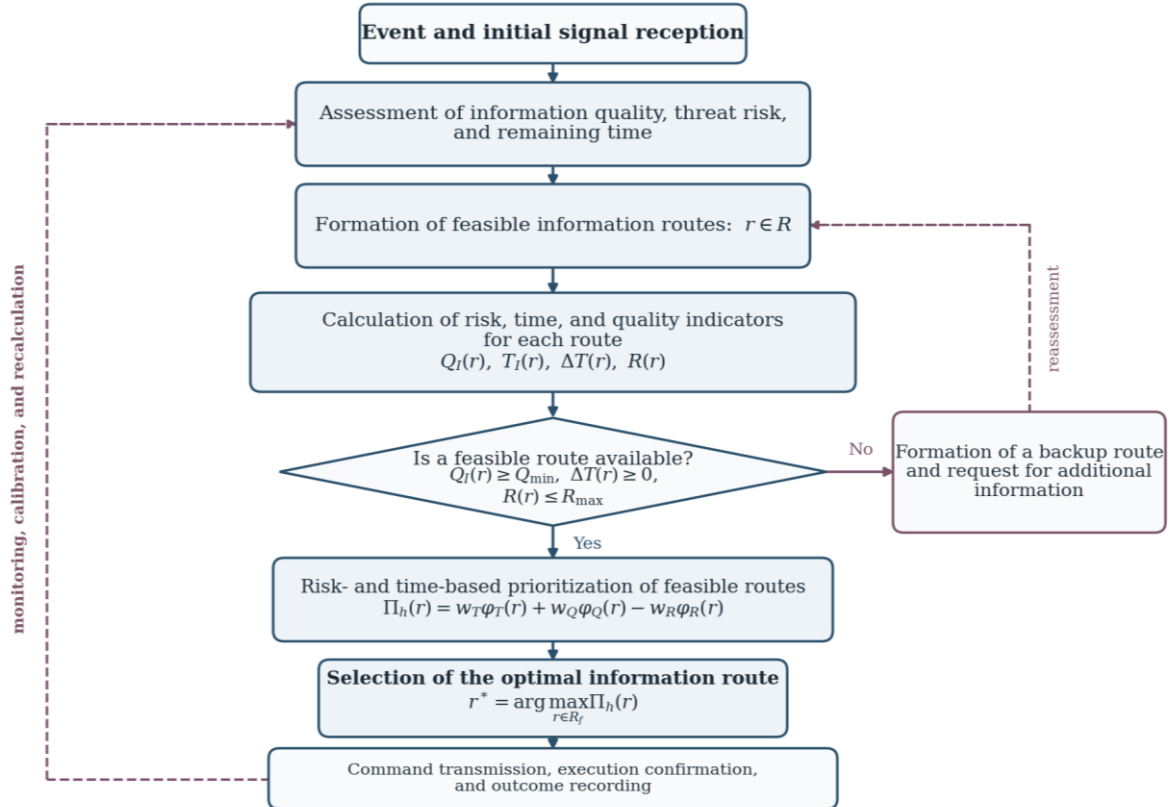


Fig. 2. Algorithm for risk-time-based prioritization and information-route selection

3.6. Computational experiment and input data

A hypothetical three-zone critical infrastructure facility is used to verify the methodology. The first zone contains perimeter radar, a vibration sensor, and daylight and thermal cameras. The second contains access control, video analytics, and a security post. The third contains critical-room sensors, an independent local controller, and the response force. Three information modes are compared: r_1 , the primary route through a central server; r_2 , a distributed backup route with local edge processing and an independent communication channel; and r_3 , a degraded route characterized by high load, packet loss, and indications of cyber interference.

As shown in Table 2, route r_1 has high completeness, but its cyber-resilience value of 0.84 lowers the minimum component. Route r_2 has slightly lower completeness at 0.89, yet its advantages in authenticity, reliability, cyber resilience, and latency

increase the integral result. For route r_3 , both the quality components and the temporal indicator deteriorate. The combined effects of exponential aging and the minimum-component penalty therefore reduce Q_I to 0.590.

Three threat scenarios are considered: s_1 , unauthorized entry through the perimeter; s_2 , a coordinated forced intrusion; and s_3 , an insider bypassing the access procedure for the critical zone (Table 3). The parameters do not reproduce classified information from an actual facility. They are engineering-realistic synthetic values selected to demonstrate the behavior of the method.

After the deterministic calculation, the effect of uncertainty is evaluated using Monte Carlo simulation (Table 4). A total of 10,000 runs is performed for each scenario and route. T_A^0 and T_R are modeled using normal distributions with standard deviations equal to 10% of their respective means. T_G and T_D are modeled using normal distributions with standard deviations equal to 15% of their means.

Table 2 – Initial and integral indicators of the information routes

Route	T_I, s	C	A	D	P	S	Q_I
r_1 , centralized primary	4.8	0.92	0.90	0.88	0.94	0.84	0.782
r_2 , distributed backup	3.1	0.89	0.93	0.91	0.97	0.93	0.849
r_3 , congested and compromised	10.6	0.86	0.81	0.77	0.83	0.70	0.590

Table 3 – Time and probability parameters of the threat scenarios

Scenario	T_A^0, s	T_G, s	T_D, s	T_R, s	P_D	P_N
s_1 , perimeter intrusion	75	25	4	65	0.94	0.90
s_2 , forced intrusion	52	18	5	54	0.90	0.88
s_3 , insider threat	38	10	8	36	0.82	0.85

Table 4 – Monte Carlo simulation results, 10,000 runs

Scenario	Route	Mean $\Delta T, s$	$P(\Delta T \geq 0)$	Mean P_{int}	5th percentile	95th percentile
s_1	r_1	26.21	0.993	0.654	0.650	0.662
s_1	r_2	27.99	0.995	0.713	0.712	0.718
s_1	r_3	20.49	0.969	0.479	0.383	0.499
s_2	r_1	6.32	0.786	0.472	0.026	0.619
s_2	r_2	7.87	0.836	0.545	0.058	0.672
s_2	r_3	0.46	0.525	0.244	0.001	0.467
s_3	r_1	-0.84	0.442	0.246	0.006	0.533
s_3	r_2	0.89	0.562	0.326	0.014	0.586
s_3	r_3	-6.61	0.138	0.075	0.000	0.337

To preserve positive values, T_I is modeled using a lognormal distribution. The coefficient of variation is set to 0.15 for r_1 and r_2 and to 0.25 for r_3 . The random-number generator seed is set to 20260816 to ensure reproducibility. A value of $\kappa_s = 0.45 s^{-1}$ is used for the sharpness of the time threshold.

4. Discussion of results

The deterministic results show that changing the information route directly affects the time balance of physical protection. In scenario s_1 , a large positive time margin remains for all routes. The values are $\Delta T_{1,1} = 26.2 s$ for the primary route, $\Delta T_{1,2} = 27.9 s$ for the distributed route, and $\Delta T_{1,3} = 20.4 s$ for the degraded route. Because P_T is close to one in this scenario, most of the difference in the integrated intervention probability is attributable to information quality. The corresponding P_{int} values are 0.662, 0.718, and 0.499. In the forced-intrusion scenario s_2 , the system approaches the time threshold. The time margin is 6.2 s for r_1 , 7.9 s for r_2 , and only 0.4 s for r_3 . The probability of timely intervention obtained from the logistic function is 0.942, 0.972, and 0.545, respectively. On the degraded route, the increase in mean latency to 10.6 s consumes almost the entire time reserve. Consequently, while the integrated intervention probability is 0.654 for r_2 , it falls to 0.255 for r_3 . In other words, with the detection and neutralization indicators unchanged, disruption of the information chain alone can reduce the protection outcome by 61%.

Scenario s_3 , involving an insider threat, demonstrates the greatest sensitivity to the information route. Because the insider has already bypassed part of

the perimeter barriers and detection occurs later, the decision window is narrow. The time margin is -0.8 s for r_1 , 0.9 s for r_2 , and -6.6 s for r_3 . A reduction of only 1.7 s in route latency changes the time condition from negative to positive compared with the primary route. This change increases P_T from 0.411 to 0.600 and P_{int} from 0.224 to 0.355. It follows that, in scenarios close to the time threshold, a small difference in information latency may have a greater operational effect than improvements in the individual quality components.

The Monte Carlo results reveal substantial distributions behind the deterministic values. In scenario s_1 , the time margin is large and the distribution of P_{int} is therefore relatively narrow. In scenarios s_2 and s_3 , variability in the response and attack times creates wide differences between the 5th and 95th percentiles. For example, the mean intervention probability for the combination of s_2 and r_2 is 0.545, whereas the 5th percentile is only 0.058. This result indicates that a decision based solely on the mean may be excessively optimistic. Route selection at critical facilities should therefore be evaluated using a lower percentile or a conditional risk measure in addition to the mean result.

Figure 3 shows that the distributed backup route performs best in all scenarios, although the scale of its advantage depends on the risk-time conditions. In the perimeter scenario, the improvement is limited because the time condition is almost always satisfied for every route. In the forced-intrusion and insider-threat scenarios, the faster and more resilient route affects the result through two channels: it increases Q_I and enlarges the time margin that determines P_T .

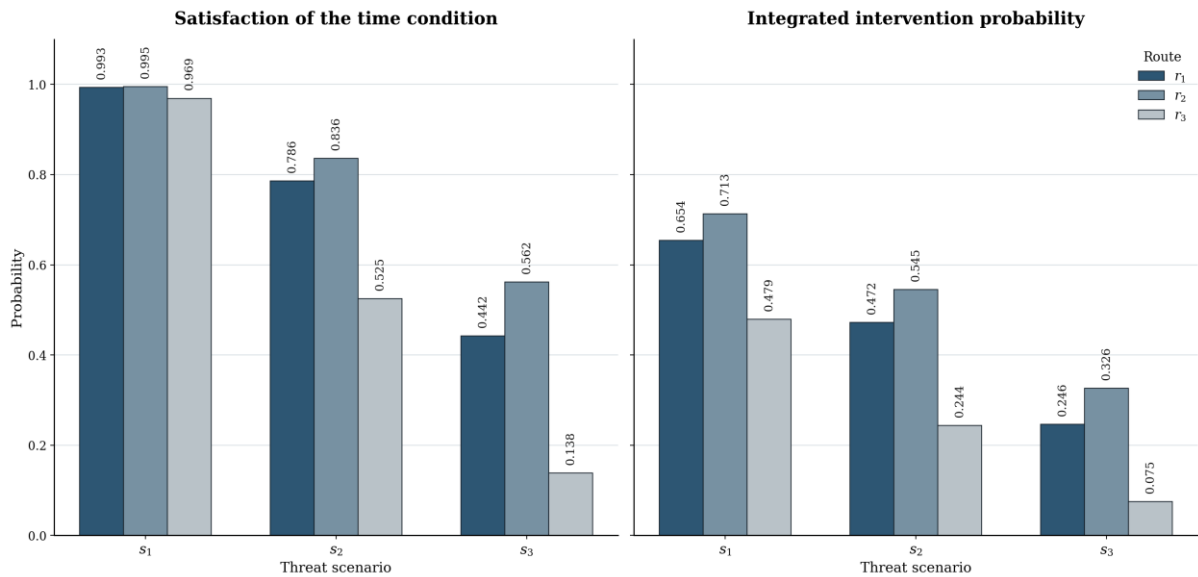


Fig. 3. Comparison of time-condition satisfaction and integrated intervention probability across scenarios

Compared with the centralized primary route, the distributed backup route increases the mean P_{int} by 8.9% for s_1 , 15.3% for s_2 , and 32.7% for s_3 . This nonlinear improvement shows that the relative importance of information latency increases as the threat approaches the critical zone.

A centralized architecture may be convenient for management and auditing, but routing all information

through the same server, switch, and analytical service introduces a common-cause failure. An edge-processing node and an independent communication channel reduce this dependency by transmitting the minimum information set required for a decision over a shorter route.

The sensitivity analysis identifies response time T_R as the most influential parameter (Table 5).

Table 5 – One-way sensitivity analysis for the main parameters

Parameter varied for s_2 and r_2	Change in P_{int} at lower value	Change in P_{int} at higher value	Interpretation
T_I , -20% and +20%	+2.25%	-2.41%	Increased latency affects the result through both Q_I and P_T
T_R , -10% and +10%	+2.60%	-22.35%	The nonlinear effect near the time threshold is strongest
λ , -20% and +20%	+1.56%	-1.54%	The information-aging rate has an approximately symmetric effect
η , 0.10 and 0.40	+0.61%	-0.60%	The weakest-link penalty is limited when the quality components are balanced
A , -0.05 and +0.05	-1.31%	+0.99%	The quality result changes because authenticity has a high weight
S , -0.05 and +0.05	-1.11%	+0.79%	Cyber resilience preserves the advantage of the backup route

A 10% increase in T_R reduces P_{int} by 22.35% for the combination of s_2 and r_2 . The fact that the decrease is much larger than the corresponding increase is explained by the nonlinear nature of the logistic function. When the system is close to the time threshold, a negative change produces a sharp loss, whereas the benefit of a positive change is limited as the probability approaches one. This finding indicates that information-route optimization should be coordinated with response-force deployment and readiness.

The relatively low sensitivity to η is attributable to the balanced quality components of the selected route r_2 . If one component, such as cyber resilience, fell to 0.50, the influence of the minimum-component factor would increase considerably. The parameter η is therefore not merely a mathematical hardening factor. It expresses the facility's risk attitude toward a weak link. At high-consequence facilities, a higher value of η and separate minimum thresholds for each component are appropriate.

A three-level management structure is proposed for practical application of the methodology. At the strategic level, acceptable risk, minimum Q_I , minimum P_I , and maximum information-latency thresholds are established. At the operational level, the state of routes, server and network loads, sensor failures, time synchronization, and cyber incidents are monitored. At the tactical level, the priority of a specific event is calculated, a route is selected, and a response command is generated. This division links technical indicators to management decisions.

The proposed model also offers a different way of addressing false alarms. Simply filtering out false alarms may increase the risk of losing a genuine event. In addition to threat probability, the priority index considers potential damage, target criticality, uncertainty, and the remaining time. Consequently, a low-confidence signal associated with a critical target and a narrow decision window is not rejected outright. Instead, it is sent for urgent verification

through an additional sensor request or a more reliable route. This approach establishes a functional balance between automation and operator oversight.

From a scientific perspective, the model connects data-fusion theory with the time sufficiency of physical protection. In the data-fusion literature, the principal objective is to improve the accuracy of a unified situation picture [10–14]. In the physical protection literature, the main objective is to enable response forces to intervene before the adversary completes the critical action [1, 2, 8]. Equation (5) integrates the quality and time value of information, while Equation (6) links that value to the physical time margin. Information-architecture improvements can therefore be assessed not merely as IT performance measures, but as investments in physical security.

The study also has several limitations. First, the computational experiment uses synthetic data and cannot be treated as a direct normative result for a specific facility. Second, Equation (3) assumes conditional independence among components, although operational systems contain power, software, and communication dependencies. Third, P_D and P_N are treated as constants, although they may correlate with weather, adversary tactics, personnel readiness, and information quality. Fourth, the parameter κ_s of the logistic function must be calibrated using exercise and incident data. Future studies could address these limitations through Bayesian networks, common-cause failure models, and adaptive weight updating based on real-time telemetry.

5. Conclusions

This study develops a model that integrates risk and time criteria for processing information flows in the physical protection of special-purpose and critical infrastructure facilities. From generation of the sensor signal to delivery of the operational command to the response force, the model separately accounts for transmission, queuing, processing, decision, and command delays. Representing the information environment as a directed graph makes it possible to identify alternative routes, bottlenecks in nodes and channels, and common-cause dependencies.

Combining completeness, authenticity, consistency, reliability, and cyber resilience through a weighted geometric mean, a minimum-component factor, and an exponential time penalty limits full compensation for a weak indicator. This property is essential for critical protection systems because distorted, unreliable, or delayed information cannot be considered safe merely because other quality indicators are strong. The proposed Q_I indicator reflects not only the technical correctness of information, but also its operational usefulness for the specific event.

Including information latency in the time-sufficiency equation shows that each unit of time lost in the communication and processing chain reduces by the same amount the reserve created by physical barriers. Near the time threshold, this effect is nonlinear. In the insider-threat scenario, switching from the primary route to the distributed backup route changes the deterministic time margin from -0.8 s to 0.9 s. In the Monte Carlo experiment, it increases the probability of satisfying the time condition from 0.442 to 0.562. The mean integrated intervention probability rises from 0.246 to 0.326, corresponding to a relative improvement of 32.7%.

The results confirm that route selection should not be treated as a static network rule. It should be implemented as a dynamic decision based on event probability, impact severity, target criticality, information uncertainty, and the remaining time. Distributed processing and an independent backup channel not only increase availability, but also widen the decision window. The design of a physical protection system should therefore specify mandatory thresholds for the latency, reliability, and cyber resilience of information routes alongside sensor and barrier performance requirements.

For practical implementation, the model parameters should be calibrated using the facility's risk register, acceptance tests, communication logs, operator-activity records, and tactical exercise results. The strategic level should define acceptable risk and minimum quality thresholds; the operational level should monitor route and node status; and the tactical level should provide real-time prioritization and route switching. Future research may validate the model using data from an operational facility, represent correlated failures through Bayesian networks, detect adversarial data manipulation, and update weight parameters adaptively.

Conflict of interest

The authors declare that they have no conflict of interest regarding this research, including financial, personal, authorship-related, or any other type of conflict that could affect the research and its results presented in this article.

Фінансування

Дослідження проводилося без фінансової підтримки.

Data availability

The manuscript has no associated data.

Use of artificial intelligence tools

The authors confirm that artificial intelligence technologies were not used in the preparation of this work.

REFERENCES

1. Garcia, M. L. (2008), *Design and Evaluation of Physical Protection Systems*, 2nd ed., Butterworth-Heinemann, Burlington, doi: <https://doi.org/10.1016/B978-0-08-055428-0.50005-1>
2. (2011), *Nuclear Security Recommendations on Physical Protection of Nuclear Material and Nuclear Facilities (INFCIRC/225/Revision 5)*, IAEA Nuclear Sec. Ser. No. 13. Vienna: IAEA, 57 p., doi: <https://doi.org/10.61092/iaea.ko2c-dc4q>
3. Endsley, M. R. (1995), "Toward a theory of situation awareness in dynamic systems", *Human Factors*, vol. 37(1), pp. 32–64, doi: <https://doi.org/10.1518/001872095779049543>

4. Tariq, S., Chhetri, M. B., and Paris, C. (2025), "Alert fatigue in security operations centres: Research challenges and opportunities", *ACM Computing Surveys*, doi: <https://doi.org/10.1145/3723158>
5. (2021), *Computer Security for Nuclear Security*, IAEA Nuclear Security Series No. 42-G., IAEA, Vienna 2021, 86 p., available at: <https://www.iaea.org/publications/13629/computer-security-for-nuclear-security>
6. Stouffer, K., Pease, M., Tang, C. Y., Zimmerman, T., Pillitteri, V. Y., Lightman, S., Hahn, A., Saravia, S., Sherule, A., and Thompson, M. (2023), *Guide to Operational Technology Security*. NIST SP 800-82 Rev. 3., National Institute of Standards and Technology, Gaithersburg, doi: <https://doi.org/10.6028/NIST.SP.800-82r3>
7. (2020), *Security and Privacy Controls for Information Systems and Organizations*, NIST SP 800-53 Rev. 5, National Institute of Standards and Technology, Gaithersburg, Joint Task Force, doi: <https://doi.org/10.6028/NIST.SP.800-53r5>
8. Wadoud, A. A., Adail, A. S., and Saleh, A. A. (2018), "Physical protection evaluation process for nuclear facility via sabotage scenarios", *Alexandria Engineering Journal*, vol. 57(2), pp. 831–839, doi: <https://doi.org/10.1016/j.aej.2017.01.045>
9. Ross, R., Pillitteri, V. Y., Graubart, R., Bodeau, D., and McQuaid, R. (2021), *Developing Cyber-Resilient Systems: A Systems Security Engineering Approach*. NIST SP 800-160 Vol. 2 Rev. 1, National Institute of Standards and Technology, Gaithersburg, doi: <https://doi.org/10.6028/NIST.SP.800-160v2r1>
10. Hall, D. L., and Llinas, J. (1997), "An introduction to multisensor data fusion", *Proceedings of the IEEE*, vol. 85(1), pp. 6–23, doi: <https://doi.org/10.1109/5.554205>
11. Steinberg, A. N., Bowman, C. L., and White, F. E. (1999), "Revisions to the JDL data fusion model", *Proceedings of SPIE*, vol. 3719, pp. 430–441, doi: <https://doi.org/10.1117/12.341367>
12. Khaleghi, B., Khamis, A., Karray, F. O., and Razavi, S. N. (2013), "Multisensor data fusion: A review of the state-of-the-art", *Information Fusion*, vol. 14(1), pp. 28–44, doi: <https://doi.org/10.1016/j.inffus.2011.08.001>
13. Castanedo, F. (2013), "A review of data fusion techniques", *The Scientific World Journal*, doi: <https://doi.org/10.1155/2013/704504>
14. Tsanousa, A., et al. (2022), "A review of multisensor data fusion solutions in smart manufacturing: systems and trends", *Sensors*, vol. 22(5), 1734, doi: <https://doi.org/10.3390/s22051734>
15. Agyepong, E., Cherdantseva, Y., Reinecke, P., and Burnap, P. (2023), "A systematic method for measuring the performance of a cyber security operations centre analyst", *Computers & Security*, vol. 124, article 102959, doi: <https://doi.org/10.1016/j.cose.2022.102959>
16. Forsberg, J., and Frantti, T. (2023), "Technical performance metrics of a security operations center", *Computers & Security*, vol. 135, 103529, doi: <https://doi.org/10.1016/j.cose.2023.103529>
17. Chen, H., Dunklee, A., Jajodia, S., Liu, R., McNamara, S., and Subrahmanian, V. S. (2022), "PCAM: A data-driven probabilistic cyber-alert management framework", *ACM Transactions on Internet Technology*, vol. 22(3), article 67, pp. 1–24, doi: <https://doi.org/10.1145/3511101>
18. Johnson, J. (2020), "Physical security and cybersecurity of energy storage systems", *Energy Storage Handbook*, Sandia National Laboratories, available at: https://www.researchgate.net/publication/348785890_Chapter_18
19. Evans, A., Russell, J. L., and Cipiti, B. B. (2023), "New security concepts for advanced reactors", *Nuclear Science and Engineering*, vol. 197(S1), pp. S70–S79, doi: <https://doi.org/10.1080/00295639.2022.2112134>
20. Akhundov, R.G., Hashimov, E.G., and Islamov, I.J. (2026), "Conceptual models of multi-level physical protection systems for special-purpose and critical infrastructure facilities", *Grail of Science*, no. 61, pp. 591–608, doi: <https://doi.org/10.36074/grail-of-science.23.01.2026.066>
21. Akhundov, R.G., Hashimov, E.G., and Islamov, I. (2026), "Methodological limitations of normative design of physical protection systems for critical and military facilities in a dynamic threat environment", *Grail of Science*, no. 62, pp. 873–889, doi: <https://doi.org/10.36074/grail-of-science.20.02.2026.096>
22. Akhundov, R.G. (2026), "Enhancing the physical protection of critical facilities through the integration of physical process models and machine learning", *Grail of Science*, no. 61, pp. 722–731, doi: <https://doi.org/10.36074/grail-of-science.23.01.2026.083>
23. Akhundov, R.G., Talibov, A.M., and Hashimov, E.G. (2026), "Temporal sufficiency of physical protection systems for critical and military facilities", *Grail of Science*, no. 68, pp. 913–930, doi: <https://doi.org/10.36074/grail-of-science.15.05.2026.102>
24. Talibov, A.M., Akhundov, R.G., and Hashimov E.G. (2026), "Integral sensitivity indicator based on multilayer indicators for critical infrastructure and special-purpose facilities and the selection of measures", *Grail of Science*, no. 65, pp. 563–579, doi: <https://doi.org/10.36074/grail-of-science.03.04.2026.063>
25. Akhundov, R.G. and Hashimov, E.G. (2026), "Modeling information processes and deriving measurable requirements in physical protection system design", *Management Information System and Devices*, no. 1(188), pp. 5–16, doi: <https://doi.org/10.30837/0135-1710.2026.188.005>
26. Gasimov, V., Mammadov, J., Islamov, I., and Hashimov, E. (2026), "Evaluation of alternative solutions for the effective structure of the cyber security system in critical information infrastructures by the hierarchical analysis method", *Advanced Information Systems*, vol. 10(2), pp. 87–99, doi: <https://doi.org/10.20998/2522-9052.2026.2.10>
27. Hashimov, E.G., Ibrahimov, B.G., Islamov, I.J., Talibov, A.M., and Akhundov, R.G. (2026), "Research of the efficiency multiservice networks using MIMO technology", *Advanced Information Systems*, vol. 10(1), pp. 66–71, doi: <https://doi.org/10.20998/2522-9052.2026.1.08>
28. Ibrahimov, B.G., Akhundov, R.G., and Hashimov E.G. (2025), "Research of multi-service communication networks taking into account dynamically emerging security threats", *Problems of Informatization*. Proc. of the 13th International Scientific and Technical Conference, vol. 2, pp. 10–11, available at: <https://repository.kpi.kharkov.ua/handle/KhPI-Press/95974>
29. Muradova, E.E., Talibov, A.M., and Hashimov E.G. (2025), "Deep learning for critical infrastructure security in energy and transportation", *Grail of Science*, no. 56, pp. 310–323, doi: <https://doi.org/10.36074/grail-of-science.19.09.2025.039>
30. Huseynov, B.S. et al. (2025), "Multisensor detection architectures on UAV platforms: methodology, indicators and implementation in Azerbaijan", *Problems of Informatization*. Proceedings of the 13th International Scientific and Technical Conference, vol. 2, pp. 132–134, available at: <https://repository.kpi.kharkov.ua/handle/KhPI-Press/96131>
31. Krikun, V.L. et al. (2025), "Accelerating decision making and operational efficiency through C4ISR architecture. *Problems of Informatization*", Proceedings of the 13th International Scientific and Technical Conference, vol. 2, pp. 120–121, available at: <https://repository.kpi.kharkov.ua/handle/KhPI-Press/96093>

32. Hazarkhanov, A.T. et al. (2025), "Visual analytics across security and systems engineering: foundations for the V-model and C4ISR", *Problems of Informatization*, Proceedings of the 13th International Scientific and Technical Conference, vol. 4, pp. 74–76, available at: <https://repository.kpi.kharkov.ua/server/api/core/bitstreams/90741da4-0b60-438f-9157-590b3b2aa6ec/content>
33. Maharramov, R.R. et al. (2025), "Application of C4ISR systems in military and security operations", *Innovative Trends in Natural Sciences, Humanities and Exact Sciences*, Proceedings of the 7th International Scientific Conference, pp. 125–134, doi: <https://doi.org/10.62731/mcnd-15.08.2025.004>
34. Talibov, A.M., and Hashimov, E.G. (2024), "Analysis of information flows in the technical support of the army", *Problems of Informatization*, Proceedings of the 12th International Scientific and Technical Conference, vol. 3, p. 6, doi: <https://doi.org/10.32620/PI.24.t3>
35. Ibrahimov, B.G., Hasanov, A.H., doi., and Hashimov, E.G. (2024), "Research and analysis of efficiency indicators of critical infrastructures in the communication system", *Advanced Information Systems*, vol. 8(2), pp. 58–64 <https://doi.org/10.20998/2522-9052.2024.2.07>

Received (Надійшла) 15.06.2026

Accepted for publication (Прийнята до друку) 26.08.2026

Publication date (Дата публікації) 25.09.2026

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Гашимов Елшан Гіяс – доктор наук з національної безпеки та військових наук, професор, професор Азербайджанського технічного університету, професор Університету національної оборони, Баку, Азербайджан;

Elşan Hashimov – Doctor in national security and military sciences, professor, professor of Azerbaijan Technical University, professor of National Defense University, Baku, Azerbaijan;

e-mail: hasimovel@gmail.com; ORCID Author ID: <http://orcid.org/0000-0001-8783-1277>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57195631270>.

Ахундов Раміль Гурбаналі – доктор філософії з національної безпеки та військових наук, професор Національного університету оборони, Баку, Азербайджан;

Ramil Akhundov – PhD in National Security and Military Sciences, Professor of National Defense University, Baku, Azerbaijan;

e-mail: mr.axundov1@gmail.com; ORCID Author ID: <https://orcid.org/0009-0001-8798-8044>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=59968793800>.

Талібов Азіз Машалла – доктор наук з національної безпеки та військових наук, професор, професор Інституту систем управління, Баку, Азербайджан;

Aziz Talibov – Doctor in National Security and Military Sciences, Professor, Professor of Institute of Control Systems, Baku, Azerbaijan;

e-mail: talibovaziz65@gmail.com; ORCID Author ID: <https://orcid.org/0000-0002-0572-7212>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=60388843600>.

Радованович Марко – Військова академія, Університет оборони, Белград, Сербія;

Marko Radovanović – Military Academy, University of Defence, Belgrade, Serbia;

e-mail: markoradovanovicgdb@yahoo.com; ORCID Author ID: <https://orcid.org/0000-0002-9866-9639>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57219384593>.

Модель оброблення інформаційних потоків з урахуванням ризику та часу в системах фізичного захисту об'єктів спеціального призначення та критичної інфраструктури

Е. Г. Гашимов, Р. Г. Ахундов, А. М. Талібов, М. Р. Радованович

Анотація. Ефективність функціонування системи фізичного захисту визначається не лише характеристиками чутливості засобів виявлення, часом затримки, що забезпечується фізичними бар'єрами, та рівнем готовності сил реагування, але й своєчасністю, повнотою та цілісністю передавання неоднорідних даних у сфері безпеки до рівня прийняття рішень. У роботі запропоновано інтегровану модель оброблення інформаційних потоків з урахуванням ризику та часових характеристик для систем фізичного захисту об'єктів спеціального призначення та критичної інфраструктури. Інформаційне середовище подано у вигляді орієнтованого графа. Часові затримки передавання, оброблення та очікування інформації в чергах інтегровано з показниками її повноти, точності, узгодженості, надійності та кіберстійкості. Для комплексного оцінювання якості інформації розроблено частково некомпенсаторний індекс, що поєднує зважене геометричне середнє, значення найслабшої складової та коефіцієнт часового згасання. Затримку інформації явно враховано в балансі часу виявлення, затримки порушника та реагування. Сформульовано аналітичні вирази для визначення ймовірності своєчасного втручання та оцінювання залишкового ризику. Для гіпотетичного об'єкта з трьома функціональними зонами проведено імітаційний експеримент методом Монте-Карло, що охоплював 10 000 реалізацій. За результатами моделювання встановлено, що розподілений резервний маршрут забезпечив інтегральний показник якості інформації на рівні 0,849, тоді як для централізованого основного маршруту цей показник становив 0,782, а для переважаного та частково скомпрометованого маршруту — 0,590. У сценарії обходу системи фізичного захисту внутрішнім порушником використання розподіленого маршруту дало змогу підвищити ймовірність виконання часової умови з 0,442 до 0,562, а середню інтегральну ймовірність переривання дій порушника — з 0,246 до 0,326. Запропонований підхід забезпечує комплексне оцінювання вразливостей інформаційних каналів у взаємозв'язку з характеристиками фізичних бар'єрів та дає змогу здійснювати динамічний вибір маршрутів передавання інформації залежно від рівня ризику події й доступного часу для реагування.

Ключові слова: система фізичного захисту; інформаційний потік; затримка інформації; інтеграція даних; ризик; достатність часу; критична інфраструктура; частково некомпенсаторна модель.

О. А. Горносталь, В. В. Челак

Національний технічний університет «Харківський політехнічний інститут», Харків, Україна

ДИВЕРСИФІКОВАНІ АНСАМБЛІ КЛАСИФІКАТОРІВ ДЛЯ ВИЯВЛЕННЯ ДЕФЕКТІВ ПРОГРАМНОГО КОДУ ЗА СТАТИЧНИМИ МЕТРИКАМИ ЯКОСТІ

Анотація. Актуальність дослідження полягає у необхідності вдосконалення існуючих підходів до класифікації дефектів вихідного коду з метою своєчасного попередження можливих проблем з його виконанням та потенційних вторгнень. **Об'єктом дослідження** є процес автоматизованого виявлення дефектів програмного коду. **Предметом дослідження** є класифікаційні методи ансамблевого машинного навчання. **Метою дослідження** є підвищення якості ансамблевих класифікаторів в задачах класифікації дефектів програмного коду. **Завдання дослідження:** проаналізувати характеристики набору даних, побудувати послідовність ансамблевих моделей від одиночних класифікаторів до диверсифікованого ансамблю з метою класифікації потенційних дефектів програмних модулів, підтвердити статистичну значущість приросту якості на кожному етапі. **Методи дослідження:** методи попередньої обробки даних, методи відбору ознак, базові методи машинного навчання, ансамблеві методи, диверсифікаційні техніки, статистичні методи оцінювання. **Результати дослідження.** Експериментально підтверджено монотонне зростання якості класифікації шляхом ускладнення ансамблевої архітектури – від найкращого базового класифікатора Gradient Boosting (CV F1 = 0,798) до запропонованого методу розширеного диверсифікованого стекінгу (CV F1 = 0,824), який перевершив інші моделі за контрольованими метриками при найнижчій дисперсії та найкращому калібруванні ймовірностей. **Висновки.** За результатами дослідження запропоновано метод класифікації дефектів програмного коду, що поєднує диверсифікований стекінг на підпросторах ознак із додатковими декорельованими гілками на основі Extra Trees та адаптивним мета-класифікатором логістичної регресії. Загальний приріст якості наблизив досягнутої точності до теоретичної верхньої межі, зумовленої наявністю шуму в розмітці набору даних.

Ключові слова: дефекти програмного коду; статичні метрики якості; машинне навчання; класифікація; ансамблеві методи; диверсифікація ознак; виявлення вразливостей; мета-класифікатор.

1. Вступ

1.1 Актуальність дослідження

Розробка програмного забезпечення сучасних інформаційних систем ведеться в рамках безперервного циклу інтеграції та доставки (CI/CD). За такого підходу кожна зміна вихідного коду дуже швидко проходить шлях від фрагменту програми до виконуваного файлу, а тому повинна пройти автоматизовану перевірку якості перед потраплянням у живе середовище, щоб гарантувати її надійність та безпечність. У цих умовах дефекти програмного коду залишаються головним джерелом вразливостей безпеки. Згідно зі звітом Національного інституту стандартів і технологій США кількість поданих записів до Національної бази даних вразливостей зросла на 263% упродовж 2020-2025 років, сформувавши стійке відставання в їх опрацюванні [1]. Крім того, за даними IBM середні світові втрати організації від одного інциденту порушення захисту даних у 2025 році склали 4,44 млн доларів за середньої тривалості життєвого циклу інциденту 241 день [2].

Провідні позиції переліку в рейтингу CWE Top 25 стабільно посідають саме дефекти рівня вихідного коду [3], тобто категорії, які можна потенційно виявити на етапі статичного аналізу коду перед його збиранням для подальшого виконання на цільових пристроях. Виявлення потенційних дефектів на цьому кроці дозволить знизити вартість усунення наслідків та зберегти репутацію.

Сукупність цих факторів зумовлює актуальність розробки автоматизованих класифікаційних систем, здатних з прийнятною точністю розрізняти дефектні та недефектні модулі на основі обчислюваних характеристик вихідного коду, а саме класичних

статичних метрик якості Чідамбера-Кемерера для об'єктно-орієнтованого коду (зокрема, складність класу WMC, глибину успадкування DIT, зчеплення CBO, відгук класу RFC, зв'язність LCOM, тощо) [4].

З урахуванням зазначених факторів стає зрозумілим, що сучасні автоматизовані системи виявлення вторгнень в комп'ютерні системи через програмне забезпечення потребують високих показників якості розпізнавання, стабільності результатів між етапами життєвого циклу ПЗ та швидкодію, яка не буде сповільнювати робочі бізнес процеси. Як показують попередні дослідження найкраще з цими задачами впораються методи машинного навчання, а саме ансамблеві класифікатори. Вони структурно краще за одиночні моделі і дозволяють поєднати прогнози кількох базових класифікаторів та знизити дисперсію помилки, підвищити здатність до узагальнення та отримати більш надійні оцінки ймовірностей [5-7].

За таких умов особливо актуальним стає відкрите питання побудови ансамблів для задач класифікації дефектів на основі статичних метрик, а також процес вибору базових моделей і процесу агрегації їх прогнозів з урахуванням складних нелінійних залежностей між статичними метриками і потенційною вразливістю. Найпоширеніший підхід, просте голосування по кількох базових моделях, не гарантує приросту якості над найкращою одиночною моделлю, оскільки структурно близькі базові класифікатори (наприклад, дерево-подібні методи) допускають корельовані помилки. Стекінг із мета-класифікатором дає кращі результати, проте типова практика обмежується невеликою кількістю гомогенних базових моделей на повному наборі ознак, що не реалізує потенціал диверсифікації [8-9].

Виходячи з наведених фактів, постає нагальна потреба у розробці методу ансамблевої класифікації дефектів програмного коду, який цілеспрямовано керує різноманітністю базових класифікаторів як на рівні набору даних та підпростору ознак, так і на рівні алгоритмічних родин базових моделей, поєднуючи це з адаптивним мета-навчанням для гарантованого монотонного підвищення якості на кожному етапі ускладнення архітектури.

1.2 Аналіз літературних джерел і визначення проблеми

Виявлення дефектів програмного коду на основі статичних метрик якості є усталеним напрямом досліджень у програмній інженерії. Класичні роботи спираються на набір метрик Чідамбера-Кемерера та похідні від них характеристики, які кількісно описують складність, зчеплення та зв'язність об'єктно-орієнтованого коду.

Перші підходи до класифікації базувались на одиночних статистичних та машинно-навчальних моделях – логістичній регресії, наївному баєсівському класифікаторі, методі опорних векторів. Такі моделі є інтерпретованими, проте демонструють обмежену точність на даних з мультиколінеарними ознаками та нелінійними залежностями, характерними для метрик складності коду. Подальший розвиток напряму пов'язаний із застосуванням методів машинного навчання, зокрема дерево-подібних методів та їх ансамблевих варіацій [10]. Сюди відносять метод випадкового лісу та методи градієнтного бустингу. Ці методи природним чином враховують нелінійні залежності та надлишковість ознак, однак, як показують емпіричні дослідження, прогнози дерево-подібних моделей у межах одного ансамблю схильні до корельованих помилок, що обмежує ефект від подальшого об'єднання таких моделей простим голосуванням.

Окремий напрям становлять гетерогенні ансамблі на основі стекінгу [11, 12], у яких прогнози кількох структурно різних базових моделей передаються мета-класифікатору, що навчається оптимально комбінувати їх. Стекінг, на відміну від голосування, здатний автоматично занижувати вагу слабких або корельованих гілок, що робить його більш стійким до різноманітності якості базових моделей. Водночас типова практика побудови стекінгу обмежується невеликою кількістю базових моделей на повному наборі ознак, не використовуючи диверсифікацію на рівні підпросторів ознак.

Питання керованої диверсифікації ансамблю активно досліджується у суміжних задачах класифікації (виявлення шахрайства, медична діагностика, виявлення мережових вторгнень [5]), де застосовуються метрики попарної розбіжності прогнозів, Q-статистика та Jaccard-відстань множин помилок для кількісної оцінки незалежності помилок базових класифікаторів. Показано, що цілеспрямоване підвищення різноманітності пулу класифікаторів – за рахунок різних підпросторів ознак, різних алгоритмічних родин або різних підвибірок навчальних даних – здатне покращити

якість ансамблю понад рівень, досяжний простим збільшенням кількості гомогенних базових моделей [13, 14]. Незважаючи на наявні результати, відкритим залишається питання систематичного поєднання трьох рівнів диверсифікації – підпросторів ознак [15–17], алгоритмічних родин та адаптивного мета-навчання [18, 19] – в єдиній архітектурі, яка б гарантувала монотонне підвищення якості класифікації дефектів програмного коду на кожному етапі ускладнення. Це й визначає напрям поточного дослідження.

1.3 Мета і завдання дослідження

Основна мета поточного дослідження полягає в розробці методу ансамблевої класифікації дефектів програмного коду на основі статичних метрик якості, який забезпечує контрольоване та монотонне підвищення якості розпізнавання при переході від одиночних класифікаторів до ансамблів зростаючої складності. Для досягнення поставленої мети необхідно дослідити наступні напрямки:

1. Аналіз характеристик набору даних – баланс класів, мультиколінеарність ознак, наявність шуму в розмітці – та визначення теоретичної стелі досяжної точності.

2. Дослідження ефективності класичних одиночних класифікаторів машинного навчання як орієнтиру для подальших ансамблевих побудов.

3. Перевірка здатності найпростіших ансамблевих методів з голосуванням перевищити якість найкращого одиночного класифікатора.

4. Розробка диверсифікованого стекінгу на основі кількох незалежних методів відбору ознак (кореляція Пірсона, взаємна інформація, LASSO, RFE) з архітектурною гарантією не гіршого результату, ніж у наївного стекінгу.

5. Розробка та обґрунтування розширеного диверсифікованого ансамблевого алгоритму на базі стекінгу, що доповнює диверсифікований стекінг алгоритмічно відібраними декорельованими гілками.

6. Експериментальна перевірка статистичної значущості приросту якості на кожному етапі за критерієм Вілкоксона та формування зведеного монотонного ланцюжка покрокового вдосконалення.

Передбачається, що виконання поставлених кроків дозволить отримати метод класифікації дефектів програмного коду, ефективність якого послідовно та статистично обґрунтовано зростає на кожному рівні ускладнення ансамблевої архітектури, а також сформулювати рекомендації щодо застосування диверсифікованих ансамблів у задачах статичного аналізу якості коду.

2. Матеріали і методи дослідження

Об'єктом дослідження є процес автоматизованого виявлення дефектів програмного коду. Предметом дослідження є класифікаційні методи ансамблевого машинного навчання.

Нехай $f(x)$ – модель класифікації, що приймає вхідний вектор статичних метрик якості $x \in \mathbb{R}^N$, де N – кількість обраних метрик, і повертає оцінку ймовірності належності фрагменту програмного

коду до класу «дефектний». В такому випадку основною задачею стає формування вирішального правила бінарної класифікації:

$$\hat{y} = 1[P(\text{defect} | x) > t], \quad (1)$$

де $\hat{y}$ – передбачена мітка класу; $1[\cdot]$ – індикаторна функція; $P(\text{defect} | x)$ – апостеріорна ймовірність дефекту; t – порогове значення (за замовчуванням $t = 0,5$). Ідея використання ансамблевих класифікаторів передбачає поєднання прогнозів декількох базових моделей за допомогою процедури агрегації:

$$F(x) = g(f_1(x), f_2(x), \dots, f_M(x); \theta), \quad (2)$$

де $f_1(x), \dots, f_M(x)$ – прогнози M базових класифікаторів для вхідного вектора x ; $g(f; \theta)$ – функція агрегації (мета-алгоритм) із параметрами θ , що поєднує прогнози базових моделей у підсумкове рішення $F(x)$; конкретний вигляд g і природа θ визначаються обраною схемою ансамблювання.

При цьому відбувається декомпозиція очікуваної помилки узагальнення на зміщення, дисперсію та незвідний шум:

$$E[(\hat{y} - y)^2] = \text{Bias}^2(\hat{y}) + \text{Var}(\hat{y}) + \sigma^2, \quad (3)$$

де математичне сподівання береться за реалізаціями навчальної вибірки D , $\hat{y} = \hat{y}(x; D)$ – прогноз моделі, побудованої на цій вибірці; σ^2 – незвідна помилка, зумовлена шумом у даних.

Агрегація прогнозів декількох слабо корельованих класифікаторів зменшує доданок дисперсії, не збільшуючи зміщення. Саме тому ключовим фактором ефективності ансамблю є взаємна незалежність помилок його членів. Для оцінки якості класифікації використовується низка класичних метрик [20]. Основною з них обрано F1-міру як збалансований показник точності за умови рівної важливості обох класів:

$$F_1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}, \quad (4)$$

Для оцінки якості ранжування незалежно від обраного порогу використовується площа під ROC-кривою, що визначається наступним чином з урахуванням частки істинно позитивних та хибно позитивних рішень:

$$\text{AUC}_{\text{ROC}} = \int_0^1 \text{TPR}(t) d\text{FPR}(t), \quad (5)$$

Крім того, в якості популярного збалансованого показника якості класифікації також прийнято брати коефіцієнт кореляції Метьюза, який є стійким до оптимістичних зміщень при незбалансованих помилках та розраховується наступним чином:

$$\begin{aligned} MCC &= \\ &= \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}. \end{aligned} \quad (6)$$

З метою визначення якості калібрування ймовірнісних оцінок, критичної для встановлення операційних порогів блокування у CI/CD-системах, використовується оцінка за допомогою метрики Brier-score, для якої менше значення свідчить про кращі результати:

$$\text{Brier} = \frac{1}{N} \sum_{i=1}^N (p_i - y_i)^2, \quad (7)$$

де p_i – передбачена ймовірність дефекту для i -го зразка, y_i – істинна мітка. Стабільність моделі між релізами оцінюється через стандартне відхилення F1-міри по групах крос-валідації.

Довірчі інтервали 95 % для метрик на ізольованій тестовій вибірці оцінено методом бутстрепа за 1000 повторними вибірками з поверненням. Набір розглянутих метрик дозволяє комплексно оцінити ефективність роботи класифікаційних моделей в процесі їх навчання та налаштування.

3. Результати розробки і дослідження методу виявлення дефектів програмного коду

3.1 Математична модель процесу формування диверсифікаційно-орієнтованого ансамблю

В рамках дослідження для вирішення сформульованої проблеми пропонується побудувати ансамблевий класифікатор з урахуванням диверсифікації базових моделей за рахунок незалежного відбору підпросторів ознак та декорельованих алгоритмічних гілок, що забезпечує зниження дисперсії помилки узагальнення без збільшення зміщення. Процес формування такого класифікатора формалізується як двоетапна задача оптимізації, яка включає вибір архітектури ансамблю та навчання параметрів агрегуювальної функції.

Нехай $P = \{f_1, f_2, \dots, f_N\}$ – пул кандидатів базових класифікаторів, отриманих навчанням різних алгоритмів на повному наборі ознак та на підпросторах ознак, відібраних незалежними методами. Задача першого етапу полягає у виборі підмножини $S \subseteq P$ потужності K , яка одночасно максимізує середню якість і середню взаємну різноманітність гілок:

$$S^* = \underset{S \subseteq P, |S|=K}{\operatorname{argmax}} \left[\frac{1}{|S|} \sum_{i \in S} Q(i) + \lambda \cdot D(S) \right], \quad (8)$$

де $Q(i)$ – нормалізована оцінка якості кандидата i (наприклад, $F1_{\text{norm}(i)}$); $D(S)$ – середня попарна розбіжність прогнозів усередині обраної підмножини S (означена нижче виразом (9)); $\lambda \geq 0$ – коефіцієнт балансу між якістю та різноманітністю на рівні підмножини (не слід плутати з коефіцієнтом α жадібного алгоритму, який діє на рівні окремого кандидата).

$$D(S) = \frac{2}{|S|(|S| - 1)} \cdot \sum_{i < j, i, j \in S} \text{dis}(i, j), \quad (9)$$

Пряме розв'язання задачі (8) вимагає перебору $C(N, K)$ підмножин $i \in \text{NP}$ -складним навіть для помірних N . Це обґрунтовує застосування жадібного наближеного алгоритму, який послідовно нарощує S , додаючи на кожному кроці кандидата з найвищим

приростом комбінованого критерію – тобто виконує локально-оптимальну апроксимацію задачі (8).

Другий етап – навчання параметрів θ агрегуювальної функції g на вже зафіксованій архітектурі S^* :

$$\theta^* = \operatorname{argmin}_{\theta} N^{-1} \times \sum_{n=1}^N L(y_n, g(f_1(x_n), \dots, f_K(x_n); \theta)) + \Omega(\theta), \quad (10)$$

де $L(\cdot; \cdot)$ – функція втрат (логістична втрата для бінарної класифікації); $\Omega(\theta)$ – регуляризаційний доданок, що обмежує складність мета-класифікатора; N – обсяг навчальної вибірки.

Підсумкова модель ансамблю визначається композицією розв'язків обох етапів:

$$F(x) = g(f_1(x), \dots, f_K(x); \theta^*), \{f_1, \dots, f_K\} = S^*. \quad (11)$$

Такий поділ на архітектурну (8) і параметричну (10) підзадачі формально відображає ключову відмінність запропонованого методу від наївних ансамблів: наївний стекинг розв'язує лише другу підзадачу на фіксованому $S = P$, тоді як диверсифікаційно-орієнтований ансамбль розв'язує обидві послідовно, що й забезпечує архітектурну гарантію не гіршого результату. Таким чином, фінальну схему запропонованого диверсифікаційно-орієнтованого ансамблю представлено на рис. 1.

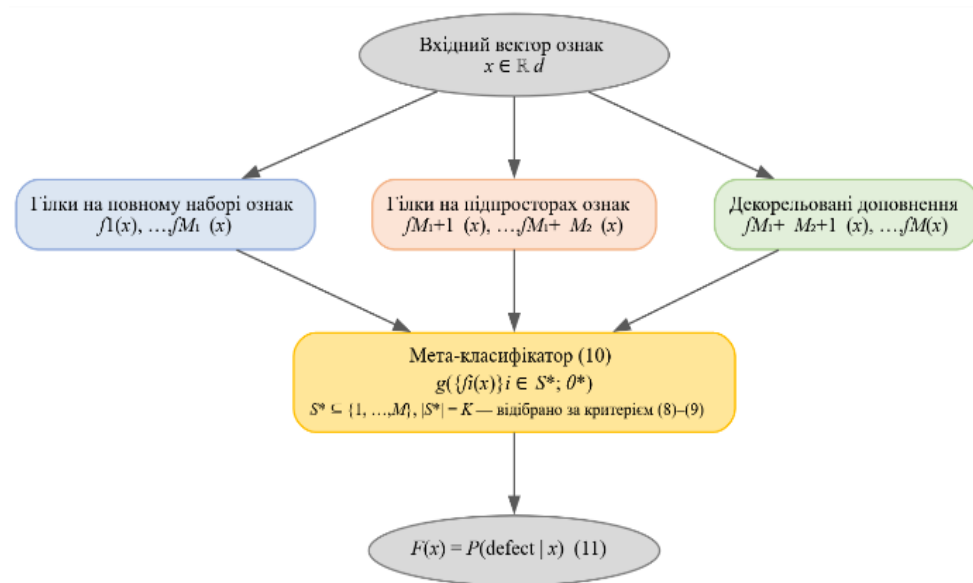


Рис. 1. Запропонована архітектура диверсифікаційно-орієнтованого ансамблю

3.2 Аналіз набору даних

Для проведення експериментального дослідження використано набір даних code_defects_identification [21], що містить 6052 спостереження та 21 статичну метрику якості коду, представлені у табл. 1, а також бінарну цільову мітку.

Важливою характеристикою набору даних є наявність шуму в розмітці: серед 6052 спостережень виявлено 286 дублікатів за вектором ознак, які мають різні мітки класу. Вони утворюють 150 конфліктних груп, що визначає теоретичну верхню межу точності, досягну будь-яким алгоритмом, незалежно від його складності [22]:

$$Accuracy_{max} = 1 - N_{conf}/N \approx 0,975, \quad (12)$$

де N_{conf} – кількість конфліктних спостережень, а $N = 6052$ – загальний обсяг вибірки.

Подальший аналіз ознак виявив суттєву мультиколінеарність: 3 пари ознак мають коефіцієнт кореляції $|r| > 0,9$, а середнє позадіагональне значення $|r|$ по всій матриці кореляцій складає 0,47, що представлено на рис. 2. Висока мультиколінеарність змусила застосувати методи відбору підпросторів ознак для зменшення надлишковості інформації, що використовується базовими класифікаторами.

Попередня обробка даних включала базові кроки з аналізу пропусків та викидів, логарифмічну трансформацію усіх 21 числових ознак для зменшення правосторонньої асиметрії розподілів з подальшою стандартизацією.

Таблиця 1 – Огляд метрик якості коду

Група	Метрики	Призначення
Складність	wmc, loc, totalMethods, totalFields, maxNestedBlocks, loopQty, comparisonsQty, tryCatchQty, returnQty	Обсяг і структурна складність класу
Зчеплення та спадковість	cbo, rfc, nosi, dit	Залежність від інших класів
Зв'язність	lcom	Узгодженість методів
Лексика та стиль коду	variablesQty, assignmentsQty, mathOperationsQty, numbersQty, stringLiteralsQty, parenthesizedExpsQty, uniqueWordsQty	Насиченість і читабельність

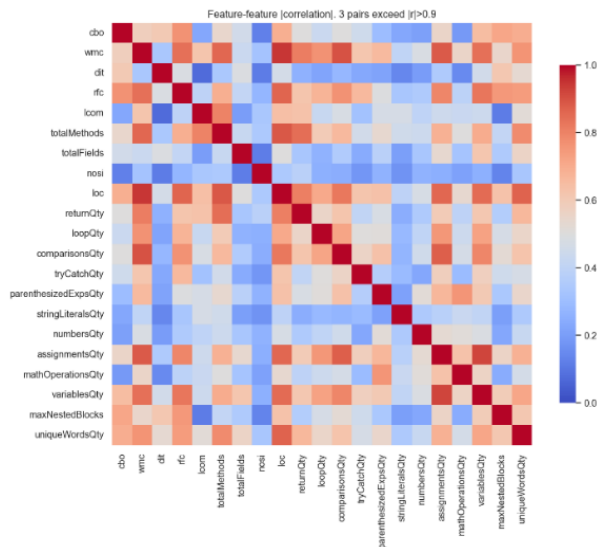


Рис. 2. Теплова карта кореляцій 21 ознаки

Після цього вибірку було розбито на навчальну (4841 зразок) та тестову (1211 зразків) частини у співвідношенні 80/20 зі збереженням пропорцій класів. Для побудови програмної реалізації та навчання моделей машинного навчання була використана мова програмування Python та бібліотека scikit-learn [23]. Аналіз результатів попередньої обробки даних показав, що найбільший приріст якості отримали моделі, які були чутливі до масштабу ознак і залежать від відстаней або градієнтів. Так F1 міра у

KNN зросла з 0,614 до 0,732, а у SVM – з 0,724 до 0,795. При цьому логістична регресія продемонструвала помірноший приріст з 0,759 до 0,785, тоді як дерево-подібний Gradient Boosting, як і очікувалося з огляду на його інваріантність до монотонних перетворень ознак, майже не відреагував, показавши приріст лише з 0,795 до 0,798.

3.3 Експериментальне дослідження ефективності пропонуваного алгоритму

Експериментальне частина дослідження складалась з п'яти послідовних етапів, які передбачали побудову різних моделей машинного навчання: від простих одиночних, до ансамблевих, в тому числі гетерогенних, структура яких була ускладнена за рахунок запропонованого методу з метою підвищення метрик якості.

Метою **першого етапу** є встановлення базового рівня якості для семи одиночних класифікаторів та визначення найкращої одиночної моделі як орієнтиру для побудови ансамблів в подальшому. Оцінювання виконувалося за допомогою 5-кратної стратифікованої крос-валідації на навчальній вибірці, у той час як тестова вибірка (20 % даних) залишалася ізольованою до фінального етапу порівняння. Найкращою одиночною моделлю виявився градієнтний бустинг (CV F1 = 0,798 ± 0,008). На другому місці була модель SVM зі значенням F1 (0,795), проте вона суттєво довше працювала в режимі тестування. Результати першого етапу наведені на рис. 3.

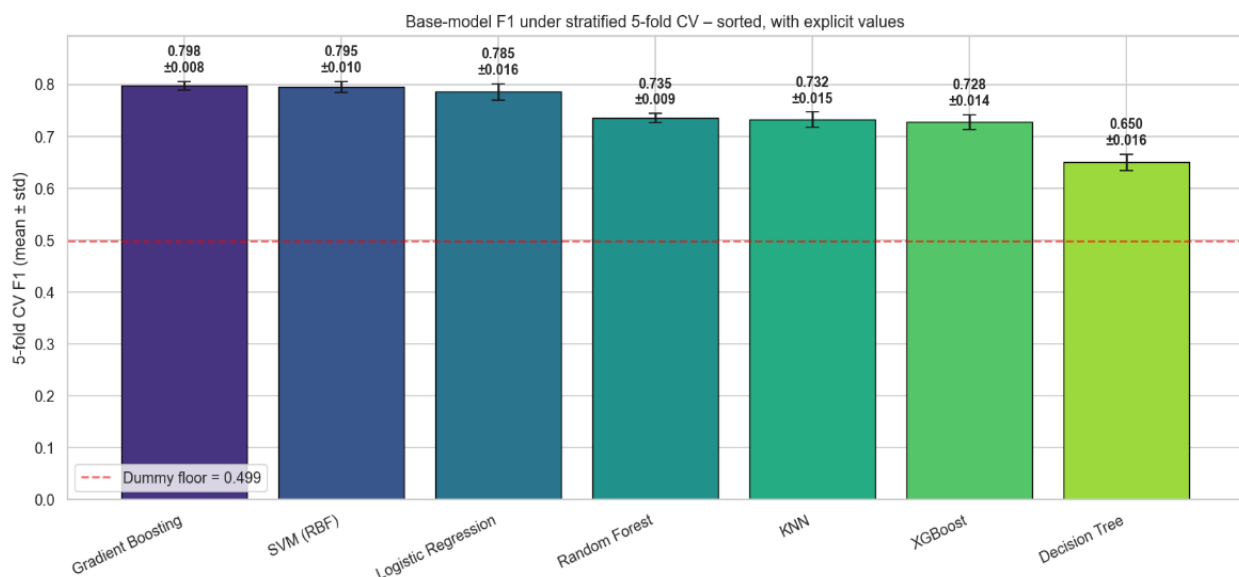


Рис. 3. Гістограма з результатами першого етапу дослідження

Метою **другого етапу** була перевірка того, чи дає проста агрегація гетерогенних базових класифікаторів приріст якості над найкращою одиночною моделлю. Для побудови наївних ансамблів обрано 6 базових класифікаторів (Logistic Regression, Random Forest, Gradient Boosting, XGBoost, SVM, KNN) на повному наборі ознак, поєднаних трьома способами: голосуванням за більшістю, усередненням ймовірностей та стекінгом із мета-класифікатором на основі логістичної регресії. За критерієм Вілкоксона

приріст Stacking-LR над Gradient Boosting є межовим за статистичною значущістю: різниця середніх значень F1 = +0,0136, $p = 0,0625$ (мінімально досяжне значення p при $n = 5$ розбиттях перехресної перевірки для одностороннього тесту). Діаграму розмаху F1 міри представлено на рис. 4.

Голосування (як просте, так і з усередненням ймовірностей) не перевищує якості найкращої базової моделі (F1 = 0,783–0,786 проти 0,798 у GB) – просте об'єднання прогнозів без мета-навчання не

гарантує покращення. Натомість стекинг із мета-класифікатором LR дає статистично межовий, але практично значущий приріст: CV F1 зростає з 0,798

до 0,811, Test F1 – з 0,780 до 0,797, ROC-AUC – з 0,811 до 0,859, а Brier score знижується з 0,170 до 0,151.

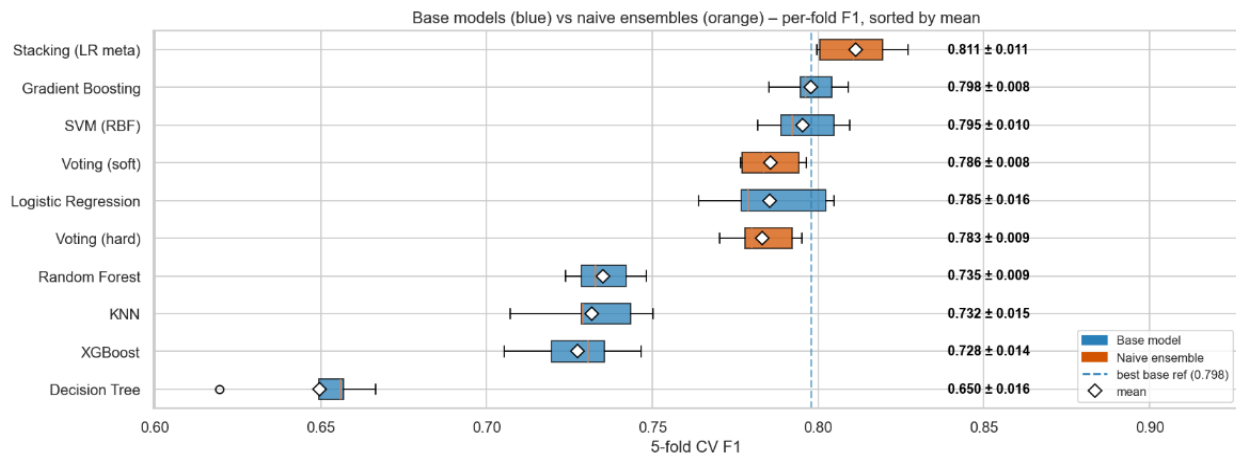


Рис. 4. Діаграма розмаху F1 міри: базові класифікатори (ліворуч) та наївні ансамблі (праворуч)

Обмежена ефективність голосування пояснюється через аналіз різноманітності базових класифікаторів: дерево-подібні методи (випадковий ліс, градієнтний бустинг, екстремальний градієнтний бустинг) кластеризуються за схожістю прогнозів, тобто допускають корельовані помилки. Це знижує ефективний приріст від простого об'єднання їхніх прогнозів. Таким чином, використання стекингу з мета-класифікатором продемонструвало перспективний приріст якості над найкращою одиночною моделлю, тоді як просте голосування такого приросту не забезпечило, що обґрунтовує доцільність подальшого розвитку саме стекингової архітектури на наступних етапах дослідження.

Метою **третього етапу** є підвищення різноманітності ансамблю шляхом поєднання кількох незалежних методів відбору ознак та побудова диверсифікованого стекингу з логістичною регресією в якості мета-класифікатора, здатного перевершити найкращий наївний метод. Для відбору інформативних підмножин ознак застосовано чотири незалежні методи: кореляцію Пірсона ($|r| \geq \text{порогу}$), взаємну інформацію, L1-регуляризацію (LASSO) та рекурсивне виключення ознак (на основі градієнтного бустингу). Для кожного методу побудовано криву CV F1 від кількості ознак k . Лінійна кореляція Пірсона та взаємна інформація ранжують ознаки по-різному, що ілюструє рис. 5.

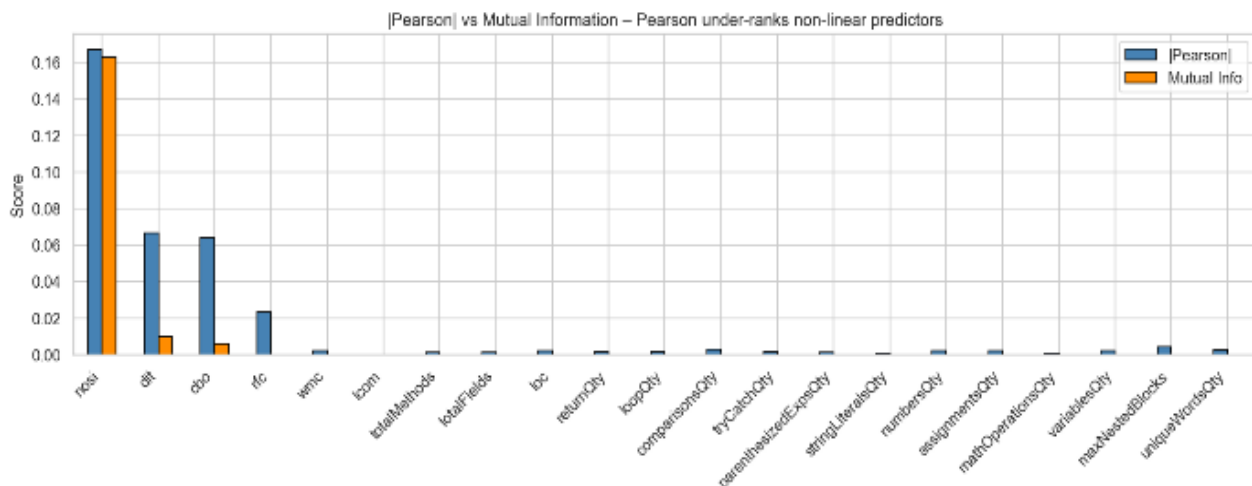


Рис. 5. Порівняння модуля коефіцієнта кореляції Пірсона $|r|$ та взаємної інформації для 21 ознаки

За взаємною інформацією домінує ознака *posi*, тоді як за коефіцієнтом Пірсона порівнянню вагу мають також *dit*, *cbo* та *rfc*. Це підтверджує доцільність використання кількох методів відбору ознак одночасно.

Усі чотири методи конвергують до схожого набору основних ознак (*posi*, *dit*, *cbo*, *rfc*), що

підтверджує їхню справжню інформативність незалежно від конкретного методу оцінки значущості.

Таким чином, було прийнято рішення на **четвертому етапі** побудувати диверсифікований стекинг. Його архітектура складається з 9 гілок: 6 моделей (Logistic Regression, Random Forest, Gradient

Boosting, XGBoost, SVM, KNN) на повному наборі ознак, що забезпечують мінімальний рівень якості не нижче Stacking-LR, та 3 моделі на підпросторах ознак – Logistic Regression на LASSO-ознаках, Random Forest на MI-ознаках, XGBoost на RFE-ознаках. Мета-класифікатором було обрано логістичну регресію. В результаті отримано значення $CV F1 = 0,813$.

Важливим методологічним спостереженням є те, що диверсифіковане голосування без мета-навчання поступається найвчому стекінгу. Це підтверджує ключову гіпотезу дослідження: диверсифікація підпросторів ознак дає приріст лише у поєднанні з мета-класифікатором, здатним оптимально зважувати різні гілки, а не сама по собі.

Перед формуванням фінальної архітектури проведено допоміжне абляційне дослідження. Воно дозволило перевірити, чи здатний алгоритмічний відбір найкомплементарніших класифікаторів із великого пулу моделей з подальшою оптимальною агрегацією голосуванням перевершити диверсифікований стекінг без залучення мета-навчання.

Для цього абляційного дослідження було сформовано пул з 8 родин класифікаторів, які в тому числі розглядалися на попередніх кроках (Logistic Regression, LDA, Naive Bayes, Decision Tree, KNN, Gradient Boosting Trees, Random Forest, XGBoost), по 10 екземплярів кожної родини за допомогою беггінг-ансамблю з різними підвбірками. З цього підготовленого пулу жадібним алгоритмом відбирається $K = 10$ моделей за критерієм, що розраховується наступним чином:

$$Q_{sel}(i) = \alpha \cdot F1_{norm}(i) + (1 - \alpha) \cdot \min_{j \in S} dis(i, j), \quad (13)$$

де $F1_{norm}(i)$ – нормалізована оцінка якості кандидата i ; S – множина вже відібраних моделей; $\alpha \in [0, 1]$ – коефіцієнт балансу між якістю та різноманітністю на рівні окремого кандидата.

Порівняння двох значень коефіцієнта α показало мінімальну різницю: відбір з урахуванням різноманітності ($\alpha = 0,5$, м'яке голосування) дає $CV F1 = 0,806 \pm 0,010$, тоді як відбір лише за якістю ($\alpha = 1,0$) дає $CV F1 = 0,807 \pm 0,009$ – різниця 0,001. Водночас диверсифікація змінює склад відібраного набору моделей (наприклад, з {GBT:8, LR:1, NB:1} на {GBT:7, LR:2, NB:1}) – невеликий, але вимірюваний ефект на структуру ансамблю. Обмеженням підходу Custom Bagging Voter є те, що пул складається з відносно слабких моделей (найкраща окрема модель пулу: $F1 \approx 0,800 < 0,813$ у Div. Stacking), а проста агрегація голосуванням не здатна суттєво перевершити найсильнішого члена пулу.

На останньому п'ятому етапі дослідження було затверджено фінальну архітектуру диверсифікованого ансамблю класифікаторів, яка складалася з 12 гілок: 9 гілок від диверсифікованого стекінгу (гарантія мінімального рівня 0,813) та 3 нові, алгоритмічно відібрані, декорельовані доповнення. Вони включали в себе моделі Extra Trees (300 дерев) на повному наборі ознак, Extra Trees (200 дерев) на підпросторі взаємної інформації та логістична

регресія зі слабшою регуляризациєю. Мета-класифікатором обрано логістичну регресію з адаптивним підбором сили регуляризації.

Обґрунтування вибору Extra Trees полягає в тому, що, на відміну від Gradient Boosting та Random Forest, цей алгоритм обирає пороги розбиття вузлів дерев цілком випадково, а не за інформаційним критерієм, що генерує суттєво відмінні поверхні рішення та зменшує кореляцію помилок з іншими дерево-подібними методами в ансамблі.

Статистична значущість приросту: критерій Вілкоксона Ext. Div. Stacking проти Div. Stacking: $\Delta mean F1 = +0,0115$, $p = 0,0625$.

Ext. Div. Stack перевершує Div. Stack за усіма контрольованими метриками одночасно. Найбільший приріст спостерігається за ROC-AUC (з 0,855 до 0,879) та за зниженням дисперсії між розбиттями перехресної перевірки на 40 %. Значення ст. відхилення = 0,006 є найнижчим серед усіх протестованих у дослідженні моделей, що робить Ext. Div. Stacking найстабільнішим класифікатором. Значення Brier = 0,140 – найкраще калібрування серед усіх моделей, тобто ймовірнісні оцінки Ext. Div. Stacking найбільш придатні для встановлення операційних порогів у CI/CD-системах.

Важливим результатом дослідження є монотонне зростання якості класифікації на кожному етапі ускладнення ансамблевої архітектури – від найкращого одиночного класифікатора до запропонованого методу Ext. Div. Stacking (табл. 2).

Таблиця 2 – Порівняння приросту якості на різних етапах дослідження

Найкраща модель	CV F1	Test F1	p-value
Gradient Boosting	0,798 ± 0,008	0,780	–
Stacking-LR	0,811 ± 0,011	0,797	0,0625
Div. Stacking	0,813 ± 0,010	0,798	0,3125
Ext. Div. Stacking	0,824 ± 0,006	0,811	0,0625

Загальний приріст на шляху від Gradient Boosting до Ext. Div. Stacking складає 0,026 за CV F1 (з 0,798 до 0,824) та 0,031 за Test F1 (з 0,780 до 0,811). Важливо, що етапи дослідження продемонстрували покращення метрик якості, що візуалізовано на рис. 6. Для остаточного підтвердження результатів чотири відібрані методи були порівняні на ізольованій тестовій вибірці. Результати цієї перевірки представлені у табл. 3. Фінальна модель розширеного диверсифікованого ансамблю продемонструвала найвищі значення F1, ROC-AUC, PR-AUC та MCC, а також найнижчі (найкращі) значення Brier score та ст. відхилення серед усіх чотирьох фіналістів, що підтверджує його переваги не лише на крос-валідації, а й на повністю незалежній тестовій вибірці. Окрім якості класифікації, для практичного застосування у CI/CD-системах важливою є обчислювальна вартість моделі, яка

характеризується часом навчання та часом тестування, що графічно представлені на рис. 7. Фінальна модель розширеного диверсифікованого ансамблю приблизно у 6 разів повільніша за

навчанням, ніж Gradient Boosting (5,45 с проти 0,90 с), проте, як вже зазначалося раніше, має кращі показники якості, а саме забезпечує зростання Test F1 з 0,780 до 0,811.

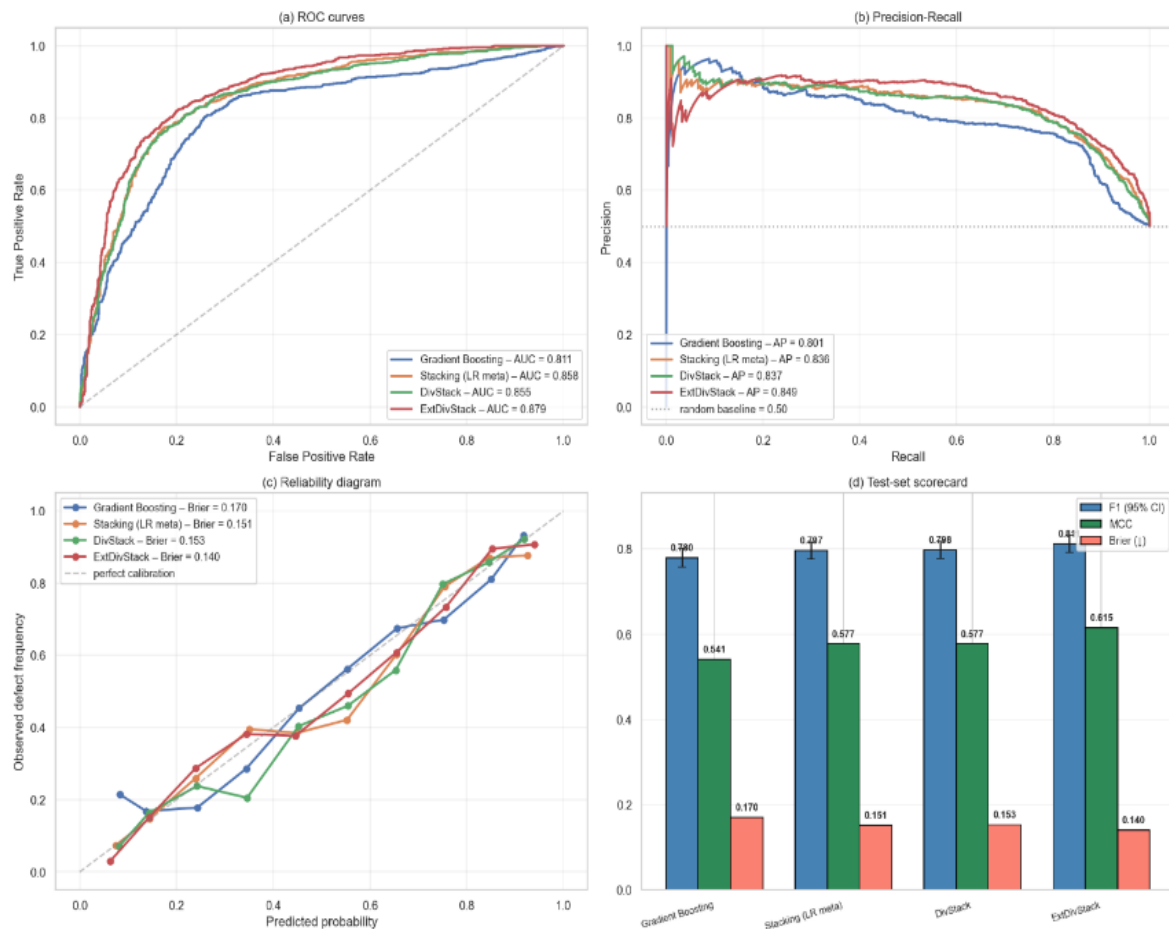


Рис. 6. Порівняння метрик якості найкращих моделей кожного з етапів:
(а) ROC-криві з AUC; (б) PR-криві; (с) крива калібровки; (д) гістограми F1, MCC та Brier

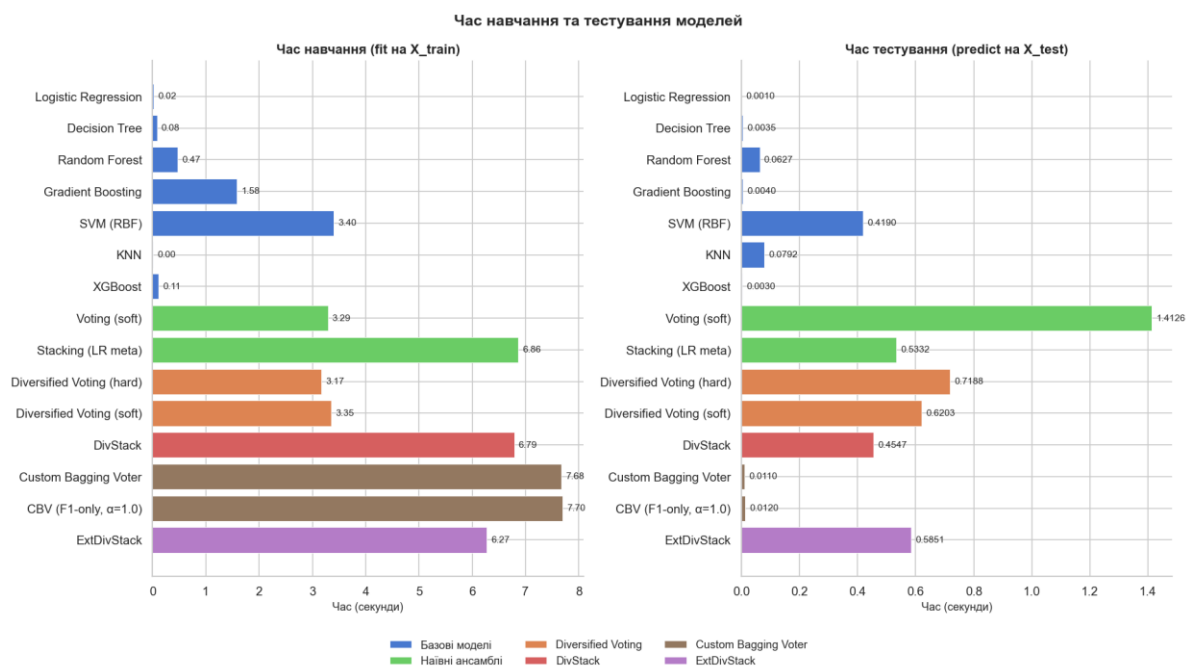


Рис. 7. Час навчання (ліворуч) та час тестування (праворуч)
для всіх досліджених моделей з кольоровим кодуванням за етапами

Таблиця 3 – Порівняння відібраних моделей на ізольованій тестовій вибірці

Метрика	GB	Stack LR	Div. Stack	Ext. Div. Stack
F1	0,780	0,797	0,798	0,811
95% ДІ	0,758-0,801	0,777-0,816	0,778-0,818	0,792-0,831
ROC-AUC	0,811	0,858	0,855	0,879
PR-AUC	0,801	0,836	0,837	0,849
MCC	0,541	0,577	0,577	0,615
Brier ↓	0,170	0,151	0,153	0,140
Ст. відхилення	0,008	0,011	0,010	0,006

Час тестування такої моделі (0,522 с) є прийнятним для асинхронної безпекової перевірки в режимі реального часу у CI/CD-конвеєрі та не блокує основний процес збірки, а тому задовольняє вимогам.

4. Обговорення результатів дослідження методу виявлення дефектів програмного коду

Отримані результати підтверджують, що цілеспрямоване керування різноманітністю ансамблю є дієвим і статистично обґрунтованим шляхом підвищення якості класифікації дефектів програмного коду.

Ключовим досягненням є демонстрація монотонного зростання якості на кожному етапі ускладнення ансамблевої архітектури: від найкращого одиночного класифікатора Gradient Boosting (CV F1 = 0,798) до запропонованого розширеного диверсифікованого ансамблю (CV F1 = 0,824).

В роботі продемонстровано, що приріст досягнуто не екстенсивним нарощуванням кількості моделей, а якісною зміною складу ансамблю за рахунок його диверсифікації. Додавання трьох декорельованих гілок на основі Extra Trees дало приріст ROC-AUC з 0,855 до 0,879 і водночас знизило дисперсію між окремими розбиттями перехресної перевірки на 40%, тобто підвищило не лише точність, а й стабільність моделі, що є важливою вимогою для практичного застосування в CI/CD-конвеєрах. Найкраще серед усіх протестованих моделей значення Brier score (0,140 проти 0,170 у Gradient Boosting) підтверджує також практичну придатність методу для встановлення операційних порогів автоматичного блокування збірки.

При аналізі отриманих результатів варто враховувати, що досяжна якість класифікації природно обмежена наявним у наборі даних шумом розмітки: 150 конфліктних груп спостережень формують верхню межу частки правильних рішень (Accuracy) $\approx 0,975$, до якої запропонована модель уже наближається, і подальше зростання точності на цьому наборі даних значною мірою залежатиме від

якості самої розмітки, а не лише від архітектури класифікатора. Крім того, приріст якості закономірно супроводжується дещо вищою обчислювальною вартістю навчання порівняно з базовим Gradient Boosting, що є звичним і очікуваним компромісом для ансамблевих методів і легко враховується при виборі конфігурації під конкретні ресурсні умови.

Перспективними напрямками подальших досліджень є перевірка узагальнюваності запропонованого методу на інших мовах програмування, типах проєктів і категоріях дефектів, зокрема вразливостях, а не лише структурних помилках, а також валідація на сучасних датасетах із суворішим контролем якості розмітки, що дозволило б точніше відокремити ефект архітектурних удосконалень від впливу шуму в даних. Не менш цікавим напрямком є перевірка стійкості диверсифікованого ансамблю до навмисних атак на класифікатор, оскільки декорельовані гілки за своєю природою ускладнюють перенесення змагальних збурень між моделями, що відкриває шлях до застосування подібної архітектури не лише для класифікації дефектів, а й для виявлення шкідливого коду, згенерованого великими мовними моделями, та розпізнавання спроб обходу автоматизованих систем контролю якості. Розвиток у цьому напрямі дозволив би поєднати задачі забезпечення якості коду та безпеки програмних систем у межах єдиної диверсифікованої ансамблевої архітектури.

5. Висновки

У рамках дослідження розроблено та експериментально обґрунтовано вдосконалений метод класифікації дефектів програмного коду на основі диверсифікованих ансамблів класифікаторів зі статичними метриками якості. Дослідження виконано на наборі даних з 6052 спостережень та 21 статичної метрики зі збалансованим розподілом класів (50/50) та теоретичною верхньою межею точності, зумовленою шумом у розмітці (Accuracy $\approx 0,975$).

Експериментально підтверджено, що проста агрегація гетерогенних базових класифікаторів (голосування) не дає приросту якості над найкращою одиночною моделлю (Gradient Boosting, CV F1 = 0,798), оскільки дерево-подібні базові класифікатори допускають корельовані помилки. Натомість стекінг із мета-класифікатором забезпечує статистично межовий, але практично значущий приріст якості.

Подальша диверсифікація стекінгу на основі підпросторів ознак, відібраних чотирма незалежними методами (кореляція Пірсона, взаємна інформація, LASSO, RFE), дозволила побудувати архітектуру з гарантованим мінімальним рівнем якості (не гірше наївного стекінгу) та незначним, проте стабільним додатковим приростом.

Найбільший приріст якості забезпечив запропонований метод розширеного диверсифікованого стекінгу, що доповнює диверсифікований стекінг алгоритмічно відібраними декорельованими гілками на основі Extra Trees.

Запропонований метод перевершує звичайний диверсифікований стекінг за усіма контрольованими метриками одночасно: CV F1 зріс з 0,813 до 0,824, Test F1 – з 0,798 до 0,811, ROC-AUC – з 0,855 до 0,879, а дисперсія знизилась на 40 %, забезпечуючи найвищу стабільність серед усіх моделей, розглянутих у дослідженні.

Загальний приріст якості на шляху від найкращого базового класифікатора до запропонованого методу Ext. Div. Stacking склав 0,026 за CV F1 та 0,031 за Test F1. Це підтверджує центральну гіпотезу дослідження: цілеспрямоване керування різноманітністю ансамблю на рівні підпросторів ознак та алгоритмічних родин, поєднане з адаптивним мета-навчанням, забезпечує монотонне підвищення якості класифікації дефектів програмного коду.

Практичне значення отриманих результатів дослідження полягає у можливості використання запропонованого метода в автоматизованих системах контролю якості програмного коду в рамках безперервного конвеєра доставки як асинхронний етап оцінки ризиків дефектності певного модуля. При цьому враховано як метрики якості, так і часові характеристики.

Наукова новизна дослідження полягає у тому, що вперше запропоновано архітектуру розширеного диверсифікованого стекінг-ансамблю, у якому диверсифікація здійснюється одночасно за рахунок

трьох компонентів, включаючи підпростір ознак, вибір алгоритмічної родини та адаптивне мета-навчання.

Крім того, набув подальшого розвитку процес застосування критеріїв різноманітності до формування складу ансамблю в задачах виявлення дефектів програмного коду на основі статистичних метрик.

Конфлікт інтересів

Автори декларують, що не мають конфлікту інтересів стосовно даного дослідження, в тому числі фінансового, особистісного характеру, авторства чи іншого характеру, що міг би вплинути на дослідження та його результати, представлені в даній статті.

Фінансування

Дослідження проводилося без фінансової підтримки.

Доступність даних

Рукопис має пов'язані дані у сховищі даних. Набір даних, використаний у дослідженні, є загальнодоступним і вказаний у тексті рукопису з посиланням на першоджерело

Використання засобів штучного інтелекту

Автори підтверджують, що не використовували технології штучного інтелекту при створенні представленої роботи.

СПИСОК ЛІТЕРАТУРИ

1. National Institute of Standards and Technology (2026), "NIST Updates NVD Operations to Address Record CVE Growth", *NVD General Announcement*, available at: <https://www.nist.gov/news-events/news/2026/04/nist-updates-nvd-operations-address-record-cve-growth> (accessed 01.08.2026).
2. IBM Security, and Ponemon Institute (2025), "Cost of a Data Breach Report 2025", *IBM Security*, available at: <https://www.ibm.com/reports/data-breach> (accessed 01.08.2026).
3. MITRE Corporation, and Cybersecurity and Infrastructure Security Agency (2024), "2024 CWE Top 25 Most Dangerous Software Weaknesses", available at: https://cwe.mitre.org/top25/archive/2024/2024_cwe_top25.html (accessed 01.08.2026).
4. Chidamber, S.R., and Kemerer, C.F. (1994), "A Metrics Suite for Object Oriented Design", *IEEE Transactions on Software Engineering*, vol. 20, no. 6, pp. 476–493, doi: <https://doi.org/10.1109/32.295895>
5. Hornostal, O., and Chelak, V. (2025), "Classification of network attacks using machine learning methods in conditions of training data imbalance", *Control, Navigation and Communication Systems*, vol. 3, no. 81, pp. 64–71, doi: <https://doi.org/10.26906/SUNZ.2025.3.064>
6. Gordieiev, O., Gordieieva, D., and Rainer, A. (2023), "Software Quality Assessment: Defect Life Cycle, Software Defect Profile, Its Types and Misalignments", *International Conference on Software Quality*, pp. 109–120, doi: https://doi.org/10.1007/978-3-031-31488-9_6
7. Setia, S., Ravulakollu, K.K., Verma, K., Garg, S., Mishra, S.K. and Sharan, B. (2024), "Software Defect Prediction using Machine Learning", *2024 11th International Conference on Computing for Sustainable Global Development (INDIACom)*, New Delhi, India, pp. 560–566, doi: <https://doi.org/10.23919/INDIACom61295.2024.10498707>
8. Alazba, A., and Aljamaan, H. (2022), "Software defect prediction using stacking generalization of optimized tree-based ensembles", *Applied Sciences*, vol. 12, no. 9, art. 4577, doi: <https://doi.org/10.3390/app12094577>
9. Meng, F., Huang, R., and Wang, J. (2023), "A survey of software defects research based on deep learning", *2023 6th International Conference on Information Systems and Computer Networks (ISCON)*, pp. 1–5, doi: <https://doi.org/10.1109/ISCON57294.2023.10112194>
10. Olivares-Galindo, J. A., Sánchez-García, Á. J., Barrientos-Martínez, R. E., and Ocharán-Hernández, J. O. (2023), "Ensemble classifiers in software defect prediction: a systematic literature review", *2023 11th International Conference in Software Engineering Research and Innovation (CONISOFT)*, pp. 1–8, doi: <https://doi.org/10.1109/CONISOFT58849.2023.00011>
11. Wolpert, D.H. (1992), "Stacked Generalization", *Neural Networks*, vol. 5, no. 2, pp. 241–259, doi: [https://doi.org/10.1016/S0893-6080\(05\)80023-1](https://doi.org/10.1016/S0893-6080(05)80023-1)
12. Mehta, A., Kaur, A., and Kaur, N. (2024), "Optimizing software fault prediction using voting ensembles in class imbalance scenarios", *International Journal of Performability Engineering*, vol. 20, no. 11, pp. 676–687, doi: <https://doi.org/10.23940/ijpe.24.11.p4.676687>
13. Isong, B., and Igo, E. (2025), "Ensemble learning for software defect prediction: performance, practicality and future directions", *Journal of Information Systems and Informatics*, vol. 7, no. 3, pp. 2245–2291, doi: <https://doi.org/10.51519/journalisi.v7i3.1171>

14. Ali, M., Mazhar, T., Al-Rasheed, A., Shahzad, T., Ghadi, Y.Y., and Khan, M.A. (2024), "Enhancing software defect prediction: a framework with improved feature selection and ensemble machine learning", *PeerJ Computer Science*, vol. 10, art. e1860, doi: <https://doi.org/10.7717/peerj-cs.1860>
15. Ho, T.K. (1998), "The Random Subspace Method for Constructing Decision Forests", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 8, pp. 832–844, doi: <https://doi.org/10.1109/34.709601>
16. Tang, Y., Dai, Q., Yang, M., Du, T., and Chen, L. (2023), "Software defect prediction ensemble learning algorithm based on adaptive variable sparrow search algorithm", *International Journal of Machine Learning and Cybernetics*, vol. 14, no. 6, pp. 1967–1987, doi: <https://doi.org/10.1007/s13042-022-01740-2>
17. Gao, Y., and Yang, C. (2016), "Software defect prediction based on manifold learning in subspace selection", *Proceedings of the 1st International Conference on Intelligent Information Processing (ICIIP)*, ACM, New York, pp. 1–6, doi: <https://doi.org/10.1145/3028842.3028859>
18. Gavrylenko, S., and Hornostal, O. (2023), "Study of methods for improving the meta-algorithm of the bagging classifier", *Proceedings of the IEEE 4th KhPI Week on Advanced Technology*, pp. 1–6, doi: <https://doi.org/10.1109/KhPIWeek61412.2023.10312977>
19. Osei, S., Masegosa, A.R., and Masegosa, A.D. (2025), "Understanding the role of diversity in ensemble-based AutoML methods for classification tasks", *IEEE Access*, vol. 13, pp. 63566–63586, doi: <https://doi.org/10.1109/ACCESS.2025.3554093>
20. Chelak, V., Hornostal, O., Chelak, Y., and Gavrylenko, S. (2025), "Advanced methods for classification quality assessment leveraging ROC analysis and multidimensional confusion matrix", *Advanced Information Systems*, vol. 9, no. 1, pp. 24–34, doi: <https://doi.org/10.20998/2522-9052.2025.1.03>
21. Xu, J., Wang, F., and Ai, J. (2021), "Defect prediction with semantics and context features of codes based on graph representation learning", *IEEE Transactions on Reliability*, vol. 70, no. 2, pp. 613–625, doi: <https://doi.org/10.1109/TR.2020.3040191>
22. Croft, R., Babar, M.A., and Kholoosi, M.M. (2023), "Data quality for software vulnerability datasets", *Proceedings of the IEEE/ACM 45th International Conference on Software Engineering (ICSE)*, pp. 121–133, doi: <https://doi.org/10.1109/ICSE48619.2023.00022>
23. Pedregosa, F., Varoquaux, G., Gramfort, A. et al. (2011), "Scikit-learn: machine learning in Python", *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, available at: <https://www.jmlr.org/papers/v12/pedregosa11a.html>

Received (Надійшла) 15.06.2026

Accepted for publication (Прийнята до друку) 26.08.2026

Publication date (Дата публікації) 25.09.2026

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Горносталь Олексій Андрійович – PhD, доцент, доцент кафедри комп'ютерної інженерії та програмування, Національний технічний університет «Харківський політехнічний інститут», Харків, Україна;

Oleksii Hornostal – PhD, Associate Professor, Associate Professor of Department of Computer Engineering and Programming, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;

e-mail: oleksii.hornostal@khpi.edu.ua; ORCID Author ID: <http://orcid.org/0000-0001-5820-9999>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57216331944>.

Челак Віктор Володимирович – PhD, доцент, доцент кафедри комп'ютерної інженерії та програмування, Національний технічний університет «Харківський політехнічний інститут», Харків, Україна;

Viktor Chelak – PhD, Associate Professor, Associate Professor of Department of Computer Engineering and Programming, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;

e-mail: viktor.chelak@khpi.edu.ua; ORCID Author ID: <https://orcid.org/0000-0001-8810-3394>

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57189040595>.

Diversified Classifier Ensembles for Software Defect Prediction Based on Static Quality Metrics

Oleksii Hornostal, Viktor Chelak

Abstract. Background. The relevance of the study lies in the need to improve existing approaches to source code defect classification in order to provide timely warning of potential execution problems and possible intrusions. **The object of the study** is the process of automated detection of software code defects. **The subject of the study** is ensemble machine learning classification methods. **The objective of the study** is to improve the quality of ensemble classifiers in software code defect classification tasks. **The research tasks are:** to analyze the characteristics of the dataset; to construct a sequence of ensemble models, from single classifiers to a diversified ensemble, for the classification of potential defects in software modules; and to confirm the statistical significance of the quality gain achieved at each stage. **The methods used** include data preprocessing techniques, feature selection methods, base machine learning methods, ensemble methods, diversification techniques, and statistical evaluation methods. **Results.** The study experimentally confirmed a monotonic increase in classification quality as the ensemble architecture was progressively complexified – from the best-performing base classifier, Gradient Boosting (CV F1 = 0.798), to the proposed extended diversified stacking method (CV F1 = 0.824), which outperformed the other models across all monitored metrics while achieving the lowest variance and the best probability calibration. **Conclusions.** Based on the results obtained, a method for software code defect classification is proposed that combines diversified stacking over feature subspaces with additional decorrelated branches based on Extra Trees and an adaptive logistic regression meta-classifier. The overall quality gain brought the achievable accuracy closer to the theoretical ceiling imposed by label noise in the dataset.

Keywords: software code defects; static quality metrics; machine learning; classification; ensemble methods; feature diversification; vulnerability detection; meta-classifier.

Tetyana Derkach¹, Andrii Dmytrenko¹, Lina Klochko²

¹National University “Yuri Kondratyuk Poltava Polytechnic”, Poltava, Ukraine

²Geological Bureau BEG SA, Switzerland

A COMPREHENSIVE APPROACH TO ENSURING THE CYBERSECURITY OF CORPORATE NETWORK INFRASTRUCTURE

Abstract. Relevance. The rapid development of digital technologies, the widespread implementation of corporate information systems, and the continuous increase in the number of cyber threats necessitate the improvement of approaches to protecting corporate network infrastructure. Traditional protection methods based on the use of separate software or hardware security tools do not provide an adequate level of resistance to modern cyberattacks, which are characterized by complexity, multiple attack vectors, and the ability to bypass individual security mechanisms. Therefore, the development of a comprehensive multi-layered cybersecurity system aimed at ensuring the confidentiality, integrity, and availability of the information resources of enterprises, institutions, and organizations is highly relevant. **Subject of the research:** modern approaches, models, protocols, and technologies for ensuring the cybersecurity of corporate network infrastructure. **Purpose of the research:** to substantiate a comprehensive approach to designing secure corporate computer networks based on multi-layered protection and the integrated application of modern network security technologies. **Research objectives.** To analyze the role of the Open Systems Interconnection reference model as a methodological basis for designing secure computer networks; to identify the main information security threats affecting corporate networks; to investigate the implementation of security mechanisms at the physical, data link, network, and application layers; to examine the principles of the Defense in Depth concept; to analyze the capabilities of the AAA, RADIUS, SSH, ACL, VLAN, Port Security, DHCP Snooping, and Dynamic ARP Inspection models, protocols, and technologies; and to determine the advantages of their integrated application for improving the cybersecurity of corporate network infrastructure. **Research methods.** The study employs the methods of systems analysis, comparison, generalization, classification, and systematization of scientific and technical information. A structural and functional approach is applied to analyze security mechanisms at different layers of the OSI model. The method of modelling a comprehensive corporate network security architecture is also used. **Research results.** The main threats to the information security of corporate networks are identified, including MAC Spoofing, MAC Flooding, ARP Spoofing, ARP Cache Poisoning, Man-in-the-Middle attacks, unauthorized device connections, the compromise of user and administrator accounts, and violations of access control policies. The Defense in Depth concept, which involves creating several complementary layers of protection against external and internal cyber threats, is analyzed. It is established that the application of a multi-layered approach makes it possible to minimize the consequences of the compromise of individual network components and ensure the continuity of information systems even when particular attacks are successfully implemented. The capabilities of the AAA model for centralized access management to information resources are examined, together with the operating principles of the RADIUS protocol as a means of centralized authentication, authorization, and accounting of user activities. The advantages of using the cryptographically protected SSH protocol for the secure administration of network equipment are identified. The application of Access Control Lists for implementing security policies and controlling communication between corporate network segments is investigated. Port Security, DHCP Snooping, and Dynamic ARP Inspection technologies aimed at countering data-link-layer attacks and unauthorized device connections are analyzed. It is established that the logical segmentation of the network environment using VLAN technology is an effective mechanism for improving cybersecurity, as it enables the localization of cyber incidents, limits the lateral movement of attacks, and supports the implementation of the principle of least privilege. **Conclusions.** The highest level of protection of corporate network infrastructure is achieved through the integrated application of AAA, RADIUS, SSH, ACL, VLAN, Port Security, DHCP Snooping, and Dynamic ARP Inspection technologies in accordance with the principles of the Defense in Depth concept. The proposed approach provides centralized access control, increases the protection of information resources, improves the manageability of the network environment, reduces the risk of cyberattack propagation, and meets modern requirements for designing secure corporate information systems.

Keywords: cybersecurity; corporate network; network infrastructure; information security; Defense in Depth; AAA; RADIUS; SSH; ACL; VLAN; Port Security; Dynamic ARP Inspection; OSI model, access control.

1. Introduction

1.1 Relevance of the Topic

The rapid development of digital technologies, the widespread implementation of corporate information systems, and the increasing number of network services significantly increase the requirements for ensuring the cybersecurity of modern computer networks. Almost all business processes of enterprises, institutions, and organizations operate on the basis of corporate network infrastructure, which provides information exchange among users, servers, databases, and cloud services, the use of which requires a separate cybersecurity risk assessment when selecting a cloud provider [1–3].

At the same time, the expansion of digitalization is accompanied by a constant increase in the number of cyber threats aimed at violating the confidentiality, integrity, and availability of information. Attacks focused on exploiting vulnerabilities in network protocols, configuration errors in active equipment, weaknesses in authentication systems, and deficiencies in access control mechanisms are particularly dangerous. Successful implementation of such attacks may result in the compromise of corporate information resources, disruption of business continuity, financial losses, and reputational risks [4–6].

Under modern conditions, information security can no longer be limited to the use of individual software or

hardware protection tools. A comprehensive approach is required, combining modern technologies for network segmentation, centralized access management, cryptographic protection of data transmission channels, monitoring systems, and control of network activity. Therefore, the study of modern threats to corporate computer networks and the development of effective mechanisms for their neutralization are among the current areas of cybersecurity and network technology development. The relevance of this approach is confirmed by modern international recommendations for building multilayered protection of network infrastructure [7–9].

1.2 Literature Review and Problem Statement

The issue of ensuring the information security of corporate computer networks is the subject of research by both Ukrainian and foreign scholars. A significant contribution to the development of the theoretical foundations of network security has been made by W. Stallings, A. S. Tanenbaum, J. Postel, D. Comer, as well as the authors of international ISO/IEC standards and IETF recommendations [10, 11].

Practical aspects of building secure corporate networks are addressed in the recommendations of the U.S. National Institute of Standards and Technology (NIST), particularly in the SP 800 series documents, which define modern approaches to building network infrastructure protection systems, applying firewalls, access control policies, and information security management [12].

Considerable attention in modern research is paid to the use of the Defense in Depth concept, centralized user authentication, AAA mechanisms, the RADIUS protocol, network segmentation using VLAN, Access Control Lists (ACL), Port Security technology, and cryptographic protection of administrative access using SSH. However, most scientific works consider individual security tools without a comprehensive analysis of their interaction within a corporate network [13, 14].

Thus, the development of a comprehensive approach to ensuring the cybersecurity of corporate network infrastructure remains relevant. Such an approach combines modern protection technologies while taking into account current threats and the practical features of operating local computer networks [11].

1.3 Aim and Objectives of the Study

The purpose of the article is to study modern threats and vulnerabilities of corporate computer networks, analyze the mechanisms of their occurrence, and substantiate a comprehensive approach to ensuring the cybersecurity of network infrastructure based on the use of modern authentication, access control, network segmentation, and cryptographic protection technologies.

To achieve this purpose, the following objectives were defined:

- to study modern threats to the information security of corporate computer networks;
- to analyze vulnerabilities of protocols at different layers of the OSI model;

- to investigate the implementation mechanisms of the most common network attacks;
- to evaluate the effectiveness of modern technologies for protecting corporate networks;
- to substantiate a comprehensive approach to building secure corporate network infrastructure.

2. Materials and Methods

2.1 Corporate Network Architecture and the OSI Model

The Role of the OSI Model in Building Secure Network Infrastructure. The design of modern corporate computer networks requires a systematic approach to organizing data transmission processes, resource administration, and information security. The conceptual basis for building network systems is the Open Systems Interconnection (OSI) reference model, standardized by the International Organization for Standardization in ISO/IEC 7498-1. This model defines the logical structure of network communications and provides a unified approach to designing, analyzing, and protecting information systems [15].

The OSI model consists of seven interconnected layers: physical, data link, network, transport, session, presentation, and application. Each layer performs a defined set of functions and interacts only with adjacent layers, which ensures architectural modularity and simplifies the implementation of protection mechanisms [10].

In the field of cybersecurity, the OSI model is used as a methodological basis for identifying potential threats, analyzing vulnerabilities, and implementing protection tools at different stages of information processing. This approach makes it possible to form a multilayered security system in which the compromise of one layer does not lead to a complete disruption of network infrastructure functioning [16].

For corporate networks, the ability to correlate specific cyber threats with the corresponding layers of the OSI model is of particular importance.

Levels of Information Protection in a Corporate Network. Effective information security requires the implementation of protective mechanisms at all layers of network interaction. Within the OSI model, each layer has its own set of potential threats and corresponding countermeasures [1].

At the physical layer, the main protection tasks include controlling access to network equipment, cable infrastructure, and server rooms. Physical security is implemented through access control systems, video surveillance, backup power supply, and protection against unauthorized network connections [9].

At the data link layer, the priority direction of protection is ensuring the security of switching equipment and controlling connected devices. The most common threats at this layer include MAC Spoofing, MAC Flooding, ARP Spoofing, and unauthorized connection of user devices. To minimize these risks, Port Security, Dynamic ARP Inspection, DHCP Snooping, and network access authentication mechanisms are used [17].

At the network and transport layers, routing mechanisms and data-flow control play a key role. Protection at these layers is implemented using Access Control Lists (ACL), firewalls, intrusion detection and prevention systems (IDS/IPS), network segmentation, and packet filtering in accordance with established security policies [18].

The application layer is of particular importance for corporate networks because it is where users interact with network services and information resources. Application-layer protection involves the use of authentication and authorization mechanisms, multifactor authentication, centralized account management, cryptographic protection of data transmission channels, and information security event monitoring tools [12].

Implementation of the Defense in Depth Principle in a Corporate Network. One of the key concepts of modern cybersecurity is the Defense in Depth principle, which involves building several complementary security layers to protect information resources from external and internal threats [14].

The main idea of the concept is that no single security mechanism can guarantee absolute protection of network infrastructure.

Therefore, a security system must consist of a set of interconnected organizational, software, and hardware tools, each of which performs its own protective function.

If one layer is compromised, an attacker encounters additional barriers, which significantly complicates the implementation of an attack [7].

2.2 Analysis of Modern Technologies for Protecting Corporate Networks

Effective cybersecurity of corporate network infrastructure requires the integrated use of modern protection technologies that cover different layers of network interaction and implement the principle of multilayered defense (Defense in Depth).

During the design of a secure corporate network, a number of technologies aimed at ensuring access control, protecting network equipment, preventing unauthorized connections, and countering common network attacks were analyzed.

These technologies include AAA, RADIUS, SSH, ACL, Port Security, and Dynamic ARP Inspection [17].

AAA Technology as the Basis of Centralized Access Control. In corporate networks, the TACACS+ protocol may also be used alongside RADIUS to implement the AAA model. It is used for centralized control of authentication, authorization, and accounting, particularly during network equipment administration [19].

One of the key concepts of modern network security is the AAA (Authentication, Authorization, Accounting) model, which provides centralized management of access to information resources.

This model involves the implementation of three interrelated processes: user authentication, authorization of user actions, and accounting of performed operations [12].

Authentication is used to confirm the identity of a user or network device before access to system resources is granted. Authorization determines the list of permitted actions after successful authentication. Accounting ensures the registration and storage of information on all login attempts, executed commands, and user network activity [10, 12].

Use of the RADIUS Protocol for Centralized Authentication. The practical implementation of the AAA model in corporate networks is often carried out using the RADIUS (Remote Authentication Dial-In User Service) protocol. This protocol provides centralized authentication, authorization, and accounting of users during access to network resources [10, 11].

The operating principle of RADIUS is based on interaction among a client, a network device, and an authentication server.

After receiving an access request, the network device transmits the user credentials to the RADIUS server, which verifies their validity and returns a decision to grant or deny access [10, 11].

2.3 Secure Administration of Network Equipment Using SSH

Secure management of network infrastructure is an important component of ensuring the cybersecurity of corporate networks.

Traditional remote administration protocols, such as Telnet, transmit credentials and service information in clear text, creating conditions for their interception and subsequent system compromise [16].

To eliminate these shortcomings, modern networks use the SSH (Secure Shell) protocol, which provides secure remote management of network devices by creating a cryptographically protected communication channel between a client and a server [20].

2.4 Access Control Lists as a Network Traffic Filtering Mechanism

Access Control Lists (ACL) are among the most widespread tools for protecting network infrastructure. This technology makes it possible to regulate the passage of network traffic in accordance with defined security rules and implement access policies among different segments of a corporate network [16].

ACLs are sets of rules that define permitted or prohibited actions for network packets based on parameters such as source IP address, destination IP address, port number, protocol type, and data transmission direction [18].

2.5 Port Security Technology for Protecting Switching Equipment

One of the most common threats at the data link layer is the connection of unauthorized devices to a corporate network. To minimize this risk, Port Security technology is used on switches.

The operating principle of Port Security consists in limiting the number of allowed MAC addresses on a specific physical switch port. The administrator may define permitted devices manually or automatically store the MAC addresses of connected equipment.

When a security policy violation is detected, the switch may perform various actions: generate an incident message, block traffic, or automatically disable the port. This approach makes it possible to effectively counter MAC Spoofing attacks, unauthorized equipment connections, and attempts to gain access to corporate resources.

2.6 Dynamic ARP Inspection as a Means of Countering ARP Attacks

ARP Spoofing and ARP Cache Poisoning attacks pose a significant threat to corporate networks, as they allow an attacker to substitute records in the ARP tables of network nodes and intercept network traffic [16].

To counter these attacks, Dynamic ARP Inspection (DAI) technology is used. It ensures verification of the authenticity of ARP messages within the network. The operating principle of DAI is based on analyzing the correspondence between IP addresses and MAC addresses using a trusted database formed by the DHCP Snooping mechanism.

When an ARP message is received, the switch checks its correctness. If a mismatch is detected between the declared IP address and the physical address of the device, the packet is blocked and the event is recorded in the security log.

3. Development and Research Results

3.1 Proposed Multilayered Security Architecture

In the proposed corporate network architecture, the Defense in Depth principle is implemented through the combination of several protection layers.

At the physical infrastructure layer, access control to server equipment and switching devices is provided. At the data link layer, Port Security and VLAN-based network segmentation are used to limit the possibility of unauthorized connections and lateral movement of attacks [2].

At the network layer, ACLs are applied to regulate interaction among individual network segments and implement the principle of least privilege.

At the application layer, a centralized authentication system based on the RADIUS protocol is implemented, while network equipment administration is performed exclusively through the secure SSH protocol [17].

Continuous monitoring of network activity, analysis of event logs, and auditing of user actions are also important components of multilayered protection. The integrated use of these mechanisms makes it possible to ensure a high level of corporate network resilience against modern cyber threats, reduce the risks of compromising information resources, and increase the overall reliability of information infrastructure functioning [12].

Thus, the OSI model serves as a methodological basis for building secure corporate networks, while the implementation of the Defense in Depth principle ensures the formation of a multilayered security system capable of effectively countering modern threats to information security.

3.2 Implementation of Centralized Access Management Based on AAA and RADIUS

Within the proposed architecture, the AAA model is used as the basis for centralized management of user and administrator access.

Authentication is performed before access to network resources is granted, authorization determines the permitted resources and operations, and accounting ensures the recording of login attempts, executed commands, and user network activity. This approach enables the consistent application of access policies and provides traceability of actions within the corporate network [9].

Within the proposed corporate network architecture, the RADIUS server functions as the central node for authenticating users and network equipment administrators, ensuring the implementation of the principle of centralized access management.

3.3 Implementation of Secure Administration via SSH

In the proposed network architecture, SSH is used as the main mechanism for remote administration of switches, routers, and server equipment, ensuring the secure execution of configuration and diagnostic operations [18].

3.4 Implementation of Inter-Segment Communication Control

In the context of a VLAN-segmented architecture, ACLs perform an important function in controlling inter-segment interaction.

They allow limiting user access to critical servers, regulating interaction among structural departments of an enterprise, and implementing the principle of least privilege [13].

3.5 Implementation of Port Security

Within the proposed architecture, Port Security is used to protect the user ports of access switches, which significantly increases the level of network connection control [17].

3.6 Results of the Comparative Analysis

The results of the comparative analysis of corporate to systematize the obtained results, the selected corporate network protection technologies were compared according to four main criteria: their functional purpose, the corresponding layer of the OSI model, the principal threats they counteract, and the practical advantages of their implementation.

This comparison makes it possible to determine the role of each security mechanism within the proposed multilayered architecture and to identify how the analyzed technologies complement one another.

Particular attention was paid to the distribution of security functions among the application, network, and data link layers.

The results of the comparative analysis are presented in Table 1.

Table 1 – Comparative Characteristics of Corporate Network Protection Technologies

Technology	Main purpose	OSI model layer	Main threats countered	Advantages of use
AAA	Authentication, authorization, and user activity auditing	Application	Unauthorized access, privilege abuse	Centralized access management and user activity auditing
RADIUS	Centralized authentication of users and devices	Application	Use of local accounts, weak access control	Single point of management for security policies
SSH	Secure remote administration	Application	Password interception, Man-in-the-Middle attacks	Data encryption and protection of administrative sessions
ACL	Control and filtering of network traffic	Network	Unauthorized access between network segments	Flexible configuration of access policies
Port Security	Control of device connections to switch ports	Data link	MAC Spoofing, unauthorized equipment connection	Protection of physical network access points
Dynamic ARP Inspection	Verification of ARP message authenticity	Data link	ARP Spoofing, ARP Poisoning, MITM attacks	Protection of ARP interaction integrity
VLAN	Logical network segmentation	Data link	Lateral movement of attacks, data leakage	Traffic isolation and incident localization
DHCP Snooping	Control of DHCP traffic	Data link	Rogue DHCP, DHCP Starvation	Formation of a trusted IP–MAC correspondence database

Network protection technologies are presented in Table 1.

As shown in Table 1, none of the considered technologies provides complete protection of a corporate network by itself.

Each mechanism is aimed at neutralizing a particular class of threats and operates at the corresponding layer of the OSI model. The highest effectiveness is achieved through the integrated use of access control, network segmentation, and network interaction protection technologies. Such an approach makes it possible to implement a multilayered cybersecurity system that meets modern requirements for corporate network infrastructure.

Thus, the results of the analysis indicate that the integrated use of AAA, RADIUS, SSH, ACL, Port Security, and Dynamic ARP Inspection forms a multilayered protection system for a corporate network.

The combination of these mechanisms provides centralized access control, secure administration of network equipment, protection against data-link-layer attacks, and effective management of information flows, which corresponds to modern requirements for building secure corporate information systems [9].

4. Discussion of Research Results

The use of the AAA model significantly increases the information security level of a corporate network due to centralized access-right management, reduced risk of unauthorized actions, and the possibility of security event auditing. In modern corporate networks, AAA is a basic

component of Identity and Access Management (IAM) systems.

The main advantage of using RADIUS is the centralization of account management and security policy processes.

An administrator is able to manage access to all network resources from a single center, which considerably simplifies administration and increases the level of control over information infrastructure. In addition, the use of RADIUS makes it possible to implement detailed auditing of user actions and ensure compliance with the information security requirements of the corporate environment [12].

The main advantages of SSH are confidentiality of transmitted data, protection against credential interception, information integrity, and authentication of the communicating parties. The use of SSH makes it possible to significantly reduce the risks of Man-in-the-Middle attacks and unauthorized access to network equipment [16].

In the proposed network architecture, SSH is used as the main mechanism for remote administration of switches, routers, and server equipment, ensuring the secure execution of configuration and diagnostic operations.

The application of ACLs significantly reduces the risks of lateral movement within a corporate network and increases the effectiveness of information flow control.

The use of Dynamic ARP Inspection makes it possible to prevent Man-in-the-Middle attacks, ensure the integrity of addressing processes in the local network,

and increase the level of trust in network interactions among nodes of the corporate infrastructure [13].

The combined use of Port Security and DHCP Snooping strengthens control at the data link layer.

Port Security restricts device connections through physical switch ports, while DHCP Snooping prevents the operation of unauthorized DHCP servers and creates a trusted IP–MAC address binding database.

The integration of VLAN segmentation and ACL-based traffic filtering is particularly important for limiting the lateral movement of attacks. VLANs isolate individual user groups and network resources, whereas ACLs regulate permitted communication between the created segments.

5. Conclusions

As a result of the study, modern threats to the information security of corporate computer networks were analyzed and the most critical vulnerabilities of network infrastructure were identified. It was established that the main sources of risk remain attacks at the data link layer of the OSI model, the use of insecure administration protocols, weaknesses in authentication mechanisms, and an insufficient level of network segmentation.

It is shown that effective cybersecurity can be ensured only through the integrated application of modern protection technologies. The most effective mechanisms include centralized user authentication based on the AAA model and the RADIUS protocol, the use of the cryptographically protected SSH protocol, the application of Port Security, Access Control Lists, Dynamic ARP Inspection technologies, and logical network segmentation using VLAN.

Conflict of Interest

The authors declare that they have no conflicts of interest related to this study, including financial, personal, authorship-related, or any other conflicts that could have influenced the research and the results presented in this article.

Funding

The study was conducted without financial support.

Data Availability

No data are associated with this manuscript.

Use of Artificial Intelligence Tools

The authors confirm that artificial intelligence technologies were not used in the preparation of this work.

REFERENCES

1. Kurose, J. F., and Ross, K. W. (2025), *Computer Networking: A Top-Down Approach*, 9th ed., Pearson, New York, 864 p. available at: <https://www.pearson.com/en-us/subject-catalog/p/computer-networking-a-top-down-approach/P200000013385>
2. Buriachok, V. L., Hulak, H. M., and Tolubko, V. B. (2024), *Information and cyberspaces: security problems, methods and means of counteraction*, textbook, Magnolia-2006, Lviv 2024. 448 p. available at: https://magnolia.lviv.ua/wp-content/uploads/2024/04/INF-TA-KIBERPROSTORY-pidruchnyk_zmist.pdf
3. Khomchak, M. (2025), “Cybersecurity risk assessment for selecting a cloud provider”, *Cybersecurity: Education, Science, Technique*, no. 27, pp. 549–559, doi: <https://doi.org/10.28925/2663-4023.2025.27.773>
4. Derkach, T. M., Holovko, H. V., Dmytrenko, A. O., and Klochko, L. A. (2026), “Analysis of threats and vulnerabilities of computer networks and substantiation of a comprehensive approach to ensuring their cybersecurity”, *Control, Navigation and Communication Systems*, no. 2, pp. 73–80, doi: <https://doi.org/10.26906/SUNZ.2026.2.073>
5. Derkach, T. M., and Lavrenko, M. (2024), “Cyberspace: analysis of threats and protection methods”, *Innovative Education: Problems and Prospects of Scientific Research*, Proceedings of the 50th International Scientific and Practical Conference, Stuttgart, December 4–6, pp. 112–115, available at: <https://reposit.nupp.edu.ua/handle/PolNTU/18104>
6. Korchenko, O., Ivanchenko, E., Bakalinsky, O., Myalkovskyi, D., and Zubkov, D. (2024), “Method for assessing the level of cybersecurity improvement of critical infrastructure facilities of the state”, *Science-Based Technologies*, no. 61(1), pp. 3–20, doi: <https://doi.org/10.18372/2310-5461.61.18509>
7. Anderson, R. (2020), *Security Engineering: A Guide to Building Dependable Distributed Systems*, Wiley, Hoboken, 1248 p., available at: <https://www.cl.cam.ac.uk/archive/rja14/book.html>
8. Easttom, C. (2024), *Network Defense and Countermeasures*, Springer, William Easttom II, 412 p. available at: https://books.google.com.ua/books/about/Network_Defense_and_Countermeasures.html?hl=ar&id=NFTDwAAQBAJ&redir_esc=y
9. Stewart, J. M., Chapple, M., and Gibson, D. (2024), *ISC2 CISSP Certified Information Systems Security Professional Official Study Guide*, 10th ed.: Wiley, Hoboken, ISBN 978-1-394-25470-5, 1248 p., available at: <https://www.wiley.com/en-us/ISC2+CISSP+Certified+Information+Systems+Security+Professional+Official+Study+Guide%2C+10th+Edition-p-9781394254705>
10. Tanenbaum, A. S., and Wetherall, D. J. (2011), *Computer Networks*, 5th ed. Upper Saddle River: Pearson, 2011. 960 p. available at: <https://www.pearson.com/en-us/subject-catalog/p/computer-networks/P200000003188>
11. Lakhno, V., Yerbolat, K., Bagdat, Y., Kryvoruchko, O., Desiatko, A., Tsiutsiura, S., and Tsiutsiura, M. (2022), “A model for protecting the local network of an educational institution”, *Cybersecurity: Education, Science, Technique*, no. 18, pp. 6–23, doi: <https://doi.org/10.28925/2663-4023.2022.18.623>
12. (2024), *NIST Cybersecurity Framework (CSF) 2.0*, National Institute of Standards and Technology, available at: <https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.29.pdf>
13. (2018), *IEEE Computer Society. IEEE Standard for Virtual Bridged Local Area Networks (IEEE 802.1Q)*, IEEE, New York, available at: https://standards.ieee.org/standard/802_1Q-2018.html
14. Derkach, T. M., and Prykhodko, R. (2025), “Intrusion detection system as a key element of multilayered cyber protection of information systems”, *Research in Science, Technology and Economics*, Proceedings of the 5th International Scientific and Practical Conference, Luxembourg, December 10–12, 2025, pp. 308–311, available at:

- <https://reposit.nupp.edu.ua/handle/PoltNTU/20866>
15. (1994), ISO/IEC 7498-1:1994. *Information Technology – Open Systems Interconnection – Basic Reference Model*, available at: <https://www.ecma-international.org/wp-content/uploads/s020269e.pdf>
 16. Stallings W. (2017), *Network Security Essentials: Applications and Standards*, 6th ed., Pearson, Boston, 464 p., available at: <https://www.pearson.com/en-us/subject-catalog/p/network-security-essentials/P200000003180>
 17. Cisco Systems. Catalyst 2960 and 2960-S Switches Software Configuration Guide, Cisco IOS Release 15.2(1)E. Cisco, 2013. available at: <https://www.cisco.com>
 18. Odom W. (2024), *CCNA 200-301*, vol. 1, Official Cert Guide, Cisco Press, 1087 p., available at: <https://faspc.com/academy/Ebook/Cisco/CCNA%20200-301%20Official%20Cert%20Guide%20Volume%201%20Second%20Edition.pdf>
 19. Ylonen T., and Lonvick C. (2006), *The Secure Shell (SSH) Protocol Architecture: RFC 4251*, IETF, available at: <https://www.rfc-editor.org/rfc/rfc4251>
 20. Dahm, T., Ota, A., Medway Gash, D.C., Carrel, D., and Grant, L. (2020), *RFC 8907: The Terminal Access Controller Access-Control System Plus (TACACS+) Protocol*, IETF, Fremont, 66 p., DOI: <https://doi.org/10.17487/RFC8907>

Received (Надійшла) 01.07.2026

Accepted for publication (Прийнята до друку) 09.09.2026

Publication date (Дата публікації) 25.09.2026

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Деркач Тетяна Миколаївна – доктор філософії, доцент, доцент кафедри комп'ютерних та інформаційних технологій і систем, Національний університет "Полтавська політехніка імені Юрія Кондратюка", Полтава, Україна

Tetyana Derkach – PhD, Associate Professor, Associate Professor of the Department of Computer and Information Technologies and Systems, National University «Yuri Kondratyuk Poltava Polytechnic», Poltava, Ukraine

e-mail: tanider@ukr.net; ORCID Author ID: <https://orcid.org/0000-0001-8062-9105>

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57204846765>

Дмитренко Андрій Олександрович – доктор філософії, доцент, доцент кафедри будівельних конструкцій, Національний університет "Полтавська політехніка імені Юрія Кондратюка", Полтава, Україна;

Andrii Dmytrenko – PhD, Associate Professor, Associate Professor of Department of building structures, National University «Yuri Kondratyuk Poltava Polytechnic», Poltava, Ukraine;

e-mail: andmyt@ukr.net; ORCID Author ID: <https://orcid.org/0000-0002-8715-7646>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57217604473>.

Ключко Ліна Андріївна – доктор філософії, інженер-геотехнік, геологічне бюро BEG SA, Швейцарія

Lina Klochko – Ph.D., Geological Bureau BEG SA, Switzerland

e-mail: lina.dmitrenko@gmail.com; ORCID Author ID: <http://orcid.org/0000-0002-6064-2887>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57217278708>.

Комплексний підхід до забезпечення кібербезпеки корпоративної мережевої інфраструктури

Т. М. Деркач, А. О. Дмитренко, Л. А. Ключко

Анотація. Актуальність. Стрімкий розвиток цифрових технологій, широке впровадження корпоративних інформаційних систем і зростання кількості кіберзагроз зумовлюють необхідність удосконалення захисту корпоративної мережевої інфраструктури. Традиційні методи, засновані на використанні окремих програмних або апаратних засобів, не забезпечують належної стійкості до сучасних багатовекторних кібератак. Тому актуальним є формування комплексної багаторівневої системи, спрямованої на забезпечення конфіденційності, цілісності та доступності корпоративних інформаційних ресурсів. **Предмет дослідження:** підходи, моделі, протоколи та технології забезпечення кібербезпеки корпоративної мережевої інфраструктури. **Мета дослідження:** обґрунтування комплексного підходу до побудови захищених корпоративних мереж на основі багаторівневого захисту та інтегрованого застосування сучасних технологій мережевої безпеки. **Завдання дослідження.** Проаналізувати роль моделі OSI у проєктуванні захищених мереж; визначити основні загрози; дослідити механізми захисту на різних рівнях мережевої взаємодії; розглянути концепцію Defense in Depth; проаналізувати технології AAA, RADIUS, SSH, ACL, VLAN, Port Security, DHCP Snooping і Dynamic ARP Inspection та переваги їх комплексного застосування. **Методи дослідження.** Використано методи системного аналізу, порівняння, узагальнення та систематизації науково-технічної інформації, структурно-функціональний підхід і моделювання комплексної архітектури безпеки корпоративної мережі. **Результати дослідження.** Визначено основні загрози корпоративним мережам, зокрема MAC Spoofing, MAC Flooding, ARP Spoofing, ARP Cache Poisoning, Man-in-the-Middle, несанкціоноване підключення пристроїв, компрометацію облікових записів і порушення політик доступу. Проаналізовано концепцію Defense in Depth та встановлено, що багаторівневий підхід обмежує наслідки компрометації окремих компонентів мережі. Розглянуто застосування AAA і RADIUS для централізованого управління доступом, SSH – для захищеного адміністрування, ACL – для контролю міжсегментної взаємодії, а Port Security, DHCP Snooping і Dynamic ARP Inspection – для протидії атакам каналного рівня. Встановлено, що VLAN-сегментація сприяє локалізації кіберінцидентів, обмеженню горизонтального поширення атак і реалізації принципу найменших привілеїв. **Висновки.** Комплексне застосування AAA, RADIUS, SSH, ACL, VLAN, Port Security, DHCP Snooping і Dynamic ARP Inspection відповідно до концепції Defense in Depth забезпечує централізований контроль доступу, захищене адміністрування, сегментацію мережі та протидію поширеним мережевим атакам.

Ключові слова: кібербезпека; корпоративна мережа; мережева інфраструктура; інформаційна безпека; Defense in Depth; AAA; RADIUS; SSH; ACL; VLAN; Port Security; Dynamic ARP Inspection; модель OSI; контроль доступу.

Ю. Г. Гічко¹, С. Ю. Мазор¹, Т. В. Храновська¹, Г. Я. Криховецький²

¹ Інститут спеціального зв'язку та захисту інформації Національного технічного університету «Київський політехнічний інститут» імені Ігоря Сікорського, Київ, Україна

² Науково-дослідний інститут воєнної розвідки, Київ, Україна

МЕТОДИКА ЕКСПЕРИМЕНТАЛЬНОГО ВИМІРЮВАННЯ ОСНОВНИХ ПАРАМЕТРІВ АНТЕН НАДВИСОКОЧАСТОТНОГО ДІАПАЗОНУ

Анотація. Актуальність. У статті розглянуто методику експериментального вимірювання основних параметрів антен надвисокочастотного (НВЧ) діапазону, зокрема діаграми спрямованості та коефіцієнта підсилення. **Об'єкт дослідження:** рупорні спіральні та вібраторні антени, які широко використовуються в НВЧ системах зв'язку, радіолокації та телеметрії. **Предмет дослідження.** Методика вимірювання основних параметрів антен НВЧ діапазону. **Мета дослідження:** розроблення, обґрунтування та експериментальна реалізація методики вимірювання основних параметрів антен НВЧ діапазону, зокрема діаграми спрямованості та коефіцієнта підсилення, з використанням лабораторної установки, оптимізованої для умов навчального і прикладного технічного середовища. **Завдання дослідження.** Розробити та експериментально реалізувати методику вимірювання діаграми спрямованості й коефіцієнта підсилення антен НВЧ-діапазону. **Методи дослідження:** експериментальні дослідження, методи вимірювання параметрів антен НВЧ-діапазону, порівняльний аналіз та математична обробка результатів вимірювань. **Результати дослідження.** У статті проведено аналіз лабораторної установки, яка дозволяє здійснювати вимірювання в діапазоні до 10 ГГц з використанням генератора НВЧ Keysight N5173B EXG X-Series Microwave Signal Generator, аналізатора спектра Keysight N9010A EXA, передавальної пірамідальної рупорної антени, приймальної антени та детекторної секції здійснюється зняття діаграм спрямованості, а також вимірювання коефіцієнта підсилення методом еталонної антени. Як еталон застосовано відкритий кінець хвильоводу, що дозволяє теоретично або експериментально визначати базові параметри випромінювання. **Висновки.** Наукова новизна роботи полягає в практичному поєднанні класичної методики вимірювань з сучасними підходами до аналізу антенних характеристик, що може бути використано як у навчальному процесі, так і при розробці ефективних антенних систем для систем зв'язку. Порівняно з аналогічними роботами, у статті детально описано процедуру визначення коефіцієнта спрямованої дії з використанням графічних залежностей, що спрощує процес оцінки антен для реальних технічних застосувань.

Ключові слова: рупорна антена; надвисокі частоти; діаграма спрямованості; коефіцієнт підсилення; еталонна антена; вимірювання параметрів антен; НВЧ випромінювання; спрямована дія; експериментальна методика.

1. Вступ

1.1 Актуальність тематики

У сучасних телекомунікаційних і радіотехнічних системах, особливо тих, що функціонують у надвисокочастотному (НВЧ) діапазоні, надзвичайно важливим є знання основних параметрів антен, зокрема діаграми спрямованості та коефіцієнта підсилення. Ці характеристики безпосередньо впливають на ефективність передачі та прийому сигналів, визначають енергетичний баланс радіоліній і якість покриття [1, 2]. Особливого значення набуває забезпечення високої точності вимірювань антенних параметрів для систем мобільного зв'язку, радіолокації, систем зв'язку з рухомими об'єктами, супутникових каналів та інших критичних сфер, де втрати сигналу можуть призвести до суттєвих функціональних обмежень або зниження надійності системи [3, 4].

Незважаючи на значну кількість відомих теоретичних моделей і числових методів аналізу антен [5, 6], саме експериментальне підтвердження характеристик випромінювання залишається незамінним етапом у процесі проектування, налаштування й експлуатації антенних пристроїв. У той же час, багато класичних лабораторних методик, особливо ті, що застосовуються в навчальному

процесі, потребують актуалізації або адаптації до сучасних вимог – зокрема щодо точності та швидкості вимірювань та простоти реалізації [7, 8].

Тому постає важливе наукове і практичне завдання: удосконалення або уточнення експериментальної методики вимірювання просторових характеристик антен НВЧ-діапазону з урахуванням вимог до точності повторюваності результатів і простоти реалізації в навчальних та прикладних технічних умовах.

1.2 Аналіз літературних джерел і визначення проблеми

Сучасні дослідження в галузі визначення параметрів антен НВЧ охоплюють розробку ширококутових конструкцій та методики вимірювань. У роботах [3, 5] описано конструкції ширококутових рупорних антен для імпульсних та сенсорних застосувань, які відзначаються високою точністю та стійкістю при калібруванні. Особливу увагу приділено оптимізації конструкцій антен для міліметрового діапазону [1]. Крім того, методи вимірювання, використані у [2] та [9], демонструють можливість оцінки коефіцієнта підсилення в цьому діапазоні навіть за умов обмеженого лабораторного оснащення. Не менш важливо: дослідження [7, 8] детально аналізують багаторежимні хвильоводи з

рупорними антенами, що підтверджує актуальність і перспективність обраної тематики в локальному академічному середовищі. Разом ці роботи демонструють високий рівень технічної розробки, але найчастіше такі методики складні або потребують дорогого обладнання, що обмежує їх використання в навчальних лабораторіях.

1.3 Мета і завдання дослідження

Метою роботи є розроблення, обґрунтування та експериментальна реалізація методики вимірювання основних параметрів антен НВЧ діапазону, зокрема діаграми спрямованості (ДС) та коефіцієнта підсилення, з використанням лабораторної установки, оптимізованої для умов навчального і прикладного технічного середовища. Завдання полягає у розробці та експериментальній реалізації методики вимірювання діаграми спрямованості й коефіцієнта підсилення антен НВЧ-діапазону та оцінити її ефективність.

2. Матеріали і методи дослідження

Антен НВЧ діапазону відіграють ключову роль у побудові сучасних систем бездротового зв'язку, радіолокації, телеметрії та сенсорних технологій. Зокрема, рупорні антени вирізняються простотою конструкції, надійністю та високою ефективністю в заданому напрямку, що робить їх широко використовуваними в НВЧ-техніці [1, 5, 6]. В останні роки посилено досліджуються широкосмугові та міліметровохвильові рупорні антени, які застосовуються для точних вимірювань імпульсних сигналів, у тому числі в умовах 5G-середовища та індустріального сенсорного зв'язку [3, 4].

Особливого значення набуває задача точного визначення параметрів випромінювання антен – таких як діаграма спрямованості та коефіцієнт підсилення. Ці характеристики є критичними при проектуванні радіоліній та оцінці якості роботи антенно-фідерних трактів [10]. У багатьох дослідженнях реалізовано складні вимірювальні схеми в безлунних камерах і систем з числовим

керуванням, які дають високу точність [2, 9]. Водночас, такі системи не завжди придатні до використання в умовах базових навчальних лабораторій або малопотужних мобільних пристроїв.

У вітчизняній практиці, зокрема в роботах [7, 8], продемонстровано приклади досліджень характеристик рупорних антен за умов багатомодової роботи хвильоводів, що підтверджує практичну цінність та складність аналізу таких структур. Це зумовлює потребу в розробці методики, яка б поєднувала доступність реалізації з достатньою точністю для проведення технічно коректних вимірювань параметрів рупорних антен у НВЧ-діапазоні.

3. Результати дослідження методики вимірювання основних параметрів антен надвисоких частот

3.1 Вимірювання діаграми спрямованості антен

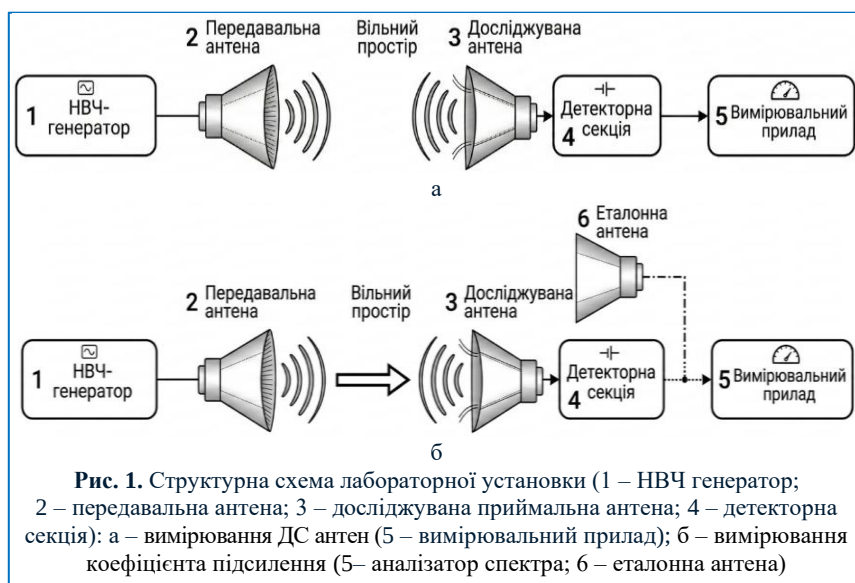
Установка для вимірювання ДС антен (рис. 1, а) складається з генератора НВЧ з передавальною антеною типу «пірамідальний рупор»; приймальних (досліджуваних) антен, в якості яких використовуються відкритий кінець хвильоводу, E та H секторіальні та пірамідальний рупори; детекторної секції; вимірювального приладу та поворотного пристрою; еталонні антени з відомим коефіцієнтом підсилення.

Лабораторна установка працює в діапазоні хвиль до 10 ГГц. НВЧ енергія генератора випромінюється передавальною рупорною антеною у вільний простір і надходить на приймальну (досліджувану) антену. Після детектування в детекторній секції сигнал надходить на вимірювальний прилад, який визначає його величину. Поворотний пристрій дозволяє змінювати орієнтацію досліджуваної антени в горизонтальній площині, чим забезпечується можливість зняття ДС у цій площині. Величина сигналу, що приймається на вимірювальному приладі, прямо пропорційна амплітуді ДС і залежить від орієнтації приймальної антени відносно передавальної.

Знята залежність відповідає ДС по потужності, оскільки вольт-амперна характеристика детектора при малих рівнях сигналу квадратична, а рівень випромінюваного сигналу генератора лабораторної установки досить малий.

3.2 Вимірювання коефіцієнта підсилення антени

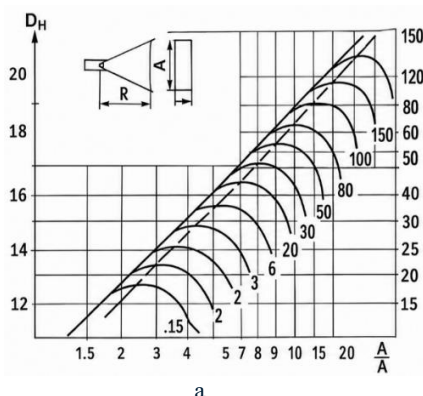
Установка для вимірювання коефіцієнта підсилення (рис. 1, б) складається з генератора НВЧ з передавальною антеною типу «пірамідальний рупор»; еталонної приймальної антени; приймальних (досліджуваних) антен; детекторної секції та вимірювального приладу.



Лабораторна установка працює в діапазоні хвиль до 10 ГГц. НВЧ енергія генератора за допомогою передавальної антени випромінюється в простір і надходить на приймальну антену. Після детектування сигнал надходить на вимірювальний прилад. Величина сигналу, що приймається, прямо пропорційна величині коефіцієнта підсилення досліджуваної антени.

Вимірювання коефіцієнта підсилення антен проводиться методом порівняння (методом еталонної антени). Цей метод є найпростішим і досить точним. В якості еталона використовується антена, коефіцієнт підсилення якої відомий або може бути визначений теоретично або експериментально. Порівняння досліджуваної антени та еталонної проводиться в режимі прийому або випромінювання. З визначення коефіцієнта підсилення випливає, що потужності випромінювання досліджуваної та еталонної антен пропорційні їх коефіцієнтам підсилення:

$$\frac{P_{IC}}{P_{ET}} = \frac{G_{IC}}{G_{ET}}, \quad G_{IC} = G_{ET} \frac{P_{IC}}{P_{ET}}. \quad (1)$$



Як видно із співвідношення (1), для визначення коефіцієнта підсилення досліджуваної антени достатньо знати коефіцієнт підсилення еталона та відношення їх потужностей випромінювання. Через те, що у виразі (1) використовується відношення потужностей, визначення абсолютних значень цих величин вимірювати не потрібно.

У даній роботі в якості еталонної антени використовується відкритий кінець хвильоводу, коефіцієнт підсилення якого є таким:

$$G_{ET} = \frac{4\pi}{\lambda^2} \cdot 0,81ab\eta_a, \quad (2)$$

a і b відповідно розміри широкої та вузької стінок хвильоводу.

При розрахунку коефіцієнта підсилення рупорних антен з достатньою для практики точністю можна знехтувати втратами та скористатися графіком залежності коефіцієнта спрямованої дії (КСД) від електричних розмірів рупора (рис. 2), вважаючи, що

$$G = D\eta_a, \quad \eta_a = 1. \quad (3)$$

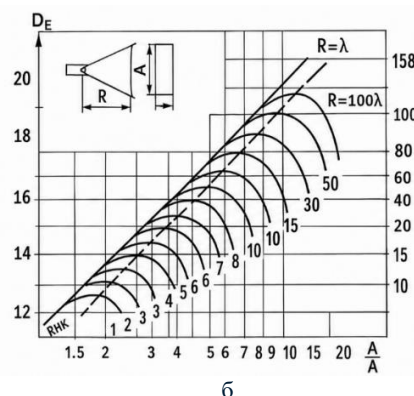


Рис. 2. Залежність КСД від розмірів розкриття рупора: а – H – секторного; б – E – секторного

Для визначення КСД необхідно заміряти висоту рупора R і з урахуванням відомої довжини хвилі вибрати відповідну криву на графіку. Потім, вимірявши величину розкриття рупора, визначити відношення величини розкриття до довжини хвилі, після чого за вертикальною шкалою знайти відповідне для даного рупора значення коефіцієнта підсилення. Коефіцієнт підсилення пірамідального рупора розраховують за формулою

$$G = 0,1 \left(\frac{\lambda}{A} D_E \right) \left(\frac{\lambda}{B} D_H \right).$$

Конфлікт інтересів

Автори декларують, що не мають конфлікту інтересів стосовно даного дослідження, в тому числі

фінансового, особистісного характеру, авторства чи іншого характеру, що міг би вплинути на дослідження та його результати, представлені в даній статті.

Фінансування

Дослідження проводилося без фінансової підтримки.

Доступність даних

Рукопис не має пов'язаних даних.

Використання засобів штучного інтелекту

Автори підтверджують, що не використовували технології штучного інтелекту при створенні представленої роботи.

СПИСОК ЛІТЕРАТУРИ

1. Zhang, Z.-Y., Lu, K., and Leung, K. W. (2023), "Gain Enhancement of Horn Antenna Using a Metal Lens", *IEEE Transactions on Antennas and Propagation*, vol. 71, no. 2, pp. 1337–1348, doi: <https://doi.org/10.1109/TAP.2022.3228631>
2. Zheng, J., Ala-Laurinaho, J., and Räsänen, A. (2019), "On the One-Antenna Gain Measurement Method in Probe Station Environment at mm-Wave Frequencies", *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 11, pp. 4510–4517, doi: <https://doi.org/10.1109/TIM.2018.2890454>
3. Ma, B., Li, J., Chen, Y., Si, Y., Gao, H., Wu, Q., and Li, M. (2024), "A 60 GHz Slotted Array Horn Antenna for Radar Sensing Applications in Future Global Industrial Scenarios", *Micromachines*, vol. 15, no. 6, doi: <https://doi.org/10.3390/mi15060728>

4. Hong, J., Zheng, S., Liu, R., and Zhao, W. (2021), "Design of mmWave Directional Antenna for Enhanced 5G Broadcasting Coverage", *Sensors*, vol. 21, no. 3, Art. 746, doi: <https://doi.org/10.3390/s21030746>
5. Zhang, T., Wei, Y., Wang, Y., Duan, C., Wang, L., Li, Z., Li, X., Li, X., & Cao, B. (2025), "Design of Segmented Ultra-Wideband TEM Horn Antenna for Calibration of Wideband Electromagnetic Pulse Sensors", *Sensors*, vol. 25, no. 12, doi: <https://doi.org/10.3390/s25123599>
6. Singh, G. I., Singh, R. K., and Lal, R. (2020), "Design and Development of 2.1 GHz Horn Antenna", *Lecture Notes in Electrical Engineering*, vol. 587, Springer, Singapore, pp. 251–259, doi: https://doi.org/10.1007/978-981-32-9775-3_25
7. Семенов, А., Гаврилюк, П., Кириченко, С., Пархоменко, О. & Литвиненко, А. (2022), "Дослідження діаграми спрямованості прямокутного рупора за умови багаторежимного поширення електромагнітних хвиль", *Технологічний аудит та резерви виробництва*, т. 64, № 2, С. 50–55, doi: <https://doi.org/10.15587/2706-5448.2022.256560>
8. Семенов, А., Гаврилюк, П., Кириченко, С. & Литвиненко, А. (2021), "Одномодовий і багатомодовий режими роботи прямокутного хвилеводу з сферичним феритовим зондом", *Матеріали 2021 IEEE 3-ї Української конференції з електротехніки та обчислювальної техніки (UKRCON)*, Львів, Україна, 26–28 серпня 2021 р., С. 100–104, doi: <https://doi.org/10.1109/UKRCON53503.2021.9575750>
9. Wang, J.-Y., Ng Mou Keh, M., and Rajo-Iglesias, E. (2020), "Design of a planar array of low profile horns at 28 GHz", *Sensors*, vol. 20, Art. 6989, doi: <https://doi.org/10.3390/s20236989>
10. Denghua, L., Fangfang, J., Cui, H., and Wenbai, C. (2018), "Study on Design and Simulation of Pyramidal Horn Antenna", *Journal of System Simulation*, vol. 30, no. 10, pp. 3933–3938, doi: <https://doi.org/10.16182/j.issn1004731x.joss.201810040>

Received (Надійшла) 29.06.2026

Accepted for publication (Прийнята до друку) 26.08.2026

Publication date (Дата публікації) 25.09.2026

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Гічко Юрій Геннадійович – аспірант, Інститут спеціального зв'язку та захисту інформації Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського», Київ, Україна;

Hichko Yurii – postgraduate student, Institute of special communication and information security of National technical university of Ukraine «Igor Sikorsky Kyiv polytechnic institute», Kyiv, Ukraine;
e-mail: gichko@gmail.com; ORCID Author ID: <https://orcid.org/0009-0004-3147-6092>.

Мазор Сергій Юрійович – кандидат технічних наук, доцент кафедри, Інститут спеціального зв'язку та захисту інформації Національного технічного університету України «КПІ імені Ігоря Сікорського», Київ, Україна;

Mazor Serhii – Candidate of Technical Sciences, Associate Professor, Institute of Special Communications and Information Protection of the National Technical University of Ukraine «Igor Sikorsky Kiev Polytechnic Institute», Kyiv, Ukraine;
e-mail: mazorsk@gmail.com; ORCID Author ID: <https://orcid.org/0009-0003-2103-4883>.

Храновська Тетяна Василівна – молодший науковий співробітник, Інститут спеціального зв'язку та захисту інформації Національного технічного університету України «КПІ імені Ігоря Сікорського», м. Київ, Україна;

Khranovska Tetyana – Junior Researcher, Institute of Special Communications and Information Protection of the National Technical University of Ukraine «Igor Sikorsky Kiev Polytechnic Institute», Kyiv, Ukraine;
e-mail: tanyakhranovskaya@gmail.com; ORCID Author ID: <https://orcid.org/0009-0007-7873-4919>.

Криховецький Георгій Яремович – кандидат технічних наук, старший науковий співробітник, Науково-дослідний інститут воєнної розвідки, м. Київ, Україна;

Heorhii Krykhovetskyi – Candidate of Technical Sciences, Senior Researcher, Defence Intelligence Research Institute, Kyiv, Ukraine;
e-mail: kgeorg@ukr.net; ORCID Author ID: <https://orcid.org/0009-0001-2981-7810>;
Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=58697828100>.

Methods of Experimental Measurement of the Main Parameters of Antennas of the Ultra-High Frequency Range

Yurii Gichko, Serhii Mazor, Tetyana Khranovska, Heorhii Krykhovetskyi

Abstract. Relevance. The article considers the method of experimental measurement of the main parameters of antennas of the ultra-high-frequency (microwave) range, in particular the radiation pattern and gain. **Object of research:** horn spiral and vibrator antennas, which are widely used in microwave communication, radar and telemetry systems. **Subject of research.** Methods for measuring the main parameters of microwave antennas. **The purpose of the study:** development, substantiation and experimental implementation of a method for measuring the main parameters of microwave antennas, in particular the radiation pattern and gain, using a laboratory installation optimized for the conditions of the educational and applied technical environment. **Research objectives.** To develop and experimentally implement a method for measuring the radiation pattern and gain of microwave antennas. **Research methods:** experimental studies, methods for measuring the parameters of microwave antennas, comparative analysis and mathematical processing of measurement results. **Research results.** The article analyzes a laboratory setup that allows measurements in the range up to 10 GHz using the Keysight N5173B EXG X-Series Microwave Signal Generator, Keysight N9010A EXA spectrum analyzer, transmitting pyramidal horn antenna, receiving antenna and detector section. The open end is used as a standard waveguide, which allows theoretically or experimentally to determine the basic parameters of radiation. **Conclusions.** The scientific novelty of the work lies in the practical combination of the classical measurement technique with modern approaches to the analysis of antenna characteristics, which can be used both in the educational process and in the development of effective antenna systems for communication systems. Compared to similar works, the article describes in detail the procedure for determining the directional action coefficient using graphical dependencies, which simplifies the process of evaluating antennas for real technical applications.

Keywords: horn antenna; microwave frequencies; radiation pattern; antenna gain; reference antenna; antenna parameter measurement; microwave radiation, directivity; experimental methodology.

Oleksandr Zhuchenko

Ukrainian State University of Railway Transport, Kharkiv, Ukraine

DISCRETE INTEGER MODEL OF COMPLEX TIME-DIVISION SIGNALS WITH A PERIODIC STRUCTURE

Abstract. Background. The determination of approaches to the analytical synthesis of ensembles of complex time-division signals without the use of a search-and-correction procedure is relevant for reducing the computational complexity of synthesis of complex time-division signals with a periodic structure. **Subject of the study:** a discrete integer model of complex time-division signals with a periodic structure and conditions for determining the cardinality of the set of discrete positions of the ensemble and achieving no more than one pulse overlap. **Objective:** development of a discrete integer model of complex time-division signals with a periodic structure and establishment of conditions based on this model for determining the cardinality of the set of discrete positions of the ensemble and achieving no more than one pulse overlap for each mutual discrete shift of signals. **Research tasks:** to develop a discrete integer model of complex time-division signals with a periodic structure; to determine the cardinality of the set of discrete positions of the ensemble; to determine the conditions for achieving no more than one pulse overlap for each mutual discrete shift of signals. **Methods.** The methods of discrete integer modeling and the apparatus of linear Diophantine equations were applied. **Results.** The mutual coprimality of the numbers of pulses in a pair of signals is a necessary and sufficient condition for the existence of no more than one integer solution of a linear Diophantine equation and the occurrence of no more than one pulse overlap for any mutual discrete shift. **Conclusions.** The minimum cardinality of the set of discrete positions for a pair of signals is determined as the least common multiple of the numbers of pulses in this pair, and for an ensemble as the maximum of such values obtained for all pairs of signals, which ensures the same duration of ensemble signals and constant integer pulse placement steps.

Keywords: complex time-division signals; periodic structure; discrete integer model; linear Diophantine equation; coprime numbers; set of discrete positions; pulse overlap.

1. Introduction

1.1 Research relevance

The development of ultra-wideband access systems for low-power applications follows several directions. Study [1] considers regulatory limitations, physical-layer specifications, and standardized impulse-radio communications. Equipment and standards compatibility across the physical layer, medium access control, and upper layers is investigated in [2]. The development of the physical layer is addressed in [3] through requirements for higher speed, lower power consumption, and fewer hardware resources. Transmitter and receiver architectures are considered in [4]. The characteristics and practical applications of ultra-wideband technology are analyzed in [5]. Study [6] focuses on advances in ranging, sensing, and data transfer in ultra-wideband systems. Study [7] analyzes the current use of IEEE 802.15.4 impulse-radio ultra-wideband technology and prospects for further adoption. Impulse-radio sensing and directions for its practical application are investigated in [8]. Study [9] considers the operating principles of ultra-wideband wireless systems, modulation schemes, and channel characteristics. Ultra-wideband radio-link quality is investigated in [10]. Distance- and frequency-dependent characterization and modeling of indoor ultra-wideband propagation channels under line-of-sight and non-line-of-sight conditions are investigated in [11]. A reconfigurable memristive pulse generator with pulse shaping for operation in the 3–5 GHz ultra-wideband range is proposed in [12]. A digitally configurable impulse-radio ultra-wideband transmitter based on a switching oscillator, with tunable pulse width and pulse position for high-data-rate and low-power operation, is proposed in [13]. IEEE 802.15.4a supplements IEEE 802.15.4 by adding alternative physical layers for low-rate wireless personal area networks [14].

IEEE 802.15.4z incorporates enhancements to ultra-wideband physical layers and ranging techniques [15]. IEEE 802.15.3f extends the millimeter-wave physical-layer specification to the 57–71 GHz band [16]. The reviewed results confirm that impulse transmission is a promising direction for further research.

Under conditions of a steady increase in the density of radio services and the number of users, a potential limitation of further development of ultra-wideband access systems based on the IEEE 802.15 family of standards is the use of a shared channel and probabilistic access mechanisms, when each end device can transmit data only if the channel is considered free, and in the case of overlapping transmissions, signals of different users become distorted and lost. The above limitation can be eliminated by transition to code-division multiple-access systems, which makes it possible to provide guaranteed simultaneous access for a certain number of users.

A significant stage in the development of synthesis of ensembles of complex impulse signals was the approach proposed in [17], which provides synthesis of ensembles of periodic complex impulse signals with minimum energy interaction (hereinafter referred to as complex time-division signals with a periodic structure), for which no more than one pulse overlap occurs under arbitrary time shifts, making such ensembles suitable for use in asynchronous ultra-wideband systems. A characteristic feature of such ensembles is the constructive difference in signal energies caused by different numbers of equal-duration pulses in each signal [17, 18]. At the same time, synthesis of such signal ensembles is performed using a search-and-correction procedure [17], which leads to high computational complexity of the synthesis procedure.

Thus, the relevant research task is to determine approaches to analytical synthesis of ensembles of

complex time-division signals without using a search-and-correction procedure, which will significantly reduce the computational complexity of synthesis and provide possibilities for software-controlled adjustment of signal parameters in ultra-wideband code-division multiple-access systems. This enables software synthesis and reconfiguration of ensembles directly in transceiver devices by changing their parameters, creating prerequisites for adaptive modification of the operating mode of an ultra-wideband access system.

1.2 Literature review and problem definition

The approach proposed in [17] provides synthesis, by an exhaustive-search-and-correction procedure, of ensembles of periodic impulse signals with minimum energy interaction, for which no more than one pulse overlap occurs under arbitrary time shifts, making such ensembles suitable for asynchronous ultra-wideband systems. A characteristic feature of these ensembles is the inherent difference in signal energies caused by different numbers of equal-duration pulses in each signal, while their improved cross-correlation properties and preservation of these properties under arbitrary time shifts make ensembles of impulse signals with minimum energy interaction promising for use in ultra-wideband code-division multiple-access systems.

The construction of an ensemble of complex impulse signals with minimum energy interaction is performed in two stages [17]. First, the parameters of the impulse sequences that must satisfy the necessary synthesis condition are determined. Then, for the determined initial parameters, the sufficient condition of minimum similarity is checked. For the impulse sequences of the ensemble, the initial parameters are defined as

$$n_1 > n_2 > \dots > n_L, \\ n_i T_i = T, i = 1, \dots, L,$$

where n_i is the number of pulses in sequence i ; T_i is the pulse repetition period of sequence i ; T is the duration of the ensemble signals.

The above relations define the initial parameters of the impulse sequences but do not guarantee satisfaction of the sufficient condition of minimum similarity. Therefore, after the initial parameters have been determined, satisfaction of the sufficient condition of minimum similarity must be verified:

$$\max_{\tau} R_{ij}(\tau) \leq \frac{1}{\sqrt{n_i n_j}}, 1 \leq i < j \leq L,$$

where $R_{ij}(\tau)$ is the cross-correlation of impulse sequences i and j at time shift τ ; τ is the time shift between the sequences; L is the ensemble size.

The minimum-similarity condition for signals in the ensemble is achieved in [17] only by a separate exhaustive-search-and-correction procedure, which consists in sequentially checking, for each pair of impulse sequences and for each mutual discrete time shift, compliance with the constraint on the maximum cross-correlation value; if the condition is violated, the parameters of one of the sequences are corrected and the check is repeated until the sufficient condition is satisfied for all pairs of sequences in the ensemble.

Thus, the synthesis procedure for ensembles of complex signals with minimum energy interaction consists in forming an initial ensemble of periodic impulse sequences in which the sequence with the largest number of pulses is assigned the smallest initial period, whereas larger initial periods are selected for the remaining sequences so that all their pulses are fully located within the specified signal duration. Then, for each pair of sequences and for each discrete time shift, the admissible cross-correlation value equivalent to the condition of no more than one pulse overlap is calculated. If the condition is not satisfied, the parameters of one of the sequences are corrected, after which the verification is repeated iteratively. This synthesis procedure is based on exhaustive search and correction and has high computational complexity. Therefore, the known synthesis method does not provide a sufficient analytical basis for computationally efficient synthesis of complex-signal ensembles.

1.3 Research objective and tasks

The objective of the study is to develop a discrete integer model of complex time-division signals with a periodic structure and to establish, based on this model, the conditions for determining the cardinality of the set of discrete positions of the ensemble and achieving no more than one pulse overlap for each mutual discrete shift of signals. To achieve this objective, the following tasks were defined:

- to develop a discrete integer model of complex time-division signals with a periodic structure;
- to determine the cardinality of the set of discrete positions of the ensemble;
- to determine the conditions for achieving no more than one pulse overlap for each mutual discrete shift of signals.

The research results will enable the synthesis of ensembles of complex time-division signals based on the analytical determination of their parameters without the use of a search-and-correction procedure, which will reduce the computational complexity of synthesizing such signal ensembles.

2. Materials and methods

The object of the study is the process of formation and analysis of complex time-division signals with a periodic structure, in which the pulse start times are specified by discrete positions. The main hypothesis of the study: the application of a discrete integer model of complex time-division signals with a periodic structure provides analytical determination of the ensemble parameters without the use of a search-and-correction procedure.

The following conditions are adopted for the construction of the discrete integer model.

Condition 1. Complex time-division signals with a periodic structure have the same duration within the ensemble.

Condition 2. The pulses of the signals have the same duration.

Condition 3. The position of a pulse is determined by the moment of its start on the set of discrete positions of the signal.

Condition 4. Pulse overlap is determined by the coincidence of the numbers of discrete positions of their starts.

Condition 5. The periodic structure of signals is determined by constant integer pulse placement steps. The methods of discrete integer modeling and the apparatus of linear Diophantine equations are applied for the construction and analysis of the mathematical model.

The construction of the discrete model involves the transition from pulse start times to discrete position numbers, which makes it possible to determine the conditions of pulse overlap through the existence of integer solutions of a linear Diophantine equation.

3. Results of development and research of the mathematical model

3.1 Construction of the mathematical model of complex time-division signals with a periodic structure based on linear Diophantine equations

To construct and analyze the discrete integer model, we use the apparatus of linear Diophantine equations [19–21].

Let us assume that pulse positions are determined by their start instants, while pulse overlap is defined by coincidence of these instants. We introduce a discrete integer model of a complex time-division signal with a periodic structure, whose pulse positions are defined on a set of discrete integer quantities—signal positions. The elements of the set of signal positions correspond to all possible pulse start instants over the signal duration (hereinafter, the discrete set of signal positions). Pulse positions are the signal positions occupied by pulses of the corresponding sequence, and pulse overlap between two signals is determined by coincidence of the numbers of these positions. We pass from time to integer position numbers so that the overlap of two pulses is determined by coincidence of the position numbers of their start instants.

For each of the signals numbered i and j , considered as a signal pair, we define a time interval T_{ij} , identical for both signals of the pair, within which the start instants of their pulses of duration τ_u are determined as pulse positions on a set of N_{ij} discrete integer quantities – signal positions:

$$T_{ij} = N_{ij}\tau_u.$$

Each discrete signal position $m = 0, \dots, N_{ij}$ corresponds to the time instant $m\tau_u$; the value $m = N_{ij}$ corresponds to the right boundary of the interval T_{ij} .

The numbers of the discrete positions at the starts of the k th pulse of signal i and the l th pulse of signal j are denoted by $m_{i,k}$ and $m_{j,l}$, respectively. Since the discrete model takes into account only pulse start instants, pulses of duration τ_u occupy the half-open intervals:

$$\begin{aligned} &[m_{i,k}\tau_u, (m_{i,k} + 1)\tau_u); \\ &[m_{j,l}\tau_u, (m_{j,l} + 1)\tau_u). \end{aligned}$$

Because the pulse start instants are multiples of their duration τ_u , pulse overlap in time is possible only when the pulse start instants correspond to a signal position with the same number within the interval

$0, \dots, N_{ij} - 1$. Therefore, in the adopted discrete model, the pulse-overlap condition is defined as equality of the pulse-position numbers: $m_{i,k} = m_{j,l}$.

It should be noted that the end instants of the last pulses in signals i and j lie outside the interval T_{ij} . This is a limitation of the adopted discrete model but must be taken into account when implementing the signal-ensemble synthesis algorithm.

The periodic structure of complex time-division signals i and j is defined through the constant integer pulse-placement steps Q_i and Q_j , which, for the numbers of pulses n_i and n_j on the set of N_{ij} discrete positions, are determined by the relations:

$$Q_i = \frac{N_{ij}}{n_i}, \quad Q_j = \frac{N_{ij}}{n_j},$$

which show that, on the set of N_{ij} discrete positions for the pair of signals i and j , the periodic structure of each signal spans an integer number of steps.

For signals i and j , denote the pulse numbers by k and l , respectively, and define the pulse-number ranges on the set of N_{ij} discrete positions as $k = 1, \dots, n_i$, $l = 1, \dots, n_j$.

Let the initial positions of the first pulses of signals i and j be the integers d_i and d_j , respectively, with $0 < d_i \leq Q_i$, $0 < d_j \leq Q_j$. Hence, the zero position $m = 0$, corresponding to time instant $t = 0$, is not the initial position of the first pulse of either signal. The discrete position numbers of pulses $m_{i,k}$ and $m_{j,l}$ can then be defined by the linear relations:

$$m_{i,k} = d_i + (k - 1)Q_i, \quad m_{j,l} = d_j + (l - 1)Q_j.$$

The mutual discrete shift between signals numbered i and j – C_{ij} is specified as an integer number of positions, with $0 \leq C_{ij} \leq N_{ij} - 1$.

The pulse-overlap condition defined by equality of their discrete position numbers $m_{i,k} = m_{j,l}$, for the mutual discrete shift C_{ij} and cyclic transmission, understood here as repetition of the periodic signal structure after completion of transmission, is written by introducing the transmission-cycle number r as an integer:

$$m_{i,k} = m_{j,l} + C_{ij} + rN_{ij},$$

where the mutual discrete shift C_{ij} defines the displacement of the signals in discrete positions, while the product of the transmission-cycle number r and the cardinality of the set of discrete positions N_{ij} is added to the right-hand side of the pulse-overlap condition to account for cyclic transmission. For further analysis within the time interval T_{ij} , it is sufficient to consider the transmission-cycle numbers $r = 0$ and $r = -1$: for $r = 0$, the term rN_{ij} is zero, whereas for $r = -1$, the term $-N_{ij}$ appears on the right-hand side of the pulse-overlap condition and accounts for positions of the preceding cycle relative to the time interval T_{ij} .

Substituting $m_{i,k} = d_i + (k - 1)Q_i$, $m_{j,l} = d_j + (l - 1)Q_j$ into Eq. $m_{i,k} = m_{j,l} + C_{ij} + rN_{ij}$ yields an equality relating pulse numbers k and l of signals i and j to the quantities d_i , d_j , C_{ij} , and rN_{ij} :

$$d_i + (k - 1)Q_i = d_j + (l - 1)Q_j + C_{ij} + rN_{ij}.$$

Rearranging the terms so that the terms depending on the pulse numbers remain on the left-hand side gives the linear equation:

$$Q_i(k - 1) - Q_j(l - 1) = C_{ij} + (d_j - d_i) + rN_{ij},$$

where $C_{ij} + (d_j - d_i) + rN_{ij}$ is the difference between pulse positions specified by the mutual shift C_{ij} , the difference between the initial positions $d_j - d_i$, and the shift by an integer number of signal positions rN_{ij} .

For convenience, in the expressions $m_{i,k} = d_i + (k - 1)Q_i$, $m_{j,l} = d_j + (l - 1)Q_j$, we replace the numbers of steps from the first pulses to pulses numbered k and l , equal to $(k - 1)$ and $(l - 1)$, respectively, with the intermediate integer variables $x = k - 1$, $k = 1, \dots, n_i$, $y = l - 1$, $l = 1, \dots, n_j$, $x = 0, \dots, n_i - 1$, $y = 0, \dots, n_j - 1$. Hence, the pulse-position-number relations are $m_{i,x+1} = d_i + xQ_i$, $m_{j,y+1} = d_j + yQ_j$. After the same substitution in the equality $Q_i(k - 1) - Q_j(l - 1) = C_{ij} + (d_j - d_i) + rN_{ij}$, we obtain the linear Diophantine equation:

$$Q_i x - Q_j y = C_{ij} + (d_j - d_i) + rN_{ij},$$

where x is the number of Q_i steps from the first pulse of signal i to pulse number $k = x + 1$, $x = 0, \dots, n_i - 1$; y is the number of Q_j steps from the first pulse of signal j to pulse number $l = y + 1$, $y = 0, \dots, n_j - 1$.

It should be noted that after the substitutions $x = k - 1$, $y = l - 1$, the intermediate variables x and y become the direct unknowns of the linear Diophantine equation. Each pair of variables (x, y) determines the pulse numbers of signals i and j considered within the time interval T_{ij} . Fig. 1 shows an example of pulse placement for two signals on the set of discrete positions for $n_i = 3$, $d_i = 5$, $n_j = 5$, $d_j = 3$, $N_{ij} = 15$, $r = 0$, $T_{ij} = N_{ij}\tau_u = 15$, $m = 0, \dots, N_{ij} = 0, \dots, 15$, $C_{ij} = 0$.

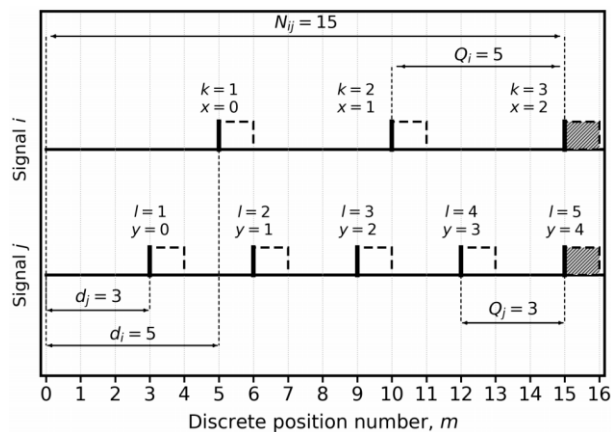


Fig. 1. Example of pulse placement for complex time-division signals with a periodic structure

The figure shows that pulse positions are determined by their start instants, whereas pulse overlap is determined by coincidence of these instants. Thus, in the constructed discrete integer model, the presence of pulse overlap between two complex time-division signals with a periodic structure is

determined by the existence of an integer solution of the linear Diophantine equation within the established ranges of the variables. The integer variables define the numbers of steps from the initial positions of the first pulses of the two signals, while the coefficients of the variables define the pulse-placement steps of the corresponding signals. Thus, the obtained model provides a basis for further determination of the number of integer solutions of the linear Diophantine equation depending on the mutual discrete shift and for establishing the condition under which the number of pulse overlaps between a pair of complex time-division signals with a periodic structure is minimal.

3.2 Determination of the cardinality of the set of discrete positions for the ensemble

Complex time-division signals with a periodic structure within an ensemble must have the same duration. In the introduced discrete integer model, each signal corresponds to a set of discrete positions at which pulses may be placed; therefore, equal duration of the ensemble signals corresponds to equal cardinality of this set.

For a pair of signals, integer pulse-placement steps are possible if the cardinality of the set of discrete positions is a common multiple of the numbers of pulses in both signals. The smallest possible interval corresponds to the least common multiple of the numbers of pulses in the pair. For the ensemble, all signal pairs must be taken into account; therefore, the cardinality of the set of discrete positions must be no smaller than the largest of the minimum values required for the individual pairs. Thus, for the ensemble as a whole, N must be selected so that the following inequality is satisfied:

$$N \geq \max_{1 \leq i < j \leq L} \text{LCM}(n_i, n_j),$$

which establishes the lower bound on the cardinality of the set of discrete positions for the ensemble: the time interval defined for all signals of the ensemble cannot be shorter than the largest among the minimum intervals required for the individual signal pairs. The equality

$$N = \max_{1 \leq i < j \leq L} \text{LCM}(n_i, n_j)$$

defines the minimum cardinality of the set of discrete positions, common to all signals of the ensemble, that provides equal duration of all ensemble signals, integer pulse-placement steps in the signals, and the maximum possible signal energy per unit time for fixed pulse duration and fixed numbers of pulses in the signals.

3.3 Determination of the condition for achieving no more than one pulse overlap under arbitrary mutual discrete shift

Based on the introduced discrete integer model, we determine the condition for achieving no more than one pulse overlap for each mutual discrete shift. For the asynchronous mode, cyclic signal transmission is taken into account, that is, repetition of the periodic signal structure after completion of transmission.

We determine the minimum cardinality of the set of discrete positions for a pair of signals that provides integer pulse-placement steps and the same time interval T_{ij} for both signals. We take into account that, for fixed pulse duration τ_u and fixed numbers of pulses n_i and n_j ,

decreasing the time interval T_{ij} increases the signal energy per unit time. Therefore, the cardinality of the set of discrete positions N_{ij} is determined from the condition of the smallest possible time interval T_{ij} .

For the cardinality of the set of discrete positions for the signal pair N_{ij} , the numbers of pulses n_i and n_j , and the pulse-placement steps Q_i and Q_j , we write the equality:

$$N_{ij} = Q_i n_i = Q_j n_j,$$

which implies that N_{ij} must be a common multiple of n_i and n_j .

Any common multiple of n_i and n_j can be expressed through their least common multiple $\text{LCM}(n_i, n_j)$ and a natural multiplicity factor q :

$$N_{ij} = q \cdot \text{LCM}(n_i, n_j).$$

If $q > 1$, the value N_{ij} exceeds the least common multiple by a factor of q , while the time interval T_{ij} exceeds the smallest possible interval by a factor of q . Therefore, the minimum value of N_{ij} is attained at $q = 1$:

$$N_{ij} = \text{LCM}(n_i, n_j).$$

The product of n_i and n_j equals the product of their least common multiple $\text{LCM}(n_i, n_j)$ and greatest common divisor $\text{GCD}(n_i, n_j)$:

$$n_i n_j = \text{LCM}(n_i, n_j) \cdot \text{GCD}(n_i, n_j),$$

hence the minimum cardinality of the set of discrete positions for the signal pair is

$$N_{ij} = \frac{n_i n_j}{\text{GCD}(n_i, n_j)},$$

which determines the shortest time interval:

$$T_{ij} = N_{ij} \tau_u.$$

It should be noted that increasing N_{ij} relative to $\text{LCM}(n_i, n_j)$ results only in additional repetitions of the periodic structure over a longer time interval.

We now generalize the results obtained for a signal pair to the ensemble, considering the ensemble as the set of all signal pairs. During synthesis of an ensemble of complex time-division signals with a periodic structure, we ensure compatibility of the periodic structures of individual signals and determine the cardinality of the set of discrete positions N , common to all signals of the ensemble. For this purpose, for each pair of signals i and j , $1 \leq i < j \leq L$, we select a value N that is a multiple of both the number of pulses n_i in signal i and the number of pulses n_j in signal j . For each signal pair, this choice ensures that the periodic structures of both signals are placed an integer number of times on the set of discrete positions. Thus, for the ensemble as a whole, N must be selected so that the inequality

$$N \geq \max_{1 \leq i < j \leq L} \text{LCM}(n_i, n_j),$$

which establishes the lower bound on the cardinality of the set of discrete positions for the ensemble: the time interval defined for all signals of the ensemble cannot be shorter than the largest among the minimum intervals required for the individual signal pairs. The equality

$$N = \max_{1 \leq i < j \leq L} \text{LCM}(n_i, n_j)$$

defines the minimum cardinality of the set of discrete positions, common to all signals of the ensemble, that provides equal duration of all ensemble signals, integer pulse-placement steps in the signals, and the maximum possible signal energy per unit time for fixed pulse duration and fixed numbers of pulses in the signals. The condition under which the number of pulse overlaps for a pair of signals does not exceed one for each value of the mutual discrete shift C_{ij} can be defined by assuming that, for a given mutual discrete shift C_{ij} , there are two distinct integer solutions of the linear Diophantine equation, each corresponding to one pulse overlap, with the sets of quantities for the first and second solutions being (x_1, y_1, r_1) and (x_2, y_2, r_2) , respectively. Subtracting the linear Diophantine equation for the set (x_2, y_2, r_2) from the linear Diophantine equation for the set (x_1, y_1, r_1) eliminates the mutual discrete shift C_{ij} and the difference between the positions $(d_j - d_i)$, yielding the relation:

$$Q_i(x_1 - x_2) - Q_j(y_1 - y_2) = (r_1 - r_2)N_{ij},$$

where variables x_1 and x_2 take values from 0 to $n_i - 1$, while variables y_1 and y_2 take values from 0 to $n_j - 1$.

For the case in which the numbers of pulses n_i and n_j are coprime, that is, $\text{GCD}(n_i, n_j) = 1$, it can be shown that the assumed two distinct integer solutions for the same mutual discrete shift C_{ij} have identical values of x , y , and r and therefore coincide. Hence, for each value of the mutual discrete shift C_{ij} , there can be no more than one integer solution of the linear Diophantine equation, and each integer solution corresponds to one pulse overlap. Thus, coprimality of the numbers of pulses n_i and n_j is a sufficient condition for the number of pulse overlaps for a signal pair not to exceed one for each value of the mutual discrete shift C_{ij} . When the greatest common divisor of the numbers of pulses n_i and n_j is greater than one, $\text{GCD}(n_i, n_j) > 1$, it can be shown that there exists a value of the mutual discrete shift C_{ij} for which the linear Diophantine equation has two distinct integer solutions that determine two separate pulse overlaps at different discrete positions $m = d_i$ and $m = d_i + N_{ij}/\text{GCD}(n_i, n_j)$ for the same value of the mutual discrete shift. Thus, if the greatest common divisor of the numbers of pulses in a signal pair is greater than one, there exists a mutual discrete shift for which at least two separate pulse overlaps occur at different positions.

Thus, taking into account the conclusion regarding the corresponding sufficient condition, coprimality of the numbers of pulses in a signal pair is a necessary and sufficient condition for limiting the number of pulse overlaps to one for each value of the mutual discrete shift.

4. Discussion of the results of the mathematical model investigation

To display all integer solutions of the linear Diophantine equation corresponding to pulse-overlap events for all values of the mutual discrete shift C_{ij} between signals i and j , and to confirm the obtained necessary and sufficient condition for limiting the number of pulse overlaps to one, we use the solution map

shown in Figure 2 for an example pair of complex time-division signals with a periodic structure having the parameters $n_i = 3$, $n_j = 5$, $N_{ij} = 15$, $d_i = 5$, $Q_i = 5$, $d_j = 3$, $Q_j = 3$. The numbers of pulses in the signals $n_i = 3$ and $n_j = 5$ satisfy the equalities $\text{GCD}(3,5) = 1$, $\text{LCM}(3,5) = 15$.

In Fig. 2, each row contains one marked cell; therefore, for each shown value of the mutual discrete shift C_{ij} , the linear Diophantine equation has one integer solution corresponding to one pulse-overlap event, and the absence of a second marked cell confirms the absence of a second integer solution for the same value of C_{ij} .

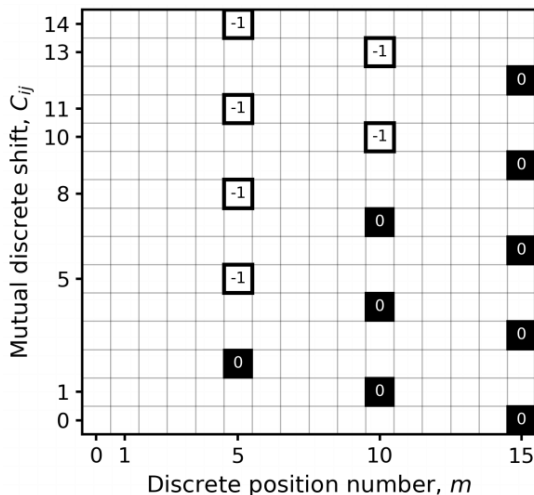


Fig. 2. Map of solutions of the linear Diophantine equation for a pair of signals with $n_i = 3$, $n_j = 5$, $N_{ij} = 15$, $d_i = 5$, $Q_i = 5$, $d_j = 3$, $Q_j = 3$. Black cells correspond to $r = 0$; white cells correspond to $r = -1$

Thus, the presence of one marked cell in each row of the solution map establishes, for the pair of signals with $n_i = 3$ and $n_j = 5$, no more than one integer solution and no more than one pulse-overlap event for each value of the mutual discrete shift, which, for the example considered, corresponds to the necessary and sufficient condition established above for limiting the number of pulse overlaps to one. Therefore, coprimality of the numbers of pulses in a signal pair is a necessary and sufficient

condition for achieving no more than one pulse overlap under an arbitrary mutual discrete shift for complex time-division signals with a periodic structure, as confirmed by the analysis of the constructed map of solutions of the linear Diophantine equation for the signal pair. A further direction of this research may be the consideration of the continuous component of the mutual time shift of complex time-division signals with a periodic structure

5. Conclusions

1. Pulse overlap for a pair of complex time-division signals with a periodic structure at a given mutual discrete shift and transmission-cycle number is determined by the existence of an integer solution of the linear Diophantine equation.

2. The minimum cardinality of the set of discrete positions for a signal pair is determined as the least common multiple of the numbers of pulses in that pair, and for the ensemble as the largest of such values obtained for all signal pairs, which provides equal duration of the ensemble signals and constant integer pulse-placement steps.

3. Coprimality of the numbers of pulses in a signal pair is a necessary and sufficient condition under which, for any mutual discrete shift, there is no more than one integer solution of the linear Diophantine equation and no more than one pulse overlap occurs.

Conflict of interest

The authors declare that they have no conflict of interest regarding this research, including financial, personal, authorship-related, or any other type of conflict that could affect the research and its results presented in this article.

Funding

The research was conducted without financial support.

Data availability

The manuscript has no associated data.

Use of artificial intelligence tools

The authors confirm that artificial intelligence technologies were not used in the preparation of this work.

REFERENCES

1. Niemelä, V., Haapola, J., Hämäläinen, M., and Iinatti, J. (2017), "An Ultra Wideband Survey: Global Regulations and Impulse Radio Research Based on Standards", *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 874–890, doi: <https://doi.org/10.1109/COMST.2016.2634593>
2. Coppens, D., Shahid, A., Lemey, S., Verhaevert, J. et al. (2022), "An Overview of Ultra-Wideband (UWB) Standards and Organizations (IEEE 802.15.4, FiRa, Apple): Interoperability Aspects and Future Research Directions", *IEEE Access*, vol. 10, pp. 70219–70241, doi: <https://doi.org/10.1109/ACCESS.2022.3187410>
3. Lv, Z., Zhang, X., Chen, D. et al. (2023), "The Development and Progress of the UWB Physical Layer", *Micromachines*, vol. 14, no. 1, art. 8, doi: <https://doi.org/10.3390/mi14010008>
4. Wang, B., Song, H., Rhee, W. and Wang, Z. (2022), "Overview of Ultra-Wideband Transceivers - System Architectures and Applications", *Tsinghua Science and Technology*, vol. 27, no. 3, pp. 481–494, doi: <https://doi.org/10.26599/TST.2021.9010044>
5. Zheng, C., Ge, Y., and Guo, A. (2023), "Ultra-Wideband Technology: Characteristics, Applications and Challenges", arXiv, doi: <https://doi.org/10.48550/arXiv.2307.13066>
6. Kabbinala, A.R., Bansal, A., Gopalan, K.S. et al. (2023), "Towards Next-generation Ultra-Wideband Technology", *IEEE Int. Conf. on Comm. Systems & Networks*, pp. 830–834, doi: <https://doi.org/10.1109/COMSNETS56262.2023.10041270>
7. Powell, C., Rolfe, B.A., Neirynck, D. and Lansford, J. (2025), "IEEE 802.15.4 IR-UWB: A Technology Precisely Positioned for Adoption", *IEEE Open Journal of Vehicular Technology*, vol. 7, pp. 237–248, doi: <https://doi.org/10.1109/OJVT.2025.3640084>

8. Ninnemann, J., Schwarzbach, P. and Michler, O. (2023), "Toward UWB Impulse Radio Sensing: Fundamentals, Potentials, and Challenges", Vargas-Bernal, R. (ed.), *UWB Technology - New Insights and Developments*, IntechOpen, doi: <https://doi.org/10.5772/intechopen.110040>
9. Donlan, B.M., Buehrer, R.M., and Reed, J.H. (2005), "Ultra-Wideband Wireless Systems", Chang, K. (ed.), *Encyclopedia of RF and Microwave Engineering*, doi: <https://doi.org/10.1002/0471654507.eme544>
10. Ansaripour, A., Heydariaan, M., and Gnawali, O. (2024), "Link characteristics study of ultra-wideband radios", *Ad Hoc Networks*, vol. 156, art. 103402, doi: <https://doi.org/10.1016/j.adhoc.2024.103402>
11. Pal, J., and Gupta, B. (2024), "Frequency Domain UWB Channel Characterization and Modeling in Indoor Static Environment", *International Journal of RF and Microwave Computer-Aided Engineering*, vol. 2024, Article ID 3742665, 21 p., doi: <https://doi.org/10.1155/2024/3742665>
12. Barraj, I., Neifar, A., Mestiri, H., and Masmoudi, M. (2025), "Reconfigurable Memristive Pulse Generator Based on Pulse Shaping for Ultra Wideband Communication", *Engineering, Technology & Applied Science Research*, vol. 15, no. 2, pp. 20668–20673, doi: <https://doi.org/10.48084/etasr.9771>
13. Albayrak, M., Dündar, G., and Batur, O.Z. (2024), "High data rate and energy efficient configurable IR-UWB transmitter with an integrated balun", *AEU - International Journal of Electronics and Communications*, vol. 179, art. 155289, doi: <https://doi.org/10.1016/j.aeue.2024.155289>
14. (2007), *IEEE Standard for Information technology – Local and metropolitan area networks—Specific requirements – Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (WPANs): Amendment 1: Add Alternate PHYs*, IEEE 802.15.4a-2007, available at: <https://standards.ieee.org/ieee/802.15.4a/3571/>
15. (2020), *IEEE Standard for Low-Rate Wireless Networks—Amendment 1: Enhanced Ultra Wideband Physical Layers and Associated Ranging Techniques*, IEEE 802.15.4z-2020, available at: <https://standards.ieee.org/ieee/802.15.4z/10230/>
16. (2017), *IEEE Standard for High Data Rate Wireless Multi-Media Networks – Amendment 3: Extending the Physical Layer (PHY) Specification for Millimeter Wave to Operate from 57.0 GHz to 71 GHz*, IEEE 802.15.3f-2017, available at: <https://standards.ieee.org/ieee/802.15.3f/7011/>
17. Lysechko, V.P., Zhuchenko, O.S., and Indyk, S.V. (2026), "Method for Synthesis of Ensembles of Complex Time-Division Signals with a Periodic Structure for Ultra-Wideband Multiple Access Systems", *Theory and Practice of Modern Science in Conditions of Transformations*, SPC, 2026, pp. 142–144, available at: <http://lib.kart.edu.ua/handle/123456789/32557>
18. Zhuchenko, O., Indyk, S. and Zablotskyi, V. (2026), "Study of Properties of Complex Time-Division Signal Ensembles with a Periodic Structure for Ultra-Wideband Code-Division Multiple Access Systems", *Computer-Integrated Technologies: Education, Science, Production*, no. 63, pp. 310–317, doi: <https://doi.org/10.36910/6775-2524-0560-2026-63-34>
19. Jones, G.A. and Jones, J.M. (1998), *Elementary Number Theory*, Springer, doi: <https://doi.org/10.1007/978-1-4471-0613-5>
20. Cohen, H. (2007), *Number Theory: Vol. I: Tools and Diophantine Equations*, doi: <https://doi.org/10.1007/978-0-387-49923-9>
21. Andreescu, T., Andrica, D. and Cucurezeanu, I. (2010), *An Introduction to Diophantine Equations: A Problem-Based Approach*, Birkhäuser Boston, doi: <https://doi.org/10.1007/978-0-8176-4549-6>

Received (Надійшла) 16.06.2026

Accepted for publication (Прийнята до друку) 02.09.2026

Publication date (Дата публікації) 25.09.2026

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Жученко Олександр Сергійович – кандидат технічних наук, доцент, доцент кафедри транспортного зв'язку, Український державний університет залізничного транспорту, Харків, Україна;

Oleksandr Zhuchenko – Candidate of Technical Sciences, Associate Professor, Associate Professor of Department of Transport Communication, Ukrainian State University of Railway Transport, Kharkiv, Ukraine;
e-mail: n030201@gmail.com; ORCID Author ID: <https://orcid.org/0000-0003-3275-810X>.

Дискретна цілочисельна модель складних часорозділювальних сигналів з періодичною структурою

О. С. Жученко

Анотація. Актуальність. Актуальним є визначення шляхів аналітичного синтезу ансамблів складних часорозділювальних сигналів без застосування перебірно-коригувальної процедури для зменшення обчислювальної складності синтезу складних часорозділювальних сигналів з періодичною структурою. **Предмет дослідження:** дискретна цілочисельна модель складних часорозділювальних сигналів з періодичною структурою та умови визначення потужності множини дискретних позицій ансамблю і досягнення не більше одного суміщення імпульсів. **Мета дослідження:** Метою дослідження є розроблення дискретної цілочисельної моделі складних часорозділювальних сигналів з періодичною структурою та встановлення на її основі умов визначення потужності множини дискретних позицій ансамблю і досягнення не більше одного суміщення імпульсів для кожного дискретного взаємного зсуву сигналів. **Завдання дослідження:** розробити дискретну цілочисельну модель складних часорозділювальних сигналів з періодичною структурою; визначити потужність множини дискретних позицій ансамблю; визначити умови досягнення не більше одного суміщення імпульсів для кожного дискретного взаємного зсуву сигналів. **Методи дослідження.** Застосовано методи дискретного цілочисельного моделювання та апарат лінійних діофантових рівнянь. **Результати дослідження.** Взаємна простота кількостей імпульсів у парі сигналів є необхідною та достатньою умовою існування не більше одного цілочисельного розв'язку лінійного діофантового рівняння та виникнення не більше одного суміщення імпульсів для будь-якого взаємного дискретного зсуву. **Висновки.** Визначено найменшу потужність множини дискретних позицій для пари сигналів як найменше спільне кратне кількостей імпульсів у цій парі, а для ансамблю – як найбільше з таких значень, одержаних для всіх пар сигналів, що забезпечує однакову тривалість сигналів ансамблю та цілочисельні сталі кроки розміщення імпульсів.

Ключові слова: складні часорозділювальні сигнали; періодична структура; дискретна цілочисельна модель; лінійне діофантове рівняння; взаємно прості числа; множина дискретних позицій; суміщення імпульсів.

О. О. Коваленко¹, В. М. Почерняєв², С. О. Івко¹

¹ Військовий інститут телекомунікацій та інформатизації імені Героїв Крут, Київ, Україна

² Національна академія Служби безпеки України, Київ, Україна

МОДИФІКОВАНИЙ МЕТОД НЕОРТОГОНАЛЬНОГО МНОЖИННОГО ДОСТУПУ З ПАРЦІАЛЬНИМ СПЕКТРАЛЬНИМ КОДУВАННЯМ ДЛЯ МОБІЛЬНИХ ВУЗЛОВИХ ЦИФРОВИХ ТРОПОСФЕРНИХ СТАНЦІЙ

Анотація. Предмет. Модифікований метод неортогонального множинного доступу з парціальним спектральним кодуванням абонентських сигналів для мобільної вузлової цифрової тропосферної станції. **Мета.** Аналітично обґрунтувати формування та розділення кількох абонентських сигналів у спільному часово-частотному ресурсі тропосферного каналу. **Завдання.** Визначити структурно-функціональну схему формування групового сигналу та розділення абонентських сигналів, формалізувати математичний опис парціального спектрального кодування, дослідити використання ортогональних поліномів для синтезу частотних вагових функцій, а також перевірити отримані аналітичні залежності шляхом схемотехнічного моделювання. **Методи.** Застосовано аналітичний опис сигнальних перетворень, z -перетворення, частотний аналіз цифрових фільтрів, розклад вагових функцій за ортогональними поліномами та схемотехнічне моделювання часткового випадку передатної функції. **Результати.** Запропоновано формувати для кожного абонентського сигналу індивідуальний спектрально-енергетичний профіль як додаткову ознаку його розділення після неортогонального суміщення. Отримано аналітичні співвідношення для передатних функцій парціальних каналів обробки, спектрально-енергетичних профілів і частотних вагових функцій. Показано можливість застосування поліномів Лежандра, Чебишова та Гегенбауера для керування формою профілів. Результати схемотехнічного моделювання підтвердили отримані аналітичні залежності формування спектрально-енергетичних профілів для розглянутого часткового випадку. **Висновки.** Запропонований підхід визначає сигнально-алгоритмічну основу багатокористувацького фізичного рівня тропосферної станції та створює передумови для підвищення ефективності використання частотного ресурсу і стійкості розділення абонентських сигналів.

Ключові слова: неортогональний множинний доступ; парціальне спектральне кодування; мобільна вузлова цифрова тропосферна станція; тропосферний канал; послідовне усунення інтерференції; фільтр з кінцевою імпульсною характеристикою; ортогональні поліноми.

1. Вступ

1.1 Актуальність тематики

У сучасних тропосферних станціях зв'язку спостерігається перехід від спеціалізованих систем великої потужності до цифрових транспортних систем із програмно визначуваними радіомодемами, адаптивним кодуванням і модуляцією, підтримкою інтернет-протоколу, твердотільними підсилювачами потужності, частотним або поляризаційним рознесенням і зменшеними масогабаритними параметрами. Розвиток цього напрямку зумовлений потребою в організації позагоризонтного зв'язку (Beyond-Line-of-Sight, BLOS) без обов'язкової залежності від супутникової або кабельної інфраструктури [1–3].

Тропосферні системи зв'язку доцільно застосовувати як альтернативні або резервні засоби за відсутності, пошкодження або економічної недоцільності використання волоконно-оптичної інфраструктури, а також за обмеженого доступу до супутникового ресурсу [4, 5]. Повідомлення про застосування засобів радіоелектронної боротьби для порушення роботи супутникових мереж під час російсько-української війни [6] додатково підтверджують необхідність резервування супутникових каналів альтернативними BLOS-засобами. За таких умов тропосферний зв'язок може забезпечувати підвищення стійкості системи управління та підтримання пропускну здатності транспортного сегмента мережі.

Тенденцію розвитку цифрових адаптивних систем зв'язку в межах прямої видимості (Line-of-Sight, LOS) і поза її межами (BLOS) підтверджують характеристики сучасних серійних тропосферних станцій [1–3], наведені в табл. 1.

Таблиця 1 – Характеристики тропосферних станцій

Станція	Швидкість передавання даних, Мбіт/с	Режим радіолінії / мережева топологія
Comtech COMET	до 60	LOS, BLOS / PtP
Cubic GATR	до 210	LOS, BLOS / PtP
Ultra I&C Archer	до 180	BLOS / H&S

Традиційно тропосферний зв'язок організовують за топологією «точка–точка» (Point-to-Point, PtP) або за ланцюговою топологією з проміжною ретрансляцією [1–3]. Водночас для побудови вузлових мереж може застосовуватися топологія «центр–абонентські станції» (hub-and-spoke, H&S) [4]. За взаємодією вузлів така топологія є функціонально близькою до схеми «точка–багатоточка» (Point-to-Multipoint, PtMP), оскільки один центральний вузол взаємодіє з кількома віддаленими абонентськими станціями. Проте H&S не є тотожною PtMP на фізичному рівні, оскільки зв'язки «центр–абонентська станція» можуть реалізовуватися як сукупність незалежних PtP-радіоліній, кожна з яких використовує окремий апаратний або часово-частотний ресурс.

Топологія PtMP визначає мережеву конфігурацію, але сама по собі не встановлює механізму одночасного використання спільного часово-частотного ресурсу кількома абонентськими станціями. Реалізація такого режиму потребує відповідного методу множинного доступу та сигнально-алгоритмічних засобів формування і розділення абонентських сигналів на фізичному рівні.

У проаналізованих відкритих описах серійних тропосферних систем недостатньо розкрито реалізацію багатостанційного доступу, за якого вузлова станція одночасно обслуговує кілька абонентських станцій у спільному часово-частотному ресурсі [1–4].

У [7] запропоновано концепцію мобільної вузлової цифрової тропосферної станції (МВЦТрС), апаратна структура якої забезпечує можливість роботи в кінцевому, проміжному та вузловому режимах і створює передумови для організації зв'язку за топологією PtMP. Зазначене технічне рішення визначає апаратно-структурну основу функціонування станції у вузловому режимі. Проте воно не визначає сигнально-алгоритмічного способу формування групового сигналу та розділення сигналів кількох абонентських станцій у спільному часово-частотному ресурсі.

Додаткову складність реалізації багатостанційного доступу та розділення абонентських сигналів створюють відмінності умов поширення радіохвиль на окремих тропосферних напрямках, що спричиняють неоднакові амплітудно-фазові спотворення, частотно-селективні завмирання та різні рівні потужності прийнятих сигналів [8, 9]. За таких умов вузлова станція має забезпечувати не лише комутацію або ретрансляцію трафіку, а й розділення абонентських сигналів на фізичному рівні.

Отже, дослідження методів спільного використання часово-частотного ресурсу та розділення абонентських сигналів у МВЦТрС є актуальними.

1.2 Аналіз літературних джерел і визначення проблеми

У системах радіозв'язку багатостанційний доступ традиційно реалізують за допомогою ортогонального множинного доступу (Orthogonal Multiple Access, OMA). У праці [10] систематизовано принципи частотного, часового та кодового розподілу ресурсів між абонентами. У [11] розглянуто теоретичні основи багатокористувацьких систем зв'язку та умови розділення сигналів у приймачі. Ортогональне виділення ресурсів спрощує приймання, однак потребує розподілу доступної смуги частот, часу або кодового простору між абонентами. У зазначених працях не розглянуто одночасне передавання кількох абонентських сигналів у спільному часово-частотному ресурсі тропосферного каналу.

У [12] запропоновано концепцію неортогонального множинного доступу (Non-

Orthogonal Multiple Access, NOMA) з розділенням за потужністю (Power-Domain NOMA, PD-NOMA). Метод передбачає суміщення сигналів кількох абонентів у спільному ресурсі з різними рівнями потужності та їх розділення за допомогою процедури послідовного усунення інтерференції (Successive Interference Cancellation, SIC). Водночас розглянуте рішення орієнтоване на стільниковий радіодоступ і не передбачає формування індивідуальних частотнозалежних спектрально-енергетичних профілів як додаткової ознаки розділення сигналів.

У технічному звіті 3GPP [13] систематизовано результати дослідження NOMA для систем радіодоступу 5G New Radio (NR), зокрема структури передавачів, процедури приймання, складність багатокористувацького детектування та можливі сигнальні послідовності. Проте цей документ визначає загальні підходи для мереж NR і не встановлює методу NOMA, адаптованого до частотно-селективного тропосферного каналу та вузлової топології МВЦТрС.

У роботі [14] наведено класифікацію методів NOMA з розділенням за потужністю та кодом, проаналізовано принципи їх передавання і приймання, а також основні напрями практичної реалізації. У методах NOMA з кодовим розділенням (Code-Domain NOMA, CD-NOMA) розділення користувачів забезпечується спеціальними кодовими книгами або розрідженими сигнальними структурами, тоді як у PD-NOMA основною ознакою є різниця рівнів потужності. Однак у [14] не розглянуто використання керованої форми спектрально-енергетичного профілю як додаткової ознаки розділення абонентських сигналів.

У [15] NOMA подано як узагальнений підхід, за якого кілька користувачів обслуговуються в одному ресурсному блоці, а їх сигнали розрізняються за потужнісними, кодовими або іншими ознаками. Авторами визначено проблеми розподілу ресурсів, багатокористувацького детектування та практичної реалізації. Проте питання формування спектральних ознак абонентських сигналів засобами цифрових фільтрів із кінцевою імпульсною характеристикою (KIX) та їх використання під час SIC не досліджено.

У [16] розглянуто розвиток NOMA до концепції множинного доступу наступного покоління (Next Generation Multiple Access, NGMA), яка передбачає гнучке поєднання NOMA з багатоантенними системами та іншими технологіями фізичного рівня. У цій роботі також визначено необхідність урахування складності приймача та практичних обмежень реалізації. Водночас не запропоновано структурно-функціонального рішення для багатостанційного доступу в тропосферній PtMP-системі та не розглянуто синтез індивідуальних спектральних профілів користувачів.

У [17] виконано критичний аналіз сучасного стану NOMA та показано, що переваги неортогонального доступу залежать від умов застосування, способу групування користувачів, складності приймача та затримки обробки. Це вказує

на необхідність окремого обґрунтування сигнальних ознак і алгоритму розділення для кожної конкретної системи. Проте в [17] не розглянуто розділення сигналів за їхніми спектрально-енергетичними профілями сигналів у частотно-селективному тропосферному каналі.

У [18] проаналізовано поєднання багатонесучих форм сигналів із методами множинного доступу та визначено їхні переваги й обмеження, однак не розглянуто формування індивідуальних спектрально-енергетичних профілів абонентських сигналів. У [19] запропоновано вдосконалену схему NOMA для зменшення міжабонентської інтерференції та пікового фактора, проте розділення сигналів не ґрунтується на керованому спектральному зважуванні. У [20] проаналізовано ймовірність символної помилки в системах NOMA із SIC, але не враховано частотно-селективний тропосферний канал і спектральні ознаки абонентських сигналів. Отже, ці роботи доповнюють аналіз відомих підходів, але не вирішують задачу парціального спектрального кодування для МВЦТРС.

Аналіз джерел [10–20] показав, що відомі методи ОМА передбачають ортогональний розподіл ресурсів, PD-NOMA використовує переважно відмінності рівнів потужності, а CD-NOMA – кодові книги або спеціальні сигнальні структури. У проаналізованих працях не визначено структурної та математичної реалізації NOMA для МВЦТРС, у якій частотнозалежні спектрально-енергетичні профілі абонентських сигналів формуються засобами цифрової КІХ-фільтрації та використовуються як додаткові ознаки їх розділення.

Таким чином, невирішеною залишається проблема розроблення структурно-функціонального та математичного методу формування й розділення кількох абонентських сигналів у спільному часово-частотному ресурсі МВЦТРС з урахуванням частотно-селективних спотворень тропосферного каналу.

1.3 Мета і завдання дослідження

Мета роботи полягає в аналітичному обґрунтуванні формування та розділення кількох абонентських сигналів у спільному часово-частотному ресурсі тропосферного каналу на основі модифікованого методу неортогонального множинного доступу з парціальним спектральним кодуванням для МВЦТРС, що функціонує за топологією PtMP. Для досягнення поставленої мети необхідно розв'язати такі завдання:

- визначити структурно-функціональну схему формування групового сигналу та розділення абонентських сигналів;
- формалізувати математичний опис парціального спектрального кодування в z -області на основі цифрової КІХ-фільтрації;
- дослідити можливість використання ортогональних поліномів для синтезу частотних вагових функцій парціальних каналів обробки;
- перевірити отримані аналітичні залежності формування спектрально-енергетичних профілів шляхом схематичного моделювання для розглянутого часткового випадку передатної функції.

2. Матеріали і методи дослідження

Об'єктом дослідження є процеси формування групового сигналу та розділення абонентських сигналів у спільному часово-частотному ресурсі МВЦТРС. Предметом дослідження є модифікований метод NOMA з парціальним спектральним кодуванням абонентських сигналів.

Основна гіпотеза дослідження полягає в тому, що індивідуальні спектрально-енергетичні профілі абонентських сигналів, сформовані засобами цифрової КІХ-фільтрації, можуть використовуватися як додаткові сигнальні ознаки під час їх розділення процедурою SIC.

Для побудови математичної моделі прийнято, що K абонентських сигналів одночасно передаються у спільному часово-частотному ресурсі. Кожен сигнал обробляється в окремому парціальному каналі, частотна характеристика якого формує індивідуальний спектрально-енергетичний профіль. У частотній області спектр групового сигналу $X(\Omega)$ визначено співвідношенням:

$$X(\Omega) = \sum_{k=1}^K X_k(\Omega),$$

де $X_k(\Omega) = S_k(\Omega)H_k(e^{j\Omega})$ – спектр k -го сигналу після парціального спектрального кодування; k – індекс парціального каналу обробки; $S_k(\Omega)$ – спектр сигналу до кодування; $H_k(e^{j\Omega})$ – частотна характеристика k -го парціального кодера.

Тропосферний канал подано узагальнено як лінійний частотно-селективний канал з адитивним шумом, у якому груповий сигнал зазнає частотно-залежних амплітудно-фазових спотворень.

Парціальне спектральне кодування формалізовано як цифрову КІХ-фільтрацію. Передатну функцію k -го парціального каналу обробки задано у вигляді:

$$H_k(z) = \sum_{m=0}^{L_k-1} h_k[m]z^{-m}, \quad (1)$$

де $h_k[m]$ – коефіцієнти імпульсної характеристики КІХ-фільтра; L_k – довжина КІХ-фільтра, тобто кількість коефіцієнтів його імпульсної характеристики; m – індекс відліку імпульсної характеристики.

Частотна характеристика фільтра визначається підстановкою $z = e^{j\Omega}$ у (1), а спектрально-енергетичний профіль абонентського сигналу, сформований відповідним парціальним каналом обробки, задається як:

$$W_k(\Omega) = |H_k(e^{j\Omega})|^2, \quad (2)$$

де $\Omega \in [\Omega_{\min}; \Omega_{\max}]$ – нормована кутова частота, рад/відлік; $\Omega_{\min}, \Omega_{\max}$ – нижня та верхня межі досліджуваного частотного інтервалу відповідно. Отже, форма профілю $W_k(\Omega)$ безпосередньо визначається функцією $H_k(z)$, заданою у (1).

Для керованого синтезу частотних вагових функцій використано їх скінченне апроксимаційне подання за системою ортогональних базисних функцій:

$$w_k(x) = d_k \sum_{r=0}^{M-1} c_{k,r} g_r(x), \quad (3)$$

де $w_k(x)$ – частотна вагова функція k -го парціального каналу обробки; d_k – нормувальний коефіцієнт; $c_{k,r}$ – коефіцієнти ортогонального розкладу; $g_r(x)$ – базисні ортогональні функції; r – індекс базисної функції; M – кількість членів скінченного розкладу; $x \in [-1;1]$ – нормована частотна змінна.

Для переходу від поліноміального подання частотної вагової функції до частотної характеристики КІХ-фільтра змінну x пов'язано з нормованою кутовою частотою Ω лінійним відображенням:

$$x(\Omega) = \frac{2\Omega - \Omega_{\max} - \Omega_{\min}}{\Omega_{\max} - \Omega_{\min}}.$$

Для першої зони Найквіста $\Omega \in [0;\pi]$ це співвідношення набуває вигляду:

$$x = \frac{2\Omega}{\pi} - 1.$$

Поліноміальну функцію $w_k(x)$ розглянуто як бажану дійсну частотну вагову функцію k -го парціального каналу. Бажану комплексну частотну характеристику лінійно-фазового КІХ-фільтра задано співвідношенням:

$$H_{k,d}(e^{j\Omega}) = w_k(x(\Omega))e^{\frac{-j\Omega(L_k-1)}{2}},$$

а відповідний спектрально-енергетичний профіль визначається:

$$W_{k,d}(\Omega) = |w_k(x(\Omega))|^2.$$

У загальній постановці коефіцієнти імпульсної характеристики $h_k[m]$ можуть бути визначені методом найменших квадратів шляхом мінімізації відхилення частотної характеристики КІХ-фільтра від $H_{k,d}(e^{j\Omega})$ на дискретній частотній сітці з урахуванням умови симетрії коефіцієнтів. Тип поліноміального базису, кількість членів розкладу M і довжина фільтра L_k можуть бути обрані за критерієм мінімуму похибки апроксимації.

Як базисні функції розглянуто поліноми Лежандра, Чебишова та Гегенбауера. Вибір типу базису, порядку поліномів і коефіцієнтів розкладу визначає форму частотної вагової функції та відповідного спектрально-енергетичного профілю.

Структурно-функціональну реалізацію методу досліджено шляхом аналізу послідовності сигнальних перетворень під час формування групового сигналу та розділення абонентських сигналів. Процедру послідовного усунення інтерференції розглянуто на структурному рівні як послідовне оцінювання, реконструкцію та віднімання внесків окремих абонентських сигналів із прийнятого групового сигналу.

Вхідними параметрами аналітичної моделі є кількість парціальних каналів, коефіцієнти та довжини імпульсних характеристик КІХ-фільтрів, тип і порядок ортогонального поліноміального

базису, коефіцієнти розкладу та нормувальні коефіцієнти. Розрахунковими величинами є передатні функції парціальних каналів, їх частотні характеристики, частотні вагові функції та спектрально-енергетичні профілі абонентських сигналів.

Для перевірки отриманих аналітичних залежностей виконано схемотехнічне моделювання часткового випадку КІХ-фільтра з двома ненульовими відліками імпульсної характеристики у модулі Circuit Design програмного середовища Ansys Electronics Desktop (AEDT) 2025 R2. Передатну функцію реалізовано сумуванням прямої складової сигналу та його затриманої копії з фазовим зсувом 180° . Фізичну затримку визначено як

$$\tau_p = pT_d = \frac{p}{f_d},$$

де p – затримка, виражена кількістю відліків; T_d – період дискретизації, с; f_d – частота дискретизації, Гц.

Частотний аналіз виконано засобами Linear Network Analysis в діапазоні 0–40 МГц для значень цілочислового параметра затримки $p = 1$ і $p = 2$. Спектрально-енергетичний профіль визначали як квадрат модуля коефіцієнта передавання схемної моделі та зіставляли з відповідною аналітичною залежністю.

Статистичні методи не застосовували, оскільки дослідження має аналітичний і схемотехнічний характер. Проведене моделювання перевіряє лише розглянутий частковий випадок передатної функції та не охоплює повної багатокористувацької моделі, впливу похибок оцінювання каналу, залишкової інтерференції й похибок синхронізації, а також не передбачає оцінювання показників ефективності модифікованого методу.

3. Результати розробки та дослідження модифікованого методу неортогонального множинного доступу з парціальним спектральним кодуванням

3.1 Структурно-функціональна схема модифікованого методу NOMA

Структурною основою реалізації запропонованого методу є базова конфігурація МВЦТРС [7], наведена на рис. 1. Антенна система станції містить фазовану антенну решітку (ФАР) і діаграмоформувальну схему (ДФС). Приймальна частина базової конфігурації має два приймачі (Прм 1, Прм 2) та гетеродин (Гет), а передавальна – передавач (Прд).

До складу модема входять суматор (Σ) і модулятор (M), а каналоутворювальна апаратура (КУА) містить апаратуру багатостанційного доступу (АБД), комутатор каналів (K) і мультиплексор-демультиплексор ($M-D$).

Базова конфігурація має два приймальні та один передавальний тракт надвисоких частот (НВЧ). Такий склад відповідає мінімальній апаратній побудові вузлової станції. У розгорнутих конфігураціях кількість приймальних і

передавальних НВЧ-трактів може бути збільшена відповідно до кількості одночасно обслуговуваних

тропосферних напрямків, вимог до рознесення, резервування та пропускної спроможності.

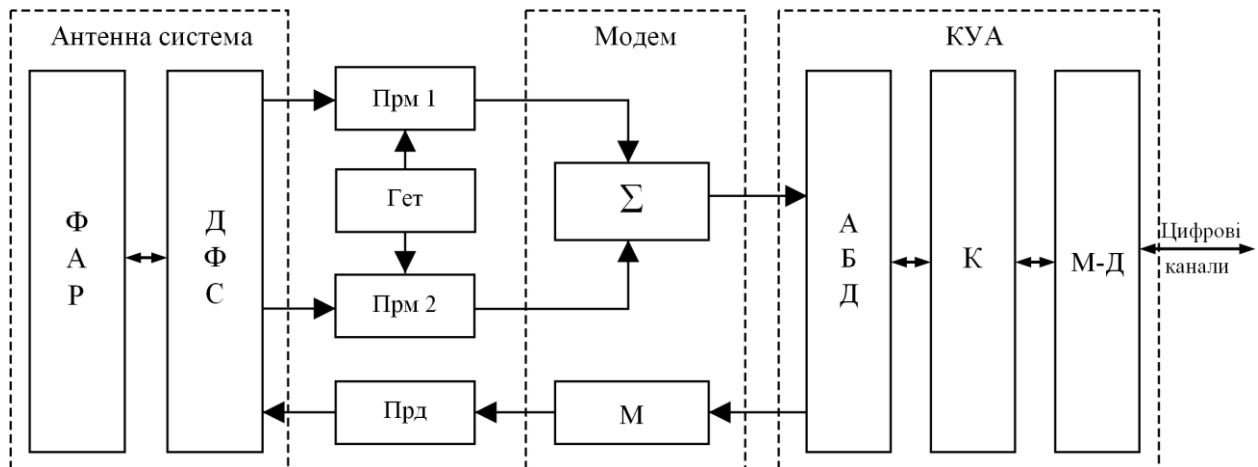


Рис. 1. Базова структурна схема МВЦТРС із двома приймальними та одним передавальним НВЧ-трактами

У запропонованій реалізації функції парціального спектрального кодування та розділення абонентських сигналів виконуються засобами АБД і цифрового модема. Узагальнену схему модифікованого методу NOMA наведено на рис. 2.

На рис. 2 використано такі позначення: C_k – парціальний спектральний кодер k -го каналу обробки; ЗШПрФ – блок зворотного швидкого

перетворення Фур'є; ЦАП – цифро-аналоговий перетворювач; М – модулятор; НК – еквівалентний неперервний тропосферний канал; ДМ – демодулятор; АЦП – аналого-цифровий перетворювач; ПУІ – блок послідовного усунення інтерференції; ШПФ – блок швидкого перетворення Фур'є; D_k – декодер k -го парціального каналу обробки.

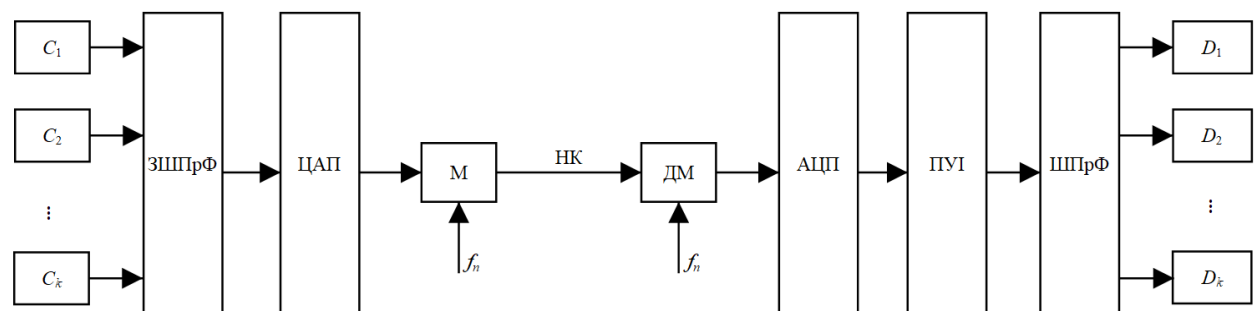


Рис. 2. Схема неортогонального множинного доступу на основі парціального спектрального кодування та послідовного усунення інтерференції

У передавальному тракті кожен абонентський сигнал надходить до відповідного парціального спектрального кодера C_k , у якому виконується цифрова КІХ-фільтрація. Така фільтрація формує для кожного абонентського сигналу індивідуальний спектрально-енергетичний профіль із заданим розподілом потужності в межах робочої смуги частот. Після цього сигнали сумуються, а сформований груповий спектр за допомогою ЗШПрФ перетворюється на груповий цифровий сигнал. Далі він надходить на ЦАП, переноситься на несучу частоту f_n і передається через НК, у якому зазнає амплітудно-фазових спотворень та частотно-селективних завмирань.

У приймальному тракті прийнятий сигнал після проходження через НК демодулюється в блоці ДМ, перетворюється в цифрову форму за допомогою АЦП та надходить до блока ПУІ. Блок ПУІ разом із ШПФ і декодерами D_k утворює ітераційний контур розділення сигналів: після оцінювання чергового абонентського сигналу його внесок реконструюється

та віднімається від прийнятого групового сигналу. Отже, структурно-функціональна реалізація доповнює апаратну основу МВЦТРС механізмом формування групового сигналу та розділення абонентських сигналів у спільному часово-частотному ресурсі. Індивідуальні спектрально-енергетичні профілі використовуються як додаткові сигнальні ознаки поряд із рівнями прийнятої потужності та можуть ураховуватися під час послідовного усунення інтерференції.

3.2 Аналітичні залежності парціального спектрального кодування

Математичну модель парціального спектрального кодування задано співвідношеннями (1)–(3) у розділі 2. Аналіз співвідношень (1) і (2) показує, що спектрально-енергетичний профіль k -го абонентського сигналу безпосередньо визначається коефіцієнтами та довжиною імпульсної характеристики відповідного КІХ-фільтра.

Зміна параметрів КІХ-фільтра забезпечує формування відмінних частотнозалежних профілів абонентських сигналів без розподілу робочої смуги на неперекривні частотні канали. Після неортогонального суміщення форма профілю може використовуватися як додаткова ознака під час оцінювання та реконструкції сигналів у процедурі SIC.

Поширені частотні вагові функції будують на основі тригонометричних складових, зокрема $\cos^2(\Omega)$ та $\sin^2(\Omega)$ [21, 22]. Для узагальненого керування формою профілів використано скінченний розклад (3) за системою ортогональних базисних функцій. Такий запис дає змогу формувати сімейство вагових функцій шляхом вибору типу базису, порядку функцій, коефіцієнтів розкладу та нормувального коефіцієнта.

3.3 Синтез частотних вагових функцій на основі ортогональних поліномів

Для синтезу частотних вагових функцій парціальних каналів обробки доцільно використовувати ортогональні поліноми, зокрема поліноми Чебишова I та II роду, Лежандра, а також поліноми Гегенбауера [23]. Надалі використовуватимемо такі позначення: $T_n(x)$ – поліном Чебишова I роду; $U_n(x)$ – поліном Чебишова II роду; $P_n(x)$ – поліном Лежандра; $C_n^{(\alpha)}(x)$ – поліном Гегенбауера з параметром α . Тут n – порядок полінома; α – параметр, що визначає конкретне сімейство поліномів Гегенбауера.

Використання ортогональних поліномів дає змогу формувати різні частотні вагові функції парціальних каналів обробки шляхом вибору типу базису, порядку поліномів і коефіцієнтів ортогонального розкладу.

Для ілюстрації властивостей поліноміальних базисів розглянуто поліноми порядків $n = 1, \dots, 5$. Використання невисоких порядків обмежує кількість членів розкладу та зменшує обчислювальну складність формування частотних вагових функцій.

Для узагальненого подання та можливості синтезу поліномів Чебишова I роду довільного порядку $n \geq 1$ може бути використано співвідношення:

$$T_n(x) = \frac{n}{2} \sum_{s=0}^q (-1)^s \frac{(n-s-1)!}{s! (n-2s)!} (2x)^{n-2s}, \quad (4)$$

де $q = \lfloor n/2 \rfloor$, s – індекс підсумовування.

Поліноми $T_n(x)$ для $n = 1, \dots, 5$, отримані з (4), мають вигляд:

$$\begin{aligned} T_1(x) &= x, \\ T_2(x) &= 2x^2 - 1, \\ T_3(x) &= 4x^3 - 3x, \\ T_4(x) &= 8x^4 - 8x^2 + 1, \\ T_5(x) &= 16x^5 - 20x^3 + 5x. \end{aligned}$$

Поліноми Чебишова II роду можуть бути задані рекурентним співвідношенням:

$$U_{n+1}(x) = 2xU_n(x) - U_{n-1}(x), \quad n \geq 1. \quad (5)$$

Поліноми $U_n(x)$ для $n = 1, \dots, 5$, отримані з (5), мають вигляд:

$$\begin{aligned} U_1(x) &= 2x, \\ U_2(x) &= 4x^2 - 1, \\ U_3(x) &= 8x^3 - 4x, \\ U_4(x) &= 16x^4 - 12x^2 + 1, \\ U_5(x) &= 32x^5 - 32x^3 + 6x. \end{aligned}$$

Поліноми $T_n(x)$ ортогональні з ваговою функцією:

$$\rho_T(x) = (1 - x^2)^{-1/2},$$

а поліноми $U_n(x)$ – з вагою:

$$\rho_U(x) = (1 - x^2)^{1/2}.$$

Поліноми Лежандра порядку $n \geq 1$ можуть бути задані рекурентним співвідношенням:

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x), \quad (6)$$

Поліноми $P_n(x)$ для $n = 1, \dots, 5$, отримані з (6), мають вигляд:

$$\begin{aligned} P_1(x) &= x, \\ P_2(x) &= \frac{3x^2 - 1}{2}, \\ P_3(x) &= \frac{5x^3 - 3x}{2}, \\ P_4(x) &= \frac{35x^4 - 30x^2 + 3}{8}, \\ P_5(x) &= \frac{63x^5 - 70x^3 + 15x}{8}. \end{aligned}$$

Поліноми $P_n(x)$ утворюють ортогональну систему з ваговою функцією:

$$\rho_P(x) = 1.$$

Поліноми Гегенбауера є параметричним сімейством ортогональних поліномів. За умови $\alpha > -\frac{1}{2}$, $\alpha \neq 0$ вони можуть бути задані формулою:

$$C_n^{(\alpha)}(x) = \sum_{s=0}^q (-1)^s \frac{\Gamma(\alpha + n - s)}{\Gamma(\alpha)s!(n-2s)!} (2x)^{n-2s}. \quad (7)$$

Із (7) для $n = 1, 2, 3$ отримуємо:

$$\begin{aligned} C_1^{(\alpha)}(x) &= 2\alpha x, \\ C_2^{(\alpha)}(x) &= 2\alpha(\alpha+1)x^2 - \alpha, \\ C_3^{(\alpha)}(x) &= \frac{4}{3}\alpha(\alpha+1)(\alpha+2)x^3 - 2\alpha(\alpha+1)x. \end{aligned}$$

Поліноми вищих порядків визначаються безпосередньо за співвідношенням (7).

Поліноми Гегенбауера утворюють ортогональну систему з ваговою функцією:

$$\rho_\alpha(x) = (1 - x^2)^{\alpha-1/2}, \quad \alpha > -1/2, \quad \alpha \neq 0.$$

Зв'язки між поліномами Гегенбауера, Лежандра та Чебишова II роду визначаються таким чином:

$$C_n^{(1/2)}(x) = P_n(x), \quad (8)$$

$$C_n^{(1)}(x) = U_n(x). \quad (9)$$

Зв'язок між поліномами Гегенбауера та Чебишова I роду отримується у граничному випадку $\alpha \rightarrow 0$ через нормовану форму:

$$\tilde{C}_n^{(0)}(x) = \lim_{\alpha \rightarrow 0} \frac{C_n^{(\alpha)}(x)}{\alpha} = \frac{2}{n} T_n(x), \quad n \geq 1. \quad (10)$$

Наведені співвідношення (8)–(10) показують, що поліноми Лежандра та Чебишова можуть розглядатися як окремі випадки або граничні форми поліномів Гегенбауера.

Це дає змогу використовувати єдиний параметризований базис для синтезу частотних вагових функцій парціальних каналів обробки.

Отже, поліноми $T_n(x)$, $U_n(x)$, $P_n(x)$ та $C_n^{(\alpha)}(x)$ можуть використовуватися як ортогональні базисні функції $g_r(x)$ у (3) для апроксимаційного подання частотної вагової функції $w_k(x)$.

Вибір типу поліноміального базису, порядку поліномів і коефіцієнтів $c_{k,r}$ дає змогу формувати різні частотні вагові функції парціальних каналів обробки та керувати спектрально-енергетичними профілями відповідних абонентських сигналів.

3.4 Схемотехнічна перевірка аналітичних залежностей для часткового випадку передатної функції

Загальна передатна функція КІХ-фільтра k -го парціального каналу обробки задана у (1). Для ілюстративного аналізу форми спектрально-енергетичного профілю можна розглядати окремий клас КІХ-фільтрів із двома ненульовими відліками імпульсної характеристики:

$$H_{k,p}(z) = 1 - z^{-p}, \quad p \geq 1, \quad (11)$$

де p – затримка у відліках між двома ненульовими коефіцієнтами імпульсної характеристики. Така

передатна функція є частковим випадком (1), для якого ненульовими є лише коефіцієнти $h_{k,p}[0] = 1$ та $h_{k,p}[p] = -1$, а для інших значень m виконується $h_{k,p}[m] = 0$.

Після підстановки $z = e^{j\Omega}$ у (11) отримуємо частотну характеристику фільтра $H_{k,p}(e^{j\Omega})$, а відповідний спектрально-енергетичний профіль задається як:

$$W_{k,p}(\Omega) = |H_{k,p}(e^{j\Omega})|^2 = |1 - e^{-jp\Omega}|^2 = 4\sin^2\left(\frac{p\Omega}{2}\right).$$

Оскільки максимальне значення спектрально-енергетичного профілю $W_{k,p}(\Omega)$ в розглянутому частотному інтервалі дорівнює 4, нормований профіль $\tilde{W}_{k,p}(\Omega)$, значення якого належать інтервалу $[0;1]$, визначається як:

$$\tilde{W}_{k,p}(\Omega) = \frac{W_{k,p}(\Omega)}{4} = \sin^2\left(\frac{p\Omega}{2}\right).$$

Аналіз наведеної залежності показує, що параметр p визначає кількість і положення екстремумів спектрально-енергетичного профілю в межах першої зони Найквіста. Для $p = 1$ нормований профіль монотонно зростає від нуля до одиниці, а для $p = 2$ досягає максимуму за значення $\Omega/\pi = 0,5$ та нульові значення на межах досліджуваного інтервалу.

За методикою, наведеною в розділі 2, у модулі Circuit Design програмного середовища AEDT 2025 R2 виконано схемотехнічне моделювання часткового випадку КІХ-фільтра для $p = 1$ і $p = 2$.

Отримані нормовані спектрально-енергетичні профілі зіставлено з аналітичними залежностями на рис. 3.

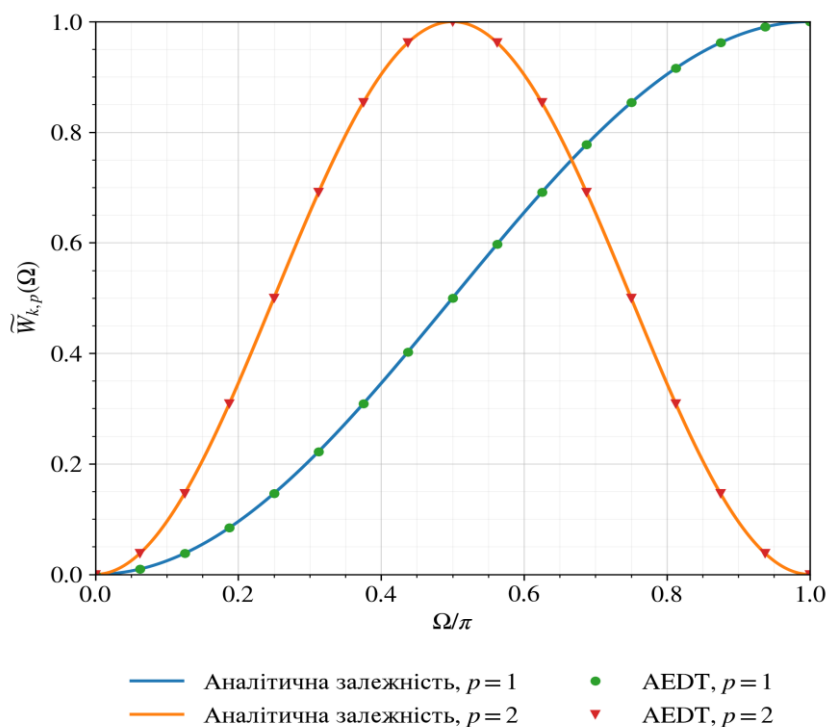


Рис. 3. Нормовані спектрально-енергетичні профілі

Зіставлення на рис. 3 демонструють узгодження змодельованих та аналітично визначених профілів для $p = 1$ і $p = 2$. Це підтверджує отриману залежність між параметром затримки КІХ-фільтра та формою спектрально-енергетичного профілю в межах розглянутого часткового випадку.

4. Обговорення результатів

Запропонована структурно-функціональна схема модифікованого методу NOMA визначає послідовність формування групового сигналу та розділення абонентських сигналів у спільному часово-частотному ресурсі МВЦТРС.

Особливістю схеми є введення парціальних спектральних кодерів і декодерів, які формують та використовують індивідуальні спектрально-енергетичні профілі абонентських сигналів. Це доповнює апаратну структуру МВЦТРС, наведену на рис. 1, сигнально-алгоритмічним механізмом багатостанційного доступу, реалізацію якого подано на рис. 2.

Співвідношення (1)–(3) встановлюють зв'язок між коефіцієнтами імпульсної характеристики КІХ-фільтра, його частотною характеристикою та спектрально-енергетичним профілем абонентського сигналу. Отже, зміна коефіцієнтів і довжини КІХ-фільтра дає змогу формувати відмінні частотнозалежні профілі без поділу робочої смуги на неперекривні частотні канали. Це забезпечує можливість використання форми спектрального розподілу енергії як додаткової ознаки під час оцінювання, реконструкції та послідовного усунення абонентських сигналів.

Подання частотних вагових функцій за системами ортогональних поліномів дає змогу параметрично керувати формою спектрально-енергетичних профілів шляхом вибору типу базису, порядку поліномів і коефіцієнтів розкладу. Співвідношення (8)–(10) показують, що поліноми Лежандра та Чебишова можуть розглядатися як окремі випадки або граничні форми поліномів Гегенбауера. Це дає змогу використовувати параметризований поліноміальний базис для синтезу частотних вагових функцій парціальних каналів обробки.

Схемотехнічне моделювання двовідлікового КІХ-фільтра, заданого співвідношенням (11), показало практичний збіг змодельованих та аналітично визначених нормованих спектрально-енергетичних профілів. Отриманий результат пояснюється безпосереднім зв'язком між параметром затримки p , частотною характеристикою КІХ-фільтра та положенням екстремумів профілю.

На відміну від PD-NOMA [12, 14, 15], де основною ознакою розділення сигналів є різниця рівнів потужності, у запропонованому методі додатково використовується форма частотнозалежного спектрально-енергетичного профілю. На відміну від CD-NOMA [14, 15], де розділення користувачів забезпечується кодовими книгами або спеціальними сигнальними структурами, запропонований підхід формує сигнальні ознаки засобами цифрової КІХ-фільтрації.

Таким чином, відмінністю запропонованого методу є поєднання неортогонального суміщення сигналів, керованого спектрального зважування та процедури SIC у структурі МВЦТРС. Запропонований метод дав змогу визначити принципи структурно-функціональної реалізації багатостанційного доступу, формалізувати парціальне спектральне кодування як цифрову КІХ-фільтрацію та сформулювати математичну основу синтезу частотних вагових функцій. Практичне значення результатів полягає в можливості програмної реалізації парціальних спектральних кодерів і декодерів у цифровому модемі МВЦТРС. Зміна коефіцієнтів КІХ-фільтрів створює передумови для адаптації спектрально-енергетичних профілів до характеристик окремих тропосферних напрямків.

Запропонований метод слід застосовувати з урахуванням таких обмежень:

- у роботі не задано параметричної або статистичної моделі тропосферного каналу;

- не досліджено вплив похибок оцінювання каналу та синхронізації сигналів, а також залишкової інтерференції після SIC;

- не визначено допустимий ступінь перекриття спектрально-енергетичних профілів і граничну кількість одночасно обслуговуваних абонентських станцій. Крім того, практичне застосування запропонованого підходу потребує частотної еквалізації та реконструкції внесків окремих сигналів у приймальному тракті.

Недоліком проведеного дослідження є відсутність повної багатокористувацької моделі та кількісного порівняння запропонованого методу з OMA, PD-NOMA і CD-NOMA за однакових початкових умов. Схемотехнічна перевірка охоплює лише частковий випадок КІХ-фільтра з двома ненульовими відліками імпульсної характеристики. Тому можливе підвищення спектральної ефективності та стійкості розділення сигналів слід розглядати як обґрунтовану передумову, а не як кількісно підтверджену перевагу.

Подальший розвиток дослідження доцільно спрямувати на розроблення алгоритму розрахунку коефіцієнтів ортогонального розкладу та критеріїв вибору поліноміального базису, дослідження впливу кількості абонентських сигналів, ступеня перекриття профілів, порядку КІХ-фільтрів і похибок оцінювання каналу на якість розділення. Наступним етапом дослідження мають бути моделювання повної багатокористувацької системи в частотно-селективному тропосферному каналі та порівняння модифікованого методу з OMA, PD-NOMA і CD-NOMA за спектральною ефективністю, імовірністю бітової помилки, відношенням сигналу до суми інтерференції та шуму, рівнем залишкової інтерференції після SIC й обчислювальною складністю приймача.

5. Висновки

У статті розроблено та аналітично описано модифікований метод неортогонального

множинного доступу з парціальним спектральним кодуванням для МВЦТРС, що функціонує за топологією PtMP. Визначено структурно-функціональну схему формування групового сигналу та розділення абонентських сигналів у спільному часово-частотному ресурсі.

Парціальне спектральне кодування формалізовано як цифрову КІХ-фільтрацію. Встановлено, що коефіцієнти та довжина імпульсної характеристики КІХ-фільтра визначають форму спектрально-енергетичного профілю абонентського сигналу. Це дає змогу використовувати частотно-залежний профіль як додаткову сигнальну ознаку під час розділення неортогонально суміщених сигналів.

Показано можливість використання поліномів Лежандра, Чебишова та Гегенбауера як ортогональних базисних функцій для параметричного синтезу частотних вагових функцій парціальних каналів обробки. Вибір поліноміального базису, порядку поліномів і коефіцієнтів розкладу визначає форму відповідного спектрально-енергетичного профілю.

Для розглянутого часткового випадку КІХ-фільтра результати схематичного моделювання за $p = 1$ і $p = 2$ узгоджуються з аналітично визначеними нормованими спектрально-енергетичними профілями. Це підтверджує отриману залежність між параметром затримки КІХ-фільтра та формою профілю в межах розглянутого часткового випадку.

Практичне значення результатів полягає в

можливості програмної реалізації парціальних спектральних кодерів у цифровому режимі МВЦТРС і використання сформованих профілів під час розділення абонентських сигналів. Водночас кількісне підвищення спектральної ефективності та стійкості розділення в межах проведеного дослідження не підтверджено.

Подальші дослідження доцільно спрямувати на розроблення алгоритму розрахунку коефіцієнтів ортогонального розкладу, моделювання повної багатокористувацької системи та порівняння модифікованого методу з ОМА, PD-NOMA і CD-NOMA за показниками ефективності.

Конфлікт інтересів

Автори заявляють про відсутність конфлікту інтересів, зокрема фінансового, авторського чи іншого характеру, який міг би вплинути на дослідження або результати, опубліковані в цій статті.

Фінансування

Дослідження проводилося без фінансової підтримки.

Доступність даних

Рукопис не має пов'язаних даних.

Використання засобів штучного інтелекту

Автори підтверджують, що не використовували технології штучного інтелекту при створенні представленої роботи.

СПИСОК ЛІТЕРАТУРИ

- Comtech (2024), *COMET Data Sheet*, technical data sheet, available at: https://comtech.com/wp-content/uploads/2024/07/Comet-Data-Sheet_19224.pdf
- Cubic (2023), *GATR 2.4m Troposcatter Data Sheet*, technical data sheet, available at: https://www.cubic.com/sites/default/files/2023-02/Cubic_Datasheet_GATR_Troposcatter.pdf
- Ultra I&C (2025), *Archer Troposcatter Terminals Data Sheet*, technical data sheet, available at: https://www.ultra-ic.com/media/caboizyk/archer-datasheet-july2025_web.pdf
- Spectra Group (UK) Limited (n.d.), *Over The Horizon Communications Networks*. Retrieved June 17, 2026, available at: <https://spectra-group.co.uk/tropo/>
- Sarbu, A., Alexandrescu, D., and Galben, C. (2025), "Prospects of troposcatter systems in military communication networks", *International Conference Knowledge-Based Organization*, vol. 31, no. 3, pp. 142–148, doi: <https://doi.org/10.2478/kbo-2025-0090>
- Colom Piella, G., and Marzal Ruano, C. (2025), "Has Russian Electronic Warfare Underperformed in the Ukrainian Conflict?", *Journal of Strategic Security*, vol. 18, no. 4, pp. 78–95, doi: <https://doi.org/10.5038/1944-0472.18.4.2369>
- Почерняєв В. М., Повхліб В. С., Наритник Т. М. Мобільна вузлова цифрова тропосферна станція: пат. 126206 Україна: МПК Н04В 7/04, Н04В 7/0413. № а202001674; заявл. 10.03.2020; опубл. 31.08.2022, Бюл. № 35.
- ITU-R (2025), *Recommendation ITU-R P.617-6: Propagation prediction techniques and data required for the design of trans-horizon radio-relay systems*, International Telecommunication Union, Geneva, available at: <https://www.itu.int/rec/R-REC-P.617-6-202509-I/en>
- Shao, J., Liu, Z., Liu, Y., and Xie, T. (2024), "Performance Analysis of Troposphere Scattering Communication Channel with Chirp-BOK Modulation", *Entropy*, vol. 26, no. 12, article 1052, doi: <https://doi.org/10.3390/e26121052>
- Goldsmith, A. (2005), *Wireless Communications*, Cambridge University Press, Cambridge, 644 p., doi: <https://doi.org/10.1017/CBO9780511841224>
- Tse, D., and Viswanath, P. (2005), *Fundamentals of Wireless Communication*, Cambridge University Press, Cambridge, 564 p., doi: <https://doi.org/10.1017/CBO9780511807213>
- Saito, Y., Kishiyama, Y., Benjebbour, A., Nakamura, T., Li, A., and Higuchi, K. (2013), "Non-Orthogonal Multiple Access (NOMA) for Cellular Future Radio Access", *2013 IEEE 77th Vehicular Technology Conference (VTC Spring)*, pp. 1–5, doi: <https://doi.org/10.1109/VTCSpring.2013.6692652>
- 3GPP (2018), *3GPP TR 38.812, version 1.0.0: Study on Non-Orthogonal Multiple Access (NOMA) for NR*, technical report, 3rd Generation Partnership Project, available at: <https://www.3gpp.org/dynareport/38812.htm>
- Dai, L., Wang, B., Ding, Z., Wang, Z., Chen, S., and Hanzo, L. (2018), "A Survey of Non-Orthogonal Multiple Access for 5G", *IEEE Communications Surveys & Tutorials*, vol. 20, no. 3, pp. 2294–2323, doi: <https://doi.org/10.1109/COMST.2018.2835558>

15. Ding, Z., Lei, X., Karagiannidis, G. K., Schober, R., Yuan, J., and Bhargava, V. K. (2017), "A Survey on Non-Orthogonal Multiple Access for 5G Networks: Research Challenges and Future Trends", *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 10, pp. 2181–2195, doi: <https://doi.org/10.1109/JSAC.2017.2725519>
16. Liu, Y., Zhang, S., Mu, X., Ding, Z., Schober, R., Al-Dhahir, N., Hossain, E., and Shen, X. (2022), "Evolution of NOMA Toward Next Generation Multiple Access (NGMA) for 6G", *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 4, pp. 1037–1071, doi: <https://doi.org/10.1109/JSAC.2022.3145234>
17. Makki, B., Chitti, K., Behravan, A., and Alouini, M.-S. (2020), "A Survey of NOMA: Current Status and Open Research Challenges", *IEEE Open Journal of the Communications Society*, vol. 1, pp. 179–189, doi: <https://doi.org/10.1109/OJCOMS.2020.2969899>
18. Kebede, T., Wondie, Y., Steinbrunn, J., Belay, H. B., and Kornegay, K. T. (2022), "Multi-Carrier Waveforms and Multiple Access Strategies in Wireless Networks: Performance, Applications, and Challenges", *IEEE Access*, vol. 10, pp. 21120–21140, doi: <https://doi.org/10.1109/ACCESS.2022.3151360>
19. Kazemian, M., Abouei, J., and Anpalagan, A. (2021), "A low complexity enhanced-NOMA scheme to reduce inter-user interference, BER and PAPR in 5G wireless systems", *Physical Communication*, vol. 48, article 101412, doi: <https://doi.org/10.1016/j.phycom.2021.101412>
20. Alkhazzar, A. M. A., and Aghaeinia, H. (2021), "Symbol error rate analysis of non-orthogonal multiple access systems", *Physical Communication*, vol. 46, article 101295, doi: <https://doi.org/10.1016/j.phycom.2021.101295>
21. Harris, F. J. (1978), "On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform", *Proceedings of the IEEE*, vol. 66, no. 1, pp. 51–83, doi: <https://doi.org/10.1109/PROC.1978.10837>
22. Nuttall, A. H. (1981), "Some Windows with Very Good Sidelobe Behavior", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 29, no. 1, pp. 84–91, doi: <https://doi.org/10.1109/TASSP.1981.1163506>
23. Korn, G. A., and Korn, T. M. (1968), *Mathematical Handbook for Scientists and Engineers: Definitions, Theorems, and Formulas for Reference and Review*, 2nd ed., McGraw-Hill, New York, 1130 p.

Received (Надійшла) 22.06.2026

Accepted for publication (Прийнята до друку) 09.09.2026

Publication date (Дата публікації) 25.09.2026

ВІДОМОСТІ ПРО АВТОРІВ/ ABOUT THE AUTHORS

Коваленко Олексій Олександрович – начальник інституту, Військовий інститут телекомунікацій та інформатизації імені Героїв Крут, Київ, Україна;

Oleksii Kovalenko – Chief of the Institute, Kruty Heroes Military Institute of Telecommunications and Information Technology, Kyiv, Ukraine;

e-mail: oleksii.kovalenko@viti.edu.ua; ORCID Author ID: <https://orcid.org/0009-0008-4087-8770>.

Почерняєв Віталій Миколайович – доктор технічних наук, професор, Національна академія Служби безпеки України, Київ, Україна;

Vitaly Pocherniaiev – Doctor of Technical Sciences, Professor, National Academy of the Security Service of Ukraine, Kyiv, Ukraine;

e-mail: vpochernyaev@gmail.com; ORCID Author ID: <https://orcid.org/0000-0001-7130-8668>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=6603517180>.

Івко Сергій Олександрович – кандидат технічних наук, старший науковий співробітник НДУ перспектив розвитку інформаційних систем, Військовий інститут телекомунікацій та інформатизації імені Героїв Крут, Київ, Україна;

Serhii Ivko – Candidate of Technical Sciences, Senior Researcher, Research Directorate for Prospective Development of Information Systems, Kruty Heroes Military Institute of Telecommunications and Information Technology, Kyiv, Ukraine;

e-mail: serhii.ivko@viti.edu.ua; ORCID Author ID: <https://orcid.org/0000-0002-8723-9035>.

Modified Method of Non-Orthogonal Multiple Access with Partial Spectral Coding for Mobile Nodal Digital Troposcatter Stations

Oleksii Kovalenko, Vitaly Pocherniaiev, Serhii Ivko

Abstract. Subject. A modified non-orthogonal multiple access method using partial spectral coding of user signals for a mobile nodal digital troposcatter station. **Objective.** To analytically substantiate the generation and separation of multiple user signals in a shared time-frequency resource of a troposcatter channel. **Research tasks.** To define the structural and functional diagram for composite-signal generation and user-signal separation; formulate a mathematical description of partial spectral coding; investigate the use of orthogonal polynomials to synthesize frequency-weighting functions; and verify the derived analytical relationships through circuit-level simulation. **Methods.** The methods included analytical modeling of signal transformations, the z-transform, frequency-response analysis of digital filters, expansion of weighting functions in orthogonal-polynomial bases, and circuit-level simulation for a particular transfer function. **Results.** An individual spectral-energy profile was proposed for each user signal as an additional discriminating feature after non-orthogonal superposition. Analytical expressions were derived for the transfer functions of partial processing channels, spectral-energy profiles, and frequency-weighting functions. Legendre, Chebyshev, and Gegenbauer polynomials were shown to be suitable for controlling the profile shape. Circuit-level simulation confirmed the derived analytical relationships for spectral-energy profile formation in the particular case considered. **Conclusions.** The proposed approach provides a signal-processing and algorithmic foundation for the multiuser physical layer of a troposcatter station and establishes a basis for improving spectrum utilization and the robustness of user-signal separation.

Keywords: non-orthogonal multiple access; partial spectral coding; mobile nodal digital troposcatter station; troposcatter channel; successive interference cancellation; finite impulse response filter; orthogonal polynomials.

Н. Г. Кучук, С. С. Пироженко

Харківський національний університет радіоелектроніки, Харків, Україна

МЕТОДИКА РОЗРАХУНКУ ОВЕРХЕДУ LIVE-МІГРАЦІЇ ВІРТУАЛЬНИХ МАШИН НА БАЗІ КОНЦЕПЦІЇ РЕСУРСНОГО КУБА NRC

Анотація. Актуальність. Актуальні вимоги до енергоефективного та адаптивного функціонування сучасних дата-центрів обумовлені стрімкою розбудовою хмарного сегмента. За таких умов інструментарій міграції ВМ у гарячому режимі стає фундаментом гнучкого розподілу апаратного фонду. Завдяки здатності перемішувати віртуальні машини в процесі їхньої роботи без критичних пауз у доступі до додатків, адміністратори успішно вирішують проблеми оптимізації живлення шляхом консолідації навантажень, усувають локальні перевантаження хостів та обслуговують технічну базу без зупинки бізнес-процесів. З іншого боку, процедура безперервної міграції неминуче породжує паразитні витрати ресурсів. Функціонування ітеративних схем (Pre-Copy) вимагає постійного перехоплення модифікованих блоків оперативної пам'яті, їхньої повторної передачі та фінального узгодження контекстів CPU, що суттєво деформує ресурсний профіль серверів. Як наслідок, виникають різкі стрибки утилізації обчислювальних ядер гіпервізором, вичерпується місткість каналів передачі даних і зростає час відгуку. Хаотичний або нерациональний запуск таких процедур створює ризики взаємного пригнічення сусідніх задач, збоїв у виконанні нормативів SLA і навіть ризик зацікнення передачі даних, коли сторінки перезаписуються швидше, ніж надсилаються мережею. **Предмет дослідження:** процеси виникнення та методики розрахунку накладних витрат (паразитного навантаження на обчислювальні ядра, пам'ять і мережеві канали) під час живої міграції віртуальних машин, а також математичні моделі й алгоритми їх просторового оцінювання на основі векторної концепції ресурсного куба NRC. **Мета дослідження:** підвищення ефективності динамічного керування ресурсами хмарного дата-центру та запобігання порушенням угод SLA шляхом розроблення моделі і методу оцінювання накладних витрат живої міграції віртуальних машин на основі векторного простору ресурсного куба NRC. **Завдання дослідження.** Сформулювати критерій мінімізації навантаження з урахуванням ємності серверів у моделі NRC та реалізувати алгоритм вибору цільових хостів у класі Load Balancing. **Методи дослідження.** Методи векторного та багатовимірного геометричного аналізу, методи системного аналізу та теорії систем, теорія графів та деревоподібних структур даних. **Результати дослідження.** Доведено, що ефективна місткість хостів зменшується на величину векторного оверхеду міграції, що унеможливорює помилкове перевантаження системи в процесі виконання балансувальних операцій. Запропонована схема формує математичну основу для розробки адаптивних алгоритмів релокації ВМ із гарантованим дотриманням граничних умов куба NRC.

Ключові слова: міграція віртуальних машин; передача даних; алгоритм імітації відпалу; генетичний алгоритм; екосистема Інтернету речей.

1. Вступ

1.1 Актуальність дослідження

Стрімка цифровізація, зростання обсягів обробки великих даних і масштабування хмарних центрів обробки даних (ЦОД) висувають жорсткі вимоги до гнучкості, енергоефективності та безперервності функціонування обчислювальної інфраструктури [1]. Фундаментальним механізмом оперативного керування ресурсами в сучасному віртуалізованому середовищі є технологія «живої» міграції віртуальних машин (Live VM Migration) [2]. Вона забезпечує динамічне усунення перевантажень на фізичних серверах, консолідацію навантаження з метою вимкнення незадіяного обладнання для енергозбереження, а також проведення технічного обслуговування апаратної частини без помітної зупинки клієнтських сервісів.

Проте практична реалізація живої міграції неминуче супроводжується формуванням вторинного навантаження — накладних витрат (overhead) [3]. Ітераційне копіювання пам'яті (зокрема за алгоритмами типу Pre-Copy), безперервна фіксація та донадсилання модифікованих сторінок (dirty pages), а також фінальна синхронізація стану процесора створюють різкі сплески утилізації CPU гіпервізором і спричиняють значну деградацію пропускну

здатності каналів зв'язку. Якщо процес міграції ініціюється без належного розрахунку поточних потужностей, виникає ефект «галасливого сусіда» (noisy neighbor), зростає латентність і виникають масові порушення угод про рівень обслуговування (SLA). У найгірших випадках висока інтенсивність перезапису пам'яті додатками призводить до розбіжності ітераційного процесу, коли міграція заціклюється, виснажуючи ресурси системи.

Більшість класичних евристик і моделей планування розглядають ресурси ізольовано (переважно орієнтуючись лише на процесорні цикли або загальний обсяг оперативної пам'яті), ігноруючи складну багатовимірну інтерференцію між обчислювальними, ємнісними та мережевими компонентами. Застосування векторної моделі ресурсного куба (Network-Resource Cube, NRC) дозволяє комплексно формалізувати стан серверів і перебіг оверхеду в єдиному координатному просторі R3 (Compute–Memory–Network). Це уможливорює точний предиктивний розрахунок супутніх втрат продуктивності, запобігання виходу навантаження за межі апаратних можливостей хостів та підвищення стабільності систем диспетчеризації.

Таким чином, розроблення методики розрахунку накладних витрат live-міграції на основі концепції ресурсного куба NRC та її алгоритмічна інтеграція в контур балансування навантаження є

актуальною науково-прикладною задачею, спрямованою на підвищення якості обслуговування, керованості та енергетичної ефективності сучасних хмарних інфраструктур.

1.2 Аналіз літературних джерел і визначення проблеми

Питання оптимізації розміщення завдань, консолідації віртуальних машин та мінімізації накладних витрат у хмарних ЦОД перебувають у центрі уваги багатьох провідних вітчизняних і закордонних дослідників.

Теоретичні засади протоколів Pre-Copy та Post-Copy live-міграції закладені в роботах [5, 6]. Розвитком аналітичних моделей для оцінювання часу міграції та часу простою залежно від темпу модифікації сторінок пам'яті займалась дослідницька група під керівництвом С. Акуша [7].

Для вирішення задач багатовимірного пакування рюкзака під час міграції ВМ широко застосовуються еволюційні алгоритми, наприклад, алгоритм кажанів [8]. Спроби узагальнити ресурсний профіль вузлів у вигляді багатовимірних векторів та гіперкубів робилися в роботах [9, 10], а також у дослідженнях векторного домінування ресурсів [11]. Концепція ресурсного куба (Resource Cube) також епізодично розглядалася в контексті статичного планування задач у середовищах OpenStack та CloudSim [12, 13].

Отже на основі проведеного аналізу робіт [5–13] є підстави стверджувати, що проблема полягає у відсутності комплексної математичної моделі та алгоритмів оптимізації розміщення віртуальних машин, які б одночасно враховували багатовимірні ресурсні обмеження вузлів, просторову топологію мережі та транзитний динамічний оверхед безпосередньо в процесі їх міграції.

2. Матеріали і методи дослідження

Об'єктом дослідження виступає процес динамічного керування та перерозподілу ресурсів хмарного центру обробки даних в умовах виконання live-міграції віртуальних машин.

Методи дослідження базуються на теорії графів та деревоподібних структур даних, методах системного аналізу та теорії систем.

У процесі розробки методу керування розподілом обчислювального ресурсу туманного шару мобільного високошвидкісного Інтернету речей була застосована низка різноманітних методів та алгоритмів.

Математична модель міграції на базі ресурсного куба NRC базується на ключових засадах теорії векторних просторів і нормалізації в n -вимірному евклідовому просторі. Векторний стан вузла та ВМ характеризується ресурсами фізичного вузла (H_i) та вимогами віртуальної машини (VM), які описуються векторами у нормованому просторі:

$$r = (u_{cpu}, u_{ram}, u_{net}), |u| \in [0, 1].$$

Об'єм і запас міцності NRC визначає ресурсний куб, також він визначає допустиму робочу область.

Запас стійкості хоста розраховується як евклідова відстань від поточної точки сумарного навантаження H_i до критичних меж гіперкуба:

$$d(H_i, 1) = \sqrt{(1 - u_{cpu})^2 + (1 - u_{ram})^2 + (1 - u_{net})^2}.$$

Цільова функція визначається балансом між зменшенням енергоспоживання, мінімізацією часу міграції і недопущенням порушення SLA:

$$F = \alpha \times Cost_{placement} + \beta \times \int_0^{T_{mig}} O_{mig}(t) dt + \gamma \times Penalty_{interference},$$

де α , β та γ – відповідні вагові коефіцієнти, T_{mig} – час міграції, $Cost_{placement}$ – узагальнена вартість міграції, O_{mig} – часові витрати варіанту міграції, $Penalty_{interference}$ – штрафна функція.

Система нерівностей-обмежень формується жорсткими обмеженнями неперевиконання сумарних векторів ресурсів над ємністю куба вузла в кожен момент часу на розглядаємому часовому інтервалу $t \in [0, T_{mig}]$.

Таким чином, розробка спеціалізованої методики розрахунку накладних витрат у координатах тривимірного простору NRC усуває зазначені прогалини, забезпечуючи перехід від реактивного усунення перевантажень до предиктивного векторного керування стійкістю системи.

3. Результати розробки і дослідження векторної моделі простору NRC та формалізація накладних витрат міграції

3.1 Математична модель векторної моделі простору NRC

Розглянемо зазначені категорії обчислювально-комунікаційних ресурсів (ОКР) як три виміри абстрактного об'єкта. Завдяки процедурі нормалізації метрик ОКР уздовж кожної координатної осі, сукупний пул доступних потужностей фізичної інфраструктури множини F відображається у формі одиничного куба [14] (рис. 1). Цей концепт визначено як нормалізований ресурсний куб (Normalized Resource Cube – NRC).

Мовою векторної алгебри гранична потужність віртуалізованих ресурсів за кожним із базових напрямів (C, M, N) задається одиничними базисними ортами простору NRC.

Аналітичні вектори наявності, резервування та запиту ресурсів відображаються безпосередньо всередині одиничного об'єму NRC.

Вектор наявності ресурсів (ВНР) характеризує інтегральний апаратний потенціал усієї множини F і спрямований із початку координат (0, 0, 0) у протилежну вершину (1, 1, 1):

$$ВНР = C + M + N.$$

Вектор виділення ресурсів (ВВР): відображає частку трьох типів ОКР, уже алоковану підмножині активних віртуальних машин $VA \subset V$.

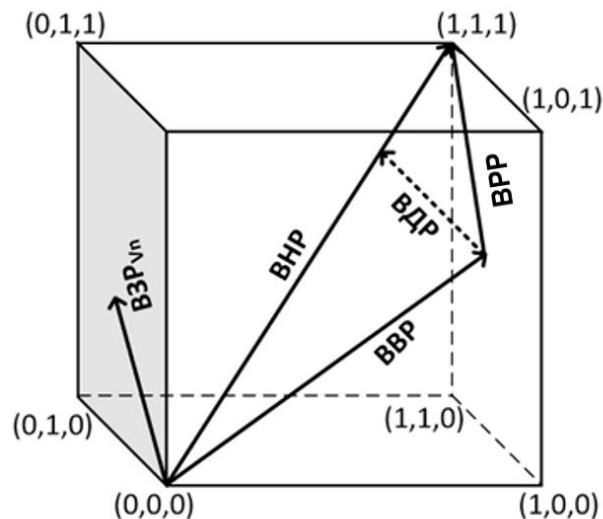


Рис. 1. Геометрична інтерпретація простору NRC та взаємне розташування векторів стану системи

Нормалізація компонентів C , M , N здійснюється безпосередньо відносно значень вектора BHP .

Вектор залишкових ресурсів (BPP): визначає невикористаний резерв системи і формується як геометрична різниця:

$$BPP = BHP - BVP$$

Вектор запиту ресурсів $B3Pv_n$: акумулює вимоги до ОКР з боку окремої віртуальної машини $v_n \in V$ і обчислюється через векторну суму нормалізованих складових Δ_j кожного типу ресурсу.

Рівень гармонійності (збалансованості) завантаження інфраструктури оцінюється ступенем просторового зближення напрямків BHP та BVP . За умови їхньої повної колінеарності спостерігається ідеально рівномірна утилізація процесорних потужностей, оперативної пам'яті та мережевих інтерфейсів. Для кількісної оцінки перекосу навантаження вводиться вектор дисбалансу ресурсів (BDR), що є векторною різницею між ортогональною проекцією BVP на головну діагональ BHP і фактичним положенням вектора BVP :

$$\begin{aligned} BDR &= \text{proj}_{BHP}(BVP) - BVP = \\ &= \left(\frac{BVP \times BHP}{|BHP|^2} \right) BHP - BVP. \end{aligned}$$

Таке просторове представлення дає змогу використовувати модуль $|BDR|$ як цільовий оптимізаційний критерій в алгоритмах перерозподілу навантаження.

Зокрема, під час вибору віртуального інстансу $v_n \in V$ для консолідації на фізичному вузлі перевага надається тому розміщенню, яке максимально мінімізує довжину вектора BDR , забезпечуючи рівновагу між осями C , M і N .

Аналітичне дослідження процедур живої міграції (як ітеративного копіювання Pre-Copy, так і передачі на вимогу Post-Copy) свідчить, що перенесення гостей систем створює вторинне навантаження на фізичні вузли [15, 16].

Оскільки ключовим процесом міграції виступає фонові синхронізація пам'яті та передача сторінок

через мережеву підсистему між сервером-джерелом і сервером-приймачем [17, 18], то головними споживачами у такому випадку стають ресурси типів M (RAM) та N (Network), доповнені обчислювальними витратами процесора на стиснення та шифрування.

Відтак, векторна модель ресурсного куба потребує обов'язкового врахування супутнього оверхеду міграції, вектора $OKP_{пр.жм}$.

Реальний обсяг доступного фонду системи в період динамічного балансування коригується відніманням паразитного навантаження:

$$BHP_{дост.}(t) = BHP - OKP_{пр.жм}(t).$$

Тобто ефективна місткість хостів зменшується на величину векторного оверхеду міграції, що унеможливує помилкове перевантаження системи в процесі виконання балансувальних операцій. Запропонована схема формує математичну основу для розробки адаптивних алгоритмів релокації ВМ із гарантованим дотриманням граничних умов куба NRC.

Умови моделювання: серія експериментів у середовищі CloudSim на вибірці з 100 фізичних вузлів (heterogeneous hosts) та 400 віртуальних машин під стандартизованим динамічним навантаженням (PlanetLab traces).

Нижче наведено зведену таблицю (табл. 1) результатів експериментального порівняння запропонованого векторного методу на основі простору NRC із традиційними алгоритмами балансування та консолідації віртуальних машин:

- BFD - Best Fit Decreasing,
- FFD - First Fit Decreasing,
- DVFS - динамічне регулювання напруги й частоти без просторового аналізу мережі.

Завдяки врахуванню вектора накладних витрат $OKP_{пр.жм}$ у формулі доступного фонду ресурсів вдалося уникнути вторинного лавиноподібного перевантаження хостів-реципієнтів, що скоротило загальну кількість міграцій із 960–1180 до 645.

Таблиця 1 – Результати експериментального порівняння запропонованого векторного методу на основі простору NRC із традиційними алгоритмами балансування та консолідації віртуальних машин

Критерій / Метрика	Традиційний FFD	Модифікований BFD	Евристика на базі CPU-порогів (IQR / MAD)	Запропонований метод на базі куба NRC	Ефект / Поліпшення
Кількість порушень SLA (SLAV, %)	9.42%	6.85%	4.12%	1.34%	Зниження у ~3.1 рази відносно базових підходів
Середній модуль дисбалансу ВДР	0.481	0.392	0.315	0.118	Вирівнювання завантаження по осях C, M, N на 62.5%
Сумарна кількість живих міграцій ВМ	1420	1180	960	645	Скорочення зайвих релокацій на 32.8%
Загальний час оверхеду міграцій $\sum T_{mig}, \text{ хв}$	312.4	248.6	194.2	116.5	Зменшення навантаження на мережу на 40.0%
Середній час простою ВМ ($T_{downtime}, \text{ с} / \text{ ВМ}$)	1.84 с	1.52 с	1.15 с	0.68 с	Зниження паузи сервісів на 40.8%
Випадки колізій мережевої осі N (викиди за межі куба)	84	61	43	0	Повне усунення за рахунок динамічного зменшення ВНР _{доступний}
Інтегральне споживання енергії (кВт·год)	184.2	162.5	145.8	138.4	Додаткова економія електроенергії на 5.1% – 14.8%

Алгоритми BFD та IQR орієнтовані переважно на утилізацію процесора, що призводило до 43–61 випадку вичерпання смуги пропускання.

Векторна модель куба NRC гарантує перебування робочої точки всередині одиничного об'єму 0, 13, звівши випадки переповнення мережі до нуля.

Комплексна мінімізація норми ВДР забезпечила синхронне вивільнення вузьких місць системи, завдяки чому метрика порушення SLA знизилася до мінімального значення 1.34%.

4. Обговорення результатів дослідження

У роботі проведено порівняльний аналіз функціонування розробленого методу та класичних аналогів (FFD, BFD, порогові евристики IQR/MAD) засвідчив якісні переваги переходу від одновимірних метрик до концепції тривимірного простору NRC.

Ключові кількісні та якісні аспекти отриманих результатів (табл. 1):

Мінімізація просторового дисбалансу: Завдяки оптимізації розміщення за критерієм мінімізації довжини вектора дисбалансу ВДР, середній показник перекошу утилізації знизився до 0.118 (поліпшення на 62.5% порівняно з BFD).

Це свідчить про близькість траєкторії сумарного навантаження до головної просторової діагоналі

вектора ВНР, що запобігає передчасному вичерпанню одного ресурсу при значному недовантаженні інших.

Ліквідація мережевих колізій і «галасливого сусіда»: у той час як евристики на базі CPU-порогів генерували від 43 до 84 випадків перевищення мережевої ємності, у моделі NRC цей показник зведено до 0.

Це досягнуто через предиктивне динамічне зменшення ефективного обсягу ресурсів:

$$\text{ВНР}_{\text{доступний}}(t) = \text{ВНР} - \text{ОКР}_{\text{пр.жм}}(t).$$

Тимчасове резервування смуги пропускання на період T_{mig} унеможливило призначення реципієнтом хоста, який не має достатнього запасу міцності для пропуску фонових трафіку міграції.

Зниження транзитного оверхеду та стабільність SLA визначає запобігання лавиноподібним повторним міграціям дозволило зменшити загальну кількість релокацій на 32.8% (до 645 випадків), що скоротило сумарний час додаткового навантаження мережі на 40.0% (до 116.5 хв).

Як наслідок, коефіцієнт деградації SLA (SLAV) знизився до 1.34% (у 3.1 раза нижче, ніж у порогових евристик, і в 7 разів нижче, ніж у класичного FFD).

Енергетична ефективність: скорочення паразитного навантаження на обчислювальні ядра, зумовлене зменшенням кількості циклів стиснення/шифрування пам'яті гіпервізором,

забезпечило додаткове зменшення енергоспоживання інфраструктури до 138.4 кВт·год (економія від 5.1% до 14.8%).

Оптимальною для перенесення або консолідації є така віртуальна машина vp , чий вектор запиту $V3Pvp$ зменшує евклідову норму ВДР. У наведеному прикладі пріоритетом є вивільнення процесорного ресурсу (C) або підселення ВМ із переважанням мережевого трафіку (N), що наближає вектор ВВР до діагоналі ВНР.

Врахування оверхеду ОКР_{нр.жр} наочно показує, що доступний простір для розміщення нових інстансів тимчасово звужується (в прикладі мережевий ліміт знижується з 1.0 до 0.50).

Ігнорування цього фактора в класичних алгоритмах призводить до миттєвого переповнення каналу та зростання затримок пакета.

5. Висновки

1. Розроблено комплексну математичну модель та методику оцінювання оверхеду live-міграції ВМ на базі нормалізованого ресурсного куба NRC, що дозволяє формалізувати взаємодію обчислювальних, емнісних і мережевих параметрів вузлів дата-центру в єдиному координатному просторі R^3 .

2. Предиктивне врахування динамічного оверхеду: Доведено, що формалізація накладних витрат у вигляді вектора ОКР_{нр.жм(t)} та його віднімання від вектора повної місткості ВНР запобігає деградації каналів зв'язку і виникненню ефекту «галасливого сусіда». Це гарантує локалізацію робочої точки системи виключно всередині допустимого об'єму куба $[0, 1]^3$ протягом усього інтервалу T_{mig} , повністю нівелюючи мережеві викиди (зниження з 43–84 до 0).

3. Ефективність критерію дисбалансу: Застосування евклідової норми вектора дисбалансу

ВВР як оптимізаційного критерію в класі задач балансування навантаження забезпечує гармонійну колінеарність векторів виділення та наявності ресурсів.

Це забезпечило скорочення дисбалансу на 62.5% та зниження кількості порушень нормативів SLA до рекордних 1.34%.

4. Запропонована методика формує надійний алгоритмічний базис для метаевристичних методів планування (зокрема генетичних алгоритмів та імітації відпалу) у хмарних і туманних середовищах, скорочуючи загальну кількість вимушених міграцій на 32.8% та заощаджуючи до 14.8% електроенергії при класичному методі, в 1,1–1,6 разів.

При великому навантаженні на туманний шар, більше 80% від максимально можливого, показники, що розглядалися, майже не відрізнялися.

Конфлікт інтересів

Автори декларують, що не мають конфлікту інтересів стосовно даного дослідження, в тому числі фінансового, особистісного характеру, авторства чи іншого характеру, що міг би вплинути на дослідження та його результати, представлені в даній статті.

Фінансування

Дослідження проводилося без фінансової підтримки.

Доступність даних

Рукопис не має пов'язаних даних.

Використання засобів штучного інтелекту

Автори підтверджують, що не використовували технології штучного інтелекту при створенні представленої роботи.

СПИСОК ЛІТЕРАТУРИ

1. Ogura, N., Duolikun, D., Enokido, T., Watanabe, R. and Takizawa, M. (2019), "A Virtual Machine Migration for Storage Processes", *Lecture Notes on Data Engineering and Communications Technologies*, vol 25. Springer, Cham, doi: https://doi.org/10.1007/978-3-030-02613-4_67
2. Sawalkar, V. and More, N. (2026), "Virtual Machine Migration Optimization in Cloud Data Centers: A Comprehensive Review", *Lecture Notes in Networks and Systems*, 1647 LNNS, pp. 146–161, doi: https://doi.org/10.1007/978-3-032-06668-8_15
3. Liu, C., Ma, L., Zhang, M. and Long, H. (2025), "Optimizing cloud resource management with an IoT-enabled optimized virtual machine migration scheme for improved efficiency", *Journal of Network and Computer Applications*, 237, 104137, doi: <https://doi.org/10.1016/j.jnca.2025.104137>
4. Kuchuk, G., Nechausov, S. and Kharchenko, V. (2015), "Two-stage optimization of resource allocation for hybrid cloud data store", *International Conference on Information and Digital Technologies*, Zilina, pp. 266–271, DOI: <http://dx.doi.org/10.1109/DT.2015.7222982>
5. Hines, M.R., Deshpande, U., and Gopalan, K. (2009), "Post-copy live migration of virtual machines", *Operating Systems Review ACM*, vol. 43(3), doi: <https://doi.org/10.1145/1618525.1618528>
6. Jalaee, N., and Safi-Esfahani, F. (2021), "VCSP: virtual CPU scheduling for post-copy live migration of virtual machines", *Int. j. inf. tecnol.*, vol. 13, pp. 239–250, doi: <https://doi.org/10.1007/s41870-020-00483-z>
7. Akoush, S., Sohan, R., Rice, A., Moore, A. W. and Hopper, A. (2010), "Predicting the Performance of Virtual Machine Migration", *2010 IEEE International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems*, Miami Beach, FL, USA, pp. 37–46, doi: <https://doi.org/10.1109/MASCOTS.2010.13>
8. Archana and Kumar, N. (2025), "A Modified Bat Mechanism for Virtual Machine Migration in a Cloud Environment", *SN Computer Science*, vol. 6, article number 74, doi: <https://doi.org/10.1007/s42979-024-03627-1>
9. Shi, Q. and Zhao, F. (2024), "Research on Computer Cloud Intelligent System Based on Intelligent Virtualization Technology", *2024 IEEE 3rd Int. Conf. on Eebda 2024*, pp. 1245–1250, doi: <https://doi.org/10.1109/EEBDA60612.2024.10485750>
10. Ferreto, T. C., Netto, M. A. S., Calheiros, R. N., and De Rose, C. A. F. (2011), "Server consolidation with migration control for virtualized data centers", *Future Generation Computer Systems*, vol. 27, issue 8, pp. 1027–1034, doi: <https://doi.org/10.1016/j.future.2011.04.016>

11. Shankar, S. and Anbarasan, M. (2025), "An Intelligent Approach for Cloud Infrastructure With Improved Multi-Objective Graywolf Optimization and Resource Allocation for Dynamic Virtual Machine Placement", *Transactions on Emerging Telecommunications Technologies*, vol. 36(6), e70172, doi: <https://doi.org/10.1002/ett.70172>
12. Ma, D., Cao, X., Hu, J., Xia, T., Zhou, Y., Liu, K., Zhu, L., Su, L. and Gao, F. (2026), "Topology-aware virtual machine placement for improving cloud servers resource utilization", *Future Generation Computer Systems*, vol. 179, article number 108361, doi: <https://doi.org/10.1016/j.future.2025.108361>
13. Zhu, X., Griffith, R., Shah, P., Smirni, E. (2014), "Application-driven dynamic vertical scaling of virtual machines in resource pools", *2014 IEEE Network Operations and Management Symp.*, doi: <https://doi.org/10.1109/NOMS.2014.6838238>
14. Wood, T., Shenoy, P., Venkataramani, A. and Yousif, M. (2009), "Sandpiper: Black-box and gray-box resource management for virtual machines", *Computer Networks*, vol. 53(17), pp. 2923–2938, doi: <https://doi.org/10.1016/j.comnet.2009.04>
15. Meng, X., Pappas, V. and Zhang, L. (2010), "Improving the scalability of data center networks with traffic-aware virtual machine placement", *2010 Proc. IEEE INFOCOM*, doi: <https://doi.org/10.1109/INFOCOM.2010.5461930>
16. Petrovska, I., Kuchuk, H. and Mozhaiev, M. (2022), "Features of the distribution of computing resources in cloud systems", *2022 IEEE 4th KhPI Week on Advanced Technology*, KhPI Week 2022 – Conference Proceedings, 03–07 October 2022, Code 183771, doi: <https://doi.org/10.1109/KhPIWeek57572.2022.9916459>
17. Thorpe, J., Swiler, L.P., Hanson, S., Cruz, G., Tarman, T., Rollins, T. and Debusschere, B.J. (2022), "Verification of Cyber Emulation Experiments Through Virtual Machine and Host Metrics", *ACM International Conference Proc. Series*, pp. 71–80, doi: <https://doi.org/10.1145/3546096.3546115>
18. Kortas, N. and Youssef, H. (2025), "A Bayesian neural network study for virtual machine migration within cloud environment", *Journal of Supercomputing*, vol. 81(16), 1505, doi: <https://doi.org/10.1007/s11227-025-07974-5>

Received (Надійшла) 29.06.2026

Accepted for publication (Прийнята до друку) 02.09.2026

Publication date (Дата публікації) 25.09.2026

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Кучук Ніна Георгіївна – доктор технічних наук, професор, професорка кафедри комп'ютерної інженерії та програмування, Національний технічний університет "Харківський політехнічний інститут", Харків, Україна;
Nina Kuchuk – Doctor of Technical Sciences, Professor, Professor of Computer Engineering and Programming Department, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;
 e-mail: nina_kuchuk@ukr.net; ORCID Author ID: <http://orcid.org/0000-0002-0784-1465>;
 Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57196006131>.

Пироженко Сергій Станіславович – аспірант кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Serhii Pyrozhenko – PhD candidate of Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;
 e-mail: serhii.pyrozhenko@nure.ua; ORCID Author ID: <http://orcid.org/0009-0006-4209-2144>;
 Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=59675302400>.

**Methodology for calculating the overhead of live migration of virtual machines
 based on the NRC resource cube concept**

Nina Kuchuk, Serhii Pyrozhenko

Abstract. Relevance. Current requirements for the energy-efficient and adaptive operation of modern data centers are driven by the rapid expansion of the cloud computing segment. Under these conditions, live virtual machine (VM) migration tools have become a fundamental component of flexible hardware resource allocation. The ability to relocate virtual machines while they are running, without critical interruptions in application availability, enables administrators to address power optimization issues through workload consolidation, eliminate local host overloads, and perform infrastructure maintenance without interrupting business processes. On the other hand, continuous migration inevitably generates resource overhead. The operation of iterative schemes (Pre-Copy) requires continuous tracking of modified memory blocks, their repeated transmission, and final synchronization of CPU contexts, which significantly distorts the resource profile of servers. As a result, sharp increases in hypervisor utilization of computational cores occur, data transmission channel capacity is exhausted, and response time increases. Chaotic or inefficient initiation of such procedures creates risks of mutual interference among neighboring tasks, violations of SLA requirements, and even the risk of transmission loops, when memory pages are rewritten faster than they can be transmitted over the network. **Subject of the study:** processes of generating and methods for calculating overhead (additional load on computational cores, memory, and network channels) during live migration of virtual machines, as well as mathematical models and algorithms for their spatial assessment based on the vector concept of the NRC resource cube. **Objective of the study:** to improve the efficiency of dynamic resource management in a cloud data center and prevent SLA violations by developing a model and method for assessing the overhead of live virtual machine migration based on the vector space of the NRC resource cube. **Research tasks.** To formulate a load-minimization criterion that takes into account server capacity within the NRC model and to implement an algorithm for selecting target hosts in the Load Balancing class. **Research methods.** Methods of vector and multidimensional geometric analysis, methods of systems analysis and systems theory, graph theory, and tree-based data structures. **Results of the study.** It has been demonstrated that the effective capacity of hosts decreases by an amount corresponding to the vector migration overhead, which prevents false system overload during load-balancing operations. The proposed framework provides a mathematical basis for developing adaptive VM relocation algorithms with guaranteed compliance with the boundary conditions of the NRC cube.

Keywords: virtual machine migration; data transmission; simulated annealing algorithm; genetic algorithm; Internet of Things ecosystem.

І. І. Самборський¹, О. В. Залужний², Н. Т. Лозова³

¹ Інститут спеціального зв'язку та захисту інформації Національного технічного університету «Київський політехнічний інститут» імені Ігоря Сікорського, Київ, Україна

² Військовий інститут телекомунікацій і інформатизації імені героїв Крут, Київ, Україна

³ Науково-дослідний інститут воєнної розвідки, Київ, Україна

МЕТОД ЛОГІКО-ДИНАМІЧНОГО УПРАВЛІННЯ ФУНКЦІОНАЛЬНОЮ СТІЙКІСТЮ ЕЛЕКТРОННОЇ КОМУНІКАЦІЙНОЇ МЕРЕЖІ НА ОСНОВІ ЦИФРОВОГО ДВІЙНИКА

Анотація. **Актуальність.** Забезпечення функціональної стійкості сучасних електронних комунікаційних мереж в умовах зростання інтенсивності кіберзагроз, структурної динаміки та збільшення кількості інцидентів інформаційної безпеки потребує розроблення нових методів адаптивного управління, здатних оперативно реагувати на зміну безпекового стану мережі. **Предмет дослідження:** метод логіко-динамічного управління функціональною стійкістю електронної комунікаційної мережі на основі цифрового двійника. **Мета дослідження:** підвищення функціональної стійкості електронно-комунікаційної мережі в умовах виникнення інцидентів інформаційної безпеки та забезпечення адаптивного управління її безпековим станом. **Завдання дослідження:** здійснити формалізацію простору можливих станів електронної комунікаційної мережі та подій інформаційної безпеки, синтезувати метод логіко-динамічного управління на основі застосування технологій цифрового двійника, визначити механізми його інтеграції із SIEM-системами та підтвердити ефективність методу експериментальними дослідженнями. **Методи дослідження:** логіко-динамічне моделювання, математичне моделювання, системний аналіз, цифрове двійникове моделювання та експериментальні дослідження. **Результати дослідження.** У статті розроблено математичну модель управління, алгоритм адаптивного реагування на інциденти інформаційної безпеки та архітектуру цифрового двійника, що забезпечує прогнозування розвитку подій, оцінювання можливих сценаріїв і формування керуючих дій у реальному масштабі часу. Запропонований в статті метод базується на формалізації множини станів мережі, подій інформаційної безпеки, правил переходів між станами та синтезі керуючих впливів із використанням логіко-динамічної моделі й цифрового двійника, інтегрованого до контуру систем управління інформацією та подіями безпеки (SIEM). **Висновки.** Експериментальна перевірка запропонованого методу підтвердила підвищення ефективності управління функціональною стійкістю електронної комунікаційної мережі за рахунок скорочення середнього часу виявлення інцидентів, зменшення часу відновлення після них та підвищення ефективності локалізації деструктивних впливів порівняно з традиційними підходами. Практичне значення роботи полягає у можливості поетапного впровадження запропонованого методу у корпоративних та операторських електронних комунікаційних мережах без суттєвої модифікації існуючої інфраструктури інформаційної безпеки.

Ключові слова: електронна комунікаційна мережа; функціональна стійкість; метод логіко-динамічного управління; цифровий двійник; логіко-динамічна модель; інциденти інформаційної безпеки; SIEM; адаптивне управління; управління подіями безпеки.

1. Вступ

1.1 Актуальність тематики

Сучасні електронні комунікаційні мережі (ЕКМ) належать до класу складних розподілених організаційно-технічних систем, функціонування яких супроводжується безперервною зміною топології, режимів роботи, інформаційних потоків даних і безпекового стану. Зростання кількості кіберзагроз, інтенсивності інформаційних впливів, складності мережевих сервісів та вимог до безперервності їх функціонування істотно підвищує вимоги до забезпечення функціональної стійкості таких мереж.

Існуючі засоби управління ЕКМ переважно орієнтовані на моніторинг, локалізацію інцидентів, резервування ресурсів, балансування навантаження або підтримання показників якості обслуговування (QoS/QoE).

Разом із тим більшість сучасних SIEM-, SOAR- та SOC-рішень реалізують переважно реактивний принцип оброблення подій безпеки, використовуючи статичні правила кореляції, що

недостатньо враховують причинно-наслідкові зв'язки між подіями, структурну динаміку мережі та зміну її функціонального стану в умовах розвитку інцидентів інформаційної безпеки. Одним із найбільш перспективних напрямів розвитку систем управління є використання технології цифрового двійника (ЦД), яка забезпечує створення динамічної цифрової моделі мережі, прогнозування розвитку подій та оцінювання наслідків керуючих впливів без ризику для реальної інфраструктури. Проте більшість існуючих підходів застосовують його переважно як засіб моніторингу або моделювання, тоді як питання використання цього віртуального засобу як активного елемента контуру управління функціональною стійкістю мережі залишаються недостатньо дослідженими.

Наукова проблема полягає у відсутності формалізованого методу, який поєднував би логіко-динамічне моделювання станів електронної комунікаційної мережі, аналіз подій інформаційної безпеки та можливості цифрового двійника для синтезу адаптивних керуючих впливів у реальному масштабі часу.

З огляду на це актуальним є розроблення методу логіко-динамічного управління функціональною стійкістю електронної комунікаційної мережі на основі цифрового двійника, який забезпечує формалізацію простору станів мережі, правил переходів між ними, прогнозування розвитку інцидентів інформаційної безпеки та синтез керуючих дій, спрямованих на підтримання функціональної стійкості мережі в умовах динамічних деструктивних впливів.

1.2 Аналіз літературних джерел і визначення проблеми

Останніми роками ЦД розглядається як один із найбільш перспективних напрямів розвитку систем управління складними електронними комунікаційними та кіберфізичними системами. Основними напрямками сучасних досліджень є побудова архітектур цифрових двійників, інтеграція їх із системами кібербезпеки, забезпечення функціональної стійкості мереж та автоматизація процесів прийняття рішень.

Одним із перших напрямів розвитку є застосування цифрового двійника для управління подіями інформаційної безпеки електронних комунікаційних мереж. У роботах [1, 2] запропоновано концептуальні підходи до інтеграції ЦД із системами управління інформацією та подіями безпеки (SIEM), формалізовано окремі елементи логіко-динамічного опису мережевих станів і визначено загальні принципи побудови цифрових моделей мобільних комунікаційних мереж.

Подальший розвиток цього напрямку пов'язаний із дослідженнями архітектури ЦД електронних мереж. Так, у роботах [3, 4] запропоновано сучасні архітектурні рішення побудови мереж ЦД, визначено принципи взаємодії цифрових моделей із телекомунікаційною інфраструктурою та окреслено перспективи їх використання у мережах нового покоління. Водночас зазначені дослідження зосереджені переважно на питаннях архітектурної організації ЦД і практично не розглядають задачі синтезу адаптивних керуючих впливів.

Окремий напрям становлять дослідження застосування цифрових двійників у сфері кібербезпеки та забезпечення кіберстійкості. У роботах [5–7] показано ефективність використання цифрових моделей для аналізу кіберзагроз, оцінювання вразливостей, моделювання сценаріїв атак та підтримки процесів реагування на інциденти інформаційної безпеки. Разом із тим більшість запропонованих рішень орієнтована на моделювання або оцінювання стану систем і не передбачає побудови замкненого контуру адаптивного управління функціональною стійкістю мереж.

Важливе значення для розвитку методів управління складними динамічними системами мають дослідження у галузі логіко-динамічного моделювання та гібридних динамічних систем. У роботах [8, 9] запропоновано математичні засади

опису систем, поведінка яких визначається поєднанням дискретних подій та безперервних процесів. Саме такі моделі є найбільш придатними для формалізації процесів зміни станів електронних комунікаційних мереж під впливом інцидентів інформаційної безпеки.

Фундаментальні принципи побудови цифрових двійників, їх переваги, обмеження та проблеми розвитку систематизовано у науковій праці [10], яка фактично стала базовою для подальших досліджень цього наукового напрямку. Подальший розвиток концепції цифрових двійників для підтримки процесів реагування на кіберінциденти представлено у роботі [11], де запропоновано інтегрувати цифрові моделі із системами реагування на інциденти. Разом із тим питання формалізації процесів управління станами мережі та синтезу керуючих впливів залишаються відкритими.

Перспективність ризик-орієнтованого підходу до управління складними організаційно-технічними системами обґрунтовано у роботі [12], де показано доцільність врахування рівня ризику під час формування управлінських рішень. Практичні аспекти синтезу моделей управління складними динамічними об'єктами з урахуванням подій безпеки досліджено у роботі [13], у якій сформовано окремі принципи побудови логіко-динамічних моделей для систем управління безпекою.

Отже, аналіз сучасних досліджень свідчить, що сьогодні достатньо ґрунтовно опрацьовано питання архітектури цифрових двійників, моделювання кіберзагроз, інтеграції цифрових моделей із SIEM-системами та використання логіко-динамічного підходу для опису складних динамічних процесів. Водночас практично відсутні комплексні методи, які поєднували б логіко-динамічне моделювання станів ЕКМ, цифровий двійник як активний елемент контуру управління та механізми синтезу адаптивних керуючих впливів, спрямованих на підтримання її функціональної стійкості в умовах розвитку інцидентів інформаційної безпеки. Саме усунення зазначеного дослідницького розриву визначає актуальність і наукову спрямованість даної роботи.

1.3 Мета і завдання дослідження

Метою роботи є розробка методу логіко-динамічного управління ЕКМ на основі використання ЦД для підвищення її функціональної стійкості в умовах виникнення інцидентів інформаційної безпеки та забезпечення адаптивного управління її безпековим станом.

Для досягнення поставленої мети доцільно провести формалізацію множини станів та подій корпоративної ЕКМ, розробити правила переходів та причинно-наслідкових залежностей, побудувати алгоритм логіко-динамічного управління, сформулювати правила інтеграції ЛДМ з існуючими SIEM-системами ЕКМ та провести експериментальні сценарії для оцінки ефективності розробленого методу. Поєднання логіко-динамічного підходу з концепцією ЦД та подієвим

конвеєром дозволить не лише аналізувати інциденти інформаційної безпеки, а й здійснювати проактивне управління (зміна правил, що визначають політику доступу до ресурсів мережі, сегментація, ізоляція вузлів). Це забезпечить можливість адаптивного управління функціональною стійкістю мережі в умовах постійної варіації динамічних загроз цій цифровій структурі. Практичне значення дослідження передбачає сумісність запропонованого методу з типовим стеком корпоративних систем безпеки (SIEM, брокери повідомлень, графові бази даних, playbooks) та можливості його поетапного впровадження без радикальної перебудови існуючої структури систем управління інформацією і подіями безпеки корпоративної ЕКМ.

2. Матеріали і методи дослідження

Відмітимо, що ЕКМ є складною організаційно-технічною динамічною системою з багатовимірним простором станів, структурною мінливістю та подієвою природою деградаційних процесів. В умовах існуючих наразі кіберзагроз, відмов елементів, перевантажень і деструктивних інформаційних впливів критичним показником її якості є функціональна стійкість, тобто здатність зберігати допустимі параметри функціонування або відновлювати їх за обмежений час після дії дестабілізуючих факторів.

Існуючі методи управління мережами переважно орієнтовані на оптимізацію окремих параметрів (маршрутизації, пропускної здатності, балансування навантаження), проте не забезпечують формалізованого опису причинно-наслідкових зв'язків між подіями безпеки, переходами станів та керуючими впливами в єдиному логіко-динамічному просторі. Поряд з цим корпоративна ЕКМ розглядається як багаторівнева система, що включає рівні доступу, транспорту, сервісів та безпеки. На кожному рівні функціонують вузли, канали, ресурси та сервіси, які забезпечують цілісність інформаційних потоків та підтримку бізнес-процесів.

Така багаторівнева структура зумовлює складність організації та реалізації управління, оскільки відмова або атака на окремих її компонент може спричинити каскадні ефекти, що впливають на всю мережу. У цьому контексті ЕКМ виступає об'єктом управління, який потребує формалізованого опису станів та подій для забезпечення функціональної стійкості.

3. Результати розробки і дослідження методу логіко-динамічного управління функціональною стійкістю ЕКМ на основі ЦД

3. 1 Математична модель, що описує стани ЕКМ, події безпеки та їх кореляцію з функціональною стійкістю мережі

Представимо математичну модель, яка формально описує стани ЕКМ, події безпеки та їх кореляцію з функціональною стійкістю мережі.

Задамо:

$S = \{s_1, s_2, \dots, s_n\}$ – множина можливих станів

ЕКМ;

$E = \{e_1, e_2, \dots, e_m\}$ – множина подій безпеки;

$U = \{u_1, u_2, \dots, u_k\}$ – множина управляючих

впливів.

Тоді оператор переходу станів визначимо як

$$F : S \times E \times U \rightarrow S, \quad (1)$$

де $\Omega \subset S$ – область функціональної стійкості.

Умову функціональної стійкості, яка полягає в тому, що ЕКМ – система повертається до області Ω за скінченний допустимий час $T \leq T_{dm}$ формалізуємо у вигляді:

$$\forall s(t_0) \in \Omega, \forall e \in E, \exists u \in U : F(s(t), e, u) \in \Omega \quad (2)$$

Управління подіями безпеки ЕКМ визначимо у вигляді функції

$$u = \Phi(S, E). \quad (3)$$

Вплив управління безпекою ЕКМ полягає в мінімізації функціонала відхилення від області стійкості:

$$J = \int_{t_0}^{t_0 + T} d(s(t), \Omega) dt \rightarrow \min. \quad (4)$$

Модель ЦД опишемо залежністю виду

$$\hat{s}(t + \Delta t) = F(\hat{s}(t), e(t), u(t)), \quad (5)$$

де $\hat{s}(t)$ – оцінений стан мережі ЦД.

Ґрунтуючись на (1)–(5) відмітимо, що використання ЦД у поєднанні з логіко-динамічним оператором $\Phi(S, E)$ забезпечує підтримання функціональної стійкості мережі шляхом прогнозування переходів між її можливими станами та синтезу управлінських впливів, що мінімізують відхилення від області Ω . Функціональна стійкість визначається як здатність мережі зберігати цільові функції та рівень сервісу при виникненні подій, атак чи відмов.

Це досягається завдяки керованій перебудові конфігурацій та адаптивному управлінню ресурсами.

Показник функціональної стійкості сучасних комунікаційних мереж має багатовимірний характер. Він включає технічні, організаційні та процедурні аспекти, а також здатність до швидкого відновлення після впливу інцидентів інформаційної безпеки.

Тому функціональну стійкість у даній роботі будемо трактувати як інтегральну властивість системи, що забезпечує безперервність сервісів навіть за умов інтенсивних дестабілізуючих впливів. Для кількісної оцінки функціональної стійкості використаємо набір метрик, що охоплюють як якість сервісу ЕКМ, так і її керованість при ліквідації деструктивних впливів безпекових інцидентів, а саме:

- *якість сервісу*: середній час відгуку (RTT), пропускна здатність (BPS/PPS), доступність критичних сервісів.

- *керованість*: середній час виявлення інциденту (MTTD), середній час відновлення (MTTR), частота хибних спрацювань систем моніторингу.

- *інтегральний індекс ризику*: показник, який враховує ймовірність виникнення подій та їхній потенційний вплив на функціонування ЕКМ.

Відмітимо, що застосування запропонованих метрик дозволяє не лише оцінити поточний стан мережі, а й порівняти ефективність різних методів управління її функціональною стійкістю. Зокрема, у багаторівневих мережах, якими є корпоративні ЕКМ, інтегральні показники дають змогу не лише моделювати каскадні ефекти в мережі, але прогнозувати наслідки управління подіями безпеки.

Вказане свідчить, що формалізація ЕКМ як об'єкта управління та критеріїв його функціональної стійкості створюють основу для побудови ЛДМ ЦД, яка забезпечує адаптивне управління корпоративною ЕКМ у динамічному середовищі безпекових загроз. ЦД корпоративної ЕКМ буде розглядати в контексті реалізації керованої система станів, що відображає як поточний функціональний стан мережі так і її реакцію на можливі зовнішні й внутрішні події.

До множини станів ЕКМ пропонується включити:

- *Normal* – штатне функціонування без інцидентів;

- *Degraded* – часткова деградація сервісів мережі;

- *UnderAttack* – активна атака на мережу або аномальний трафік;

- *Isolated* – сегментація чи ізоляція вузлів для локалізації інциденту;

- *Recovery* – процес відновлення після інциденту.

Для прикладу відмітимо, що, подія «аномальний трафік» у стані *Normal* може призвести до переходу у стан *UnderAttack* з виконанням дії «активація *rate-limit*». Для зменшення кількості хибних спрацювань застосовується механізм підтвердження подій (кворум або часові вікна), що дозволяє враховувати контекст та кореляцію сигналів

При цьому будемо враховувати деструктивні впливи таких подій:

- *мережевих* (зміни трафіку, відмови каналів);

- *безпекових* (спроби несанкціонованого доступу, DoS-атаки, виявлені вразливості);

- *керуючих* (зміна ACL, сегментація, ізоляція вузла, застосування *rate-limit*- обмеження частоти запитів).

Для кожного стану визначається набір правил переходів, що описують реакцію системи на події. Формалізація здійснюється у вигляді (табл. 1).

Таблиця 1 – Елементарні функції операцій, керованих інформацією

Поточний стан (state)	Подія (event)	Контекст (context)	Новий стан (new_state)	Дія (action)
Idle	Отримано подію безпеки	Система в нормальному режимі	Monitoring	Логування події, сповіщення оператора
Monitoring	Виявлено аномалію	Критичність середня	Alert	Надсилання попередження, збір додаткових даних
Monitoring	Виявлено аномалію	Критичність висока	Critical	Блокування доступу, сповіщення адміністратора, активація протоколу реагування
Alert	Подія підтверджена	Система у стані Alert	Mitigation	Запуск процедур пом'якшення, резервне копіювання, активізація безпеки
Alert	Подія помилкова	Система у стані Alert	Monitoring	Логування інциденту як хибного спрацювання
Critical	Загроза ліквідована	Система у стані Critical	Monitoring	Відновлення нормального режиму, генерація звіту
Mitigation	Пом'якшення успішне	Система у стані Mitigation	Monitoring	Повернення до моніторингу, оновлення аналітики загроз
Mitigation	Пом'якшення неуспішне	Система у стані Mitigation	Critical	Ескалація, повідомлення, активація аварійного протоколу

Для реалізації правил переходу (табл.1) розроблена архітектура ЦД як керуючої надбудови ЕКМ. Вона реалізується як багатокomпонентна

система, що включає: модуль збору подій; брокер повідомлень для маршрутизації потоків; аналітичне ядро з логіко-динамічною моделлю; графову базу

даних для збереження станів та переходів; інтеграцію з SIEM для кореляції подій; оркестратор для виконання керуючих дій. Такий підхід

забезпечує сумісність із типовим корпоративним стеком ЕКМ без необхідності радикальної перебудови інфраструктури (рис. 1).



Рис. 1. Архітектура ЦД як керуючої надбудови ЕКМ

У цій керуючій надбудові політику доступу (ACL – Access Control Lists) розглядатимемо як ключові управлінські дії та одночасно як об'єкти контролю за станами ЕКМ. Управлінські дії визначають конфігураційний стан мережі та можуть бути змінені у відповідь на події безпеки. ACL застосовуються на різних рівнях, а саме:

edge/CE-router - контроль доступу до периферійних пристроїв;

фаєрвол - фільтрація трафіку;

SG/NSG у віртуалізованих середовищах - сегментація у хмарних інфраструктурах корпоративної ЕКМ;

мікросегментація - ізоляція на рівні окремих сервісів чи контейнерів.

Інтеграція ACL у логіко-динамічний контур дозволяє реалізувати адаптивне управління, коли ЦД не лише відображає стан системи, а й активно змінює її конфігурацію для збереження заданого рівня функціональної стійкості.

Основним завданням адаптивного управління є мінімізація часу деградації та інциденту, а також обмеження каскадних ефектів у корпоративній ЕКМ. Ціль полягає у збереженні критичних сервісів навіть за умов атак чи відмов. Для цього необхідно забезпечити адаптивне реагування на події, що включає не лише їхнє виявлення, а й активні управляючі впливи на ЕКМ, а саме: зміну конфігурацій мережі, сегментацію або ізоляцію мережевих вузлів. Метод логіко-динамічного управління пропонується реалізувати у вигляді алгоритму, що поєднує обробку подій, оцінку контексту та виконання керуючих дій. Алгоритм представимо у вигляді псевдокоду:

1. *Отримати подію з потоку (мережеву, безпекову або керуючу).*

2. *Виконати нормалізацію події (уніфікація формату, усунення дублювань).*

3. *Оцінити контекст:*

- *топологія мережі;*
- *поточний стан цифрового двійника;*
- *кореляція з іншими подіями.*

4. *Застосувати правила переходів: (state, event, context) → new_state + action.*

5. *Оновити граф станів цифрового двійника.*

6. *Виконати керуючу дію (playbook):*

- *зміна ACL;*
- *сегментація або ізоляція вузла;*
- *активація rate-limit;*
- *інші конфігураційні зміни.*

7. *Записати результат у журнал подій та передати у SIEM – систему для кореляції з ЦД.*

Розроблений алгоритм забезпечує циклічний процес управління, де кожна нова подія оновлює модель станів і запускає відповідні реакції.

Запропонований метод інтегруємо з існуючими системами безпеки ЕКМ.

Врахуємо, що SIEM – система реалізує функції збору та кореляції подій, тоді як ЦД забезпечує логіко-динамічне управління станами мережі та прогнозування наслідків безпекових інцидентів деструктивних впливів. Це створює зворотний зв'язок: система не лише детектує інциденти, а й активно змінює конфігурацію мережі для локалізації загроз.

Такий підхід дозволяє перейти від реактивного моніторингу до проактивного управління функціональною стійкістю.

При цьому реалізація методу логіко-динамічного управління забезпечує адаптивне реагування на інциденти у корпоративних мережах, поєднуючи аналіз подій, прогнозування станів та виконання керуючих дій у реальному часі.

Запропонований метод логіко-динамічного управління функціональною стійкістю може бути реалізований у вигляді модульної системи, що працює у середовищі *Linux* та підтримує контейнеризацію (*Docker/Kubernetes*).

Мінімальні вимоги включають:

- серверні ресурси для обробки подій,
- брокер повідомлень (наприклад, *Apache Kafka* або *RabbitMQ*),
- графову базу даних (*Neo4j*) для моделювання станів та інтеграцію з SIEM-системами.

Така архітектура дозволяє масштабувати рішення відповідно до навантаження та забезпечує сумісність із типовим корпоративним стеком ЕКМ [1, 2, 13].

3. 2 Етапи впровадження методу логіко-динамічного управління функціональною стійкістю ЕКМ та реалізація ЦД

Впровадження методу пропонується здійснювати поетапно, а саме:

Етап 1: ЦД функціонує як інструмент моніторингу, що відображає стан мережі та корелює події без виконання керуючих дій.

Етап 2: Система формує рекомендації щодо керуючих дій (зміна ACL, сегментація), які можуть бути підтверджені оператором.

Етап 3: Цифровий двійник здійснює автоматичне виконання керуючих дій у реальному часі, забезпечуючи адаптивне управління стійкістю.

Запропонований поетапний підхід втілення методу суттєво мінімізує ризики деструктивних впливів інцидентів інформаційної безпеки та дозволяє поступово інтегрувати метод у критичну інформаційну інфраструктуру – корпоративні ЕКМ.

У даному методі передбачено, що окрім функцій управління, ЦД може використовуватися як віртуальна лабораторія для тестування сценаріїв інцидентів в ЕКМ, навчання обслуговуючого персоналу та валідації їх підготовки. Це дозволяє безпечно перевіряти нові політики доступу та алгоритми реагування на інциденти без впливу на реальну мережу цифрового зв'язку. Такий підхід особливо важливий для мобільних стільникових операторських мереж, де навіть короткочасні відмови можуть призвести до значних фінансових втрат.

Реалізація ЦД ґрунтується на розробленій методиці та будується по принципу модельно-модульної топології.

ЦД складається з таких складових:

1. *Графова модель топології* - відображає вузли, канали та їхні стани.

2. *Модуль прийому подій* - забезпечує збір та нормалізацію даних із SIEM-систем ЕКМ.

3. *Модуль логіко-динамічних правил* - реалізує таблицю переходів та причинно-наслідкові залежності.

4. *Модуль реакцій* — реалізує відповідні управляючі впливи дії.

Система може запускатися модульно, що дозволяє інтегрувати окремі компоненти поступово. Наприклад, початкове впровадження може обмежуватися лише графовою моделлю та модулем прийому подій, а згодом доповнюватися керуючими механізмами.

Відмітимо, що реалізація та інтеграція методу у корпоративну ЕКМ забезпечить поетапне впровадження, сумісність із існуючими системами безпеки та можливість використання ЦД як ефективного інструменту управління подіями безпеки.

Для перевірки ефективності запропонованого методу було проведено серію експериментальних сценаріїв у тестовому середовищі корпоративної мережі.

З врахуванням рекомендацій [1, 2, 13] було змодельовано три типові ситуації:

Нормальний режим. Мережа функціонує без інцидентів, усі сервіси доступні, події безпеки відсутні. Цей сценарій використовується як базовий для порівняння.

DoS/аномальний трафік. На мережу спрямовується високий потік запитів, що призводить до перевантаження каналів та деградації сервісів.

Комбінований сценарій. Поєднання атаки на рівні безпеки (несанкціонований доступ) із деградацією сервісу (зниження пропускної здатності), що створює складні причинно-наслідкові зв'язки.

3. 3 Оцінка ефективності методу логіко-динамічного управління функціональною стійкістю ЕКМ

Для оцінки ефективності методу застосовувалися такі метрики [13]:

- *MTTD (Mean Time to Detect)* – середній час виявлення інциденту.

- *MTTR (Mean Time to Recovery)* – середній час відновлення після інциденту інформаційної безпеки.

- *RTT (Round Trip Time)* – середній час відгуку сервісів.

- *PPS/BPS* – пропускна здатність у пакетах та бітах за секунду.

- *Частка локалізованих інцидентів* - відсоток випадків, коли інцидент був ізольований без впливу на всю ЕКМ.

Результати показали, що застосування логіко-динамічного методу з ЦД дозволяє скоротити MTTD у середньому на 30%, а MTTR — на 25% порівняно з традиційними підходами.

Крім того, кількість хибних спрацювань системи моніторингу зменшилася завдяки використанню механізму підтвердження подій.

Практичний ефект від реалізації розробленого методу для оператора корпоративної мережі полягає у наступному:

- швидше виявлення інцидентів та скорочення часу реагування на інциденти;

- можливість контрольованої ізоляції сегмента без повного падіння сервісів;

- збереження критичних функцій мережі навіть у випадку комбінованих атак-джерел подій безпеки;

- використання динамічного ЦД як інструменту для тестування сценаріїв атак та навчання обслуговуючого корпоративну ЕКМ персоналу.

Таким чином, експериментальні результати підтверджують ефективність запропонованого методу та його практичну доцільність для корпоративних і операторських мереж.

4. Обговорення результатів дослідження методу логіко-динамічного управління функціональною стійкістю ЕКМ на основі ЦД

Запропонований метод логіко-динамічного управління функціональною стійкістю електронної комунікаційної мережі найбільш ефективний у безпечових середовищах із високим потоком подій та складними причинно-наслідковими зв'язками.

У вказаних умовах динамічний ЦД забезпечує швидке виявлення інцидентів, кореляцію сигналів та адаптивне реагування.

Водночас у мережах із низьким навантаженням або відсутністю корельованих загроз ефект від застосування методу може бути менш вираженим, оскільки традиційні механізми моніторингу та резервування наразі ще здатні забезпечити достатній рівень функціональної стійкості.

Особливо відмітимо, що інтеграція динамічного ЦД у корпоративну мережу супроводжується низькою ризиків.

Суть цих ризиків полягає в наступному:

Ресурсні обмеження. Виконання алгоритмів логіко-динамічного управління потребує додаткових обчислювальних ресурсів, що може призвести до перевантаження системи у випадку масштабних інцидентів.

Ризик self-DoS. Некоректна конфігурація або надмірна кількість керуючих дій може спричинити небажану деградацію сервісів електронної комунікаційної мережі.

Сумісність із існуючими системами. Інтеграція з SIEM-системами вимагає ретельного узгодження форматів подій безпеки та політик доступу, що може бути складним у гетерогенних інформаційних інфраструктурах.

Людський фактор. Автоматизація керуючих дій потребує довіри з боку операторів; у випадку недостатнього контролю можливі помилки у прийнятті ними рішень щодо ліквідації наслідків впливів інцидентів інформаційної безпеки.

Сформульовані умови застосовності методу для мінімізації ризиків та забезпечення його ефективності полягають в наступному:

Сегментація середовища. Впровадження цифрового двійника має здійснюватися поетапно, починаючи з окремих сегментів електронної комунікаційної мережі.

Контроль доступу. Використання ACL та політик безпеки повинно бути інтегроване у логіко-динамічний контур для запобігання неконтрольованим змінам.

Ліміти ресурсів. Встановлення обмежень на кількість одночасних керуючих дій дозволяє уникнути перевантаження мережі.

Валідація сценаріїв. Перед впровадженням у продуктивне середовище необхідно тестувати сценарії у «віртуальній лабораторії» ЦД.

Таким чином, запропонований метод має значний потенціал для підвищення функціональної

стійкості корпоративних та операторських мереж, проте його ефективність залежить від умов застосування та правильного управління подіями безпеки.

5. Висновки

У статті вирішено актуальне науково-прикладне завдання розроблення методу логіко-динамічного управління функціональною стійкістю електронної комунікаційної мережі на основі цифрового двійника, який забезпечує адаптивне реагування на інциденти інформаційної безпеки шляхом формалізації простору станів мережі, правил переходів між ними та синтезу керуючих впливів.

Наукова новизна запропонованого методу полягає у поєднанні логіко-динамічного моделювання, цифрового двійника та механізмів адаптивного управління в єдиному контурі підтримання функціональної стійкості електронної комунікаційної мережі. На відміну від існуючих підходів, запропонований метод забезпечує не лише моніторинг і кореляцію подій інформаційної безпеки, а й прогнозування розвитку інцидентів, формування керуючих впливів та підтримання функціональної стійкості мережі в реальному масштабі часу.

У процесі дослідження розроблено математичну модель логіко-динамічного управління, сформовано множину станів електронної комунікаційної мережі, визначено правила переходів між ними, запропоновано алгоритм адаптивного управління та архітектуру цифрового двійника, інтегрованого із SIEM-системами.

Експериментальна перевірка підтвердила ефективність запропонованого методу. Його застосування забезпечує скорочення середнього часу виявлення та ліквідації інцидентів інформаційної безпеки, підвищує ефективність локалізації деструктивних впливів і сприяє підтриманню функціональної стійкості електронної комунікаційної мережі порівняно з традиційними підходами.

Практична цінність роботи полягає у можливості поетапного впровадження запропонованого методу в корпоративних та операторських електронних комунікаційних мережах без істотної модернізації існуючої інфраструктури інформаційної безпеки, а також у використанні цифрового двійника як середовища для моделювання сценаріїв інцидентів, перевірки політик безпеки та підтримки прийняття управлінських рішень.

Подальші дослідження доцільно спрямувати на таке:

- масштабування запропонованого методу для багаторівневих електронних комунікаційних мереж,
- інтеграцію методів машинного навчання для прогнозування розвитку інцидентів інформаційної безпеки,

• автоматичний синтез політик управління та розширення функціональних можливостей цифрового двійника для підтримки автономного управління мережами наступних поколінь.

Конфлікт інтересів

Автори декларують, що не мають конфлікту інтересів стосовно даного дослідження, в тому числі фінансового, особистісного характеру, авторства чи іншого характеру, що міг би вплинути на дослідження та його результати, представлені в даній статті.

Фінансування

Дослідження проводилося без фінансової підтримки.

Доступність даних

Рукопис не має пов'язаних даних.

Використання засобів штучного інтелекту

Автори підтверджують, що не використовували технології штучного інтелекту при створенні представленої роботи.

СПИСОК ЛІТЕРАТУРИ

1. Samborskyi Ie. Management of information and security events of a mobile cellular network using a digital twin. *Information Technology and Security*. 2025. Vol. 13, No. 2. P. 346–356. DOI: [10.20535/2411-1031.2025.13.2.344719](https://doi.org/10.20535/2411-1031.2025.13.2.344719).
2. Samborskyi Ye., Krykhovetskyi H. Synthesis of the digital twin of the logical-dynamic information and events management system for the security of computer systems of the mobile cellular information and communication network. *Control, Navigation and Communication Systems*. 2025. Vol. 4, No. 82. DOI: [10.26906/SUNZ.2025.4.207](https://doi.org/10.26906/SUNZ.2025.4.207).
3. Wieme J., Baert M., Hoebeke J. Architectural design for digital twin networks. *Network*. 2025. Vol. 5, No. 3. Art. 24. DOI: [10.3390/network5030024](https://doi.org/10.3390/network5030024).
4. Afzal M. K., Naeem M., Ejaz W. (Eds.). *Digital Twins for Wireless Networks: Overview, Architecture, and Challenges*. Cham : Springer Nature Switzerland AG, 2025. DOI: [10.1007/978-3-031-73679-7](https://doi.org/10.1007/978-3-031-73679-7).
5. Faleiro R., Pan L., Pokhrel S. et al. Digital Twin for Cybersecurity: Towards Enhancing Cyber Resilience. In: *Broadband Communications, Networks, and Systems (BROADNETS)*. Cham : Springer, 2022. DOI: [10.1007/978-3-030-93479-8_4](https://doi.org/10.1007/978-3-030-93479-8_4).
6. Ren Q., Xiong X., Liu J., Yao Q. Research on digital twins technology in cyberspace security. *Journal of System Simulation*. 2024. Vol. 36, No. 8. P. 1944–1957. DOI: [10.16182/j.issn1004731x.joss.23-0853](https://doi.org/10.16182/j.issn1004731x.joss.23-0853).
7. Jiang Y., Wang W., Ding J. et al. Leveraging digital twin technology for enhanced cybersecurity in cyber–physical production systems. *Future Internet*. 2024. Vol. 16, No. 4. Art. 134. DOI: [10.3390/fi16040134](https://doi.org/10.3390/fi16040134).
8. Platzer A. Hybrid dynamical systems logic and its refinements. *Science of Computer Programming*. 2025. Vol. 239. Art. 103179. DOI: [10.1016/j.scico.2024.103179](https://doi.org/10.1016/j.scico.2024.103179).
9. Zhao H., Xing Y., Zhou C., Wang Y., Duan H., Liu K., Jiang S. Research on Hybrid Logic Dynamic Model and Voltage Predictive Control of Photovoltaic Storage System. *Energies*. 2024. Vol. 17, No. 17. Art. 4285. DOI: [10.3390/en17174285](https://doi.org/10.3390/en17174285).
10. Rasheed A., San O., Kvamsdal T. Digital Twin: Values, Challenges and Enablers from a Modeling Perspective. *IEEE Access*. 2020. Vol. 8. P. 21980–22012. DOI: [10.1109/ACCESS.2020.2970143](https://doi.org/10.1109/ACCESS.2020.2970143).
11. Suhail S., Iqbal M., McLaughlin K., Lee B., Imtiaz B. A framework for enhancing cyber incident response with Security-Enhancing Digital Twins in Cyber–Physical Systems. *Internet of Things*. 2025. Vol. 31. Art. 101547. DOI: [10.1016/j.iot.2025.101547](https://doi.org/10.1016/j.iot.2025.101547).
12. Гобела В. В., Леськів Г. З., Фляк В. М. Моделювання системи ризик-орієнтованого управління підприємством. *Науковий вісник Львівського державного університету внутрішніх справ. Серія економічна*. 2023. № 2. С. 10–20. DOI: [10.32782/2311-844X/2023-2-2](https://doi.org/10.32782/2311-844X/2023-2-2).
13. Samborskyi I., Samborskyi I., Hol V., Peleshok Y., Sholokhov S. Synthesis of the model of management of complex dynamic objects taking into account the events of their security. *Information Technology and Security*. 2024. Vol. 12, No. 1. P. 4–16. DOI: [10.20535/2411-1031.2024.12.1.306254](https://doi.org/10.20535/2411-1031.2024.12.1.306254).

Received (Надійшла) 29.06.2026

Accepted for publication (Прийнята до друку) 26.08.2026

Publication date (Дата публікації) 25.09.2026

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Самборський Іван Іванович – кандидат технічних наук, старший науковий співробітник, професор кафедри Інституту спеціального зв'язку та захисту інформації Національного технічного університету «Київський політехнічний інститут імені Ігоря Сікорського», Київ, Україна;

Ivan Samborskyi – Candidate of Technical Sciences, Senior Researcher, Professor of the Department Institute of Special Communications and Information Protection, Igor Sikorsky National Technical University «Kyiv Polytechnic Institute», Kyiv, Ukraine;

e-mail: i.samborskyi@gmail.com; ORCID Author ID: <https://orcid.org/0000-0001-5579-8740>.

Залужний Олексій Вікторович – кандидат технічних наук, доцент кафедри кібербезпеки Військового інституту телекомунікацій та інформатизації імені Героїв Крут, Київ, Україна;

Oleksii Zaluzhnyi – Candidate of Technical Sciences, Associate Professor of the Department of Cybersecurity, Heroes of Kruty Military Institute of Telecommunications and Informatization, Kyiv, Ukraine;

e-mail: oleksii.zaluzhnyi@viti.edu.ua; ORCID Author ID: <https://orcid.org/0000-0002-8722-4087>.

Лозова Надія Теодорівна – старший науковий співробітник, Науково-дослідний інститут воєнної розвідки, Київ, Україна;

Nadiia Lozova – Senior Researcher Fellow Defence Intelligence Research Institute, Kyiv, Ukraine;

e-mail: lozova245@ukr.net; ORCID Author ID: <https://orcid.org/0009-0005-4335-6810>.

Method of logical-dynamic control of functional stability of an electronic communication network based on a digital twin

Ivan Samborskyi, Oleksii Zaluzhnyi, Nadiia Lozova

Abstract. Relevance. Ensuring the functional stability of modern electronic communication networks in the context of increasing intensity of cyber threats, structural dynamics and an increase in the number of information security incidents requires the development of new methods of adaptive management capable of promptly responding to changes in the security state of the network. **Subject of research:** the method of logical-dynamic control of the functional stability of an electronic communication network based on a digital twin. **The purpose of the study:** to increase the functional stability of the electronic communication network in the context of information security incidents and to ensure adaptive management of its security state. **The objectives of the study:** to formalize the space of possible states of the electronic communication network and information security events, to synthesize the method of logical-dynamic control based on the use of digital twin technologies, to determine the mechanisms of its integration with SIEM systems and to confirm the effectiveness of the method by experimental research. **Research methods:** logical-dynamic modeling, mathematical modeling, system analysis, digital binary modeling and experimental research. **Research results.** The article develops a mathematical model of management, an algorithm for adaptive response to information security incidents and a digital twin architecture, which provides forecasting of the development of events, assessment of possible scenarios and formation of control actions in real time. The method proposed in the article is based on the formalization of a set of network states, information security events, rules of transitions between states and the synthesis of control influences using a logical-dynamic model and a digital twin integrated into the circuit of information and security event management systems (SIEM). **Conclusions.** Experimental verification of the proposed method confirmed an increase in the efficiency of managing the functional stability of the electronic communication network by reducing the average time of incident detection, reducing the recovery time after them and increasing the efficiency of localization of destructive influences compared to traditional approaches. The practical significance of the work lies in the possibility of phased implementation of the proposed method in corporate and operator electronic communication networks without significant modification of the existing information security infrastructure.

Keywords: electronic communication network; functional stability; method of logical-dynamic control; digital twin; logical-dynamic model; information security incidents; SIEM; adaptive control; Security event management.

АЛФАВІТНИЙ ПОКАЖЧИК

Ахундов Р. Г.	139	Захарченко Р. В.	78	Почерняєв В. М.	179
Барковська О. Ю.	57	Іващенко Г. С.	68	Радованович М. Р.	139
Биба В. В.	47	Івко С. О.	179	Резнік Д. В.	37
Бойчун С. О.	5	Капітон А. М.	100	Рудаков І. С.	22
Бондаренко М. Е.	68	Ключко Л. А.	161	Самборський І. І.	195
Бунякіна Н. В.	78	Коваленко А. А.	106	Скорлупін О. В.	123
Войтко О. В.	115	Коваленко О. О.	179	Слободяник О. Ю.	31
Галай В. М.	78	Криховецький Г. Я.	168	Сукач С. В.	37
Гашимов Е. Г.	139	Кучук Г. А.	106	Талібов А. М.	139
Гічко Ю. Г.	168	Кучук Н. Г.	189	Тищенко Д. О.	100
Горносталь О. А.	150	Лашин Я. О.	115	Усенко Д. В.	47
Горовий К. Ю.	11	Лашко Є. Є.	37	Федорієнко В. А.	115
Деркач Т. М.	161	Левицький С. М.	5	Франчук Т. М.	100
Дмитренко А. О.	161	Лозова Н. Т.	195	Храновська Т. В.	168
Дрючко О. Г.	78	Мазор С. Ю.	168	Чала О. О.	11
Дяченко Д. О.	88	Невлюдов І. Ш.	11	Челак В. В.	150
Євсєєв В. В.	11	Петренко І. С.	37	Ченчева О. О.	37
Жученко О. С.	172	Пироженко С. С.	189	Шефер О. В.	78
Залужний О. В.	195	Пінчук Н. М.	47	Шостак А. В.	131
Захаров Р. Г.	100	Подорожняк А. О.	123	Шостак М. В.	88

Наукове видання

СИСТЕМИ УПРАВЛІННЯ, НАВІГАЦІЇ ТА ЗВ'ЯЗКУ

Збірник наукових праць

Випуск 3 (85)

Відповідальний за випуск *О. В. Шефер*

Ідентифікатор медіа R30-04135 згідно з рішенням Національної ради України
з питань телебачення і радіомовлення від 25.04.2024 № 1416

Підписано до друку 21.09.2026. Формат 60×84/8. Ум.-друк. арк. 25,5. Тираж 100 прим. Зам. 921-26
Адреса редакції: Україна, 36011, м. Полтава, проспект Віталія Грицаєнка, 24, тел. (050) 302-20-71
Національний університет «Полтавська політехніка імені Юрія Кондратюка»

Віддруковано з готових оригінал-макетів у цифровій друкарні Impress
61002, м. Харків, вул. Пушкінська, 56, тел. + 38 (057) 714-52-11
e-mail: irina@impress.biz.ua