

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 2 (80)

Issue 2 (80)

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:

kuchuk_nina@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:

kuchuk_nina@ukr.net

Information site:

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

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

*Журнал індексується міжнародними наукометричними базами: Index Copernicus (ICV = **82.05**),
General Impact Factor, Google Scholar, Academic Resource Index, Scientific Indexed Service*

Затверджений до друку Вченою Радою Національного університету

“Полтавська політехніка імені Юрія Кондратюка” (протокол від 19 червня 2025 року № 9).

*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 – наказами МОН України від 17.03.2020 № 409 та від 09.02.2021 № 157

Полтава • 2025

Редакційна колегія

Головний редактор:

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

Заступники головного редактора:

НЕСТЕРЕНКО Катерина Сергіївна
(*д-р техн. наук, проф., Київ, Україна*);
ШЕФЕР Олександр Віталійович
(*д-р техн. наук, проф., Полтава, Україна*).

Члени редакційної колегії:

БЛАУНШТЕЙН Натан Олександрович
(*д-р техн. наук, проф., Ізраїль*);
БОГОМ'Я Володимир Іванович
(*д-р техн. наук, проф., Київ, Україна*);
ВАРБАНЕЦЬ Роман Анатолійович
(*д-р техн. наук, проф., Одеса, Україна*);
ВЕСОЛОВСЬКИЙ Кшиштоф
(*д-р техн. наук, проф., Польща*);
ГАВРИЛКО Євген Володимирович
(*д-р техн. наук, проф., Київ, Україна*);
ГАШИМОВ Ельшан Гіяс огли
(*д-р наук, проф., Баку, Азербайджан*);
ГЛИВА Валентин Анатолійович
(*д-р техн. наук, проф., Київ, Україна*);
ДАКІ Олена Анатоліївна
(*д-р техн. наук, доц., Ізмаїл, Україна*);
КОВАЛЕНКО Андрій Анатолійович
(*д-р техн. наук, проф., Харків, Україна*);
КОЛОМІЙЦЕВ Олексій Володимирович
(*д-р техн. наук, проф., Харків, Україна*);
КОРОБКО Богдан Олегович
(*д-р техн. наук, доц., Полтава, Україна*);
КРАСНОБАЄВ Віктор 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, доц., Київ, Україна*);
ЛУКОВА-ЧУЙКО Наталія Вікторівна
(*д-р техн. наук, проф., Київ, Україна*);
ЛУНТОВСЬКИЙ Андрій Олегович
(*д-р техн. наук, проф., Німеччина*);
МИРОНЦОВ Микита Леонідович
(*д-р ф.-м. наук, с.н.с., Київ, Україна*);
ПЄВНЄВ Володимир Яковлевич
(*д-р техн. наук, доц., Харків, Україна*);
ПИСАРЧУК Олексій Олександрович
(*д-р техн. наук, проф., Київ, Україна*);
ПОДКОПАЄВ Сергій Вікторович
(*д-р техн. наук, проф., Покровськ, Україна*);
СЕМЕНОВ Сергій Геннадійович
(*д-р техн. наук, проф., Краків, Польща*);
ТРИСТАН Андрій Вікторович
(*д-р техн. наук, проф., Черкаси, Україна*);
ФРОЛОВ Євгеній Андрійович
(*д-р техн. наук, проф., Полтава, Україна*);
ЧОРНИЙ Олексій Петрович
(*д-р техн. наук, проф., Кременчук, Україна*);
Відповідальний секретар:
КУЧУК Ніна Георгіївна
(*д-р техн. наук, проф., Харків, Україна*).

Технічні секретарі:

ЗАХАРЧЕНКО Руслан Володимирович
(*канд. техн. наук, доц., Полтава, Україна*);
ПЕТРОВСЬКА Інна Юріївна
(*д-р філософії, Харків, Україна*).

Editorial board

Editor-in-Chief:

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

Associates editor:

Katerina NESTERENKO
(*Dr. Sc. (Tech.), Prof., Kyiv, Ukraine*);
Oleksandr SHEFER
(*Dr. Sc. (Tech.), Prof., Poltava, Ukraine*).

Editorial board members:

Nathan BLAUNSTEIN
(*Dr. Sc. (Tech.), Prof., Israel*);
Volodymyr BOHOMYA
(*Dr. Sc. (Tech.), Prof., Kyiv, Ukraine*);
Roman VARBANETS
(*Dr. Sc. (Tech.), Prof., Odesa, Ukraine*);
Krzysztof WESOŁOWSKI
(*Dr. Sc. (Tech.), Prof., Poland*);
Yevhen HAVRILKO
(*Dr. Sc. (Tech.), Prof., Kyiv, Ukraine*);
Elshan Giyas oglu HASHIMOV
(*Dr. Sc., Prof., Baku, Azerbaijan*);
Valentyn GLYVA
(*Dr. Sc. (Tech.), Prof., Kyiv, Ukraine*);
Olena DAKI
(*Dr. Sc. (Tech.), Ass. Prof., Izmail, Ukraine*);
Andriy KOVALENKO
(*Dr. Sc. (Tech.), Prof., Kharkiv, Ukraine*);
Oleksii KOLOMIITSEV
(*Dr. Sc. (Tech.), Prof., Kharkiv, Ukraine*);
Bohdan KOROBKO
(*Dr. Sc. (Tech.), Ass. Prof., Poltava, Ukraine*);
Viktor KRASNOBAYEV
(*Dr. Sc. (Tech.), Prof., Kharkiv, Ukraine*);
Heorhii KUCHUK
(*Dr. Sc. (Tech.), Prof., Kharkiv, Ukraine*);
Larysa LEVCHENKO
(*Dr. Sc. (Tech.), Ass. Prof., Kyiv, Ukraine*);
Natalia LUKOVA-CHUIKO
(*Dr. Sc. (Tech.), Prof., Kyiv, Ukraine*);
Andryy LUNTOVSKYY
(*Dr. Sc. (Tech.), Prof., Kharkiv, Ukraine*);
Mykyta MYRONTSOV
(*Dr. Sc. (Ph.&M.), Senior Res., Kyiv, Ukraine*);
Volodymyr PEVNEV
(*Dr. Sc. (Tech.), Ass. Prof., Kharkiv, Ukraine*);
Oleksii PYSARCHUK
(*Dr. Sc. (Tech.), Prof., Kyiv, Ukraine*);
Serhii PODKOPAIEV
(*Dr. Sc. (Tech.), Prof., Pokrovsk, Ukraine*);
Serhii SEMENOV
(*Dr. Sc. (Tech.), Prof., Krakow, Poland*);
Andrii TRYSTAN
(*Dr. Sc. (Tech.), Prof., Cherkasy, Ukraine*);
Yevhen FROLOV
(*Dr. Sc. (Tech.), Prof., Poltava, Ukraine*);
Oleksii CHORNYI
(*Dr. Sc. (Tech.), Prof., Kremenchuk, Ukraine*);
Responsible secretary:
Nina KUCHUK
(*Dr. Sc. (Tech.), Prof., Kharkiv, Ukraine*).

Technical secretaries:

Ruslan ZAKHARCHENKO
(*PhD (Tech.), Ass. Prof., Poltava, Ukraine*);
Inna PETROVSKA
(*PhD (Tech.), Kharkiv, Ukraine*).

З М І С Т

НАВІГАЦІЯ ТА ГЕОІНФОРМАЦІЙНІ СИСТЕМИ

<i>Гридасов І. Ю., Худов Г. В., Хижняк І. А., Яловега М. О., Бурков В. В.</i> Оптимізація маршруту руху безпілотного літального апарату літакового типу при пошуку повітряного судна, яке зазнало лиха із використанням рівномірно-оптимальної стратегії пошуку	5
<i>Leha O., Martovytskyi V., Severin I.</i> Analysis of methods for detecting unmanned aerial vehicles	13

АВТОМОБІЛЬНИЙ, РІЧКОВИЙ, МОРСЬКИЙ ТА АВІАЦІЙНИЙ ТРАНСПОРТ

<i>Никитюк П. В., Мельник О. М.</i> Розрахункова модель оцінки експлуатаційної безпеки судна з урахуванням деградації бар'єрів та прогнозу ризику	20
<i>Орисенко О. В., Криворот А. І., Шаповал М. В., Розозін І. А.</i> Дослідження коефіцієнта жорсткості пружинного та пневматичного елементів підвіски автомобіля Opel Vivaro	27
<i>Шантала С. О., Шмельова Т. Ф.</i> Дослідження завдання формування потоків повітряного руху в інтегрованому повітряному просторі	35

УПРАВЛІННЯ В СКЛАДНИХ СИСТЕМАХ

<i>Авакін С. М., Довгополий С. О., Запорожець О. В., Мощенко І. О.</i> Коригування функції перетворення засобів вимірювання з використанням нейромережевого підходу	41
<i>Бовчалюк С. Я., Бовчалюк Н. І., Дрозд О. Ю.</i> Концепція «модульної архітектури» керуючих пристроїв паралельної дії	47
<i>Harmash V., Diachenko V., Mikhal O., Znaidiuk V.</i> Energy-saving method in wireless sensor networks	54
<i>Дрючко О. Г., Шефер О. В., Кислиця С. Г., Галай В. М., Сухоребрий О. В.</i> Розробка адаптивних систем управління термічною поризацією лужно-силікатних теплоізоляційних виробів та з'ясування особливостей режимів її відтворення	59
<i>Качанов П. О., Тарасенко М. В.</i> Система автоматичного регулювання розрідження в топці котла	66
<i>Янко А. С., Михайліченко О. В., Крук О. І.</i> Оптиміальні алгоритми автоматичного керування робототехнічною платформою, які забезпечують її високу продуктивність та надійність функціонування	70

ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

<i>Barkovska O., Shapiro A., Mavrynskyi O., Zhebin P.</i> Research on the sensitivity of the distance measurement method based on the facemesh algorithm	76
<i>Бологова Н. М., Білоусов І. А.</i> Метод розпізнавання жестів для інтерактивного керування комп'ютером з урахуванням контекстної адаптації	83
<i>Бондаренко М. Е., Іващенко Г. С.</i> Використання послідовності методів попередньої обробки в системах голосової ідентифікації	90
<i>Golovko G., Taranenko I., Kushch O.</i> Zero trust: why traditional security models no longer work	97
<i>Hulevykh M., Kolomiitsev O.</i> Automated test generation techniques for C++ software	102
<i>Demydenko M., Fesenko A.</i> Development of a Web API for an online car parts store using N-Tier architecture	108
<i>Derkach T., Dmytrenko A., Klochko L.</i> Modeling and implementation of a distance learning system: stages, requirements, functionality	113
<i>Do K., Klymova I., Naumova E., Herevykh M., Yankovskyi O.</i> Data processing and analysis methods in IOT using machine learning	119
<i>Заковоротний О. Ю., Хулап А. В.</i> Аналіз моделей виявлення вторгнень на основі нейромереж у системах Інтернету речей	125
<i>Льїна І. В., Артюх Р. В., Зимогляд М. М.</i> Аналіз даних та машинне навчання у хмарних та туманних платформах для ефективної передачі даних	132
<i>Коваленко А. А., Чхеїдзе В. О., Севостьянова О. М., Фомічов О. О.</i> Підвищення точності аналізу та обробки складноструктурних зображень	137
<i>Кожевніков Г. К., Кравченко А. О., Стадник А. О., Гриньов Д. В.</i> Формування стійкого мобільного кластеру туманного шару системи підтримки Інтернету речей	141
<i>Кравченко О. С., Галуза О. А., Савченко А. О., Погорелов С. В., Роговий А. І.</i> Моделювання топографії мікроподрій на дзеркальних поверхнях	145
<i>Ліпчанський М. В., Ліпчанська О. В., Молчанов Г. І., Христофоров К. А.</i> Розробка API для адаптивного тестування в системах управління навчальними курсами	154
<i>Носко С. В., Бульба С. С., Черних О. П., Скородєлов В. В., Панченко В. І.</i> Спосіб високоефективної реалізації сайдкар компонента з мінімальними витратами системних ресурсів	161
<i>Первеев В. Д., Філімончук Т. В., Бугрій А. М., Партика С. О.</i> Модель мобільного застосунку для ОС Android	169

<i>Помелуйко Д. А., Єрьоміна Н. С., Петров С. В., Іщенко Н. В., Волощук О. Б.</i> Методи масштабування Kubernetes кластеру в хмарному середовищі	175
<i>Пономаренко О. Є., Горбачов В. О.</i> Метод пошуку вузьких місць в складних мережах для підвищення ефективності обчислювальних ресурсів	180
<i>Samborskyi Ie., Peleshok Ye.</i> Synthesis of logical-dynamic information and event management systems for computer security structures	185
<i>Тереник Д., Харченко В. С.</i> Автоматизація створення класифікаторів: алгоритми і програмне забезпечення для формування множини стратегій розгортання та забезпечення надійності рою БПЛА	195
<i>Шандиба А. С., Сітніков В. І., Ніколаєв О. С., Кібірев Д. О.</i> Нанесення цифрових водяних знаків з використанням хаотичних карт	203
<i>Шматко О. В., Рагулін О. Є., Кравченко П. О., Буслов П. В.</i> Дослідження архітектурних рішень для побудови безпечної системи зберігання та передачі конфіденційних даних	210

ЦИВІЛЬНА БЕЗПЕКА

<i>Бірук Я. І., Шаманський С. Й., Климчук А. В.</i> Дослідження ефективності електромагнітних екранів резонансного типу	218
<i>Глива В. А., Кашлев М. С.</i> Екранування шуму будівельного обладнання	222
<i>Левченко Л. О., Ашуева Н. М., Ткаченко Т. М.</i> Моделювання зміни температури на поверхнях теплоізоляційних покриттів	226
<i>Токарський О. І., Зозуля Л. А., Галонько Я. О., Москаленко О. О.</i> Засоби нормалізації фізичних факторів середовища приміщень спеціального призначення	230

ЗВ'ЯЗОК, ТЕЛЕКОМУНІКАЦІЇ ТА РАДІОТЕХНІКА

<i>Bikchentyayev M., Boriak B.</i> Integration of SDR into UAV systems	235
<i>Perets K., Komar O.</i> Assessment of the impact of sparsity and Geman-McClure regularization on signal reconstruction accuracy	239
<i>Руденко В. В.</i> Огляд протоколу LoRa та перспективи використання	244
<i>Sadovnykov B., Lysechko V.</i> A method for searching and recognising objects in a video stream by calculating interframe deltas	249
<i>Syvolovskyi I., Zhuchenko O.</i> A method of hierarchical clustering of nodes in distributed telecommunication systems using graph algorithms	255

АЛФАВІТНИЙ ПОКАЖЧИК

.....	263
-------	-----

Організації авторів

Геологічне бюро BEG SA, Швейцарія
 Державна служба України з надзвичайних ситуацій у Закарпатській області, Мукачеве, Україна
 Державний біотехнологічний університет, Харків, Україна
 Державний університет «Київський авіаційний інститут», Київ, Україна
 Київський національний університет будівництва і архітектури, Київ, Україна
 Науково-дослідний інститут воєнної розвідки, Київ, Україна
 Національна академія Національної гвардії України, Харків, Україна
 Національний аерокосмічний університет «ХАІ», Харків, Україна
 Національний технічний університет «Харківський політехнічний інститут», Харків, Україна
 Національний технічний університет України «КПІ імені Ігоря Сікорського», Київ, Україна
 Національний університет «Полтавська політехніка імені Юрія Кондратюка», Полтава, Україна
 Одеський національний морський університет, Україна
 Український державний університет залізничного транспорту, Харків, Україна
 Харківський національний університет Повітряних Сил імені Івана Кожедуба, Харків, Україна
 Харківський національний університет радіоелектроніки, Харків, Україна
 Харківський національний університет імені В.Н. Каразіна, Харків, Україна

Authors affiliation

Geological Bureau BEG SA, Switzerland
 State Emergency Service of Ukraine in the Zakarpattia Oblast, Mukacheve, Ukraine
 State Biotechnological University, Ukraine
 State University "Kyiv Aviation Institute," Kyiv, Ukraine
 Kyiv National University of Construction and Architecture, Kyiv, Ukraine
 Research Institute of Military Intelligence, Ukraine
 National Academy National Guard of Ukraine, Kharkiv, Ukraine
 National Aerospace University "KHAI", Kharkiv, Ukraine
 National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine
 National Technical University "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine
 National University "Yuri Kondratyuk Poltava Polytechnic", Poltava, Ukraine
 Odesa National Maritime University, , Ukraine
 Ukrainian State University of Railway Transport, Kharkiv, Ukraine
 Ivan Kozhedub Kharkiv National Air Force University, Kharkiv, Ukraine
 Kharkiv National University of Radio Electronics, Kharkiv, Ukraine
 V. N. Karazin Kharkiv National University, Kharkiv, Ukraine

Навігація та геоінформаційні системи

УДК 355.588

doi: 10.26906/SUNZ.2025.2.5-12

І. Ю. Грідасов¹, Г. В. Худов¹, І. А. Хижняк¹, М. О. Яловега¹, В. В. Бурков²

¹ Харківський національний університет Повітряних Сил імені Івана Кожедуба, Харків, Україна

² Національна академія Національної гвардії України, Харків, Україна

ОПТИМІЗАЦІЯ МАРШРУТУ РУХУ БЕЗПІЛОТНОГО ЛІТАЛЬНОГО АПАРАТУ ЛІТАКОВОГО ТИПУ ПРИ ПОШУКУ ПОВІТРЯНОГО СУДНА, ЯКЕ ЗАЗНАЛО ЛИХА ІЗ ВИКОРИСТАННЯМ РІВНОМІРНО-ОПТИМАЛЬНОЇ СТРАТЕГІЇ ПОШУКУ

Анотація. Предметом вивчення в статті є маршрут руху безпілотної літальної апарату літакового типу при проведенні пошуку повітряного судна, яке зазнало лиха. **Завдання:** У даній статті пропонується провести аналіз існуючих методів пошуку повітряного судна, оптимізацію маршруту руху для безпілотної літальної апарату літакового типу при пошуку повітряного судна, яке зазнало лиха з використанням рівномірно-оптимальних стратегій пошуку. На першому етапі роботи представлено вирішення оптимізаційної задачі пошуку для рівномірно-оптимальної стратегії пошуку. На другому етапі роботи розроблена блок-схема алгоритму для пошуково-рятувальної операції із використанням рівномірно-оптимальної стратегії пошуку. На третьому етапі проілюстровані результати застосування рівномірно-оптимальної стратегії пошуку повітряного судна, яке зазнало лиха з використанням мови програмування Python. Розглянуті шляхи використання отриманих наукових результатів. Запропоновані можливі шляхи подальших наукових досліджень, які будуть ґрунтуватися на моделюванні авіаційних катастроф та розрахунку кількісних показників за результатами маршруту руху безпілотної літальної апарату літакового типу, із застосуванням для даної задачі тренажеру (симулятора), у якому безпілотний літальний апарат буде використовуватися, як засіб пошуку. Використовуваними методами є: теоретичні методи аналізу та синтезу, методи теорії імовірності, методи математичної статистики, методи комбінаторики, методи математичного моделювання, математичний апарат теорії матриць, методи диференційного числення, методи теорії оптимізації, методи математичної логіки, метаевристичні методи оптимізації, генетичні алгоритми, аналітичні та емпіричні методи порівняльного дослідження. **Висновки.** У роботі проведена оптимізація маршруту руху безпілотної літальної апарату літакового типу при пошуку повітряного судна, яке зазнало лиха із використанням рівномірно-оптимальної стратегії пошуку. Новими науковими результатами є: виконана задача оптимізації маршруту руху безпілотної літальної апарату літакового типу при пошуку повітряного судна, яке зазнало лиха, розроблена блок-схема алгоритму реалізації методу пошуку повітряного судна, яке зазнало лиха із використанням рівномірно-оптимальної стратегії пошуку, отриманні результати застосування рівномірно-оптимальної стратегії пошуку повітряного судна, яке зазнало лиха із використанням мови програмування Python, які в свою чергу підтверджують ефективність застосування удосконаленого методу пошуку. Сформульовані можливі напрями для подальших досліджень в контексті виявлення повітряних суден, які зазнали лиха. Дослідження були проведені за рахунок грантової підтримки Національного фонду досліджень України в рамках конкурсу “Наука для зміцнення обороноздатності України”, назва проєкту “Інформаційна технологія автоматизованого сегментування зображень об’єктів в системах націлювання ударних FPV-дронів на основі алгоритмів ройового інтелекту”, реєстраційний номер 2023.04/0153.

Ключові слова: аварійна ситуація; безпілотний літальний апарат; маршрут руху; повітряне судно; пошук; оптимізація; пошуково-рятувальне забезпечення; пошуково-рятувальна операція; рівномірно-оптимальна стратегія.

Вступ

Постановка проблеми у загальному вигляді. Актуальність проведення наукових досліджень в області пошуково-рятувальних операцій екіпажів повітряних суден, які зазнали лиха, є особливо важливою задачею в сучасному світі з огляду на деякі ключові фактори.

По-перше, в сучасному світі відбулося суттєве зростання авіаційного трафіку, особливо в комерційному секторі, що в свою чергу збільшило ймовірність виникнення авіаційних подій [1]. Зазначене потребує проведення ефективних і швидких пошуково-рятувальних операцій. Незважаючи на відносно високий рівень безпеки сучасних авіаційних технологій, ризик авіаційних подій залишається актуальним через такі фактори, як несприятливі погодні

умови, технічні несправності повітряних суден, людський фактор тощо.

По-друге, збільшення виникнення авіаційних подій та втрати повітряних суден пов’язані із збільшенням інтенсивності сучасних збройних конфліктів (війн) у ХХІ (двадцять першому) столітті, зокрема російсько-українська війна, палістино-ізраїльська війна, війна в Сирії тощо. Інтенсивне застосування в сучасних війнах широкої номенклатури різних засобів протиповітряної оборони (ППО) робить бойові винишувачі та інші типи літаків доволі вразливою ціллю [2]. Зазначений фактор потребує проведення добре спланованих пошуково-рятувальних операцій, які обмежені за часом та мають свої особливості у зв’язку з великими ризиками для рятувальних підрозділів. Вони пов’язані із протидією противника, у разі якщо операції з пошуку та рятування відбуваються на

території, яка контролюється противником. Додатковими викликами можуть стати: застосування противником засобів радіоелектронної боротьби (РЕБ), засобів ППО, що в свою чергу значно ускладнює проведення пошуку та порятунку екіпажу, який зазнав лиха. Зазначені фактори призводять до необхідності удосконалення існуючих методів пошуку в умовах ведення бойових дій.

По-третє, сучасні виклики, пов'язані із змінами кліматичних умов, вимагають також удосконалення методів пошуку і рятування. Авіакатастрофи, пов'язані із екстремальними погодними явищами, такими як тайфуни, бурі та аномальні погодні умови, суттєво ускладнюють існуючі методи пошуку повітряних суден, які зазнали лиха. Прикладом таких пошукових операцій є авіатроща, яка відбулася 19 травня 2024 року, в якій загинув президент Ірану Раїсі та всі члени екіпажу, які перебували в гелікоптері. Масштабна пошуково-рятувальна операція, яка тривала майже добу, була ускладнена погодними умовами (густим туманом) [3].

Особливо важливим чинником, який надає можливості для удосконалення методів пошуку є технологічний прогрес у галузі безпілотних літальних апаратів (БпЛА), штучного інтелекту та супутникових систем зв'язку. Слід зазначити, що традиційні методи пошуку, які використовуються в міжнародній практиці не передбачають специфіку застосування БпЛА, супутників в операціях пошуку та рятування [4]. Наукові дослідження, спрямовані на інтеграцію цих технологій у світовій системі пошуку та рятування мають вагомий потенціал для підвищення ефективності операцій, що ймовірно у майбутньому може зменшити час реагування та збільшити шанси на виживання пасажирів та екіпажу повітряного судна, яке зазнало лиха.

Крім цього, глобалізація авіаційних перевезень означає, що авіакатастрофи можуть відбуватися у віддалених або важкодоступних регіонах світу, де існуючі методи пошуку можуть мати певні недоліки та бути менш ефективними. Це підкреслює необхідність у міжнародному співробітництві та стандартизації підходів до проведення пошуково-рятувальних операцій, що також потребує наукових досліджень.

У даній науковій роботі запропонована оптимізація маршруту руху БпЛА літакового типу при пошуку повітряного судна, яке зазнало лиха з використанням рівномірно-оптимальної стратегії пошуку. Використання БпЛА літакового типу в пошуково-рятувальних операціях перш за все пов'язане з їхніми унікальними тактико-технічними характеристиками, які дозволяють ефективно покривати великі площі в короткі проміжки часу. На відміну від вертольотів або літаків, БпЛА літакового типу здатні тривалий час перебувати в повітрі, здійснювати моніторинг з великої висоти та передавати зображення в реальному часі, що особливо важливо при обмеженій видимості або складних погодних умовах. Такий підхід дозволяє не лише мінімізувати витрати часу на обстеження ймовірних зон перебування повітряного судна, яке зазнало лиха, але й забезпечує раціональне використання наявних

ресурсів у рятувальних підрозділах. Врахування ймовірностей місцезнаходження об'єкта пошуку, швидкості та дальності польоту БпЛА, а також обмежень зон польоту дозволяє оптимізувати маршрут, який забезпечує максимальне покриття території з найвищим очікуваним результатом. Запропонований метод може бути реалізований у реальних умовах і доповнений інтеграцією із геоінформаційними системами, супутниковими даними та алгоритмами машинного навчання для більш точної оцінки ймовірності знаходження об'єкта пошуку.

Таким чином, проведення наукових досліджень в області пошуково-рятувального забезпечення, а саме застосування БпЛА літакового типу для пошуково-рятувальних операцій є актуальною задачею. Отримані результати наукових досліджень в майбутньому можуть бути використані в пошуково-рятувальних операціях, з урахуванням сучасних викликів та загроз, які виникають перед рятувальними підрозділами.

Аналіз останніх досліджень і публікацій.

Аналіз досліджень останніх років за напрямом проведення пошуково-рятувальних операцій, екіпажів повітряних суден, які зазнали лиха демонструють значний прогрес у застосуванні новітніх технологій, зокрема БпЛА, супутникових навігаційних систем, штучного інтелекту тощо [5]–[11]. У даних наукових роботах відображається тенденція до інтеграції цих технологій у загальну (міжнародну) систему рятувальних заходів, що дозволяє суттєво підвищити ефективність і точність виявлення повітряних суден, які зазнали лиха, особливо в складних умовах проведення пошуково-рятувальної операції. Одним із ключових аспектів останніх досліджень є акцент на розвиток автоматичного аналізу великих обсягів даних, які надходять із різних джерел (радіолокаційних станцій, супутників тощо), для швидкого визначення можливого місцезнаходження повітряного судна, яке зазнало лиха.

У роботі [5] розроблено обґрунтування можливостей нарощування спроможностей авіаційної компоненти оперативно-рятувальної служби Державної служби надзвичайних ситуацій України шляхом впровадження БпЛА для виконання завдань оперативного моніторингу природних екосистем і технологічних об'єктів з високим рівнем ризику виникнення надзвичайних ситуацій. Результатом роботи є наведена класифікація БпЛА, обґрунтування вимог до БпЛА в інтересах цивільного захисту та обґрунтування вимог до льотно-технічних характеристик БпЛА, які використовуються у пошуково-рятувальних операціях. У роботі проведені розрахунки величин радіусів розвороту, кутових швидкостей розвороту та можливих перевантажень БпЛА під час обстеження районів в який відбулася надзвичайна ситуація.

Суть [6] полягає в дослідженні архітектури, яка може бути застосована для керування кількома БпЛА, задіяними у пошуково-рятувальних операціях. У запропонованій системі один оператор здатний координувати дії набору роботів для отримання відповідної інформації про навколишнє середовище. Ця

робота виконана з метою розробки змішаної наземної та повітряної роботизованої платформи для підтримки пошуково-рятувальних операцій в умовах високогір'я. На відміну від типових налаштувань взаємодії людини та БпЛА, тут оператор не повністю управляє БпЛА, але бере участь у пошуково-рятувальних заходах, тому може давати роботам лише розрізнені та неповні інструкції. У даній роботі обговорюються домен, структура взаємодії та виконавча система для автономного виконання пошуково-рятувальних заходів. Зазначена система була апробована у реальній пошуковій операції із залученням двох БпЛА, оснащеними бортовими камерами.

Наукова робота [7] присвячена дослідженням, які спрямовані на порятунок заблукалих, хворих або поранених людей за допомогою автономних БпЛА. У статті розглянута проблема виявлення постраждалих людей в густо-лісистій місцевості, яка має певні складності, через оклюзію, запропоновані дієві механізми їх виявлення. Автори доказали, що автоматичне виявлення постраждалої людини в умовах оклюзії можна значно покращити шляхом поєднання багато перспективних зображень перед класифікацією. Для вирішення проблемного питання автори здійснили інтеграцію зображення за допомогою бортового оптичного секціонування (AOS) – методу створення зображень із синтетичною апертурою, який використовують безпілотні камери для захоплення неструктурованих теплових світлових полів, для того щоб досягти цього з точністю та запам'ятовуванням 96 % і 93 % відповідно.

Перевагою роботи [8] є запропонований алгоритм розподілу завдань для кількох роботизованих систем для різних підходів БпЛА із використанням ринкового підходу для оптимізації ресурсів пошуково-рятувальної місії. Досліджений алгоритм перевіряє наявність транспортних засобів, характеристики кожного завдання та корисне навантаження, необхідні вхідні дані, а потім надає кожному транспортному засобу план завдань і команди для подальших дій. Запропонований алгоритм було перевірено шляхом виконання кількох пошуково-рятувальних операцій в моделюванні гірської місцевості з реальними розмірами, отримані результати показали надійність та ефективність запропонованої системи порятунку потерпілих осіб.

У науковій публікації [9] представлені результати початкової оцінки комп'ютеризованого методу інтерпретації, перевіреного в порівнянні з ручними методами під час імітації пошуково-рятувальних операцій типу Search and Rescue (SAR). Проведена оцінка зосереджена на використанні ресурсів, доступних для рятувальних підрозділів, які виконують операції типу SAR, зокрема: даних від свідків, програмних засобів тощо. Проте, результати цього випробування є неоднозначними, оскільки традиційна група виявила більше об'єктів, але вимагала більше часу, у людино-годинах, щоб виявити об'єкти. Автори зауважили про необхідність проведення подальших експериментів, які будуть засновані на можливостях рятувальних підрозділів, з метою визначення, наскільки комп'ютерні методи пошуку ефек-

тивні в операціях типу SAR. Також заслуговує на увагу зростаюча роль у стандартизації підходів до рятувальних операцій. У роботі [10] автори наголошують на важливості розробки єдиних протоколів зв'язку та використання спільних платформ для обміну інформацією між країнами, що значно покращує координацію і ефективність міжнародних пошуково-рятувальних операцій.

Аналіз наукової публікації [11], яка розглядає перші випадки застосування БпЛА у пошуково-рятувальних операціях, має важливе значення для розуміння перспектив використання цих технологій у надзвичайних ситуаціях. У роботі детально описано два випадки застосування БпЛА під час рятувальних операцій, що дозволяє провести порівняльний аналіз переваг та недоліків цих технологій. Аналіз наукової публікації дозволяє зробити висновок, що БпЛА мають значний потенціал для пошуково-рятувальних операцій, однак їх ефективність значною мірою залежить від специфіки операції, технічних характеристик та умов експлуатації.

Таким чином, проведений аналіз сучасних наукових публікацій за тематикою проведення наукових досліджень, демонструє значний прогрес у впровадженні новітніх технологій, зокрема БпЛА, супутникових систем, штучного інтелекту, роботизованих платформ тощо у процесі проведення пошуково-рятувальних операцій. Перевагою існуючих наукових публікацій є те, що дослідження застосування БпЛА в пошуково-рятувальних операціях показали свою ефективність, як у міських умовах, так і в важкодоступних природних зонах. Вони дозволяють швидко і безпечно проводити огляд великих територій, мінімізуючи ризики для рятувальних підрозділів. Також, сучасні дослідження вказують на значний потенціал автоматизації та роботизації пошуково-рятувальних операцій, що забезпечує покращену координацію дій і можливість більш точного та швидкого реагування на надзвичайні ситуації.

На теперішній час, значна увага приділяється інтеграції великих обсягів даних з різних джерел, таких як, радіолокаційні системи та супутникові зображення, що дозволяє суттєво підвищити точність визначення місцезнаходження повітряного судна, яке зазнало лиха.

Проте, слід констатувати, що технології, які залучаються до пошуково-рятувальних операцій, досі стикаються з певними обмеженнями, такими як залежність від погодних умов, тривалість польоту БпЛА, а також необхідності удосконалення маршруту руху БпЛА в районі пошуку. Також слід зауважити, що досліджень в області пошуку повітряних суден, які зазнали лиха із застосуванням сучасних засобів пошуку обмежена кількість. Таким чином, напрям наукових досліджень даної наукової публікації, а саме в контексті пошуково-рятувальних операцій при пошуку повітряних суден, які зазнали лиха є актуальним, враховуючи сучасні виклики та загрози.

Мета статті – оптимізація маршруту руху БпЛА літакового типу при пошуку повітряного судна, яке зазнало лиха із використанням рівномірно-оптимальних стратегій пошуку.

Основна частина

Задача пошуку та виявлення об'єктів в повітрі та на воді відіграє вирішальну роль у забезпеченні успішного проведення авіаційних і морських пошуково-рятувальних операцій. Важливість цього аспекту стала особливо очевидною під час пошукових заходів, спрямованих на виявлення рейсу Malaysia Airlines MH370. Незважаючи на масштабні зусилля, що включали близько 40 суден та авіаційних засобів, якими було обстежено понад 4.6 мільйона квадратних кілометрів акваторії Індійського океану, пошукова операція завершилася без отримання підтверджених результатів щодо виявлення уламків літака або пасажирів [12].

Проектування пошукових операцій на земній поверхні (океану) стикається з двома основними труднощами. Перша пов'язана з високим рівнем невизначеності при визначенні району пошуку. Для прикладу, у випадку рейсу MH370 цільові області постійно переглядалися через постійний потік нової інформації, включаючи супутникові дані, випадкові спостереження, сміття та сигнали, які могли походити від аварійного радіомаяка [12]. Дана невизначеність може ускладнюватися динамічними процесами дрейфу на поверхні океану. Океанічні течії нестабільні та можуть спричиняти помилки в оцінці ймовірного місця падіння об'єкта, особливо в ті дні, що проходять між самим інцидентом та початком пошуково-рятувальної операції. Окрім цього, через неточність моделей дрейфу остаточні оцінки цільових зон пошуку, як правило, залишаються з великою часткою невизначеності.

Другим джерелом труднощів є те, що навіть коли район пошуку відомий і невизначеність щодо його меж добре описана, наприклад, через розподіл ймовірностей, проблема проектування оптимальних маршрутів для рятувальних підрозділів залишається нерозв'язаною [13]. На практиці, для вирішення цього питання район пошуку розподіляють на кілька підрайонів, кожний з яких призначається певному рятувальному підрозділу. Таким чином, рятувальні підрозділи обстежують підрайони організованими маршрутами, зазвичай, виконуючи пошук паралельними галсами (рис. 1) або вздовж лінії шляху (рис. 2) повітряного судна, яке зазнало лиха [4], [13–16].

Метод пошуку з паралельними галсами є простим у плануванні та реалізації для районів з нескладною геометрією, однак він не може бути ефективно застосований до складних районів пошуку, які виникають через дрейфи на поверхні океану або нерівномірний розподіл ймовірності знаходження цілей. Цей метод також не є достатньо гнучким для реальних операцій, наприклад, у разі, якщо один з рятувальників збивається з маршруту, необхідно буде назначати нові пошукові області та рятувальні підрозділи, що ускладнює проведення пошуково-рятувальних операцій.

Метод пошуку вздовж лінії шляху передбачає проведення пошукових заходів уздовж передбачуваного маршруту об'єкта, який зазнав лиха. Даний метод ґрунтується на ідеї, що літак, який зазнав лиха, ймовірно знаходиться поблизу свого останнього відомого місцезнаходження або вздовж запланованого курсу.

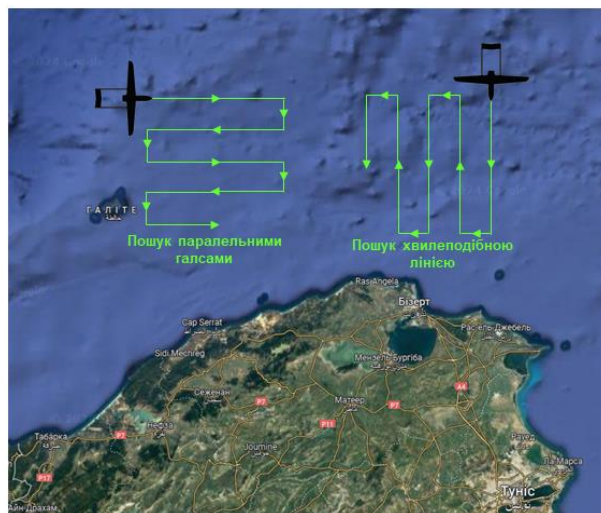


Рис. 1. Метод пошуку повітряного судна, яке зазнало лиха паралельними галсами



Рис. 2. Метод пошуку вздовж лінії шляху

Рятувальні підрозділи організовують пошук, слідуючи вздовж цієї лінії, що дозволяє ефективно охопити територію ймовірного місця інциденту. Метод часто використовується на початкових етапах пошуково-рятувальних операцій для швидкого визначення потенційного місця лиха.

Проектування траєкторії руху засобів пошуку являє собою багатокomпонентне завдання, що пов'язане з управлінням кількома рятувальними засобами, де цільова функція включає аналіз змінних у часі розподілів ймовірностей. Важливим елементом успішних пошуково-рятувальних операцій є врахування зміни динаміки навколишнього середовища. Водночас, у наукових роботах [17]–[18], присвячених як статичним, так і динамічним підходам до задач пошуку, надається багато корисних і значущих деталей, що розширюють можливості пошукових методик в умовах постійно змінюваного середовища, в якому здійснюється пошук.

Розглянувши існуючі методи пошуку, які використовуються на практиці, далі пропонується оптимізувати маршрут руху БПЛА літакового типу, при пошуку повітряного судна, яке зазнало лиха із використанням рівномірно-оптимальних стратегій пошуку. Засобами пошуку на практиці можуть бути літаки, кораблі, БПЛА тощо. У контексті даної роботи, будемо вважати, що засобом пошуку є БПЛА літакового типу, який здійснює послідовний, закономірний пошук. Основною задачею удосконалення маршруту руху БПЛА із використанням рівномірно-

оптимальних стратегій пошуку є зменшення витраченого часу в пошуково-рятувальній операції від моменту початку пошуку до виявлення повітряного судна, яке зазнало лиха.

По-перше, в рамках вирішення поставленої наукової задачі постає необхідність вирішення оптимізаційної задачі пошуку для рівномірно-оптимальної стратегії пошуку. Оптимізаційна задача необхідна для прийняття найкращого рішення в умовах обмежених ресурсів та змінних умов. У контексті пошуково-рятувальних операцій, зокрема пошуку повітряних суден, які зазнали лиха, оптимізаційна задача спрямована на мінімізацію часу і ресурсів, необхідних для виявлення об'єкту пошуку.

Оптимізація пошуку базується на мінімізації ймовірності не виявлення об'єкту пошуку за час відведений на пошук, із врахуванням ймовірних моделей пошуку. Функція ймовірності того, що об'єкт пошуку буде виявлений до моменту часу t розраховується за формулою (1):

$$K(t) = 1 - \int_0^t \int_{\Omega} \lambda(x, t) y(x, t) dt dx, \quad (1)$$

де $\lambda(x, t)$ – щільність ймовірності виявлення об'єкту пошуку в точці x в момент часу t ; $y(x, t)$ – функція інтенсивності пошуку.

Функція ймовірності визначається через інтеграл за площею району пошуку Ω , із врахуванням ймовірності виявлення об'єкту пошуку в залежності від часу та місцезнаходження. Оптимізаційна задача полягає у мінімізації функції ймовірності, математично вона буде мати вигляд:

$$\int_{\Omega} y(x, t) dx \rightarrow \min. \quad (2)$$

Наведений вираз означає, що пошуково-рятувальні підрозділи повинні підібрати таку стратегію пошуку, при якій час до моменту виявлення об'єкту пошуку буде мінімальним. При цьому функції $\lambda(x, t)$ та $y(x, t)$ повинні задовольняти наступним умовам:

Функція інтенсивності пошуку $y(x, t)$ пов'язана з функцією щільності ймовірності виявлення $\lambda(x, t)$ за допомогою виразу (3):

$$y_t(x, t) = -\lambda(x, t) y(x, t), \quad x \in \Omega, t > 0, \quad (3)$$

де x – координата в області пошуку; Ω – область пошуку.

Функція щільності ймовірності виявлення об'єкту пошуку $\lambda(x, t)$ має бути невід'ємною вираз (4):

$$\lambda(x, t) \geq 0, \quad x \in \Omega, t > 0. \quad (4)$$

Інтенсивність пошуку в області Ω задана функцією $L(t)$ вираз (5):

$$\int_{\Omega} \lambda(x, t) dx = L(t), \quad t > 0, \quad (5)$$

де $L(t)$ – інтенсивність пошуку в момент часу t (пошукові зусилля);

Початкова умова для інтенсивності пошуку $y(x, 0)$ задається функцією $u(x)$ вираз (6):

$$y(x, 0) = u(x), \quad (6)$$

Рівномірно-оптимальна стратегія пошуку має наступну властивість: існує область пошуку $\Omega_+(t) \subset \Omega$, яка постійно розширюється за часом, таким чином, що стратегія пошуку позитивна в межах області $\Omega_+(t)$ та негативна в $\Omega \setminus \Omega_+(t)$. Крім цього, стратегія пошуку є часовою функцією. З самого початку є за необхідним вважати, $\lambda(x, t)$ задовольняє наступним вимогам (вираз (7)):

$$\begin{aligned} \lambda(x, t) &> \text{for } x \in \Omega_+(t), \\ \lambda(x, t) &= 0 \text{ for } x \in \Omega \setminus \Omega_+(t). \end{aligned} \quad (7)$$

Вираз (8) демонструє, що на кожному $x \in \Omega$ існує певний момент часу $t(x)$, такий що $\lambda(x, t)$ позитивна для $t \in [t(x), T]$ і негативна для $t \in [0, t(x)]$:

$$\begin{aligned} \lambda(x, t) &> 0 \text{ for } t \in [t(x), T]; \\ \lambda(x, t) &= 0 \text{ for } t \in [0, t(x)]. \end{aligned} \quad (8)$$

На основі виразів (7) та (8) можливо зробити висновки, що стратегія пошуку є поступово зростаючою функцією, яка поступово охоплює всю область Ω у часі t , тобто $t(x)$ збільшується разом з x .

Властивості виразу (8) можливо отримати в результаті вирішення задачі оптимізації (вирази (2)–(6)). Далі в результаті вирішення задачі оптимізації, функція $\lambda(x, t)$ на практиці буде відповідати цим властивостям. Отже, для побудови такої стратегії пошуку, яка є рівномірно-оптимальною, тобто охоплює всю задану область пошуку рівномірно за часом, виведемо формулу (9):

$$\lambda(x, t) = \lambda_0(x) \exp \left(- \int_{t(x)}^t L(t) dt \right), \quad (9)$$

де $\lambda_0(x)$ – початкова функція пошуку в точці x .

Даний вираз дозволяє мінімізувати час пошуку всередині області покриваючи якомога більшу площу Ω_t в найкоротший час. Для практичної реалізації, рівномірно-оптимальної стратегії пошуку повітряного судна, яке зазнало лиха, авторами розроблена блок-схема, яка відображає ключові етапи запропонованого алгоритму (рис. 3).

Блок-схема алгоритму включає основні етапи процесу пошуку. Процес пошуку розпочинається після отримання оповіщення про авіаційну подію. На етапі введення вихідних даних в програмні засоби заноситься апріорна інформація про подію, межі району пошуку, кількість засобів пошуку, час відведений на пошук тощо.

Наступним кроком є розподіл засобів пошуку, на якому визначений район пошуку поділяється на підрайони та проводиться рівномірне розподілення засобів пошуку в них.

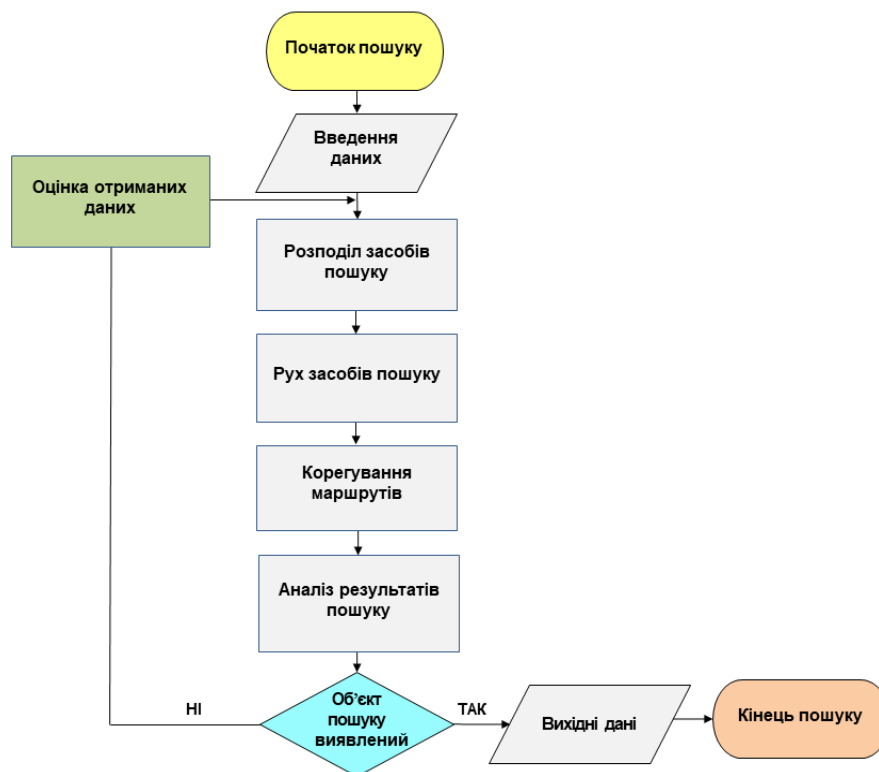


Рис. 3. Блок-схема алгоритму для практичної реалізації рівномірно-оптимальної стратегії пошуку

Далі організовується рух засобів пошуку (закономірний пошук) галсами, який здійснюється рівномірно у всьому районі пошуку.

Наступним етапом є корегування маршрутів засобів пошуку, який проводиться шляхом перерозподілу районів пошуку із врахуванням отриманої нової інформації про подію. Після обстеження підрайонів пошуку проводиться аналіз отриманих результатів. У разі, якщо об'єкт пошуку був виявлений – пошук завершується. У протилежному випадку, пошук повторюється в інших підрайонах із врахуванням оцінки отриманих нових даних.

Наведений алгоритм допомагає ефективно організувати пошук за допомогою декількох засобів пошуку, таких як, БпЛА літакового типу, які сліду-

ють за визначеними маршрутами з можливістю адаптації у процесі пошуково-рятувальної операції.

Рис. 4 ілюструє результати застосування рівномірно-оптимальної стратегії пошуку повітряного судна, яке зазнало лиха, реалізованої з використанням мови програмування Python. Три авіаційних засоби пошуку (БпЛА літакового типу), починають пошук із різних точок і рухаються до ймовірної точки місцезнаходження об'єкту пошуку (позначено червоним), використовуючи паралельні траєкторії або так звані галси. Дані галси пошуку (рис. 1) є стандартною частиною стратегії для забезпечення рівномірного покриття досліджуваної області. Ключова ідея алгоритму полягає у мінімізації часу пошуку при максимальному покритті території.

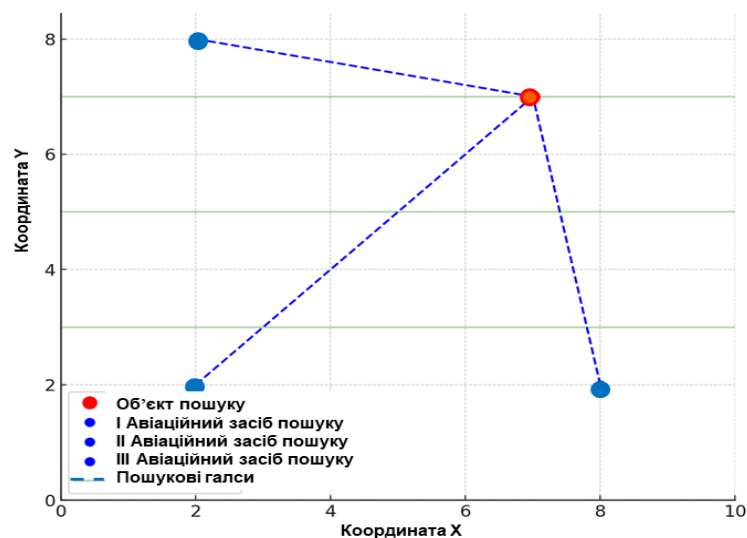


Рис. 4. Результат застосування рівномірно-оптимальної стратегії пошуку

Авіаційні засоби пошуку рухаються узгодженими траєкторіями руху, які адаптуються відповідно до динамічних змін або додаткової інформації, отриманої в процесі пошуку. Зображена схема демонструє важливість організації галсів для пошукових операцій і наочність рівномірно-оптимальної стратегії пошуку.

Таким чином, практична реалізація методу визначення маршруту руху для БпЛА літакового типу при пошуку повітряного судна, яке зазнало лиха із використанням рівномірно-оптимальної стратегії пошуку, може бути широко застосована в різних типах пошуково-рятувальних операцій. Основною перевагою даного маршруту руху БпЛА є необхідність відносно невеликого часу на виявлення об'єкту пошуку, перекриття усього району пошуку та простота реалізації.

Подальші дослідження доцільно спрямувати на практичне моделювання різних авіаційних катастроф та розрахунків ймовірностей виявлення, часу необхідного для пошуку повітряного судна, яке зазнало лиха, використовуючи рівномірно-оптимальну стратегію пошуку із застосуванням для даної задачі тренажеру (симулятора) БпЛА літакового типу, в якості засобу пошуку.

Висновки

У роботі проведена оптимізація маршруту руху БпЛА літакового типу при пошуку повітряного суд-

на, яке зазнало лиха із використанням рівномірно-оптимальної стратегії пошуку.

Новими науковими результатами є:

– виконана задача оптимізації маршруту руху БпЛА літакового типу при пошуку повітряного судна, яке зазнало лиха;

– розроблена блок-схема алгоритму реалізації методу пошуку повітряного судна, яке зазнало лиха із використанням рівномірно-оптимальної стратегії пошуку;

– отриманні результати застосування рівномірно-оптимальної стратегії пошуку повітряного судна, яке зазнало лиха із використанням мови програмування Python, які в свою чергу підтверджують ефективність застосування удосконаленого методу пошуку.

Сформульовані можливі напрями для подальших досліджень в контексті виявлення повітряних суден, які зазнали лиха.

Подяка

Дослідження були проведені за рахунок грантової підтримки Національного фонду досліджень України в рамках конкурсу “Наука для зміцнення обороноздатності України”, назва проєкту “Інформаційна технологія автоматизованого сегментування зображень об'єктів в системах націлювання ударних FPV-дронів на основі алгоритмів ройового інтелекту”, реєстраційний номер 2023.04/0153.

СПИСОК ЛІТЕРАТУРИ

1. ACI World Airport Traffic Forecasts 2023–2052. *ACI World*. URL: <https://store.aci.aero/product/aci-world-airport-traffic-forecasts-2023-2052>
2. The combat statistics for all the aircraft currently in use. *MiGFlug*. URL: <https://migflug.com/jetflights/the-combat-statistics-for-all-the-aircraft-currently-in-use>
3. Helicopter carrying Iranian President Raisi crashes, prompting massive search operation. *CNN World*. URL: <https://edition.cnn.com/2024/05/19/middleeast/iranian-president-raisi-helicopter-intl/index.html>
4. International Aeronautical and Maritime Search and Rescue Manual, vol. III. Mobile Facilities (Doc. 9731) 12 th Ed. 2022. 288 p.
5. Застосування безпілотних авіаційних систем у сфері цивільного захисту: монографія / Д.В. Бондар, А.В. Гурник, А.О. Литовченко, В.В. Хижняк, В.Л. Шевченко, Д.М. Ядченко. Київ. 2022. 312 с.
6. J. Cacace, A. Finzi, V. Lippiello, M. Furci, N. Mimmo, L. Marconi. A control architecture for multiple drones operated via multimodal interaction in search & rescue mission. *International Symposium on Safety, Security, and Rescue Robotics. Lausanne, Switzerland: IEEE; 2016*. p. 233–239. DOI: 10.1109/SSRR.2016.7784304.
7. DC. Schedl, I. Kurmi, O. Bimber. Search and rescue with airborne optical sectioning. *Nature Machine Intelligence*. vol. 2(12). 2020. p. 783–790.
8. M. Rodriguez, A. Al-Kaff, A. Madridano, D. Martin, A. De La Escalera. Wilderness search and rescue with heterogeneous multi-robot systems. *International Conference on Unmanned Aircraft Systems*. Athens, Greece: Divani Caravel Hotel, 2020. p.110–116.
9. WT. Weldon, J. Hupy. Investigating methods for integrating unmanned aerial systems in search and rescue operations. *Drones*. 2020. vol. 4(3). p. 1–16. DOI: 10.3390/drones4030038.
10. JN. McRae, BM. Nielsen, CJ. Gay, AP. Hunt, AD Nigh. Utilizing drones to restore and maintain radio communication during search and rescue operations. *Wilderness Environ Medicine*. 2021. vol. 32, issue 1. p. 41–46. DOI: 10.1016/j.wem.2020.11.002.
11. C. Van Tilburg. First report of using portable unmanned aircraft systems (drones) for search and rescue. *Wilderness Environ Medicine*. 2017. vol. 28, issue 2. p. 116–118. DOI: 10.1016/j.wem.2016.12.010.
12. Stefan Ivić, Bojan Crnković, Hassan Arbabi, Sophie Loire, Patrick Clary and Igor Mezić. Search strategy in a complex and dynamic environment: the MH370 case. *Scientific reports*. 2020. vol. 10, issue 1. p. 1-15. DOI: 10.1038/s41598-020-76274-0.
13. Худов Г.В., Грідасов І.Ю., Жилін Є.І. Математична формалізація завдання пошуку повітряного судна, яке зазнало лиха, авіаційними засобами. *Системи озброєння і військова техніка*. 2024. №1 (77). С. 117-123. DOI: 10.30748/soivt.2024.77.17.
14. Грідасов І.Ю., Худов Г.В. Аналіз методів пошуку повітряного судна, яке зазнало лиха, авіаційними засобами. *Збірник наукових праць Харківського національного університету Повітряних Сил імені Івана Кожедуба*. 2024. №2(80). С. 54-62. <https://doi.org/10.30748/zhups.2024.80.07>.
15. Грідасов І.Ю., Худов Г.В., Хижняк І.А. Визначення маршруту руху підрозділів, які задіяні до порятунку екіпажу повітряного судна, яке зазнало лиха. *Перспективи розвитку озброєння та військової техніки Сухопутних військ: зб. тез доп. міжн. наук.-техніч. конф.*, м. Львів, 15-16 травня 2024 р. Львів, 2024. С.257.

16. Грідасов І. Ю., Худов Г.В., Хижняк І. А. Обстеження району пошуку повітряного судна ,яке зазнало лиха підрозділом спеціального призначення Національної гвардії України з використанням рівномірно-оптимальної стратегії пошуку. *Актуальні проблеми діяльності складових сектору безпеки і оборони України (до 10-ї річниці створення Національної гвардії України)* : тези доп. Всеукр. наук.-практ. конф. (Харків, 24 жовт. 2024 р.). Харків, 2024. С. 296–298.
17. D. Stone. Theory of Optimal Search. New York, USA: *Mathematics in science and engineering*. 1975. p.288.
18. B. O. Koopman. Search and Screening. Operations evaluation group office of the chief of naval operations navy department, Washington. 1946. p. 179.

Received (Надійшла) 25.02.2025

Accepted for publication (Прийнята до друку) 14.05.2025

ВІДОМОСТІ ПРО АВТОРІВ/ ABOUT THE AUTHORS

- Грідасов Ілля Юрійович** – провідний науковий співробітник НМВ забезпечення якості освітньої діяльності та вищої освіти, Харківський національний університет Повітряних Сил імені Івана Кожедуба, Харків, Україна;
Illia Hridasov – Leading Researcher of Scientific and Methodological Department for Quality Assurance of Educational Process and Higher Education, Ivan Kozhedub Kharkiv National Air Force University, Kharkiv, Ukraine;
 e-mail: gridasovillya91@gmail.com; ORCID Author ID: <https://orcid.org/0000-0001-5689-0311>;
 Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=58623833600>.
- Худов Геннадій Володимирович** – доктор технічних наук, професор, начальник кафедри тактики радіотехнічних військ, Харківський національний університет Повітряних Сил імені Івана Кожедуба, Харків, Україна;
Hennadii Khudov – Doctor of Technical Sciences, Professor, Head of Department of Radar Troops Tactic, Ivan Kozhedub Kharkiv National Air Force University, Kharkiv, Ukraine;
 e-mail: 2345kh_hg@ukr.net; ORCID Author ID: <https://orcid.org/0000-0002-3311-2848>;
 Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57196079841>.
- Хижняк Ірина Анатоліївна** – кандидат технічних наук, начальник НМВ забезпечення якості освітньої діяльності та вищої освіти, Харківський національний університет Повітряних Сил імені Івана Кожедуба, Харків, Україна;
Irina Khizhnyak – Candidate of Technical Sciences, Head of Scientific and Methodological Department for Quality Assurance of Educational Process and Higher Education, Ivan Kozhedub Kharkiv National Air Force University, Ukraine;
 e-mail: khizh_ia@ukr.net; ORCID Author ID: <https://orcid.org/0000-0003-3431-7631>;
 Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57196075236>.
- Яловега Микола Олексійович** – завідувач навчального кабінету Харківського національного університету Повітряних Сил імені Івана Кожедуба, Харків, Україна
Mykola Yalovega – Head of the Training Office of Ivan Kozhedub Kharkiv National Air Force University, Kharkiv, Ukraine;
 e-mail: yalovegamykola@gmail.com; ORCID Author ID: <https://orcid.org/0009-0000-7720-3111>.
- Бурков Віталій Валерійович** – старший викладач кафедри логістики підрозділів Національної академії Національної гвардії України, Харків, Україна
Vitalii Burkov – Senior Lecturer Department of Logistics Units of the National Academy National Guard of Ukraine, Kharkiv, Ukraine
 e-mail: burkovvitalii@gmail.com; ORCID Author ID: <https://orcid.org/0009-0000-6300-7830>.

Optimization the route of an unmanned aerial vehicle during the search for an aircraft that has crashed using an uniformly optimal search strategy

Illia Hridasov, Hennadii Khudov, Irina Khizhnyak, Mykola Yalovega, Vitalii Burkov

Abstract. Unmanned aerial vehicles (UAV) have been a significant advancement in SAR operations, ushering in an era of increased technological efficiency. The seamless integration of UAV into various tasks of the aviation search and rescue process highlights their high ability to perform tasks that were once burdensome or even impossible for traditional aircraft search methods, which human factors can negatively impact them. The primary objective of SAR operations remains the rapid and accurate location of a crashed aircraft. This is followed by such important actions as handover, information exchange, and timely delivery of various supplies to victims. All this is done within time constraints. The inclusion of UAV in SAR operations offers numerous notable advantages, including ease of deployment, cost-effective maintenance, exceptional mobility, and the ability to hover in areas where human intervention may be dangerous, limited, or require rapid decision making. Also, many UAV tasks can be performed automatically, such as image segmentation, which can help -locate victims in the surveyed area. The methodology for using UAV systems in SAR type search and rescue operations covers several key components, such as operation management, risk assessment, sensor technology, and UAV route selection. This article proposes a study to substantiate an optimization the route of a fixed-wing UAV during the search for an aircraft that has crashed, using uniformly optimal search strategies. The article analyzes existing search methods, highlighting their advantages and disadvantages. In the first stage of the study, the solution to the optimization problem for a uniformly optimal search strategy is presented. At the second stage, the authors developed a flowchart for a search and rescue operation algorithm utilizing the uniformly optimal search strategy. In the third stage, the results of applying the uniformly optimal search strategy to locate an aircraft in distress are illustrated, with implementation demonstrated using the Python programming language. The study discusses the potential application of the scientific findings. Additionally, possible directions for further research are suggested, focusing on modeling aviation accidents and calculating quantitative indicators for the search method examined in the scientific publication. These directions include employing fixed-wing UAV simulator as a search tool for the given task.

Keywords: aircraft; emergency situation; search and rescue operation; search; optimization; personnel recovery; route planning; uniformly optimal strategy; unmanned aerial vehicles.

O. Leha, V. Martovytskyi, I. Severin

Kharkiv National University of Radio Electronics, Kharkiv, Ukraine

ANALYSIS OF METHODS FOR DETECTING UNMANNED AERIAL VEHICLES

Abstract. In the modern era of increasing use of unmanned aerial vehicles (UAVs) in civilian, commercial, and military sectors, the issue of UAV detection has become critically important. Moreover, the development of artificial intelligence and machine learning algorithms opens new opportunities for improving identification accuracy and reducing false alarms. The use of advanced data processing methods allows detection systems to adapt to changing conditions and enhance their effectiveness in real-world environments. Therefore, research on UAV detection methods is relevant not only from a scientific perspective but also from a practical standpoint, as its findings can be applied to ensure security in urban areas, combat zones, and other critical fields. An unmanned aerial vehicle is an aircraft designed to operate without a pilot on board, with flight control carried out either by a pre-programmed system or via a remote control station located outside the aircraft. UAV detection is the process of locating and, in some cases, identifying an unmanned aerial vehicle. This article examines the primary UAV detection methods, analyzing their advantages and limitations. The study has demonstrated that none of the reviewed detection methods is universal. Each method has its own strengths and weaknesses, which can significantly impact the effectiveness of the system in real-world conditions. Based on the obtained results, promising directions for further research include: development of combined detection systems, application of deep learning methods, exploration of new physical principles for UAV detection. The results of this study lay the foundation for further advancement in UAV detection technologies, contributing to the development of more effective and adaptive security solutions.

Keywords: unmanned aerial vehicle, UAV, detection, RF Spectrum Analysis, Real-Time Drone Monitoring, UAV classification.

Introduction

In today's world, UAVs are increasingly being used across various areas of life, ranging from toys to the military sector. At the same time, the need for effective UAV detection methods is growing. Developing such methods is a challenging technical task, as drones often have small sizes, low visibility, and the ability to operate in complex conditions. This article provides a detailed analysis of the main technologies and solutions for UAV detection. The advantages and disadvantages of these methods are also described. Examining the strengths and weaknesses of each approach allows for a better understanding of their limitations and potential applications under different conditions.

Research Objective. The objective of this research is to analyze and systematize modern methods of UAV detection, taking into account their technical features, effectiveness, and limitations. The research aims to identify the advantages and disadvantages of each method, including radar, acoustic, optical, and radio frequency systems. The results of this research will contribute to building a knowledge base that will support the development of new, more effective technologies for UAV detection.

Research results

Four primary methods of UAV detection can be identified: radar, radio frequency, visual, and acoustic. Each of these methods includes several solutions, which will also be discussed further.

Radar Method. Radar monitoring is a widely used method for airspace surveillance. Radars use radio waves to detect objects in the air. They transmit signals that bounce off the object, and the receiver analyzes these reflected signals to determine the object's position, speed, and flight trajectory. This technology can be utilized to detect fully autonomous UAVs and those that neither send nor receive radio signals.

A Doppler radar can be used for this purpose. Doppler radar is a powerful technology that uses the Doppler effect to measure the velocity of moving objects. The micro-Doppler radar is an advanced version of Doppler radar, capable of detecting differences in speed and motion within objects. It is particularly suited for UAV detection as it can distinguish the propeller rotation of UAVs from, for example, the wing movements of birds [1]. Other radar technologies that can theoretically be applied for UAV detection include:

- over-the-horizon radar;
- ultra-wideband radar;
- millimeter-wave radar [2];
- chaotic monostatic radars;
- bistatic [3] and multistatic radars [4].

However, due to the low radar cross-section [5] and low speed of UAVs, their detection using these radars can be a challenging task [6].

Advantages:

- can detect UAVs over long distances (up to several kilometers);
- operate independently of weather conditions (rain, fog, snow);
- effective for monitoring large areas.

Disadvantages:

- difficulty detecting small UAVs due to their size and low radar reflectivity;
- high equipment costs and complexity of integration;
- possibility of false positives caused by other flying objects (birds, aircraft).

Radio Frequency Method. UAVs emitting radio frequency (RF) signals can be intercepted and analyzed to track and determine their location [7]. RF systems scan the spectrum to detect signals transmitted between the UAV and its operator. These systems can identify control and data transmission frequencies, determine the type of drone, and even locate the operator.

In many cases, manually controlled UAVs connect to a ground station and GNSS (Global Navigation Satellite System) for operation, enabling interception of signals and access to information such as coordinates and video feeds. RF monitoring has demonstrated the capability to detect UAVs within a range of up to 5000 meters.

Advantages:

- ability to detect UAVs regardless of weather conditions;
- can identify the location of both the UAV and the operator;
- effective over long distances (up to several kilometers).

Disadvantages:

- ineffective against autonomous UAVs that do not rely on radio communication;
- difficulty in distinguishing signals in dense RF environments;
- high equipment costs.

The process of UAV detection based on radio frequency (RF) typically involves analyzing RF signals to extract characteristic features, which are then compared with a database of UAV RF profiles for identification. In recent years, machine learning-based approaches have gained significant traction, profoundly influencing the development of RF-based UAV detection methods. To enhance the efficiency of these methods, researchers employ various approaches that can be categorized into two main groups.

The first category involves the use of signal processing techniques, such as Fourier Transform, to extract features, followed by classification using algorithms like SVM, decision trees, etc. [8,9]. The second category relies on minimal pre-processing of signals, utilizing neural networks to automatically extract features and identify UAV signals [10–14].

The following presents an overview of both traditional RF-based UAV detection methods and modern machine learning approaches.

Traditional Radio Frequency Detection

Radio frequency detection involves identifying control signals and video transmission signals exchanged between the ground station and the UAV. The primary principle of this approach is based on analyzing frequency, symbol rate, modulation type, channel bandwidth, and other parameters that create a unique "fingerprint" of the UAV signal.

Below are the main methodologies for extracting the spectral characteristics of UAVs:

1. Fractal Dimension [8]: This statistical indicator characterizes the complexity, roughness, and irregularity of a signal. Methods from time and phase domains are used to calculate the fractal dimension. Popular approaches include the correlation dimension, Hausdorff dimension [15], and Higuchi's algorithm [16], which demonstrates about 100% accuracy and effectiveness in identification.

2. Bispectrum: A powerful tool for analyzing nonlinear and non-Gaussian signals, it reduces the influence of white Gaussian noise [17]. Integrated bispectral methods, such as Axially Integrated Bispectra

(AIB) [18] and Square Integrated Bispectra (SIB) [19], achieve identification accuracies of 98% and 96%, respectively.

3. Signal Spectrum [9]: Frequency analysis of signals using Short-Time Fourier Transform (STFT) distinguishes UAV signals based on the characteristic vibrations caused by propeller rotation. The accuracy of this method is 97–98%.

4. Wavelet Energy Entropy (WEE) [9]: This approach analyzes sudden short-term signal variations associated with UAV movement in three-dimensional space or wind effects. The accuracy of this method is 93–94%.

5. Spectral Power Entropy (SPE) [9]: This metric describes the relationship between the power spectrum and entropy. UAV signals generally show higher SPE values compared to other signals. The accuracy of this method is 83–84%.

A comparison of these methods shows that fractal dimension, AIB, and SIB offer higher accuracy but are limited to identifying only known UAV types. Methods like signal spectrum, WEE, and SPE can detect unknown UAVs, although their accuracy is slightly lower. A combined application of these approaches can significantly improve identification efficiency.

RF-Based Detection Using Deep Learning

With the development of machine learning methods, researchers are actively exploring new ways to enhance the efficiency of signal feature extraction for more accurate detection and classification. Numerous studies have investigated a wide range of approaches for extracting signal features using deep learning algorithms. Notably, recent research [20, 21] demonstrates that neural networks are highly effective for tasks involving the detection and classification of radio frequency signals, highlighting the immense potential of deep learning in this field.

The primary advantage of deep learning methods lies in the ability of neural networks to automatically extract relevant features directly from raw signals, eliminating the need for complex preprocessing algorithms or signal reconstruction. For instance, the use of convolutional neural networks (CNNs) allows for the efficient extraction of features from compressed probing signals, bypassing the process of restoring the original signal. This significantly simplifies signal processing and reduces computational resource requirements.

Compared to traditional feature extraction methods, the deep learning-based approach demonstrates superior efficiency and accuracy. In certain experiments, the accuracy of UAV signal detection and classification reached 99%. Moreover, deep neural networks have proven adaptable, enabling them to work with various signal types and improve accuracy even in cases of noisy data.

However, despite their high effectiveness, the use of deep learning methods requires large volumes of training data and significant computational power. This remains one of the key challenges that researchers are addressing through neural network architecture optimization and the adoption of approaches for learning from small datasets.

As result, deep learning significantly expands the capabilities of UAV detection and classification based on radio frequency signals, offering more automated, accurate, and adaptive approaches to data processing.

Optical Method. Optical systems use cameras (conventional, infrared, or thermal imaging) for visual detection and tracking of UAVs. Modern systems are often equipped with computer vision algorithms for automatic object recognition. Using machine learning techniques, models can be trained to automatically extract appearance and motion features from UAV datasets, enabling UAV identification and tracking [22]. Additionally, infrared cameras can be used to detect and identify UAVs under low-light conditions [23]. By combining computer vision methods with infrared cameras, UAVs can be detected at different times of the day, both during daylight and nighttime.

However, visual identification methods face challenges. For instance, it is difficult to differentiate UAVs from birds, especially when UAVs operate at significant altitudes. In fact, detecting UAVs beyond 1000 meters becomes exceedingly difficult or nearly impossible.

Advantages:

- provide precise information about UAV location;
- can operate at night using infrared or thermal imaging cameras;
- suitable for operation within the line of sight.

Disadvantages:

- dependence on weather conditions (fog, rain, and snow reduce effectiveness);
- limited range (typically only a few hundred meters);
- require high resolution to identify small drones.

Object detection is a core task in computer vision, which involves locating and classifying specific types of visual objects (e.g., animals, humans, or vehicles) in images or videos. Thanks to development in computational models and methods, computer vision systems can address key questions: What object is present in the scene, and where is it located? This task serves as a foundation for many related tasks, such as instance segmentation, generating image captions, scene analysis, and object tracking in video streams.

From an application perspective, research in object detection is generally divided into two main directions:

1. General object detection – focuses on universal systems capable of detecting various object types within a unified framework, mimicking the human vision system's ability to recognize multiple objects.

2. Specialized object detection – targets the development of approaches for specific scenarios, such as pedestrian, face, or vehicle detection, or detecting objects in challenging conditions like nighttime or fast-moving environments.

Recent years have been revolutionary for this field due to rapid progress in deep learning methods. The use of neural networks has significantly improved the accuracy, speed, and adaptability of object detection systems. This has led to breakthroughs that outperform traditional methods in both performance and the ability

to process large volumes of data.

The following sections will present common object detection methods, including traditional approaches and deep learning-based techniques.

Traditional Object Detection.

Viola Jones Detector. The Viola Jones detector [24] is based on a simple but effective object detection approach known as "sliding windows." In this method, each image is scanned through different positions and scales of a window to determine whether a face is present. Although this approach seems simple, at the time of its development, the computational resources of computers were insufficient for the efficient execution of such operations. To address this issue, the Viola Jones detector integrates three key techniques:

- Integral Image – for fast computation of pixel sums in rectangular regions;
- Feature Selection – for the effective selection of relevant features;
- Detection Cascade – for step-by-step rejection of negative windows and focusing on potentially relevant areas.

These improvements significantly increased the speed and performance of detection, making the method popular.

Histogram of Oriented Gradients (HOG) Detector. The Histogram of Oriented Gradients (HOG) method [25] represented a significant step forward compared to previous approaches, such as Scale-Invariant Feature Transform (SIFT) and contextual shape analysis. The main goal of HOG is to compute image features on a dense grid of uniformly spaced cells. This approach provides a balance between feature invariance (translation, scale, lighting) and the ability to distinguish between object classes.

To improve accuracy, the method uses local contrast normalization in overlapping "blocks," making the analysis more robust to lighting changes. Although HOG can be used for detecting objects of various classes, its primary goal is pedestrian detection.

The detection process in HOG involves resizing the input image while keeping the detection window size fixed. This allows for successful various sizes objects recognition.

Thanks to its efficiency and versatility, HOG has become a key component in many computer vision tasks. Over the years, this method has been widely applied in fields such as security (video surveillance) and the automotive industry (pedestrian detection).

Object Detection Based on Deep Learning. RCNN uses a selective search approach to extract a set of candidate regions from an image [26]. Then for each region uses a convolutional neural network (CNN), pre-trained on ImageNet, which scales the region to a fixed size (e.g., as in the AlexNet model) for feature extraction. The final stage involves classifying these regions using a linear SVM classifier to determine the presence of objects and their class. The RCNN method significantly improved the average precision (mAP) of detection, increasing it by 8.5%.

SPPNet, introduced by K. He in 2014 [27], addresses the issue of fixed-size input data, which is

typical of CNNs. The key innovation in this architecture is the Spatial Pyramid Pooling (SPP) layer, which allows generation of fixed-length features for images of any size. This eliminates the need to resize images or regions of interest, as feature computation is performed only once for the entire image.

The SPPNet approach greatly accelerates the object detection process, reducing processing time by 20 times compared to RCNN, without significant loss in accuracy.

In 2015, R. Girshick proposed the Fast RCNN method [28], which improved the RCNN and SPPNet architectures. Using the Fast RCNN detector, both the bounding box regressor and object detector are trained with the same network configuration, which is an improvement over the previous two-stage approach used in RCNN and SPPNet. The streamlined Fast RCNN architecture significantly accelerated processing, making it more than 200 times faster than RCNN, while saving high detection accuracy. This method became a key step towards efficient object detection, providing a balance between performance and accuracy.

In 2015, S. Ren introduced Faster RCNN [29], which was a further improvement over Fast RCNN. It was the first detector to demonstrate end-to-end deep learning capable of almost real-time performance. Faster RCNN integrated detection proposals, feature extraction, and bounding box regression into a single architecture. Due to its high speed and accuracy, this method became widely used in object detection tasks.

YOLO, proposed by R. Joseph in 2015 [30], became a revolutionary single-stage approach to object detection. Its main advantage is extremely high speed. For example, the fast version of YOLO achieves a performance of 155 frames per second. Unlike traditional methods, which used the "object detection + classification" approach, YOLO applies a single neural network to the entire image. The image is divided into a grid of cells, each predicting its own bounding boxes and probabilities. Despite the high speed, YOLO's accuracy decreases when localizing small objects compared to two-stage detectors. Nevertheless, YOLO is one of the most popular algorithms for UAV detection due to its speed and acceptable accuracy. In recent years, the algorithm has become the basis for many studies focused on UAV detection [31, 32].

SSD, developed by W. Liu in 2015 [33], became the second single-stage detector in deep learning. This model introduced a multi-scale approach and detection with varying resolutions, significantly improving accuracy, especially for small objects. SSD performs object detection at five different scales at different network levels, which distinguishes it from previous methods that worked only with higher levels. SSD outperforms many previous detectors in both speed and accuracy.

RetinaNet, proposed by Lin in 2018 [34], is a single-stage detector that addresses the problem of class imbalance between positive and negative examples, which is typical for single-stage models. Unlike two-stage methods that achieve high accuracy through dense region proposal definition, single-stage approaches face difficulties in training due to the large number of negative examples. To overcome this issue, RetinaNet introduced

a new loss function — the focal loss. This function modifies the standard cross-entropy loss to enable the detector to focus more on hard-to-classify examples during training, reducing the impact of well-classified or easy examples. The use of focal loss significantly improves RetinaNet's accuracy while maintaining the speed typical of single-stage detectors.

Acoustic Detection. Acoustic systems use microphones and sound processing algorithms to detect specific noises created by the engines and propellers of UAVs. Sound waves generate a unique "audio fingerprint" for each UAV, enabling individual identification. However, detecting these sound waves can be challenging in practice due to factors such as environmental noise and sound wave attenuation. Machine learning models have been applied to improve the effectiveness of this approach. By training these models on datasets of acoustic signatures, they can learn to distinguish UAV engine or propeller sounds from noise and other interference, improving accuracy and reliability [35]. Nonetheless, acoustic monitoring is highly sensitive to background noise and has a limited detection range. Advantages:

- easily detects drones at low altitudes;
- relatively inexpensive and simple to use;
- operates in real-time.
- Disadvantages:
- efficiency decreases in noisy environments (wind, urban noise);
- limited detection range (up to 500 meters).

Acoustic detection of UAVs is another popular technical solution for their identification. Despite differences between sound and electromagnetic waves, audio detection methods share similarities with radio frequency methods, as both are based on wave characteristic analysis.

Sound-based detection methods can be divided into two main categories: traditional acoustic detection and acoustic detection using deep learning.

Traditional Acoustic Detection

This approach involves processing audio signals using methods such as spectral analysis, short-time Fourier transform (STFT), wavelet transform, and other sound processing algorithms. The main goal is to extract characteristic features such as frequencies, amplitudes, or harmonics that may be specific to UAV engines or propellers. Traditional methods are less resource-intensive; however, their accuracy and robustness to external noise are often limited. They work well in conditions where the acoustic signal is clearly separated from background noise. Cepstral analysis is one of the most common methods for processing audio signals, based on the inverse Fourier transform of the logarithmic power spectrum. This technique allows the extraction of unique features inherent to audio signals, which is particularly useful for UAV detection tasks. Common cepstral coefficients include linear predictive cepstral coefficients (LPCC) and Mel-frequency cepstral coefficients (MFCC) [36].

The Mel scale is a nonlinear frequency scale that models the human ear's perception of sound. It does not have a linear correspondence to frequencies in Hertz,

making it better suited for analyzing acoustic characteristics. MFCC are spectral features calculated using the Mel scale. This method has found widespread use in speech recognition tasks and is also employed in UAV detection due to its ability to extract unique acoustic features characteristic of propeller or engine sounds [37,3 8–41]. The process of calculating MFCC includes the following steps:

- audio signal segmentation into frames: The signal is broken into short time intervals;
- application of window functions: A window function is applied to each frame to reduce spectral leakage;
- fast Fourier transform (FFT): The linear spectrum of each frame is computed;
- Mel filterbank processing: A set of filters distributed according to the Mel scale is applied;
- logarithmic spectrum calculation: The filtered spectrum is logarithmically transformed to obtain a spectrogram;
- discrete cosine transform (DCT): The spectrogram is transformed, retaining the required number of coefficients to form MFCC.

Despite MFCC's limitations in the high-frequency range (due to reduced accuracy at higher frequencies), this method demonstrates high efficiency in UAV detection tasks, achieving an accuracy of over 97%.

Linear Predictive Cepstral Coefficients (LPCC)

LPCC is a variant of MFCC and is used for audio signal analysis based on linear prediction methods [42]. In the calculation of LPCC, linear predictive filters are used instead of Mel filters, enabling more accurate identification of resonance peaks (formants) of the signals. This method has certain advantages over MFCC, especially in the presence of Gaussian white noise, as LPCC better highlights important signal characteristics. However, the high computational complexity of LP filters is a disadvantage, requiring a balanced approach in practical applications. LPCC is usually used in combination with MFCC, and the fusion of these two features is often employed to improve UAV detection performance [39]. The combination of MFCC and LPCC helps compensate for the weaknesses of each method, enhancing overall UAV identification accuracy. This combined approach ensures robustness to external noise and high performance even in challenging acoustic conditions.

Acoustic Detection Using Deep Learning. This modern approach utilizes deep learning algorithms, specifically Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN), to automatically extract features and classify audio signals. By using raw

audio data, neural networks can learn to identify different sound characteristics such as frequency spectrum, propeller or engine noise, even in complex background noise conditions. These methods demonstrate higher robustness and accuracy, achieving 95-98% accuracy depending on the quality of the training data. Acoustic detection of UAVs has two primary research directions. The first focuses on precise and fast separation of sound sources and identification of UAV acoustic signals in noisy environments. The second focuses on detecting UAVs in time-varying scenarios, which is important due to the rapid movement of UAVs and the need to track their acoustic signatures over time [43].

Although acoustic detection is less commonly used in commercial and military counter-UAV systems due to its limited detection range, this approach still holds significant potential for application in urban environments, such as residential areas, schools, and business districts. By using unique UAV acoustic signatures, this approach can help identify and track UAVs in real-time.

In addition to traditional signal processing methods, learning-based methods are also widely used for feature extraction in sound-based UAV detection. Recent studies [37, 44] have shown that deep learning models can also be employed to extract features from UAV acoustic signals for detection purposes.

Conclusions

Research on UAV detection methods demonstrates that each of the approaches examined has its advantages and limitations. Radar systems provide long-range detection and effectiveness in challenging weather conditions, but struggle with detecting small UAVs. Acoustic sensors are affordable and easy to use, but their effectiveness significantly decreases in noisy environments. Optical systems offer high visual tracking accuracy but rely on weather conditions and line of sight. Radio frequency systems not only detect UAVs but also locate their operators; however, they are ineffective against autonomous devices.

The optimal solution is the integration of these approaches into combined systems, which help mitigate the drawbacks of each method. Such systems provide high accuracy, reliability, and versatility in various conditions, which is crucial for effective UAV detection.

The results of this research lay the foundation for further advancements in UAV detection technologies. The development of integrated systems using the latest data processing algorithms and artificial intelligence will significantly improve the efficiency of these systems.

REFERENCES

1. Chamola, V.; Kotes, P.; Agarwal, A.; Naren; Gupta, N.; Guizani, M. A Comprehensive Review of Unmanned Aerial Vehicle Attacks and Neutralization Techniques. *Ad Hoc Netw.* 2021, 111, 102324. doi: <https://doi.org/10.1016/j.adhoc.2020.102324>
2. Li, S.; Chai, Y.; Guo, M.; Liu, Y. Research on Detection Method of UAV Based on micro-Doppler Effect. In *Proceedings of the 2020 39th Chinese Control Conference (CCC)*, Shenyang, China, 27–29 July 2020; pp. 3118–3122. doi: [10.23919/CCC50068.2020.9189414](https://doi.org/10.23919/CCC50068.2020.9189414).
3. Pappu, C.S.; Beal, A.N.; Flores, B.C. Chaos based frequency modulation for joint monostatic and bistatic radar-communication systems. *Remote Sens.* 2021, 13, 4113. doi: <https://doi.org/10.3390/rs13204113>
4. Abd, M.H.; Al-Suhail, G.A.; Tahir, F.R.; Ali Ali, A.M.; Abbood, H.A.; Dashtipour, K.; Jamal, S.S.; Ahmad, J. Synchronization of monostatic radar using a time-delayed chaos-based FM waveform. *Remote Sens.* 2022, 14, 1984. doi: <https://doi.org/10.3390/rs14091984>

5. Mehta, S.; Rastegari, M.; Caspi, A.; Shapiro, L.; Hajishirzi, H. Espnet: Efficient spatial pyramid of dilated convolutions for semantic segmentation. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018, pp. 552–568. doi: <https://doi.org/10.48550/arXiv.1803.06815>
6. Shi, X.; Yang, C.; Xie, W.; Liang, C.; Shi, Z.; Chen, J. Anti-Drone System with Multiple Surveillance Technologies: Architecture, Implementation, and Challenges. *IEEE Commun. Mag.* 2018, 56, 68–74. doi: <https://doi.org/10.1109/MCOM.2018.1700430>
7. Zhang, J.; Liu, M.; Zhao, N.; Chen, Y.; Yang, Q.; Ding, Z. Spectrum and energy efficient multi-antenna spectrum sensing for green UAV communication. *Digital Commun. Netw.* 2022, 9, 846–855. doi: <https://doi.org/10.1016/j.dcan.2022.09.017>
8. Nie, W.; Han, Z.C.; Zhou, M.; Xie, L.B.; Jiang, Q. UAV detection and identification based on WiFi signal and RF fingerprint. *IEEE Sensors J.* 2021, 21, 13540–13550. doi: <https://doi.org/10.1109/JSEN.2021.3068444>
9. Nie, W.; Han, Z.C.; Li, Y.; He, W.; Xie, L.B.; Yang, X.L.; Zhou, M. UAV detection and localization based on multi-dimensional signal features. *IEEE Sensors J.* 2021, 22, 5150–5162. doi: <https://doi.org/10.1109/JSEN.2021.3105229>
10. Mo, Y.; Huang, J.; Qian, G. UAV Tracking by Identification Using Deep Convolutional Neural Network. In Proceedings of the 2022 IEEE 8th International Conference on Computer and Communications (ICCC), Chengdu, China, 9–12 December 2022; pp. 1887–1892. doi: <https://doi.org/10.1109/ICCC56324.2022.10065721>
11. Swinney, C.J.; Woods, J.C. Low-Cost Raspberry-Pi-Based UAS Detection and Classification System Using Machine Learning. *Aerospace* 2022, 9, 738. doi: <https://doi.org/10.3390/aerospace9120738>
12. Lu, S.; Wang, W.; Zhang, M.; Li, B.; Han, Y.; Sun, D. Detect the Video Recording Act of UAV through Spectrum Recognition. In Proceedings of the 2022 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA), Dalian, China, 24–26 June 2022; pp. 559–564. doi: <https://doi.org/10.1109/ICAICA54878.2022.9844471>
13. He, Z.; Huang, J.; Qian, G. UAV Detection and Identification Based on Radio Frequency Using Transfer Learning. In Proceedings of the 2022 IEEE 8th International Conference on Computer and Communications (ICCC), Virtual, 9–12 December 2022; pp. 1812–1817. doi: <https://doi.org/10.1109/ICCC56324.2022.10065628>
14. Li, T.; Hong, Z.; Cai, Q.; Yu, L.; Wen, Z.; Yang, R. Bissiam: Bispectrum siamese network based contrastive learning for uav anomaly detection. *IEEE Trans. Knowl. Data Eng.* 2021. doi: <https://doi.org/10.1109/TKDE.2021.3118727>
15. Schweinhart, B. Persistent homology and the upper box dimension. *Discret. Comput. Geom.* 2021, 65, 331–364. doi: <https://doi.org/10.1007/s00454-019-00145-3>
16. Higuchi, T. Approach to an irregular time series on the basis of the fractal theory. *Phys. D Nonlinear Phenom.* 1988, 31, 277–283. doi: [https://doi.org/10.1016/0167-2789\(88\)90081-4](https://doi.org/10.1016/0167-2789(88)90081-4)
17. Zhang, X.D.; Shi, Y.; Bao, Z. A new feature vector using selected bispectra for signal classification with application in radar target recognition. *IEEE Trans. Signal Process.* 2001, 49, 1875–1885. doi: <https://doi.org/10.1109/78.942617>
18. Tugnait, J.K. Detection of non-Gaussian signals using integrated polyspectrum. *IEEE Trans. Signal Process.* 1994, 42, 3137–3149. doi: <https://doi.org/10.1109/78.330373>
19. Yao, Y.; Yu, L.; Chen, Y. Specific Emitter Identification Based on Square Integral Bispectrum Features. In Proc. of the 2020 IEEE 20th Int. Conf. on Comm. Techn., Nanning, China, 2020; pp. 1311–1314. doi: <https://doi.org/10.1109/ICCT50939.2020.9295681>
20. Al-Sa'd, M.F.; Al-Ali, A.; Mohamed, A.; Khattab, T.; Erbad, A. RF-based drone detection and identification using deep learning approaches: An initiative towards a large open source drone database. *Future Gener. Comput. Syst.* 2019, 100, 86–97. doi: <https://doi.org/10.1016/j.future.2019.05.007>
21. Mo, Y.; Huang, J.; Qian, G. Deep Learning Approach to UAV Detection and Classification by Using Compressively Sensed RF Signal. *Sensors* 2022, 22, 3072. doi: <https://doi.org/10.3390/s22083072>
22. Unlu, E.; Zenou, E.; Riviere, N.; Dupouy, P.E. Deep learning-based strategies for the detection and tracking of drones using several cameras. *IPSN Trans. Comput. Vis. Appl.* 2019, 11, 7. doi: <https://doi.org/10.1186/s41074-019-0059-x>
23. Fang, H.; Xia, M.; Zhou, G.; Chang, Y.; Yan, L. Infrared small UAV target detection based on residual image prediction via global and local dilated residual networks. *IEEE Geosci. Remote. Sens. Lett.* 2021, 19, 7002305.
24. Viola, P.; Jones, M.J. Robust real-time face detection. *Int. J. Comput. Vis.* 2004, 57, 137–154. doi: <https://doi.org/10.1109/LGRS.2021.3085495>
25. Dalal, N.; Triggs, B. Histograms of oriented gradients for human detection. In Proceedings of the CVPR'05, San Diego, CA, USA, 20–26 June 2005; Volume 1, pp. 886–893. doi: <https://doi.org/10.1109/CVPR.2005.177>
26. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the CVPR, Columbus, OH, USA, 23–28 June 2014; pp. 580–587. url: <https://clgiles.ist.psu.edu/IST597/materials/papers-cnn/rcnn.pdf>
27. He, K.; Zhang, X.; Ren, S.; Sun, J. Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* 2015, 37, 1904–1916. doi: <https://doi.org/10.1109/TPAMI.2015.2389824>
28. Girshick, R. Fast r-cnn. In Proceedings of the ICCV, Santiago, Chile, 13–16 December 2015; pp. 1440–1448.
29. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-CNN: Towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* 2017, 39, 1137–1149. doi: <https://doi.org/10.48550/arXiv.1506.01497>
30. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the CVPR, Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788. url: https://openaccess.thecvf.com/content_cvpr_2016/html/Redmon_You_Only_Look_CVPR_2016_paper.html
31. Ajakwe, S.O.; Ihekoronye, V.U.; Kim, D.S.; Lee, J.M. DRONET: Multi-Tasking Framework for Real-Time Industrial Facility Aerial Surveillance and Safety. *Drones* 2022, 6, 46. doi: <https://doi.org/10.3390/drones6020046>
32. Wang, J.; Hongjun, W.; Liu, J.; Zhou, R.; Chen, C.; Liu, C. Fast and Accurate Detection of UAV Objects Based on Mobile-Yolo Network. In Proceedings of the 2022 14th International Conference on Wireless Communications and Signal Processing (WCSP), Virtually, 1–3 November 2022; pp. 1–5. doi: <https://doi.org/10.1109/WCSP55476.2022.10039216>
33. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. Ssd: Single shot multibox detector. In Proceedings of the ECCV, Amsterdam, The Netherlands, 11–14 October 2016; pp. 21–37. doi: https://doi.org/10.1007/978-3-319-46448-0_2
34. Lin, T.Y.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal loss for dense object detection. In Proceedings of the ICCV, Venice, Italy, 22–29 October 2017; pp. 2980–2988. doi: <https://doi.org/10.48550/arXiv.1708.02002>

35. Al-Emadi, S.; Al-Ali, A.; Mohammad, A.; Al-Ali, A. Audio based drone detection and identification using deep learning. In Proceedings of the 2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC), Tangier, Morocco, 24–28 June 2019; pp. 459–464. doi: <https://doi.org/10.1109/IWCMC.2019.8766732>
36. Yan, X.; Fu, T.; Lin, H.; Xuan, F.; Huang, Y.; Cao, Y.; Hu, H.; Liu, P. UAV Detection and Tracking in Urban Environments Using Passive Sensors: A Survey. Appl. Sci. 2023, 13, 11320. doi: <https://doi.org/10.3390/app132011320>
37. Dong, Q.; Liu, Y.; Liu, X. Drone sound detection system based on feature result-level fusion using deep learning. Multimed. Tools Appl. 2023, 82, 149–171. doi: <https://doi.org/10.1007/s11042-022-12964-3>
38. Svanström, F.; Alonso-Fernandez, F.; Englund, C. Drone Detection and Tracking in Real-Time by Fusion of Different Sensing Modalities. Drones 2022, 6, 317. doi: <https://doi.org/10.3390/drones6110317>
39. Uddin, Z.; Qamar, A.; Alharbi, A.G.; Orakzai, F.A.; Ahmad, A. Detection of Multiple Drones in a Time-Varying Scenario Using Acoustic Signals. Sustainability 2022, 14, 4041. doi: <https://doi.org/10.3390/su14074041>
40. Jamil, S.; Rahman, M.; Ullah, A.; Badnava, S.; Forsat, M.; Mirjavadi, S.S. Malicious UAV detection using integrated audio and visual features for public safety applications. Sensors 2020, 20, 3923. doi: <https://doi.org/10.3390/s20143923>
41. Guo, J.; Ahmad, I.; Chang, K. Classification, positioning, and tracking of drones by HMM using acoustic circular microphone array beamforming. EURASIP J. Wirel. Commun. Netw. 2020, 2020, 1–19. doi: <https://doi.org/10.1186/s13638-019-1632-9>
42. Gupta, H.; Gupta, D. LPC and LPCC method of feature extraction in Speech Recognition System. In Proceedings of the 2016 6th International Conference-Cloud System and Big Data Engineering (Confluence), Noida, India, 14–15 January 2016; pp. 498–502. doi: <https://doi.org/10.1109/CONFLUENCE.2016.7508171>
43. Uddin, Z.; Altaf, M.; Bilal, M.; Nkenyereye, L.; Bashir, A.K. Amateur Drones Detection: A machine learning approach utilizing the acoustic signals in the presence of strong interference. Comput. Commun. 2020, 154, 236–245. doi: <https://doi.org/10.1016/j.comcom.2020.02.065>
44. Aydın, I.; Kızılay, E. Development of a new Light-Weight Convolutional Neural Network for acoustic-based amateur drone detection. Appl. Acoust. 2022, 193, 108773. doi: <https://doi.org/10.1016/j.apacoust.2022.108773>

Received (Надійшла) 12.02.2025

Accepted for publication (Прийнята до друку) 14.05.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Лега Олексій Сергійович – аспірант кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Oleksii Leha – PhD student, Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: oleksii.leha@nure.ua; ORCID Author ID: <https://orcid.org/0009-0004-6417-6491>.

Мартовицький Віталій Олександрович – кандидат технічних наук, доцент, доцент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Vitalii Martovytskyi – candidate of technical sciences, Associate Professor, Associate Professor of Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: vitalii.martovytskyi@nure.ua; ORCID Author ID: <https://orcid.org/0000-0003-2349-0578>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57196940070>.

Северін Ігор Сергійович – студент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Igor Severin – student, Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: igor.severin@nure.ua; ORCID Author ID <https://orcid.org/0009-0005-2397-9631>.

Аналіз методів виявлення безпілотних літальних апаратів

О. С. Лега, В. О. Мартовицький, І. С. Северін

Анотація. У сучасних умовах зростаючого використання безпілотних літальних апаратів (БПЛА) у цивільних, комерційних та військових сферах питання їх виявлення набуває критичної важливості. Крім того, із розвитком штучного інтелекту та алгоритмів машинного навчання відкриваються нові можливості для підвищення точності ідентифікації та зменшення кількості хибних спрацювань. Використання сучасних методів обробки даних дозволяє адаптувати системи виявлення до змінних умов та підвищити їхню ефективність у реальному середовищі. Таким чином, дослідження методів виявлення БПЛА є актуальним не лише з наукової точки зору, а й з практичної, оскільки його результати можуть бути застосовані для забезпечення безпеки в містах, у зонах бойових дій та в інших важливих сферах. Безпілотний літальний апарат — повітряне судно, призначене для виконання польоту без пілота на борту, керування польотом якого здійснюються відповідною програмою або за допомогою спеціальної станції керування, що знаходиться поза повітряним судном. Виявлення БПЛА – процес під час якого йде знаходження і, в деяких випадках, ідентифікація безпілотного літального апарата. У статті розглянуто основні методи виявлення безпілотних літальних апаратів. Проаналізовано переваги та недоліки таких методів. Проведене дослідження продемонструвало, що жоден із розглянутих методів виявлення БПЛА не є універсальним. Кожен з них має як переваги, так і обмеження, які можуть суттєво впливати на ефективність системи в реальних умовах. З огляду на отримані результати, перспективними напрямками подальших досліджень є: розвиток комбінованих систем виявлення, застосування методів глибокого навчання, дослідження нових фізичних принципів виявлення. Результати цього дослідження формують основу для подальшого вдосконалення технологій виявлення БПЛА.

Ключові слова: безпілотний літальний апарат, БПЛА, виявлення, аналіз радіочастотного спектра, моніторинг дронів у реальному часі, класифікація БПЛА.

Автомобільний, річковий, морський та авіаційний транспорт

УДК 629.5.017.2:004.942

doi: 10.26906/SUNZ.2025.2.20-26

П. В. Никитюк, О. М. Мельник

Одеський національний морський університет, Одеса, Україна

РОЗРАХУНКОВА МОДЕЛЬ ОЦІНКИ ЕКСПЛУАТАЦІЙНОЇ БЕЗПЕКИ СУДНА З УРАХУВАННЯМ ДЕГРАДАЦІЇ БАР'ЄРІВ ТА ПРОГНОЗУ РИЗИКУ

Анотація. Актуальність. В умовах зростання складності технічного забезпечення морських суден та впровадження елементів автономного управління, зростає потреба у створенні моделей, здатних адаптивно реагувати на зміну технічного стану підсистем і зовнішніх загроз. Ключовим елементом у системах забезпечення безпеки виступає узагальнений індекс ризику, який повинен динамічно змінюватися в залежності від технічної деградації та комбінованого впливу середовища. **Об'єкт дослідження:** інтегрована модель оцінки експлуатаційної безпеки судна. **Мета статті:** розробка методу розрахункової оцінки експлуатаційної безпеки судна з урахуванням деградації бар'єрів безпеки, технічного стану підсистем і прогнозу динаміки ризику. **Результати дослідження.** У статті запропоновано багаторівневу ймовірнісну модель, яка дозволяє розраховувати ризики окремих підсистем, оцінювати ефективність захисних бар'єрів, а також формувати інтегрований індекс безпеки (SIRI). Модель доповнена блоком прогнозування часу до критичного стану та алгоритмом активації захисного реагування. Проведено числове моделювання шести підсистем судна для восьми типових сценаріїв експлуатації. **Висновки.** Реалізація моделі дозволяє забезпечити раннє виявлення ризикових ситуацій, скорочення часу реагування та підвищення інформативності систем прийняття рішень. Модель може бути використана як ядро цифрових платформ управління безпекою морських суден. Сфера використання отриманих результатів: інтелектуальні системи управління безпекою морських суден, автономне судноплавство, автоматизовані навігаційні платформи.

Ключові слова: експлуатаційна безпека судна; інтегрована модель ризику; деградація бар'єрів; технічний стан підсистем; індекс ризику SIRI; управління морською безпекою; цифрове прогнозування; ймовірнісна модель; оцінка запасу безпеки; прогноз часу реагування; багатофакторний аналіз; сценарне моделювання; критичні стани; автоматизоване реагування; інформаційна підтримка рішень.

Вступ

Постановка проблеми. У сучасних умовах інтенсифікації морських перевезень, автоматизації керування суднами та ускладнення морських маршрутів зростає потреба у створенні систем, здатних не лише фіксувати порушення безпеки, але й прогнозувати розвиток аварійних ситуацій. Існуючі підходи до оцінки безпеки здебільшого базуються на статичних або детермінованих моделях, що не дозволяють врахувати динаміку деградації технічних бар'єрів, зміну зовнішніх факторів та багатофакторну взаємодію ризиків. У зв'язку з цим актуальним є розроблення інтегрованої розрахункової моделі, яка дозволяє кількісно оцінити рівень експлуатаційної безпеки судна з урахуванням ймовірнісних характеристик підсистем, ефективності бар'єрів та часу до настання критичних умов.

Недостатність сучасних моделей експлуатаційної безпеки морських суден полягає у їхній обмеженості щодо врахування часової динаміки ризиків, кумулятивного ефекту їхнього накопичення та взаємодії множинних факторів технічного характеру і середовища. У зв'язку з цим постає потреба в такій моделі, яка дозволяє одночасно враховувати ймовірність виникнення ризиків у кожній підсистемі, оцінювати вплив деградації технічних і інформаційних бар'єрів на загальний стан системи, відстежувати динаміку зміни ризику в часі та інтегрувати отримані результати у вигляді єди-

ного узагальненого індексу ризику для підтримки прийняття оперативних рішень.

Аналіз останніх досліджень і публікацій. Сучасні дослідження в галузі морської безпеки та управління експлуатаційним ризиком активно фокусуються на розробці адаптивних моделей аналізу ризиків, які враховують технічні, інформаційні та чинники середовища. Зокрема, в роботі [1] представлено фреймворк для оцінки операційного ризику при навігації автономного судна в обмежених акваторіях, що ілюструє актуальність теми багатофакторного впливу в умовах динамічного середовища. У дослідженні [2] запропоновано структурований підхід до аналізу системного ризику автономної навігації на основі STPA і параметрів допустимості операційного середовища. Подібні завдання прогнозування ризиків також досліджуються в [3], де реалізовано модель на базі машинного навчання для раннього виявлення аварій у перевезеннях небезпечних вантажів.

Технічні аспекти суднових енергосистем, зокрема електрифікація та гібридні рішення, розглядаються в [4-6], що прямо впливає на структуру потенційних ризиків і бар'єрних механізмів безпеки. Інші автори [7, 8, 9] приділяють увагу локальним сценаріям ризику, включаючи прогнозування затоплень і безпеку переробки суден. Значну увагу привертають підходи до формалізації безпеки: FRAM-аналіз бар'єрів [10], бібліометричні огляди [11] та впровадження цифрових двійників для врахування

людського фактору [13]. У вітчизняному та міжнародному контексті [15-17] висвітлено методи оцінювання експертної якості управління ризиком та побудови моделей проектного управління на базі людського елементу. Усі ці підходи створюють передумови для формування цілісних інтегрованих моделей, подібних до запропонованої у даній роботі.

Метою роботи Метою дослідження є створення розрахункової моделі оцінки експлуатаційної безпеки судна, яка дозволяє прогнозувати ймовірність критичних станів на основі деградації бар'єрів, технічного стану та зовнішніх умов. Завдання дослідження включають формулювання математичної моделі на основі ймовірнісного підходу та побудову функції деградації ефективності бар'єрів. Проведення розрахунків агрегованих показників безпеки підсистем та судна загалом, розроблення алгоритму динамічного моніторингу та прогнозування та проведення моделювання та оцінити ефективність запропонованого підходу.

Основний матеріал

Розрахункова модель оцінки експлуатаційної безпеки судна. Оцінка рівня експлуатаційної безпеки судна вимагає врахування множини взаємозалежних чинників, серед яких ключову роль відіграють технічний стан підсистем, ефективність захисних бар'єрів і вплив зовнішніх умов середовища. У цьому контексті виникає потреба в побудові гнучкої та адаптивної математичної моделі, здатної відображати як поточний стан системи, так і його прогнозовану динаміку. Запропонована модель базується на ймовірнісному підході, де кожен підсистемний ризик розглядається як функція відмови компонента та часткової втрати ефективності відповідного бар'єра. Схема, подана на рис. 1, ілюструє логіку роботи інтегрованої системи оцінки та прогнозування ризику на основі запропонованої розрахункової моделі.

Робота системи починається з моніторингу поточного стану судна в реальному часі, включаючи технічний стан підсистем і зовнішні фактори. Паралельно проводиться оцінка ефективності бар'єрів безпеки, що дозволяє визначити рівень ризику у кожній підсистемі. На наступному етапі ці дані агрегуються, формуючи загальний індекс ризику, який доповнюється прогнозом ризику за допомогою неймережевої моделі LSTM.

Система ухвалення рішень створює адаптивні рекомендації для зміни режиму руху судна, його маршруту або швидкості. Вона також регулярно оновлює вагові коефіцієнти моделі, спираючись на отриманий зворотний зв'язок.

У разі, якщо інтегрований індекс ризику перевищує критичний поріг, активується протокол реагування, що передбачає фіксацію інциденту, переведення судна в захищений режим і ініціацію дій екіпажу. Якщо ж ризик залишається допустимим, система продовжує безперервне спостереження та оновлення оцінок.

Особливістю моделі є врахування часової деградації бар'єрів, що моделюється за експоненційним законом з урахуванням коефіцієнта зносу.

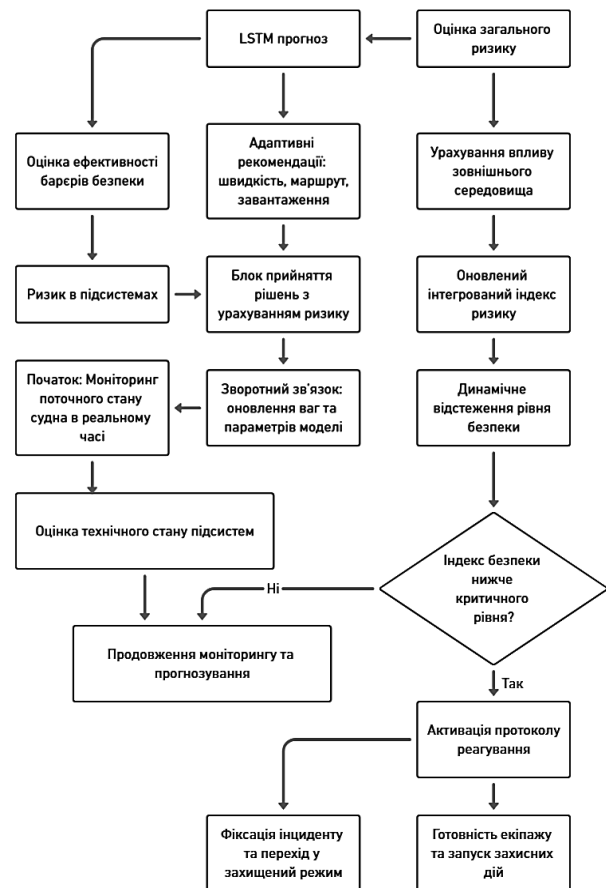


Рис. 1. Функціональна схема моделі управління експлуатаційною безпекою судна

Особливістю моделі є врахування часової деградації бар'єрів, що моделюється за експоненційним законом з урахуванням коефіцієнта зносу. Це дозволяє оцінити ризик не лише як статичну величину, а як динамічний індикатор, що змінюється в залежності від часу та сценарію експлуатації. У результаті така модель дозволяє не лише ідентифікувати потенційно небезпечні стани, але й прогнозувати час до досягнення критичного рівня безпеки та ініціювати відповідну реакцію системи (рис. 2).

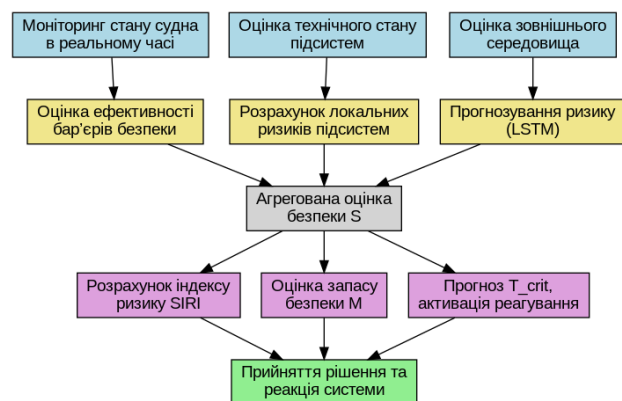


Рис. 2. Етапи розрахункової оцінки експлуатаційної безпеки судна

В основі моделі лежить припущення, що рівень експлуатаційної безпеки судна формується як результат взаємодії зовнішніх ризиків F , технічного

стану критичних підсистем C та ефективності запобіжних бар'єрів B . Для врахування комплексного впливу цих факторів використовується багаторівнева ймовірнісна модель з модулем деградації.

Умовна ймовірність аварійної події буде:

$$P(R) = \sum_{i=1}^n P(R|F_i, C_i, B_i) \cdot P(F_i) \cdot P(C_i) \cdot (1 - E_{B_i}(t)). \quad (1)$$

Такий підхід дозволяє враховувати вплив кожної компоненти на загальну ймовірність аварійного розвитку подій.

Модель деградації бар'єра безпеки:

$$E_{B_i}(t) = E_{01} \cdot \exp(-\lambda_i t), \quad (2)$$

де λ_i – індивідуальний коефіцієнт деградації, що відображає технічні особливості системи та умови експлуатації.

Ймовірність повної неефективності бар'єра:

$$P(B_i^{fail}) = 1 - E_{B_i}(t). \quad (3)$$

Підсистемна функція ризику:

$$R_j(t) = \alpha_j \cdot P(C_j)^2 + \beta_j \cdot (1 - E_{B_j}(t)), \quad (4)$$

де α_j β_j – вагові параметри залежності ризику від технічного стану і бар'єра, можливі варіації.

Агрегована оцінка ризику судна:

$$S(t) = \sum_{j=1}^m w_j \cdot R_j(t), \quad (5)$$

де w_j – коефіцієнт критичності кожної підсистеми.

Інтегрований індекс ризику судна ($SIRI$):

$$SIRI(t) = S(t) / \sum_{j=1}^m w_j. \quad (6)$$

Цей показник нормалізує загальний ризик і дозволяє порівнювати ситуації з різною конфігурацією підсистем.

Динаміка приросту ризику, яка використовується для оцінки швидкості наближення до критичного стану системи:

$$M(t) = S_{\alpha} - SIRI(t). \quad (7)$$

Запас безпеки:

$$\frac{dS}{dt} \approx \frac{S(t+) - S(t)}{\Delta t}, \quad (8)$$

де S_{cr} – заданий критичний поріг безпеки, наприклад 0.05, ф час до досягнення критичного стану:

$$T_{crit} = M(t) / \frac{dS}{dt}. \quad (9)$$

Це дозволяє системі проактивно оцінювати ризику на горизонті часу.

Каскадний ризик поширення збоїв:

$$P_{cascade} = 1 - \prod_{j=1}^m (1 - q_i \cdot (1 - E_{B_j}(t))), \quad (10)$$

де q_j – ймовірність того, що збій у j -й підсистемі пошириться на інші.

Умова активації реагування:

$$(SIRI(t) > S_{\alpha}) \vee (T_{crit} < T_{reaction}) \Rightarrow \Rightarrow (\text{Активация протоколу}). \quad (11)$$

Очікувана втрата функціональної безпеки:

$$L = \sum_{j=1}^m p_j \cdot \Delta S_j(t). \quad (12)$$

Цей показник інтегрується в модуль прийняття рішень для оптимального вибору сценарію реагування.

Запропонована розрахункова модель дозволяє комплексно оцінювати рівень експлуатаційної безпеки судна на основі багаторівневої ймовірнісної структури з урахуванням технічного стану підсистем, ефективності бар'єрів безпеки та зовнішніх факторів. Включення модуля деградації бар'єрів,

обчислення агрегованого $L = \sum_{j=1}^m p_j \cdot \Delta S_j(t)$, ризику,

нормалізованого індексу $SIRI$ та прогнозного часу до критичного стану забезпечує високу гнучкість і точність системи. Модель є адаптивною до різних сценаріїв експлуатації і може використовуватись як основа для автоматизованих рішень щодо безпеки судноплавства, включаючи динамічне прогнозування та активацію захисних дій.

Методика реалізації моделі. Методика реалізації розробленої розрахункової моделі спрямована на її адаптацію до практичних завдань оцінки експлуатаційної безпеки судна за умов технічної невизначеності та змінного середовища. Реалізація передбачає структурований підхід, що поєднує вхідну параметризацію, аналітичні розрахунки, алгоритмічну логіку та інтелектуальне відтворення сценаріїв. На початковому етапі здійснюється формалізація вхідних даних для ключових підсистем судна, зокрема системи керування, енергетичної установки, навігаційного комплексу та модуля зв'язку. Для кожної підсистеми задаються базові параметри: ймовірність відмови, початкова ефективність бар'єру, темп деградації, а також коефіцієнт критичності, що відображає вагу підсистеми у загальній структурі безпеки.

У моделі розглядається шість ключових підсистем судна, що забезпечують його критичні функції під час експлуатації. До переліку включено: систему керування, енергетичну установку, навігаційну систему, комунікаційний модуль, систему життєзабезпечення та протипожежну систему. Для кожної з них задаються базові параметри: початкова ймовірність відмови $P(C_j)$, початкова ефективність бар'єру E_{0j} , коефіцієнт деградації λ_j , а також ваговий коефіцієнт критичності w_j , який відображає відносну важливість підсистеми у загальній структурі безпеки судна (табл. 1).

Подальша процедура реалізації включає етапи поетапного обчислення ключових показників. Спершу виконується розрахунок деградації ефективності бар'єрів згідно з експоненціальним законом, після

чого формується ризикова оцінка кожної підсистеми з урахуванням технічного стану та втрати захисної функціональності.

Таблиця 1 – Розширена таблиця вхідних параметрів підсистем

Підсистема	$P(C_j)$	E_{0j}	λ_j	w_j
Система керування	0.08	0.95	0.010	0.30
Енергетична установка	0.10	0.90	0.008	0.25
Навігаційна система	0.06	0.85	0.007	0.20
Комунікаційний модуль	0.05	0.80	0.006	0.15
Система життєзабезпечення	0.04	0.88	0.005	0.05
Протипожежна система	0.03	0.92	0.004	0.05

Згодом ці оцінки агрегуються в єдиний інтегрований індекс S , який нормалізується до показника $SIRI$. На наступному кроці визначається запас безпеки як різниця між критичним порогом і поточним значенням індексу, а також обчислюється швидкість зміни ризику та прогнозується час до настання критичного стану. Модель також включає елемент каскадного аналізу ризику та механізм прийняття рішень на основі умови активації реагування, що вмикається за перевищення критичного порогу або недовостатнього часу для запобігання інциденту.

У табл. 2 представлено покроковий алгоритм реалізації моделі оцінки експлуатаційної безпеки судна, який включає етапи обробки вхідних даних, розрахунку ризику, прогнозування критичних станів та генерації ключових індикаторів.

Таблиця 2 – Алгоритм реалізації моделі оцінки експлуатаційної безпеки

Крок	Опис дії
1. Ініціалізація параметрів	Зчитування параметрів $P(C_j)$, E_{0j} , λ_j , w_j для кожної підсистеми;
2. Розрахунок ефективності бар'єрів	Застосування експоненційної формули: $E_{vj}(t) = E_{0j} \cdot e^{-\lambda_j t}$;
3. Визначення локального ризику підсистем	Розрахунок R_j на основі технічного стану та ефективності бар'єра;
4. Агрегація загального ризику $S(t)$	Обчислення $S(t)$ як зваженої суми ризиків усіх підсистем;
5. Розрахунок $SIRI$	Нормалізація ризику: $SIRI = \sum w_j \cdot S$;
6. Обчислення запасу безпеки та dS/dt	Визначення запасу: $M = S_{cr} - SIRI$; розрахунок похідної зміни ризику;
7. Прогноз T_{crit}	Обчислення часу до досягнення критичного рівня безпеки;
8. Перевірка умови реагування	Активізація захисту, якщо виконується умова: $SIRI > S_{cr}$ або $T_{crit} < T_{реакції}$;
9. Формування індикаторів	Виведення фінальних метрик: індексів ризику, запасу безпеки, прогнозованого часу, очікуваних втрат.

Технічна реалізація моделі здійснена у середовищі Python 3.11 з використанням бібліотек NumPy та Pandas для обчислень і роботи з даними, Matplotlib

для візуалізації, SciPy для статистичних оцінок, а також TensorFlow для реалізації прогновної нейронної мережі компоненти. Такий підхід дозволяє легко адаптувати модель до різних типів суден, конфігурацій підсистем та умов експлуатації, а також створює можливості для її інтеграції в інформаційні системи суден нового покоління.

З метою перевірки ефективності запропонованої моделі було проведено симуляцію п'яти типових сценаріїв експлуатації судна. Кожен сценарій відображає різний рівень складності та ризикового навантаження: базовий режим, погіршення погодних умов, часткову втрату ефективності бар'єрів, комбінований вплив зовнішніх та технічних факторів, а також критичне експлуатаційне навантаження. Для кожного з них розраховано ключові індикатори: інтегрований індекс ризику $SIRI$ та запас безпеки M .

На графіку (рис. 3) наведено динаміку зміни $SIRI$ та M в залежності від сценарію.

Як видно, із ускладненням ситуації індекс ризику поступово зменшується, що вказує на зниження рівня залишкової безпеки системи. Водночас, значення запасу безпеки M зростає, демонструючи, що судно швидше наближається до критичного стану, особливо у випадках комбінованого впливу та критичного навантаження.

Результати моделювання підтверджують чутливість моделі до змін параметрів сценаріїв та її здатність до раннього виявлення ризику. Найменше значення $SIRI$ спостерігається у сценарії з максимальним навантаженням, що свідчить про ефективну диференціацію ризикових станів. Зростання запасу безпеки M вказує на підвищену потребу в активізації заходів реагування.

Таким чином, модель демонструє високий потенціал для використання в системах попереджувального управління безпекою суден.

Динамічний аналіз ризику та часу реагування. Результати моделювання свідчать про високу інформативність інтегрованих індексів, зокрема $SIRI$, запасу безпеки M та прогнозованого часу до критичного стану T_{crit} для оцінки ризику при змінних сценаріях експлуатації.

На рис. 4 подано графік зміни індексу ризику та запасу безпеки при варіації восьми типових сценаріїв, що імітують як поодинокі, так і комбіновані порушення.

Зі зростанням інтенсивності впливу зовнішніх і технічних факторів рівень безпеки системи суттєво знижується, що виражається у зменшенні $SIRI$ та відповідному зростанні запасу M . Найнижче значення $SIRI$ зафіксовано у сценарії критичного навантаження - 0.0391, що свідчить про необхідність негайної активації захисних протоколів.

На рис. 5 подано графік зміни прогнозованого часу до критичного рівня безпеки T_{crit} .

Видно, що при переході від базових до екстремальних сценаріїв доступний час на реагування зменшується з 55 до 12 хвилин на кожній точці графіка (вказано час у хвилинах).

Видно, що при переході від базових до екстремальних сценаріїв доступний час на реагування зме-

ншується з 55 до 12 хвилин на кожній точці графіка вказано час у хвилинах).

Отже, результати динамічного моделювання демонструють здатність запропонованої моделі ефективно диференціювати сценарії експлуатації за рівнем ризику, запасом безпеки та часом досягнення критичних станів. Графічні залежності підтверджу-

ють чутливість системи до впливу зовнішніх і технічних факторів, а також до змін у ефективності бар'єрів безпеки.

Отримані індекси дають змогу формувати адаптивні рекомендації щодо реагування та можуть слугувати основою для моніторингу ризиків у реальному часі.

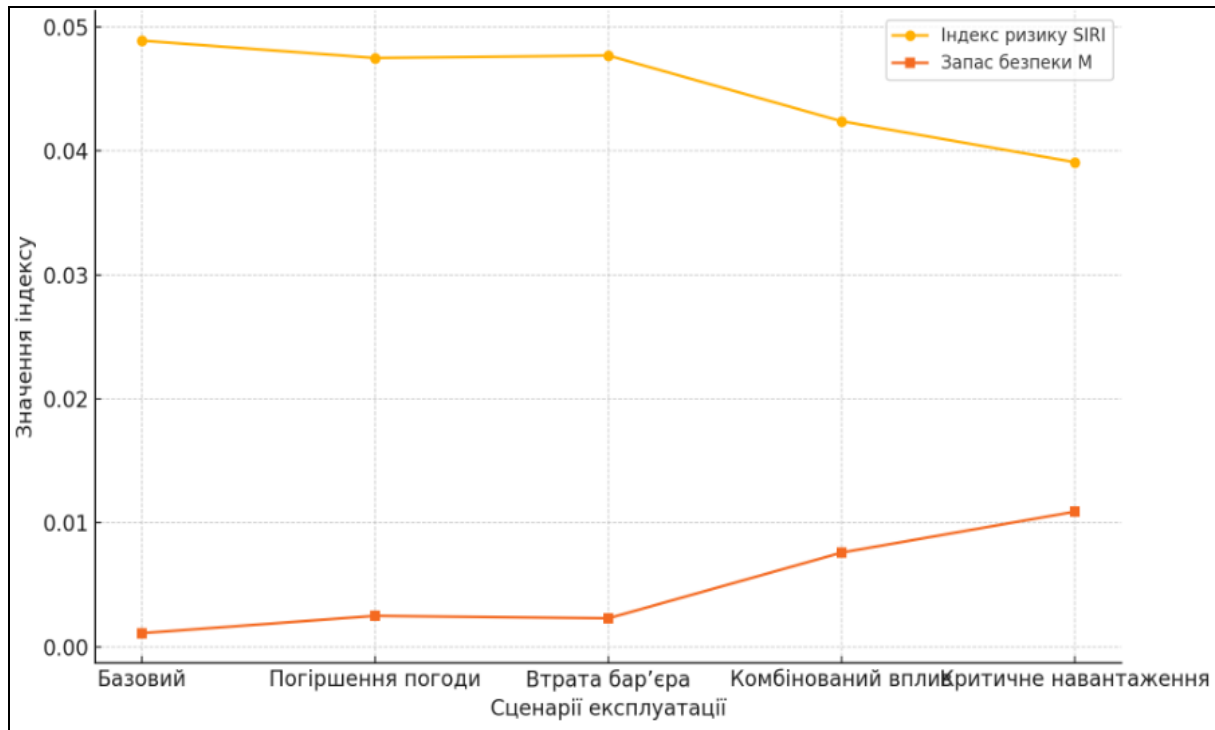


Рис. 3. Зміна індексу ризику (SIRI) та запасу безпеки (M) за сценаріями

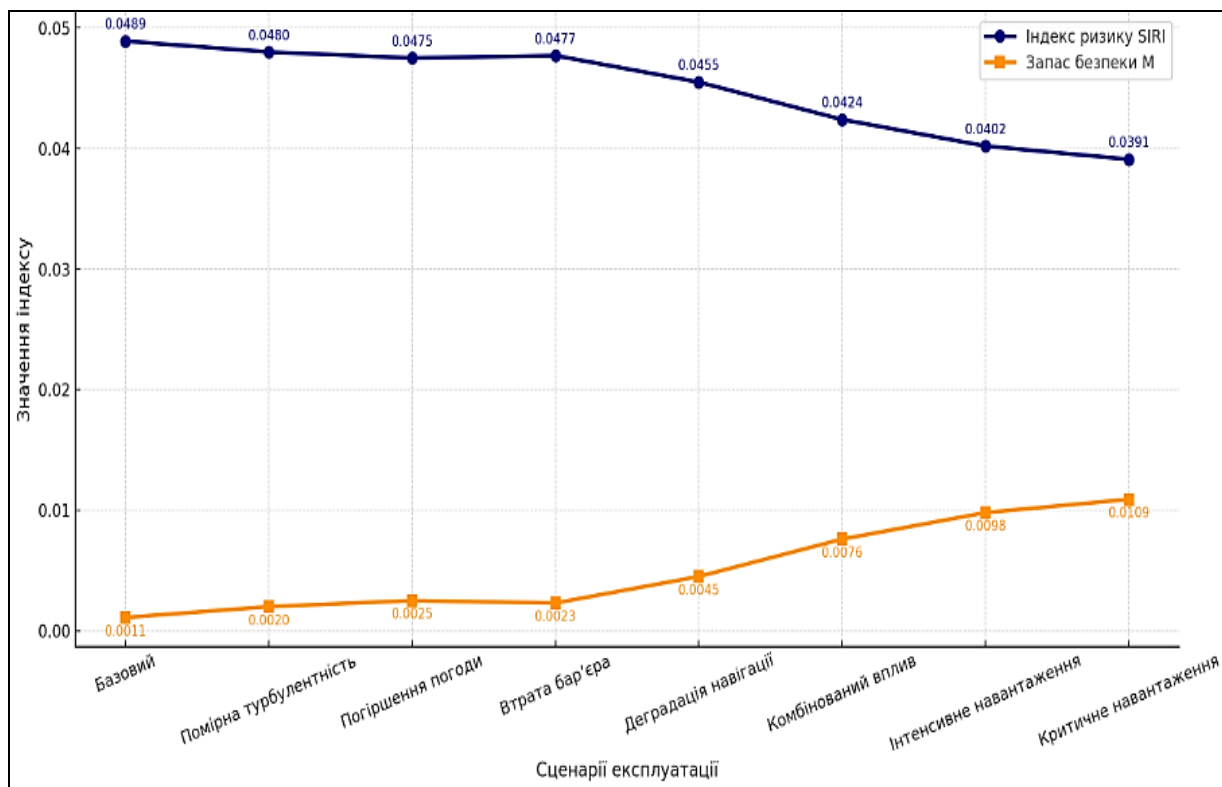


Рис. 4. Динаміка зміни індексу SIRI та запасу безпеки M у різних сценаріях експлуатації із точними значеннями індексів

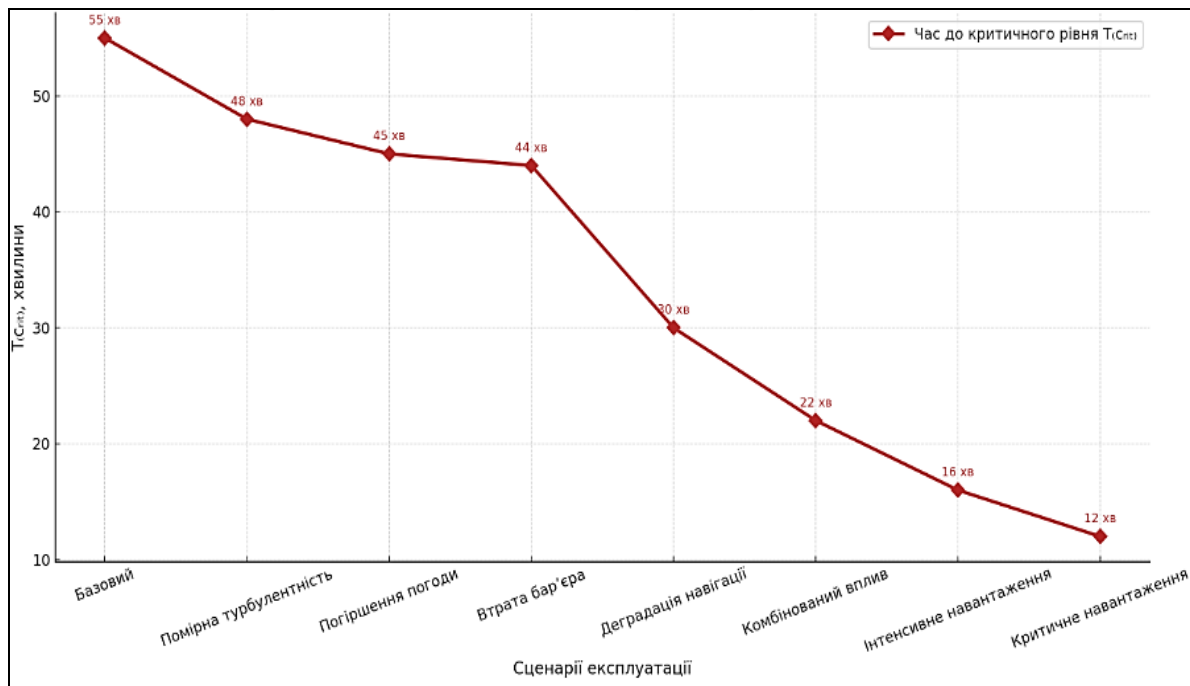


Рис. 5. Прогнозований час до досягнення критичного рівня безпеки $T_{ср}$

Висновки

У статті запропоновано та реалізовано інтегровану розрахункову модель оцінки експлуатаційної безпеки судна, яка враховує технічний стан підсистем, ефективність захисних бар'єрів, зовнішні впливи та часову динаміку ризику. Модель ґрунтується на багаторівневому ймовірнісному підході з урахуванням експоненційної деградації бар'єрів безпеки та механізму прогнозування за допомогою LSTM-архітектури.

Методика реалізації моделі охоплює повний цикл обчислень: від ініціалізації параметрів до формування узагальнених індексів ризику (SIRI), запасу безпеки, часу до досягнення критичного стану та

очікуваних втрат функціональної безпеки. Розроблений алгоритм продемонстрував ефективну реакцію на зміну сценарних умов, зокрема підвищену чутливість до комбінованих впливів і технічної деградації.

Результати моделювання показали, що зниження ефективності навіть одного бар'єра або одночасне погіршення зовнішніх умов значно скорочують прогнозований час до критичного рівня безпеки - з 55 до 12 хвилин.

Інтегровані індекси дозволяють здійснювати передчасне виявлення ризикових станів, що є ключовим фактором у забезпеченні своєчасної реакції екіпажу на або систем автоматизованого контролю на аварійні ситуації.

СПИСОК ЛІТЕРАТУРИ

1. Fan C., Montewka J., Bolbot V., Zhang Y., Qiu Y., Hu S. Towards an analysis framework for operational risk coupling mode: A case from MASS navigating in restricted waters. *Reliability Engineering & System Safety*. 2024. Vol. 248. Article ID 110176. <https://doi.org/10.1016/j.res.2024.110176>
2. Nakashima T., Kureta R., Khastgir S. Addressing systemic risks in autonomous maritime navigation: A structured STPA and ODD-based methodology. *Reliability Engineering & System Safety*. 2025. Vol. 261. Article ID 111041. <https://doi.org/10.1016/j.res.2025.111041>
3. Chai H., Dong K., Liang Y., Han Z., He R. Machine learning-based accidents analysis and risk early warning of hazardous materials transportation. *Journal of Loss Prevention in the Process Industries*. 2025. Vol. 95. Article ID 105594. <https://doi.org/10.1016/j.jlp.2025.105594>
4. Inal O. B., Charpentier J., Deniz C. Hybrid power and propulsion systems for ships: Current status and future challenges. *Renewable and Sustainable Energy Reviews*. 2022. Vol. 156. Article ID 111965. <https://doi.org/10.1016/j.rser.2021.111965>
5. Guo S., Wang Y., Dai L., Hu H. All-electric ship operations and management: Overview and future research directions. *ETransportation*. 2023. Vol. 17. Article ID 100251. <https://doi.org/10.1016/j.etrans.2023.100251>
6. Sharifzadeh M., Cooper N., Van't Noordende H., Shah N. Operational strategies and integrated design for producing green hydrogen from wind electricity. *International Journal of Hydrogen Energy*. 2024. Vol. 64. P. 650–675. <https://doi.org/10.1016/j.ijhydene.2024.03.237>
7. Onsay E. A., Bulao R. J. G., Rabajante J. F. Bagyong Kristine (TS Trami) in Bicol, Philippines: Flood Risk Forecasting, Disaster Risk Preparedness Predictions and Lived Experiences through Machine Learning (ML), Econometrics, and Hermeneutic Analysis. *Natural Hazards Research*. 2025. <https://doi.org/10.1016/j.nhres.2025.02.004>
8. Mentis A. Risk analysis of on-field and on-board activities and resilience investigation of Izmir Aliaga Ship Recycling Facilities. *Ocean Engineering*. 2023. Vol. 287. Article ID 115891. <https://doi.org/10.1016/j.oceaneng.2023.115891>

9. Ruponen P., Montewka J., Tompuri M., Manderbacka T., Hirdaris S. A framework for onboard assessment and monitoring of flooding risk due to open watertight doors for passenger ships. *Reliability Engineering & System Safety*. 2022. Vol. 226. Article ID 108666. <https://doi.org/10.1016/j.res.2022.108666>
10. Sultana S., Haugen S. An extended FRAM method to check the adequacy of safety barriers and to assess the safety of a socio-technical system. *Safety Science*. 2022. Vol. 157. Article ID 105930. <https://doi.org/10.1016/j.ssci.2022.105930>
11. Moussa A. A., Farag Y. B., Gunbeyaz S. A., Fahim N. S., Kurt R. E. Development and research directions in ship recycling: A systematic literature review with bibliometric analysis. *Marine Pollution Bulletin*. 2024. Vol. 201. Article ID 116247. <https://doi.org/10.1016/j.marpolbul.2024.116247>
12. Hasan Imran M., Khan M. I., Jamaludin S., Hasan I., Bin Ahmad M. F., Mohamad Ayob A. F., Norsani bin Wan Nik W. M., Russtam Suhrab M. I., Ridwan Bin Zulkifli M. F., Afrizal N. B., Abidin Bin Syed Ahmad S. Z. A critical analysis of machine learning in ship, offshore, and oil & gas corrosion research, part I: Corrosion detection and classification. *Ocean Engineering*. 2024. Vol. 313. Article ID 119600. <https://doi.org/10.1016/j.oceaneng.2024.119600>
13. Han S., Li F., Lee C., Wang T., Diaconeasa M. A. Mirror the mind of crew: Maritime risk analysis with explicit cognitive processes in a human digital twin. *Advanced Engineering Informatics*. 2024. Vol. 62. Article ID 102746. <https://doi.org/10.1016/j.aei.2024.102746>
14. Alsuwian T., Ansari S., Zainuri M. A. A. M., Ayob A., Abdolrasol M. G., Sudaryanto S., Alhawari A. R., Almajwani A., Almasabi S., Hindi A. T. A review of hybrid methods based remaining useful life prediction framework and SWOT analysis for energy storage systems in electric vehicle application. *Journal of Energy Storage*. 2025. Vol. 117. Article ID 116152. <https://doi.org/10.1016/j.est.2025.116152>
15. Melnyk O., Bychkovsky Y., Onishchenko O., Onyshchenko S., Volianska Y. Development the Method of Shipboard Operations Risk Assessment Quality Evaluation Based on Experts Review. *Studies in Systems, Decision and Control*. 2023. Vol. 481. P. 695–710. https://doi.org/10.1007/978-3-031-35088-7_4
16. Onyshchenko S., Bychkovsky Y., Melnyk O., Onishchenko O., Jurkovič M., Rubskyi V., Liashenko K. A model for assessing shipping safety within project-orientated risk management based on human element. *Scientific Journal of Silesian University of Technology. Series Transport*. 2024. Vol. 123. P. 319–334. <https://doi.org/10.20858/sjsutst.2024.123.16>
17. Melnyk O., Onyshchenko S., Shumylo O., Ocheretna V., Kononova O. Analysis of Factors Affecting the Ship Safety on the Basis of Six-Stage Risk Management Model. In: Babak V., Zaporozhets A. (eds) *Systems, Decision and Control in Energy VI. Studies in Systems, Decision and Control*. 2024. Vol. 561. P. 421–434. https://doi.org/10.1007/978-3-031-68372-5_23.

Received (Надійшла) 11.02.2025

Accepted for publication (Прийнята до друку) 24.04.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Мельник Олексій Миколайович – доктор технічних наук, професор, професор кафедри судноводіння і морської безпеки, Одеський національний морський університет, Одеса, Україна;

Oleksiy Melnyk – Doctor of Technical Sciences, Professor, Professor at the Department of Navigation and Maritime Safety, Odesa National Maritime University, Odesa, Ukraine;

e-mail: m.onmu@ukr.net; ORCID Author ID: <https://orcid.org/0000-0001-9228-8459>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57216657937>

Никитюк Петро Володимирович – старший викладач, аспірант кафедри судноводіння і морської безпеки, Одеський національний морський університет, Одеса, Україна;

Petro Nykytuik – Senior Lecturer, postgraduate of Department of Navigation and Maritime Safety, Odesa National Maritime University, Odesa, Ukraine;

e-mail: gelaevnyk@gmail.com; ORCID Author ID: <https://orcid.org/0000-0002-5905-3807>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=59680170900>.

Calculation model for assessing ship operational safety considering the degradation of barriers and risk forecasting

Petro Nykytuik, Oleksiy Melnyk

Abstract. Relevance. With the increasing complexity of the technical support of marine vessels and the introduction of autonomous control elements, there is a growing need to create models that can adaptively respond to changes in the technical condition of subsystems and external threats. A key element in safety systems is a generalized risk index, which should dynamically change depending on technical degradation and the combined impact of the environment. **Object of research:** an integrated model for assessing the operational safety of a ship. **Purpose:** to develop a method for calculating the operational safety of a ship, taking into account the degradation of safety barriers, the technical condition of subsystems and the forecast of risk dynamics. **Research results.** The article proposes a multilevel probabilistic model that allows calculating the risks of individual subsystems, evaluating the effectiveness of safety barriers, and forming an integrated safety index (SIRI). The model is supplemented with a block for predicting the time to a critical state and an algorithm for activating a protective response. Numerical modeling of six ship subsystems for eight typical operating scenarios was carried out. **Conclusions.** Implementation of the model allows for early detection of risky situations, reduction of response time, and increase of decision-making systems' information content. The model can be used as the core of digital platforms for managing the safety of ships. Scope of application of the results: intelligent ship safety management systems, autonomous navigation, automated navigation platforms.

Keywords: vessel operational safety; integrated risk model; barrier degradation; technical condition of subsystems; SIRI risk index; maritime safety management; digital forecasting; probabilistic model; safety margin assessment; response time forecast; multifactorial analysis; scenario modeling; critical states; automated response; decision support.

О. В. Орисенко, А. І. Криворот, М. В. Шаповал, І. А. Рогозін

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

ДОСЛІДЖЕННЯ КОЕФІЦІЄНТА ЖОРСТКОСТІ ПРУЖИННОГО ТА ПНЕВМАТИЧНОГО ЕЛЕМЕНТІВ ПІДВІСКИ АВТОМОБІЛЯ OPEL VIVARO

Анотація. В роботі досліджено жорсткісні характеристики пружинного та пневматичного елементів підвіски автомобіля категорії N1 Opel Vivaro. Актуальність теми зумовлена необхідністю підвищення плавності ходу транспортного засобу, що має безпосередній вплив на комфорт і безпеку водія та пасажирів. Особлива увага приділена порівнянню традиційної гвинтової пружини та пневматичного елемента типу рукав закручування, як альтернативних варіантів елементів підвіски при модернізації задньої підвіски автомобіля Opel Vivaro. Проведено аналіз чинних нормативів щодо впливу вібрації на людину (ДСТУ, ISO, EN тощо), а також розглянуто фізіологічні критерії оцінки плавності ходу автомобіля. Встановлено, що оптимальна жорсткість підвіски повинна забезпечувати мінімізацію вертикальних прискорень кузова, відповідно до гранично допустимих порогів вібраційного навантаження. Методично дослідження базується на навантаженні пружного елемента за допомогою зовнішніх сил, які при дослідженні створюються гідравлічним домкратом із вмонтованим манометром, де зміна жорсткості розглядається як функція тиску, висоти та ефективної площі робочих елементів. Усі вихідні параметри були отримані експериментальним шляхом. Встановлено, що для металевого і пневматичного пружного елементів залежність сили пружності від деформації в межах робочого ходу описується параболічним законом. Було проведено апроксимацію отриманих даних за квадратичною регресією з високою точністю, про що свідчать значення коефіцієнту детермінації R^2 . У результаті порівняльного аналізу встановлено, що пневматичний елемент має змінну жорсткість, яка забезпечує кращу адаптацію підвіски до умов руху. Це дозволяє покращити плавність ходу, зменшити прискорення, що передаються на кузов та забезпечити регулювання висоти автомобіля залежно від навантаження. Металевий елемент має простішу конструкцію, але поступається за створюваним рівнем комфорту для водія та пасажирів, особливо при змінних навантаженнях і в процесі старіння металу. Отримані результати можуть бути використані при модернізації існуючих транспортних засобів, а також у розробці нових конструкцій підвісок із використанням адаптивних пневматичних систем. Застосування пневматичних елементів особливо доцільне для транспортних засобів, де комфорт, стабільність і гнучкість у налаштуванні ходової частини мають першочергове значення.

Ключові слова: металевий пружний елемент, пневматичний пружний елемент, коефіцієнт жорсткості, автомобіль Opel Vivaro, підвіска автомобіля.

Вступ

Дослідження забезпечення необхідної плавності руху транспортних засобів (ТЗ) (з точки зору мінімізації вертикальних прискорень і вібрацій, що діють на водія і пасажирів під час руху) власне і є основою проектування автомобільної підвіски, розпочате в кінці 1940-х років. Це знайшло відображення у виділенні в окремий розділ теорії динаміки транспортних засобів. Починаючи з 1970-х років, плавність руху автомобіля, що характеризується впливом вібрації ТЗ на водія та пасажирів, стала важливим фактором комфорту водіння, і це також відображено в законодавчій базі: системи стандартів охорони здоров'я, стандарти безпеки праці, наприклад ОСТ, ДСТУ, ISO, EN, FSSM.

З точки зору теорії руху ТЗ, можна створити критерій плавності руху, тобто показник для оцінки впливу коливань і вібрацій на організм людини шляхом кількісного визначення відповідних порогів гранично допустимих значень. І автомобілісти, і лікарі використовують такі показники та відповідні їм порогові рівні ще до формування відповідної нормативної бази, тобто так звані «фізіологічні» критерії плавності ходу ТЗ почали працювати над розв'язком підвісок.

Паралельно проводилися теоретичні дослідження впливу вібрації на організм людини під час їздових досліджень в салоні автомобіля. На цьому етапі вже почали враховувати плавність руху та експлуатаційні вимоги.

Плавність ходу є однією з експлуатаційних характеристик і порівнюється з іншими елементами комфорту під час руху пов'язаними з мікрокліматом салону, зручністю управління, ергономікою та розташуванням сидінь, шумом транспорту тощо.

В даний час існує чотири основні категорії методик моделювання пневматичних пружин. Однією з таких моделей є термодинамічна модель, яка спирається на механічні та термодинамічні закони для опису фізичних явищ у системі підвіски.

Варіації об'єму пневматичної пружини визначаються за допомогою закону збереження енергії та маси. Крім того, варіації жорсткості пневмопружини прямо пропорційні її різній ефективній площі, так само, як її об'єм пропорційний висоті. При цьому вхідні параметри моделі часто збираються експериментальним шляхом. Ця категорія моделей зазвичай ігнорує поперечні характеристики пневмобалону і концентрується в основному на її вертикальних характеристиках.

Клас мембран моделей в першу чергу визначає жорсткість пневматичної пружини та інші відповідні характеристики, такі як варіації форми повітряного сильфона та пов'язану з цим втому пневматичної пружини, використовуючи характеристики матеріалу для визначення викривлення діафрагми.

В останніх дослідженнях, які проводилися в нашій країні, не було знайдено ні теоретичних ні експериментальних даних саме про зрівняння робочих процесів пружинної та пневматичної підвіски транспортних засобів категорії N1.

Тому виникає необхідність досліджень спрямованих на пружних елементів конкретного автомобіля.

Постановка проблеми. Аналіз процедур сертифікаційних випробувань та впровадження в українській автомобільній промисловості, а також відповідної законодавчої бази дозволяє виявити чинні законодавчі вимоги щодо комфорту (плавності руху), окрім чинних санітарних стандартів та правил. Національне управління безпеки дорожнього руху дотримуючись стандартів радянських часів, висунуло думку про прийнятні рівні вібрації на робочих місцях водіїв [1]. Справді, навіть порівняно з радянською епохою, коли обсяг національних випробувань для автобусів включав обов'язкові вимоги, такі як галузеві стандарти [2] для забезпечення нормативних стандартів щодо плавності пасажирських перевезень, наразі законодавчих норм не існує. Тому відсутні вимоги і для вітчизняних виробників автобусів щодо вдосконалення підвісок ТЗ з точки зору комфорту водіння (просто відбувається встановлення сидіння водія з власною вібропоглинаючою основою), а звичайним пасажиром встановлюють сидіння із мінімальним рівнем віброзахисту.

Існуючий нормативний наказ Мінтрансу про оцінку комфортності автобусів [3] ґрунтується на відповідному нормативному законі про міжнародну оцінку перевізників IRU, але не враховує вимог до плавності руху автобусів. У той же час в країнах ЄС ці вимоги враховані в автомобільній нормативній базі.

Фактично основою нормативної бази, що діє в ЄС та нашій країні для оцінки допустимого рівня вібраційного навантаження на тіло людини в транспортних засобах, є пороговий рівень ISO 2631 [4], який обмежений трьома критеріями:

1) Поріг зниження комфорту А. (можливість читати, писати, їсти);

2) Поріг В зниження продуктивності праці внаслідок втоми, який є межею працездатності водія, перевищення якого може викликати втому і суттєво погіршити безпеку руху;

3) Поріг значення максимально допустимого впливу вібрації С для забезпечення здоров'я та безпеки водія. Це відповідає приблизно половині рівня з появою відчуття болю.

Загалом, з огляду на інтенсивний розвиток нормативно-правової бази у сфері вібробезпеки за останні 15-20 років, слід зазначити, що стандарти в предметній галузі мають тенденцію структуруватись за трьома типами:

– стандарти типу А (базові), що формують загальні поняття та правила безпеки для вібраційних навантажень, вимірних значень, факторів впливу (включаючи значення корекції частоти для конкретних умов), а також загальні методи вимірювання та оцінки. Крім них, стандарти ISO 2631, ISO 5349-1 (включаючи і ДСТУ EN ISO 5349-1 [5]);

– стандарти типу В (повна група проблем вібраційної безпеки) визначають як вимірюються та оцінюються вібрації в конкретних умовах експлуатації, або в конкретних стандартах, зокрема: ISO

5349 1 (ДСТУ EN ISO 5349-1, частина 2); EN 14536 2003; ISO 2631-4; ISO 20643; EN 1032: 2003; ISO 10326-1;

– стандарти групи С стосуються кодів і методів тестування.

У той же час слід зазначити, що в частині частотного спектру, що розглядається, стандарти ISO 9611 та ISO 13332 фактично охоплюють лише шум з акцентом на дуже високих частотах та вібрацію з частотою вище за 20 Гц.

Огляд і аналіз історії виникнення дослідження вібрацій на ТЗ і нормативної бази у сфері оцінки плавності руху дозволяє констатувати, що дослідження у даному напрямку є актуальними. Також, при пересуванні ТЗ на рівень комфорту пасажирів, окрім плавності руху та вібрації, впливають ще й інші фактори, такі як шум, мікроклімат всередині автомобіля (вентиляція та опалення), властивості сидіння. Законодавчі та нормативні вимоги, чинні в Україні, ґрунтуються на відповідних національних або міжнародних стандартах, таких як ДСТУ, правила ЄЕК ООН. Нормативні вимоги щодо вібрації та допустимих порогових рівнів вібрації в Україні діють лише по відношенню до робочого місця водія у вигляді фактичних галузевих стандартів, тобто санітарних норм і правил. Обов'язок дотримання вимог стандарту ISO 2631-1, нещодавно прийнятого як ДСТУ, стосовно робочого місця водіїв у процесі сертифікації типів транспортних засобів не регулюється законодавством.

Метою роботи є проведення оригінального експериментального дослідження, а саме визначення жорсткості пружних елементів двох типів пружин: металевих та пневматичного рукава закріплення. Саме останній використовують при переобладнанні підвісок автомобілів категорії N1 Opel Vivaro.

Аналіз останніх досліджень і публікацій. Підвіска відіграє надзвичайно важливу роль в конструкції будь-якого транспортного засобу. Основними її складовими є пружні та вібропоглинаючі елементи, а також система тяг, які забезпечують з'єднання між шасі, віссю та колесами що контактують з дорогою. Основним призначенням підвіски є передача навантаження від транспортного засобу на опорну поверхню та забезпечення плавності його руху при подоланні нерівностей дорожнього покриття, яке супроводжується вібраціями та поштовхами. Також, крім виконання вказаних функцій підвіска відіграє важливу роль у забезпеченні стійкості автомобіля. Утримуючи шини в контакті з дорожнім покриттям, система підвіски (рис. 1) забезпечує стійкість до поперечного переміщення, хитаючі та перекидання. Це гарантує, що автомобіль залишається стійким і легким в управлінні за будь-яких умов руху, в тому числі під час різких маневрів або при несприятливих погодних умовах. Пружини в системі підвіски (рис. 1, а) поглинають удари, викликані нерівностями дороги, такими як горби або вибоїни, і забезпечують стабілізацію центра мас автомобіля. Тим часом амортизатори контролюють рух пружин і запобігають надмірному підстрибуванню автомобіля. Важелі з'єднують різні компоне-

нти системи підвіски і забезпечують розподіл ваги автомобіля на всі колеса для рівномірного контакту з опорною поверхнею [6].



а



б

Рис. 1. Підвіска: а – з пружинними елементами; б – з пневматичними елементами

Пневматична система (рис. 1, б) підвіски на відміну від пружинної (рис. 1, а) – це, по суті, тип підвіски, де пневматичні сильфони (пневматичні пружини), виготовлені в основному з армованої текстилем гуми, використовуються як основне або єдине віброізоляційне середовище. Пневмобалони заповнюються повітрям від повітряного компресора, що приводиться в дію двигуном внутрішнього згоряння або електричним приводом. Завдяки низькій власній частоті коливань пневматичні пружини забезпечують більш комфортну їзду і постійну висоту дорожнього просвіту незалежно від навантаження автомобіля [7].

Пневматична підвіска в основному використовувалася у вантажних ТЗ і автобусах, але за останнє десятиліття вона значно еволюціонувала до прецизійної системи з мікропроцесорним керуванням, що здатна виконувати свої функції на набагато краще, аніж звичайна пружинна підвіска. На відміну від металевих пружин, хід пневматичного пружного елемента підвіски не пропорційний її навантаженню, що дозволяє їй функціонувати як пружина і в той же час використовуватися для підйому та зниження дорожнього просвіту автомобіля. Пневматична підвіска може використовуватися з будь-якою системою амортизації, включаючи магнітореологічні адаптивні амортизатори (рис. 2) та стабілізатори поперечної стійкості змінної жорсткості. Крім того, пневматичні пружини можуть використовуватися незалежно від типу амортизаторів.

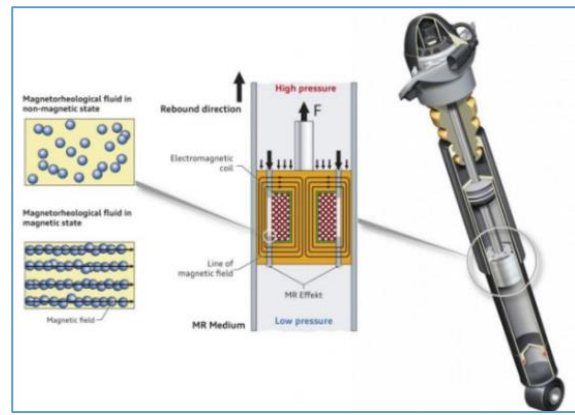


Рис. 2. Магнітореологічний адаптивний амортизатор

Пневматичні пружні елементи (пневматичну пружину) можна описати як еластомерний сильфон, який може розширюватися і стискатися для поглинання збуджених зовні вібрацій відповідно до кількості стисненого повітря, накачаного або вилученого з нього. Як показано на рис. 3, пневмопружини бувають різних розмірів і форм, наприклад, звивисті пневмобалони (одинарні, подвійні і потрійні) [8].

Структура звивистого пневмобалона включає тороподібний елемент, який постійно кріпиться до нього, що робить його більш піддатливим у бічному напрямку деформації. Конічний рукав забезпечує можливість ширшого регулювання висоти дорожнього просвіту та розроблений таким чином, що його можна встановити у місцях з обмеженим простором. Пневматична пружина з рукавом заковчування складається в основному з сильфона, поршня і верхньої пластини, яка в деяких випадках може змінюватися в залежності від деформації діафрагми, що додатково впливає на поперечну жорсткість пневматичної пружини [9].

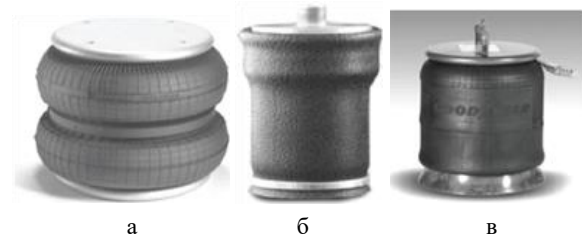


Рис. 3. Типи пневматичних пружин: а – подвійна звивина; б – конічна втулка; в – рукав заковчування

Таким чином, пневматична пружина з рухомою втулкою забезпечує поперечну гнучкість і регулювання висоти дорожнього просвіту для конкретних умов застосування. Різниця між конічними та рухомими рукавами полягає в їх можливості забезпечення висоти дорожнього просвіту, і вони розроблені для різних транспортних засобів [10]. Пневматичні пружини оцінюються на основі їх несучої здатності, яка безпосередньо пов'язана з ефективною площею поверхні (A_w) [11]. Внутрішній тиск пневматичної пружини (P_i) є основним визначальним фактором її власної частоти і жорсткості, що визначає загальну жорсткість пневмопідвіски (опорне зусилля, F), де

$F=P_i \times A_w$ [12]. Ефективна площа поверхні A_w визначається ефективним діаметром (d_w). Хоча моделювання пневматичних пружин значною мірою є аналогічним, відмінність їх застосування вимагає, щоб вони були різноманітними за конструкцією. Крім того, відмінність конструкції може вплинути на характеристики пружини. Яскравим прикладом є ефективна площа, яка є критичним параметром, що визначає кореляцію між внутрішнім тиском пневматичної пружини та силою пружинної пластини. Це може бути змінено зміщенням у деяких конструкціях, таких як пневматичні пружини з U-подібним сильфоном, де найнижча точка згину визначає ефективну площу, отже, впливає на жорсткість пружини (позитивно чи негативно) навіть при постійному внутрішньому тиску [13].

Основний матеріал

Для досліджень було вибрано підвіску автомобіля Opel Vivaro (рис. 4), як такого, що має високу масовість випуску та випускався в різних модифікаціях, наприклад, мікроавтобуси та фургони [14].

Випуск даних модифікацій розпочався в 1997 році, американська компанія General Motors уклала контракт з Renault на виробництво дещо модифікованого Renault Trafic першого покоління під власним брендом на заміну Isuzu Midi. Ця модель пропонувалася у Великобританії як Opel Vauxhall Arena, а в Бразилії як Chevrolet Space Van.



Рис. 4. Opel Vivaro III (з 2019 року і по теперішній час)

Було вирішено виконати оригінальне експериментальне дослідження, а саме провести визначення жорсткості двох типів пружних елементів: металевої пружини та пневматичного рукава закручування, який масово використовується при переобладнанні пружинної підвіски Opel Vivaro.

Під час проведення експериментальних досліджень використовувалися наступне обладнання:

- стенд для перевірки жорсткості пружних елементів автомобільної підвіски (рис. 5);
- домкрат гідравлічний;
- манометр гідравлічний;
- манометр пневматичний.

Процес дослідження виконувався так, спочатку було виміряно початкові лінійні розміри пружних елементів (для металевого пружного елемента – $l_0=350$ мм, для пневматичного пружного елемента $l_0=323$ мм) (рис. 6). Під час дослідів проводилося навантаження пружного елемента за допомогою гідравлічного домкрата з вмонтованим манометром.



Рис. 5. Стенд для перевірки жорсткості пружних елементів



Рис. 6. Визначення початкових лінійних розмірів пружних елементів

Лінійні розміри стискування знімалися через крок показника манометра у 0,2 МПа починаючи з 0,4 МПа (це початковий показник манометра). Виміри знімалися до тиску у 4 МПа (це максимальне значення показника манометра). Для того щоб перевести тиск у зусилля, яке необхідне для визначення коефіцієнта жорсткості, необхідно знати діаметр плунжера (в нашому випадку $d=28$ мм). Виміри для пневматичного пружного елемента проводилися до внутрішнього тиску у 0,25 МПа. Кількість повторів при проведенні дослідів за однакових умов дорівнювало трьом. Дані знятих характеристик жорсткості наведено у табл. 1 – 6.

Таблиця 1 – Експериментальні дані коефіцієнта жорсткості металевого пружного елемента задньої підвіски автомобіля Opel Vivaro

№ пп	Показники тиску манометра p , МПа	Значення зусилля пружного елемента F , Н	Зміна вертикальних лінійних розмірів за вимірювання l , мм			Середнє значення лінійного розміру l , мм	Коефіцієнт жорсткості k , Н/мм
			1-е	2-е	3-е		
1	0,4	246	350	350	350	350	82,1
2	0,6	369	351	348	350	350	1230,9
3	0,8	492	348	346	346	347	3774,7
4	1	615	343	342	342	342	7590,4
5	1,2	739	339	337	337	338	12062,6
6	1,4	862	335	333	333	334	16945,1
7	1,6	985	331	330	330	330	23304,7
8	1,8	1108	327	326	326	326	30648,9
9	2	1231	323	322	322	322	39388,2
10	2,2	1354	318	318	318	318	48291,5
11	2,4	1477	315	314	314	314	59082,2
12	2,6	1600	310	310	310	310	69339,6
13	2,8	1723	307	306	307	307	82140,7
14	3	1846	303	302	302	302	95393,2
15	3,2	1969	299	298	298	298	108973,9
16	3,4	2092	296	294	294	295	122759,8
17	3,6	2216	292	291	291	291	140320,3
18	3,8	2339	288	286	286	287	157470,6
19	4	2462	284	282	282	283	165758,5

Для наочності зміни коефіцієнта жорсткості, відтворимо його у графічному вигляді (рис. 7), по осі абсцис відкладаємо значення показника манометра, а по осі ординат коефіцієнт жорсткості.

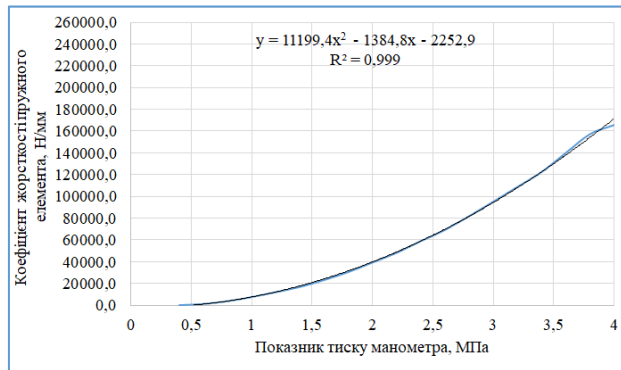


Рис. 7. Залежність зміни коефіцієнта жорсткості металевих пружних елементів задньої підвіски автомобіля Opel Vivaro

З рис. 7 бачимо, що коефіцієнт жорсткості змінюється за параболическою функцією, тому для підтвердження якості проведених вимірювань було прийнято рішення провести апроксимацію отриманих даних за квадратичною регресією.

Рівняння регресії має вигляд

$$y_i = a \cdot x^2 + b \cdot x + c \quad (1)$$

де a , b , c – коефіцієнти рівняння.

Для визначення коефіцієнтів використано систему рівнянь

$$\begin{cases} a \cdot \sum x_i^2 + b \cdot \sum x_i + n \cdot c = \sum y_i; \\ a \cdot \sum x_i^3 + b \cdot \sum x_i^2 + c \cdot \sum x_i = \sum x_i \cdot y_i; \\ a \cdot \sum x_i^4 + b \cdot \sum x_i^3 + c \cdot \sum x_i^2 = \sum x_i^2 \cdot y_i. \end{cases} \quad (2)$$

За (1)

$$a = 11199,4; b = -1384,8; c = -2252,9,$$

а отже рівняння набуде вигляду

$$y = 11199,4 \cdot x^2 - 1384,8 \cdot x - 2252,9. \quad (3)$$

Коефіцієнт кореляції знаходиться за формулою

$$R = \sqrt{1 - \frac{\sum (y_i - \bar{y})^2}{\sum (y_i - \bar{y})^2}}, \quad (4)$$

де

$$\bar{y} = \frac{1}{n} \sum y_i. \quad (5)$$

Найпоширенішою мірою залежності між двома величинами є коефіцієнт кореляції Пірсона, який зазвичай називають просто «коефіцієнтом кореляції». Його знаходять шляхом взяття співвідношення коваріації двох розглянутих змінних числового набору даних і нормалізації квадратним коренем їх дисперсії. Математично коваріація цих двох змінних просто ділиться на добуток їхніх стандартних відхилень. Карл Пірсон розробив це співвідношення на основі подібних, але дещо відмінних ідей Френсіса

Гальтона. [15] Коефіцієнт кореляції Пірсона намагається встановити лінію, яка найкраще відповідає набору даних двох змінних, по суті, відображаючи очікувані значення. Отриманий коефіцієнт кореляції Пірсона [15] показує, наскільки фактичний набір даних відхиляється від очікуваного значення. Якщо існує певний зв'язок між змінними в наборі даних, це буде або негативна, або позитивна кореляція, залежно від знаку коефіцієнта кореляції Пірсона. Для значень таблиці 1 коефіцієнт кореляції дорівнює 0,9995. Коефіцієнт детермінації, R^2 , – статистичний показник, який використовується в статистичних моделях як міра залежності варіації залежної змінної від варіації незалежної змінної. Іншими словами, це чисельно вказує, скільки варіацій залежної змінної пояснюється моделлю. Вказує, наскільки добре отримані спостереження підтверджують модель, що відображено на рис. 7 – 12. Для рис. 7. коефіцієнт детермінації дорівнює 0,999, що говорить про високе співпадіння функції отриманої апроксимацією експериментальним даним.

Таблиця 2 – Експериментальні дані коефіцієнта жорсткості пневматичного пружного елемента ($P_b=0,05$ МПа) задньої підвіски автомобіля Opel Vivaro

№ пп	Показники тиску манометра p , МПа	Значення зусилля пружного елемента F , Н	Зміна вертикальних лінійних розмірів за вимірювання l , мм			Середнє значення лінійного розміру l , мм	Коефіцієнт жорсткості k , Н/мм
			1-е	2-е	3-е		
1	0,4	246	322	323	322	322	1230,88
2	0,6	369	317	318	317	317	3692,64
3	0,8	492	312	313	312	312	7385,28
4	1	615	307	308	307	307	12924,24
5	1,2	739	301	302	301	301	19201,728
6	1,4	862	296	297	296	296	25848,48
7	1,6	985	292	293	292	292	34464,64
8	1,8	1108	287	288	287	287	44311,68
9	2	1231	282	283	282	282	55389,6
10	2,2	1354	277	278	277	277	69052,368
11	2,4	1477	271	272	271	271	84192,192
12	2,6	1600	265	266	265	265	99208,928
13	2,8	1723	260	261	260	260	115456,544
14	3	1846	255	256	255	255	131088,72
15	3,2	1969	251	252	251	251	151644,416
16	3,4	2092	245	246	245	245	177862,16
17	3,6	2216	237	238	237	237	194971,392
18	3,8	2339	234	235	234	234	215157,824
19	4	2462	230	231	230	230	226481,92

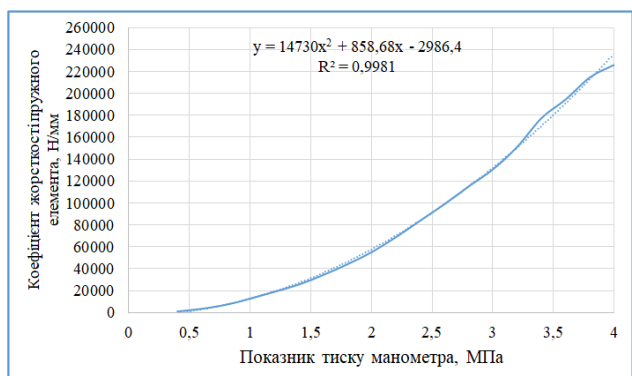


Рис. 8. Залежність зміни коефіцієнта жорсткості пневматичного пружного елемента ($P_b=0,05$ МПа) задньої підвіски автомобіля Opel Vivaro

Таблиця 3 – Експериментальні дані коефіцієнта жорсткості пневматичного пружного елемента ($P_Б=0,1$ МПа) задньої підвіски автомобіля Opel Vivaro

№ пп	Показники тиску манометра p , МПа	Значення зусилля пружного елемента F , Н	Зміна вертикальних лінійних розмірів за вимірювання l , мм			Середнє значення лінійного розміру l , мм	Коефіцієнт жорсткості k , Н/мм
			1-е	2-е	3-е		
1	0,4	246	322	323	322	322	0
2	0,6	369	322	323	322	322	0
3	0,8	492	322	323	322	322	984,704
4	1	615	320	321	320	320	3077,2
5	1,2	739	317	318	317	317	6646,752
6	1,4	862	313	314	313	313	17232,32
7	1,6	985	302	303	302	302	26587,008
8	1,8	1108	295	296	295	295	34341,552
9	2	1231	291	292	291	291	43080,8
10	2,2	1354	287	288	287	287	54158,72
11	2,4	1477	282	283	282	282	66467,52
12	2,6	1600	277	278	277	277	78407,056
13	2,8	1723	273	274	273	273	94777,76
14	3	1846	267	268	267	267	108932,88
15	3,2	1969	263	264	263	263	128011,52
16	3,4	2092	257	258	257	257	144382,224
17	3,6	2216	253	254	253	253	166168,8
18	3,8	2339	247	248	247	247	184755,088
19	4	2462	243	244	243	243	194479,04

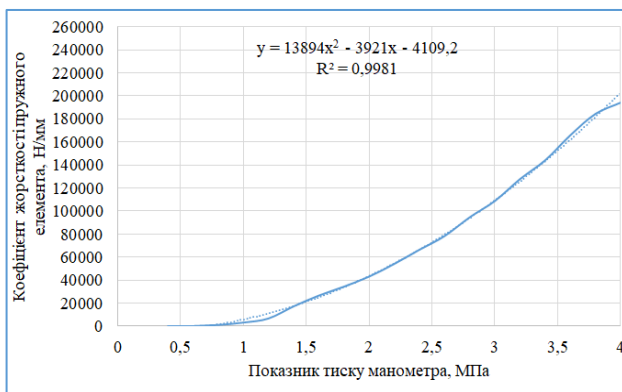


Рис. 9. Залежність зміни коефіцієнта жорсткості пневматичного пружного елемента ($P_Б=0,1$ МПа) задньої підвіски автомобіля Opel Vivaro

Таблиця 4 – Експериментальні дані коефіцієнта пружності пневматичного пружного елемента ($P_Б=0,15$ МПа) задньої підвіски автомобіля Opel Vivaro

№ пп	Показники тиску манометра p , МПа	Значення зусилля пружного елемента F , Н	Зміна вертикальних лінійних розмірів за вимірювання l , мм			Середнє значення лінійного розміру l , мм	Коефіцієнт жорсткості k , Н/мм
			1-е	2-е	3-е		
1	0,4	246	323	323	323	323	0,00
2	0,6	369	323	323	323	323	0,00
3	0,8	492	323	323	323	323	820,59
4	1	615	321	322	321	321	2256,61
5	1,2	739	319	320	319	319	4184,99
6	1,4	862	317	318	317	317	6605,72
7	1,6	985	315	316	315	315	10503,51
8	1,8	1108	312	313	312	312	14032,03
9	2	1231	310	311	310	310	20514,67
10	2,2	1354	306	307	306	306	27982,01
11	2,4	1477	302	303	302	302	36434,05
12	2,6	1600	298	299	298	298	44270,65
13	2,8	1723	295	296	295	295	52845,78
14	3	1846	292	293	292	292	62159,44
15	3,2	1969	289	290	289	289	74181,03
16	3,4	2092	285	286	285	285	85094,84
17	3,6	2216	282	283	282	282	94531,58
18	3,8	2339	280	281	280	280	111476,70
19	4	2462	275	276	275	275	117343,89

Для пневматичного пружного елемента, при різних тисках повітря в балоні, наведено лише табличні дані із графічними залежностями коефіцієнта жорсткості.

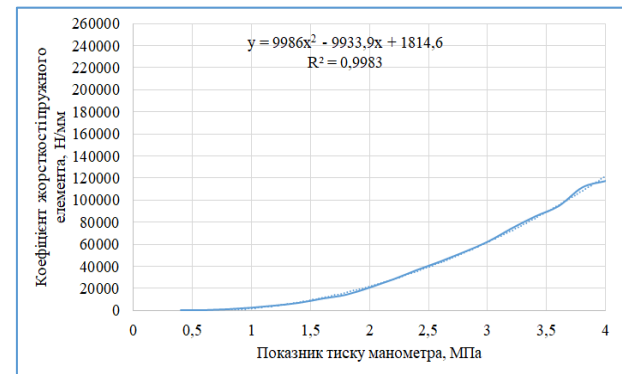


Рис. 10. Залежність зміни коефіцієнта жорсткості пневматичного пружного елемента ($P_Б=0,15$ МПа) задньої підвіски автомобіля Opel Vivaro

Таблиця 5 – Експериментальні дані коефіцієнта жорсткості пневматичного пружного елемента ($P_Б=0,2$ МПа) задньої підвіски автомобіля Opel Vivaro

№ пп	Показники тиску манометра p , МПа	Значення зусилля пружного елемента F , Н	Зміна вертикальних лінійних розмірів за вимірювання l , мм			Середнє значення лінійного розміру l , мм	Коефіцієнт жорсткості k , Н/мм
			1-е	2-е	3-е		
1	0,4	246	323	324	323	323	0,00
2	0,6	369	323	324	323	323	0,00
3	0,8	492	323	324	323	323	0,00
4	1	615	323	324	323	323	615,44
5	1,2	739	322	323	322	322	2215,58
6	1,4	862	320	321	320	320	3446,46
7	1,6	985	319	320	319	319	5908,22
8	1,8	1108	317	318	317	317	8862,34
9	2	1231	315	316	315	315	12308,80
10	2,2	1354	313	314	313	313	16247,62
11	2,4	1477	311	312	311	311	22155,84
12	2,6	1600	308	309	308	308	30402,74
13	2,8	1723	304	305	304	304	39634,34
14	3	1846	300	301	300	300	46158,00
15	3,2	1969	298	299	298	298	55143,42
16	3,4	2092	295	296	295	295	64867,38
17	3,6	2216	292	293	292	292	75329,86
18	3,8	2339	289	290	289	289	86530,86
19	4	2462	286	287	286	286	91085,12

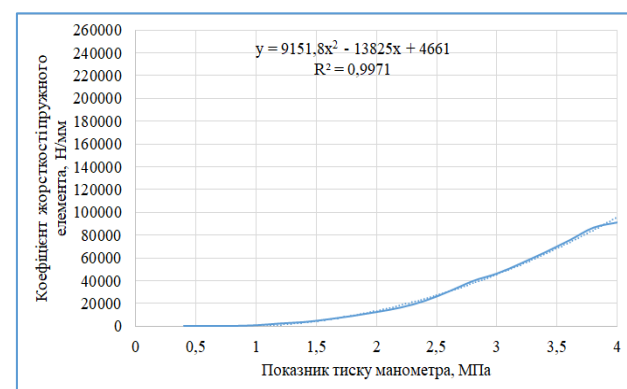


Рис. 11. Залежність зміни коефіцієнта жорсткості пневматичного пружного елемента ($P_Б=0,2$ МПа) задньої підвіски автомобіля Opel Vivaro

Таблиця 6 – Експериментальні дані коефіцієнта пружності пневматичного пружного елемента ($P_B=0,25$ МПа) задньої підвіски автомобіля Opel Vivaro

№ пп	Показники тиску манометра p , МПа	Значення зусилля пружного елемента F , Н	Зміна вертикальних лінійних розмірів за вимірювання l , мм			Середнє значення лінійного розміру l , мм	Коефіцієнт жорсткості k , Н/мм
			1-е	2-е	3-е		
1	0,4	246	323	324	323	323	0,00
2	0,6	369	323	324	323	323	0,00
3	0,8	492	323	324	323	323	0,00
4	1	615	323	324	323	323	0,00
5	1,2	739	323	324	323	323	738,53
6	1,4	862	322	323	322	322	1723,23
7	1,6	985	321	322	321	321	2954,11
8	1,8	1108	320	321	320	320	5538,96
9	2	1231	318	319	318	318	8616,16
10	2,2	1354	316	317	316	316	12185,71
11	2,4	1477	314	315	314	314	14770,56
12	2,6	1600	313	314	313	313	19201,73
13	2,8	1723	311	312	311	311	24125,25
14	3	1846	309	310	309	309	29541,12
15	3,2	1969	307	308	307	307	41357,57
16	3,4	2092	302	303	302	302	48127,41
17	3,6	2216	300	301	300	300	55389,60
18	3,8	2339	298	299	298	298	63144,14
19	4	2462	296	297	296	296	66467,52

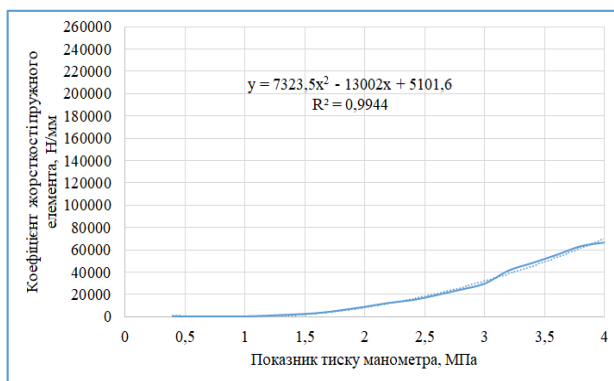


Рис. 12. Залежність зміни коефіцієнта жорсткості пневматичного пружного елемента ($P_B=0,25$ МПа) задньої підвіски автомобіля Opel Vivaro

Аналізуючи рис. 7–12 приходимо до висновку, що експериментальні дослідження виконано з високою точністю, про що свідчить значення коефіцієнта детермінації. Досліджуваний максимальний коефіцієнт жорсткості пневматичного пружного елемента

при відповідних показниках тисків на манометрі домкрату змінюється у довільній формі при будь-яких тисках у пневмобалоні і складає, наприклад, при тиску у 4 МПа на манометрі домкрату:

- при внутрішньому тиску пневматичного пружного елемента у 0,05 МПа $k=226481,9$ Н/мм;
- при внутрішньому тиску пневматичного пружного елемента у 0,1 МПа $k=194479$ Н/мм;
- при внутрішньому тиску пневматичного пружного елемента у 0,15 МПа $k=117343,9$ Н/мм;
- при внутрішньому тиску пневматичного пружного елемента у 0,2 МПа $k=91085,12$ Н/мм;
- при внутрішньому тиску пневматичного пружного елемента у 0,25 МПа $k=66467,52$ Н/мм;

На це безсумнівно впливають гістерезисні властивості гуми з якої виготовлено пневматичний пружний елемент та термодинамічні процеси стиску і розширення повітря в ньому.

Висновки

1. Отримані результати досліджень підтверджують ефективність використання пневматичних підвісок для всіх категорій автомобілів. Переобладнання підвісок автомобілів Opel Vivaro на пневматичні пружні елементи зумовлено старінням металевих пружних елементів, які під час експлуатації з часом просідають від навантаження і не забезпечують постійний дорожній просвіт. Окрім того, при використанні пневматичного пружного елемента типу рукав закручування з'являється можливість змінювати: кліренс, жорсткість та пружність підвіски лише зміною внутрішнього тиску пневмоелемента, що додатково створює комфортне перебування водія та пасажирів в автомобілі.

2. Характер зміни коефіцієнта жорсткості всіх досліджуваних пружних елементів задньої підвіски Opel Vivaro, мають не лінійну залежність. Зміна вказаного параметра відбувається за параболічною залежністю.

3. Досліджуваний максимальний коефіцієнт жорсткості пневматичного пружного елемента при відповідних показниках тисків на манометрі домкрату змінюється у довільній формі для всіх тисків у пневмобалоні.

СПИСОК ЛІТЕРАТУРИ

- Бур'ян М.В. *Плавність руху автобусів у взаємозв'язку з характеристиками підвіски та сидінь*: автореферат дис. на здобуття наукового ступеня кандидата технічних наук: спец. 05.22.02 Автомобілі та трактори. Львів, 2020. 21 с
- Наказ Мін-ва транспорту та зв'язку України від 12.04.2007 № 285 «Про затвердження Порядку визначення класу комфортності автобусів, сфери їхнього використання за видами сполучення та режимами сполучення та режимами руху».
- ДСТУ EN ISO 5349-2:2005 *Вібрація механічна. Вимірювання та оцінювання впливу на людину локальної вібрації. Частина 2. Практична настанова з вимірювання на робочому місці* (EN ISO 5349-2:2001, IDT)
- ISO 2631: *Mechanical vibration and shock-evaluation of human exposure to whole-body vibration, second ed.*, ISO – International Organization for Standardization, 1997.
- ДСТУ EN ISO 5349-1: 2005 *Вібрація механічна. Вимірювання та оцінювання впливу на людину вібрації. Частина 1: Загальні вимоги* (EN ISO 5349-1:2001, IDT).
- A. Goodarzi and A. Khajepour, "Vehicle suspension system technology and design," Synthesis Lectures on Advances in Automotive Technology, Morgan and Claypool Publishers, vol. 1, no. 1, pp i–77, California, CA, USA, 2017.
- P. Karimi Eskandary, A. Khajepour, A. Wong, and M. Ansari, "Analysis and optimization of air suspension system with independent height and sti9ness tuning," Int. Journal of Automotive Technology, vol. 17, no. 5, pp. 807–816, 2016.
- J.-J. Chen, Z.-H. Yin, X.-J. Yuan, G.-Q. Qiu, K.-H. Guo, and X.-L. Wang, "A re8ned sti9ness model of rolling lobe air spring with structural parameters and the sti9ness characteristics of rubber bellows," Measurement, vol. 169, Article ID 108355, 2021.

9. F. Lücken and M. Welsch, "The dynamic characteristic and hysteresis effect of an air spring," *International Journal of Applied Mechanics and Engineering*, vol. 20, no. 1, pp. 127–145, 2015.
10. R. Zhou, B. Zhang, and Z. Li, "Dynamic modeling and computer simulation analysis of the air spring suspension," *Journal of Mechanical Science and Technology*, vol. 36, no. 4, pp. 1719–1727, 2022.
11. L. Q. Sun, Z. X. Li, X. F. Shen, and J. Y. Zhu, "Simulation and test study on dynamic characteristic of air spring with auxiliary chamber," *Applied Mechanics and Materials*, vol. 341, pp. 391–394, 2013.
12. J.-J. Chen, Z.-H. Yin, S. Rakheja, J.-H. He, and K.-H. Guo, "Theoretical modelling and experimental analysis of the vertical stiffness of a convoluted air spring including the effect of the stiffness of the bellows," *Proceedings of the Institution of Mechanical Engineers- Part D: Journal of Automobile Engineering*, vol. 232, no. 4, pp. 547–561, 2018.
13. Z. Yin, J. Jiang, and W.-B. Shangguan, "Complex stiffness model of an air spring with auxiliary chamber considering inertial effects of gas in connecting pipeline," *Proceedings of the Institution of Mechanical Engineers- Part D: Journal of Automobile Engineering*, vol. 237, no. 1, pp. 145–158, 2023.
14. Opel Vivaro: веб-сайт. URL: https://uk.wikipedia.org/wiki/Opel_Vivaro.
15. Карташов М. В. Імовірність, процеси, статистика. – Київ : ВПЦ Київський університет, 2007. – 504 с.

Received (Надійшла) 28.02.2025

Accepted for publication (Прийнята до друку) 07.05.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Орисенко Олександр Вікторович – кандидат технічних наук, доцент, завідувач кафедри галузевого машинобудування та мехатроніки, Національний університет «Полтавська політехніка імені Юрія Кондратюка», Полтава, Україна;
Oleksandr Orysenko – Candidate of Technical Sciences, Associate Professor, Head of the Department of Mechanical Engineering and Mechatronics, National University «Yuri Kondratyuk Poltava Polytechnic», Poltava, Ukraine;
 e-mail: oleksandr.orysenko@gmail.com; ORCID Author ID: <https://orcid.org/0000-0003-3103-0096>;
 Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57215916720>.

Криворот Анатолій Ігорович – кандидат технічних наук, доцент, доцент кафедри галузевого машинобудування та мехатроніки, Національний університет «Полтавська політехніка імені Юрія Кондратюка», Полтава, Україна;
Anatolii Kryvorot – Candidate of Technical Sciences, Associate Professor, Associate Professor of the Department of Mechanical Engineering and Mechatronics, National University «Yuri Kondratyuk Poltava Polytechnic», Poltava, Ukraine;
 e-mail: anatoliikryvorot@gmail.com; ORCID Author ID: <https://orcid.org/0000-0001-5919-7352>;
 Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57217279936>.

Шаповал Микола Віталійович – кандидат технічних наук, доцент кафедри галузевого машинобудування та мехатроніки, Національний університет «Полтавська політехніка імені Юрія Кондратюка», Полтава, Україна;
Mykola Shapoval – Candidate of Technical Sciences, Associate Professor of the Department of Mechanical Engineering and Mechatronics, National University «Yuri Kondratyuk Poltava Polytechnic», Poltava, Ukraine;
 e-mail: myshapoval75@ukr.net; ORCID Author ID: <https://orcid.org/0000-0002-6943-7687>;
 Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57209207132>.

Рогозін Іван Анатолійович – кандидат технічних наук, доцент кафедри галузевого машинобудування та мехатроніки, Національний університет «Полтавська політехніка імені Юрія Кондратюка», Полтава, Україна;
Ivan Rohozin – Candidate of Technical Sciences, Associate Professor of the Department of Mechanical Engineering and Mechatronics, National University «Yuri Kondratyuk Poltava Polytechnic», Poltava, Ukraine;
 e-mail: ria.workbox@gmail.com; ORCID Author ID: <https://orcid.org/0000-0002-9052-4806>;
 Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57216482840>.

The spring and pneumatic suspension vehicle Opel Vivaro elements stiffness coefficient research

Oleksandr Orysenko, Anatolii Kryvorot, Mykola Shapoval, Ivan Rohozin

Abstract. The study investigates the stiffness characteristics of the spring and pneumatic suspension elements of an N1 category vehicle - Opel Vivaro. The relevance of the topic is driven by the need to improve the ride comfort of the vehicle, which directly affects the comfort and safety of both the driver and passengers. Particular attention is paid to the comparison between a traditional coil spring and a rolling-lobe type pneumatic element, as alternative suspension components in the modernization of the rear suspension of the Opel Vivaro. An analysis of existing regulations regarding human exposure to vibration (DSTU, ISO, EN, etc.) is conducted, along with consideration of physiological criteria for assessing vehicle ride comfort. It has been established that optimal suspension stiffness should minimize vertical body accelerations in accordance with permissible vibration exposure thresholds. The methodology of the study is based on loading the elastic element using external forces, which are applied during testing via a hydraulic jack equipped with a built-in pressure gauge. In this context, stiffness is examined as a function of pressure, height, and the effective area of the working elements. All input parameters were obtained experimentally. It was determined that for both the metallic and pneumatic elastic elements, the relationship between the restoring force and deformation within the working range can be described by a parabolic law. The obtained data were approximated using quadratic regression with high accuracy, as indicated by the R^2 coefficient. As a result of the comparative analysis, it was found that the pneumatic element exhibits variable stiffness, allowing the suspension to better adapt to driving conditions. This improves ride comfort, reduces accelerations transmitted to the vehicle body, and enables ride height adjustment based on load. The metallic element has a simpler design but provides a lower level of comfort, particularly under variable loads and as the metal ages. The obtained results can be used in the modernization of existing vehicles, as well as in the development of new suspension designs incorporating adaptive pneumatic systems. The use of pneumatic elements is especially advisable for vehicles where comfort, stability, and flexibility in chassis tuning are of primary importance.

Keywords: metal elastic element, pneumatic elastic element, stiffness coefficient, Opel Vivaro, vehicle suspension.

С. О. Шаптала, Т. Ф. Шмельова

Державний університет «Київський авіаційний інститут», Київ, Україна

ДОСЛІДЖЕННЯ ЗАВДАННЯ ФОРМУВАННЯ ПОТОКІВ ПОВІТРЯНОГО РУХУ В ІНТЕГРОВАНОМУ ПОВІТРЯНОМУ ПРОСТОРИ

Анотація. Актуальність. Сучасні тенденції розвитку безпілотної авіації, зокрема можливість її експлуатації поза межами прямої видимості, вимагають інтеграції безпілотних систем у загальний повітряний простір. Існуючі системи управління були орієнтовані на пілотовану авіацію, що обмежує ефективність їх використання в нових умовах. **Об'єкт дослідження:** процес формування потоків повітряного руху в умовах інтегрованого повітряного простору. **Мета статті:** визначення напрямків наукових досліджень щодо адаптації, координації та ефективності прийняття рішень в умовах змінної повітряної обстановки, основних принципів планування та організації повітряного простору, розробка математичної моделі постановки задачі оптимізації потоків повітряного руху в інтегрованому повітряному просторі. **Результати дослідження.** У статті визначено основні принципи планування та організації повітряного простору, охарактеризовано роль системи U-space, розроблено математичну модель постановки задачі оптимізації потоків повітряного руху. Запропоновано напрямки наукових досліджень щодо адаптації, координації та ефективності прийняття рішень в умовах змінної повітряної обстановки. **Висновки.** Запропоновані підходи дозволяють забезпечити комплексну автоматизацію процесів формування потоків повітряного руху, моніторингу та навігаційної безпеки в інтегрованому повітряному просторі. **Сфера використання отриманих результатів:** системи управління повітряним рухом, що включають пілотовані та безпілотні літальні апарати, зокрема для національних та міжнародних авіаційних регуляторів.

Ключові слова: транспортні технології, навігація, управління повітряним рухом, безпека, повітряний простір, планування повітряного простору, потоки повітряного простору, адаптація, координація, пілотована авіація, безпілотна авіація, інтеграція.

Вступ

Управління використанням повітряного простору здійснюється за принципами його гнучкого використання, визначеними законодавством Європейського Союзу, документами Євроконтролю [1] та відіграє найважливішу роль в координації і упорядкуванню використання повітряного простору всіх без винятку користувачів не залежно від підпорядкування і належності [1, 2, 3]. Одним із головних видів діяльності при використанні повітряного простору є організація потоків повітряного руху [1], яка призначена для сприяння організації безпечного, упорядкованого та прискореного потоку повітряного руху, забезпечення максимального використання та оптимізації пропускної спроможності європейської мережі організації повітряного руху [4].

Існуюча система використання повітряного простору була розроблена в інтересах і з урахуванням пілотованої авіації. Разом з тим, новий тип споживачів повітряного простору – безпілотна авіація, що впевнено увійшла в наше життя, робить її значною мірою неадекватною для безпілотних операцій як за масштабом так і за розумінням шляхів подальшого розвитку. Сектор безпілотної авіації розвивається швидкими темпами. Широке використання безпілотних літальних апаратів призвело до істотних інновацій, що стало можливим використанням безпілотних літальних апаратів (далі – БПЛА) за межами прямої видимості (далі – BVLOS).

Аналіз літературних джерел. Обсяг світового ринку БПЛА оцінюють у 17,9 млрд доларів 2021 року, і очікується, що середньорічний темп зростання становитиме 11,17%, досягнувши 33,8 млрд доларів до 2027 року [5]. У найбільш динамічно розвинутому ринку БПЛА в Китайській Народній Республіці на тлі

спаду загального обороту авіаційної діяльності пілотованої авіації в період 2020-2022 років на 32,7% спостерігалось зростання кількості польотів БПЛА в 2022 році і в цілому склав 20,67 млн. льотних годин, що на 6,17% більше порівняно з 2021 роком [6].

Завдяки значним інвестиціям і досягнутому прогресу в правовому регулюванні використання повітряного простору [1, 7, 8, 9], безпілотний сектор впевнено займає своє місце в системі повітряних комунікацій, перевезень та обслуговування. Незважаючи на бажання урегулювати незворотній процес зростаючого застосування БПЛА в авіаційній галузі, існуючі методи вже не приводять до бажаного результату. Пошуком нових шляхів подальшого розвитку безпілотної авіації у загальному і організації потоків інтегрованого повітряного руху зокрема займається низка організацій та регуляторів авіаційної діяльності світового рівня [10, 11, 12].

Висновки з аналізу літературних джерел щодо вказують на те, що поділ повітряного простору на застосування пілотованих і безпілотних повітряних суден, є неефективним використанням повітряного простору. Повітряний простір повинен бути повністю інтегрований для того, щоб переваги BVLOS були повністю реалізовані [5].

Таким чином, **метою** роботи є проведення аналізу існуючих систем та методів організації потоків повітряного простору пілотованої авіації та на їх основі визначення основних напрямків наукових досліджень щодо формування потоків пілотованих і безпілотних повітряних суден в інтегрованому (єдиному) повітряному просторі.

Основний матеріал

Планування потоків повітряного руху. Планування потоків повітряного руху є одним із

основних елементів в організації повітряного руху при використанні повітряного простору. Першим міжнародно-правовим документом, що встановив основні принципи використання повітряного простору стала Паризька Конвенція 1919 року. Постійне і динамічне збільшення використання авіаційного транспорту для авіаційних перевезень [13, 14, 15] вимагало удосконалення використання повітряного простору з метою економії коштів і забезпечення безпеки авіаційних перевезень. Не стала виключенням і Чиказька конференція 1944 року, за результатами якої була прийнята Конвенція, де закріплені основні положення, що регулюють відносини держав-учасниць в галузі міжнародної цивільної авіації та створення Міжнародної організації цивільної авіації – ICAO. Положення Чиказької конвенції 1944 р. вплинули на розвиток сучасного міжнародного повітряного права. Інституції, що створилися за її рішенням почали практично втілювати в життя основні положення щодо ефективної організації повітряного простору.

Організація повітряного простору (ПП) як і організація повітряного руху (ПР) – це загальний термін для опису спільної діяльності відповідних повноважних органів, що займаються питаннями планування й організації ПП, планування потоків ПР у межах своїх районів відповідальності з метою найбільш ефективного їхнього використання.

Для визначення основних принципів плануванні організації ПП, що є основою при плануванні потоків ПР були прийняті наступні правила [16]:

1. Виключення переваг одних експлуатантів над іншими.
2. Справедливий поділ використання ПП між всіма експлуатантами.
3. ПП над конкретною державою повинен розглядатися не ізольовано, а як складова частина регіону ICAO.

Грунтуючись на цих міркуваннях були сформувані основні принципи поділу повітряного простору, якими користуються і до сьогодні [16, 17]:

- повний і винятковий суверенітет над повітряним простором (ПП) надає право державі встановлювати правила й процедури виконання польотів;
- виконання положень SARPs ICAO – класифікація ПП, стандарт диспетчерських районів і зон, стандарт маршрутів ОПР, аеронавігаційне оснащення, стандарт позначень;
- розподіл ПП визначається потребою в ОПР (розподіл відповідальності, дотримання принципу CNS/ATM);
- виключення переваг експлуатантів при плануванні ПП і його використанні;
- координація використання ПП (між військовими і цивільними користувачами, а також між органами ОПР суміжних держав регіону).

Кожна країна у встановленому порядку повідомляє в ICAO принципи розподілу свого ПП. Межі районів польотної інформації (РПІ) розглядаються на регіональній нараді суміжних держав і встановлюються з метою забезпечення безперервності ОПР.

Стрімкий розвиток безпілотної авіаційної складової вимагає від регуляторів повітряного простору

постійного пошуку адаптивних підходів для забезпечення юридичної легітимності використання повітряного простору безпілотними літальними апаратами. Розроблено ряд концепцій по правилам застосування БПЛА, однією з яких є виділення окремих зон (районів) для використання виключно БПЛА, так звані U-space [7, 18].

Відповідно до Регламенту [7] повітряний простір U-space означає географічну зону, визначену державами-членами, в якій польоти БПЛА дозволяються лише за підтримки служб U-space, а послуга U-space означає послугу, що спирається на цифрові послуги та автоматизацію функцій, призначених для підтримки безпечного, надійного та ефективного доступу до повітряного простору U-space для великої кількості БПЛА.

U-space – це європейська система [7], яка розробляється для управління рухом та регулювання потоків безпілотних літальних апаратів (БПЛА). Це набір спеціальних послуг і процедур, призначених для забезпечення безпечного та ефективного доступу до повітряного простору для БПЛА.

Метою U-space є досягнення автоматизованого управління та інтеграції БПЛА, що дозволить виконувати велику кількість операцій, багато з яких є одночасними. Процес має бути інтегрований з існуючою системою управління повітряним рухом.

Першим кроком створення U-space є визначення меж і позначення їх як зон U-space. Це обсяги повітряного простору, в яких, як мінімум, будуть надаватися обов'язкові послуги U-space. Процес виділення повітряного простору для U-space здійснюється регуляторами країн-членів ЄС.

Якщо повітряний простір U-space виділено в межах контрольованого повітряного простору, він повинен підлягати динамічній реконфігурації. Підрозділ управління повітряним рухом може тимчасово обмежити зону, в якій можуть відбуватися операції БПЛА, щоб задовольнити попит на пілотовані перевезення. Це робиться для того, щоб переконатися, що пілотовані повітряні судна, яким надають послуги з управління повітряним рухом (далі – УПР), та БПЛА залишаються відокремленими. Динамічна реконфігурація повітряного простору не є необхідною для експлуатації повітряних суден за правилами візуальних польотів у повітряному просторі класу E і для всіх польотів у класах F і G. У цих випадках інформація про трафік надається відповідним пілотованим рейсам [19].

Послуги в повітряному просторі U-space надаються провайдером послуг U-space. На відміну від надання послуг з управління повітряним рухом, де тільки один провайдер може відповідати за певний обсяг повітряного простору, може існувати декілька провайдерів послуг в U-space, які обслуговують один і той самий повітряний простір.

Процедура розподілу повітряного простору включає в себе крім розподілу на різні категорії зони, ще і формування маршрутів.

В ідеальному випадку, екіпажі повітряних суден бажають виконувати польоти з використанням найбільш прямих маршрутів між точками вильоту й

призначення, крім випадків попадання в складні метеорологічні умови. Однак, через наявність конфліктуючих запитів різних користувачів повітряного простору, звичайно не можливо встановити найбільш спрямлений маршрут, і тому виникає необхідність у знаходженні компромісного рішення між запитами на використання повітряного простору й можливостями їхнього задоволення.

Проведений аналіз тенденцій розвитку системи планування потоків повітряного руху за стандартами ІКАО свідчить, що світова спільнота постійно приділяє велику увагу її розвитку [17]. Постійно вдосконалювалися існуючі й розроблялися нові методи планування потоків повітряного руху, а також знаходилися шляхи підвищення їх ефективного використання з позиції експлуатантів повітряного простору.

Протиріччя між експлуатантами повітряного простору, включаючи тих, хто активно користується безпіотною авіацією, стосовно раціонального використання ресурсів у польотах та належного управління потоками в повітрі, з точки зору забезпечення безпеки повітряної навігації, вимагають знаходження компромісів між ефективністю польотів та їх економічною доцільністю.

Поділ повітряного простору, який розділяє літаки з екіпажем і без екіпажу, є неефективним використанням повітряного простору. Допоки повітряний простір не буде повністю інтегрований, переваги безпілотної авіації не можуть бути реалізованими.

Проте, впровадження інтегрованого повітряного простору на теперішній час, обмежується тільки визначеними зонами через відсутність з практичної точки зору належного обладнання району засобами навігації, що забезпечують необхідну точність польоту; не досконало відпрацьовану систему формування сумісних потоків пілотованої та безпілотної авіації; не відпрацьовану систему координації між суміжними районами в інтегрованому повітряному просторі; не визначену систему динамічного формування області безпеки БПЛА та пілотованих повітряних суден.

Організації потоків повітряного руху. Організація повітряного руху нерозривно пов'язана з такими елементами системи як пропускна здатність ОПР та організації потоків повітряного руху (ОППР).

Пропускна здатність системи ОПР залежить від багатьох факторів, включаючи структуру маршрутів ОПР, навігаційну точність повітряних суден, що використовують даний повітряний простір, метеорологічні фактори і робоче навантаження на диспетчера. Системою ОПР постійно вживаються заходи для забезпечення достатньої пропускної здатності, розрахованої на нормальні та пікові рівні повітряного руху з урахуванням недопущення зниження рівня безпеки польотів. Пропускна здатність ОПР виражається максимальною кількістю повітряних суден, яка може бути прийнята під управління за певний період у даному повітряному просторі або на відповідному аеродромі.

Там, де потреби у русі перевищують пропускну спроможність системи УПР створюється Служба ОППР.

Служба ОППР створюється на основі регіональної аеронавігаційної угоди або, у окремих випадках, багатосторонньої угоди.

Правила організації потоків ОППР здійснюється в три етапи [19]:

а) стратегічне планування, якщо рішення приймається за кілька днів до набуття чинності плану. Стратегічне планування, як правило, здійснюється завчасно за 2-6 міс;

б) передтактичне планування, якщо рішення приймається за день до набуття чинності плану;

с) тактичні операції, якщо рішення приймаються в день набуття чинності плану.

Стратегічне планування

Стратегічне планування здійснюється в координації зі службою УПР та експлуатантами повітряних суден. Воно полягає в аналізі потреб на наступний період, оцінці районів і часу, коли потреби можуть перевищити наявну пропускну здатність системи УПР, і вжиття заходів для усунення незбалансованості за допомогою:

а) вжиття заходів спільно з повноважним органом УПР для забезпечення належної пропускної здатності в необхідному місці і в необхідний час;

б) зміни деяких потоків повітряного руху (орієнтування маршрутів руху);

с) у разі необхідності, планування або перепланування рейсів;

д) визначення необхідності тактичних заходів ОППР.

Там, де вводиться система орієнтації повітряного руху (TOS), маршрути формуються оптимальними за часом і відстанню для відповідних рейсів з допущенням певної гнучкості у виборі маршрутів, зокрема в разі польотів великої протяжності.

У разі узгодження TOS детальна інформація публікується всім відповідним державам у єдиному форматі.

Передтактичне планування

На етапі передтактичного планування стратегічний план приводиться в точну відповідність із поточними потребами. На цьому етапі:

а) можуть бути змінені деякі потоки повітряного руху;

б) можуть бути скоординовані розвантажувальні маршрути;

с) буде прийнято подальше рішення щодо тактичних заходів;

д) для всіх зацікавлених сторін публікуються і надаються докладні дані плану ОППР на наступний день.

Тактичні операції

Тактичні операції ОППР включають:

а) вжиття узгоджених тактичних заходів, для забезпечення розвантажувального і впорядкованого потоку руху, коли потреби можуть перевищити пропускну здатність;

б) контроль за розвитком ситуації в повітряному русі для забезпечення заходів ОППР з метою бажаної віддачі, і для вжиття або ініціювання коригувальних дій, коли повідомляється про тривалі затримки, включаючи зміну маршрутів руху і виділення

ешелонів польоту з метою максимального використання наявної пропускної спроможності УПР.

Якщо потреби в повітряному русі перевищують або передбачається, що перевищать пропускну здатність конкретного сектора або аеродрому, відповідальний орган УПР інформує відповідальний орган ОППР, де такий орган створено, та інші відповідні органи УПР. Льотні екіпажі повітряних суден, які планують політ у цей район, і експлуатанти повинні за можливості негайно інформуватися про очікувані затримки або обмеження, які будуть введені.

Як вже зазначалося вище, формування потоків нерозривно пов'язане з пропускну здатністю управління повітряним рухом. Розробкам методів розрахунку пропускної здатності приділяється постійно висока увага. Особливий інтерес становлять роботи: Директоратом оперативних досліджень та аналізу Сполученого Королівства ("Методика DORATASK для оцінки пропускної спроможності секторів УПР на маршрутах" – проміжна доповідь DORA 8818;

"Застосування методики в існуючих секторах вузлового диспетчерського району Лондона" – проміжна доповідь DORA 8918;

"Калібрування моделі DORATASK в двох секторах на маршруті в Лондонському центрі управління повітряним рухом" – доповідь DORA 8927);

дослідження німецької фірми "Messerschmidt, Bölkow und Blohm (MBB), яке завершилося розробкою методу кількісної оцінки пропускної здатності робочих місць диспетчерів УПР, відомого як "метод MBB". Суть обох методів полягає у вимірюванні часу, необхідного для всіх дій диспетчера, і співвіднесенні його із сумарним наявним часом.

Принципи застосування методики DORATASK для розрахунку пропускної спроможності секторів повітряного простору на маршруті та у вузлового диспетчерському районі (ТМА) по суті однакові. Однак при застосуванні цього методу в секторах ТМА потрібні три істотні зміни. Слід враховувати робоче навантаження, пов'язане із забезпеченням вертикального ешелонування повітряних суден при заході на посадку. Необхідно використовувати інші параметри для визначення потенційних конфліктних ситуацій з урахуванням додаткової складності конкретної структури маршрутів в ТМА. І нарешті, слід застосовувати інший метод моделювання робочого навантаження, пов'язаного з плануванням, з урахуванням того факту, що диспетчер для виявлення конфліктних ситуацій в основному покладається на екран індикатора радіолокатора, а не на табло стрипів.

В основі "методу MBB" для розрахунку пропускної спроможності робочого місця диспетчера УПР лежить кількісна оцінка робочого навантаження на диспетчера радіолокаційного контролю.

Результати досліджень протягом кількох періодів спостереження свідчать про те, що різні умови руху або типи розподілу руху зумовлюють різне робоче навантаження. Таким чином, необхідний час не можна безпосередньо перетворити на кількість повітряних суден, яку можна контролювати. У зв'язку з цим необхідно додатково оцінити кожне повітряне судно з погляду робочого навантаження на

диспетчера і відповідного необхідного йому часу під час керування цими повітряними суднами з урахуванням типу та етапу польоту.

Постановка наукового завдання

Описані вище методи є трудомісткими і, що більш важливо, дають значення пропускної здатності, які застосовні тільки в тих умовах, що визначаються наявністю обладнання, укомплектованістю кадрами, схемами руху тощо, які превальювали під час спостережень. Методи не можуть безпосередньо використовуватися для оцінки пропускної спроможності в умовах майбутньої організації повітряного простору як інтегрованого зі спільним використанням безпілотною та пілотованою авіацією, з різним обладнанням або процедурами, різною інтенсивністю руху або чисельністю кадрів.

Для системи планування повітряних потоків, що є базовою для організації повітряного руху, характерними рисами, що знижують ефективність її функціонування, є велика затратність часу на координацію процесу прийняття рішень, на використання повітряного простору ланок управління різного рівня та виникнення й існування ряду проблемних ситуацій, які виникають при зміні обстановки, що впливає як на необхідне прийняте рішення, так і на вхідні дані першого та наступних етапів планування.

Із проведеного аналізу можна зробити висновок про існування протиріччя між вимогою до системи планування потоків повітряного руху щодо зменшення часу для своєчасного прийняття рішення з одного боку і збільшенням часу на прийняття рішення з іншого через постійну зміну обстановки і збільшення об'єму інформації, яке при введенні в дію інтегрованого повітряного простору буде тільки загострюватися. У такому разі не можливо вирішити це протиріччя без планування потоку повітряного руху як цілісної системи, яка охоплює всі рівні управління, що беруть участь у плануванні.

У кінцевому результаті метою дослідження є забезпечення обґрунтованості планування потоків повітряного руху в умовах впровадження інтегрованого повітряного простору.

Якість функціонування системи планування, ефективність процесу виявлення проблем як при безпосередньому плануванні, так і в системі планування у цілому, а також процесу формування рішень по виявлених проблемах, повинні оцінюватися на основі певної сукупності показників ефективності.

Узагальненими показниками ефективності системи планування потоків повітряного руху вважаємо ефективність прийнятого рішення на застосування та час прийняття рішення через функціонал якості роботи всієї системи планування.

Формалізована постановка завдання щодо синтезу системи планування потоків повітряного руху в інтегрованому повітряному просторі до умов обстановки, що постійно змінюється на основі принципів адаптації та механізму координації за часом різних рівнів управління в умовах інтегрованого повітряного простору буде виглядати наступним чином. Маємо завдання, що стоять перед системою планування потоків повітряного руху на день D . В наявності

директивні значення часу $T_{\text{дир}}$, що відводиться на процес прийняття рішення та величини, що характеризують повітряний рух: щільність повітряного руху $P_S(t)$, інтенсивність повітряного руху $\lambda(t)$, регулярність повітряного руху $p^{(i)}(t_k)$, пропускну здатність системи ОПР $\mu(t_k)$ та коефіцієнт, що характеризує безпеку польотів при управлінні повітряним рухом K . Система огляду повітряного простору постійно надає сукупність інформації h , яка додатково впливає на прийняття рішення. Необхідно визначити таке рішення для кожного із споживачів інтегрованого повітряного руху, яке буде адаптоване до обстановки, що складається на момент прийняття ефективного рішення в мінімально короткі строки.

Математична модель постановки наукового завдання виглядає так:

$$\begin{aligned} \max_{F_{\text{нлППР}}} \min_{t_{\text{ПрР}}} F_{\text{нлППР}} = \\ = f \left[P_S(t), \lambda(t), p^{(i)}(t_k), \mu(t_k), K, h, t_{\text{ПрР}}, P \right] \end{aligned} \quad (1)$$

при таких обмеженнях:

$$P_S(t) \leq P_S(t)_{\text{зад}};$$

$$\lambda(t) \leq \lambda(t)_{\text{зад}};$$

$$t_{\text{ПрР}} \leq T_{\text{дир}};$$

$$p^{(i)}(t_k) \geq p^{(i)}(t_k)_{\text{зад}};$$

$$\mu(t_k) \geq \mu(t_k)_{\text{зад}};$$

$$(\exists_\gamma)(\exists_x) \left[P(x, \overline{D}(\gamma)) \text{ і } P(\pi_M(x), D) \right].$$

де $F_{\text{нлППР}}$ – функціонал якості планування потоків повітряного руху в інтегрованому повітряному просторі; $P_S(t), P_S(t)_{\text{зад}}$ – відповідно щільність повітряного руху на s -му рівні маршруту та її задана величина; $\lambda(t), \lambda(t)_{\text{зад}}$ – відповідно інтенсивність повітряного руху та її задана величина; $p^{(i)}(t_k), p^{(i)}(t_k)_{\text{зад}}$ – відповідно регулярність повітряного руху в i -й зоні ОПР в період часу $t_1 < t_k < t_2$ та її задана величина; $\mu(t_k), \mu(t_k)_{\text{зад}}$ – відповідно пропускну здатність системи ОПР в період часу $t_1 < t_k < t_2$ та її задана величина; K – коефіцієнт, що характеризує безпеку польотів при управлінні повітряним рухом; h – об'єм інформації, що передається на різні ланки управління щодо обстановки, яка склалася; $t_{\text{ПрР}}, T_{\text{дир}}$ – відповідно час на прийняття рішення та її директивна величина; γ – розпорядження по координаті рішення; x –

рішення, що приймається; P – процес прийняття рішення; $\overline{D}(\gamma)$ – завдання, що вирішується при координаті рішення; $\pi_M(x)$ – розпорядження по управлінню в процесі прийняття рішення; D – завдання, що вирішують при плануванні повітряного простору (глобальне завдання).

Таким чином після вирішення задачі оптимізації буде знайдено рішення щодо плану потоку повітряного руху, що буде задовольняти критерію оптимізації (1) та обмеженням.

Висновки

Таким чином, обґрунтування методів, моделей прийняття рішень при формуванні потоків повітряного руху в інтегрованому повітряному просторі спрямоване на комплексне вирішення завдань автоматизації процесів формування потоків повітряного руху, його моніторингу, забезпечення безпечних умов навігації у широкому спектрі можливих ситуацій повітряного руху.

На першому етапі рекомендується визначити, що формування єдиного повітряного простору для пілотованої і безпілотної авіації є кінцевим станом досягнення інтеграції повітряного простору, включно з перехідними кроками, необхідними для досягнення цієї мети.

На другому етапі, для реалізації цієї мети необхідно вирішити важливе наукове завдання щодо синтезу методів та моделей прийняття рішень при формуванні потоків повітряного руху в інтегрованому повітряному просторі.

Частковими завданнями дослідження для вирішення поставленого наукового завдання визначено таке:

- розробка моделі прийняття рішень з управління польотами пілотованої та безпілотної авіації в системі формування потоків повітряного руху;
- розробка методу координаті в системі прийняття рішень з управління потоками пілотованої та безпілотної авіації;
- розробка методу адаптації до поточної повітряної обстановки в системі формування потоків повітряного руху в інтегрованому повітряному просторі;
- проведення оцінки ефективності методів та моделі прийняття рішень при формуванні потоків повітряного руху;
- розробка рекомендації з реалізації методів та моделі прийняття рішень при формуванні потоків повітряного руху.

СПИСОК ЛІТЕРАТУРИ

1. Державна авіаційна служба України, Міністерство оборони України, «Наказ № 430/210 Про затвердження Авіаційних правил України «Правила використання повітряного простору України»,» 11 травня 2018. . Available: <https://zakon.rada.gov.ua/laws/show/z1056-18#Text>
2. Кабінет Міністрів України, «Законодавство України», 06 грудня 2017. . Available: <https://zakon.rada.gov.ua/laws/show/954-2017-%D0%BF#Text>
3. Верховна Рада України, «Законодавство України», 19 травня 2011. . Available: <https://zakon.rada.gov.ua/laws/show/3393-17#Text>
4. Державна авіаційна служба України, «Наказ № 567 Про затвердження Авіаційних правил України "Правила організації потоків повітряного руху",» 20 липня 2016. . Available: <https://zakon.rada.gov.ua/laws/show/z1354-16#n16>

5. P. Nilson, «Integrated airspace: a new UAV vision,» Airport Technology, 20 March 2023. . Available: <https://www.airport-technology.com/features/integrated-airspace-a-new-uav-vision/?cf-view&cf-closed>
6. Civil Aviation Administration of China, «Statistical Bulletin of Civil Aviation Industry Development in 2022,» 25 December 2023. Available: <https://www.caac.gov.cn/English/Research/Reports/Statistical/202312/P020231225495691527532.pdf>.
7. European Union Aviation Safety Agency, «Acceptable Means of Compliance and Guidance Material to Regulation (EU) 2021/664 on a regulatory framework for the U-space,» 16 December 2022. . Available: <https://www.easa.europa.eu/en/document-library/acceptable-means-of-compliance-and-guidance-materials/amc-and-gm-implementing>
8. European Union Aviation Safety Agency, «Easy Access Rules for Unmanned Aircraft Systems,» 28 September 2022. . Available: <https://www.easa.europa.eu/en/document-library/easy-access-rules/easy-access-rules-unmanned-aircraft-systems-regulations-eu#group-publications>
9. Міністерство оборони України, «Наказ «Про затвердження Правил виконання польотів безпілотними авіаційними комплексами державної авіації України» № 661,» 08 Грудня 2016. . Available: https://www.mil.gov.ua/content/yrdday/661_08122016.pdf
10. The UK Drone Association, «BVLOS Operations Forum unites innovators to unlock the next generation of aviation,» ARPAS-UK, 28 July 2022. . Available: <https://www.arpas.uk/bvlos-operations-forum-press-release>
11. European Union Aviation Safety Agency, «Drones & Air Mobility,» . Available: <https://www.easa.europa.eu/en/domains/civil-drones>
12. Federal Aviation Administration, «UAS Information Papers,» Research & Development, . Available: https://www.faa.gov/uas/research_development/information_papers. [Дата звернення: 10 Січня 2024].
13. European Union Aviation Safety Agency, «Annual Safety Review 2022,» 2022. . Available: <https://www.easa.europa.eu/en/document-library/general-publications/annual-safety-review-2022>
14. М. З. Згуровський, «Витоки української авіації,» Київський Політехнік, № 38 (2652), pp. 1-3, 2003.
15. J. G. Leishman, «Introduction to Aerospace Flight Vehicles,» EMBRY-RIDDLE Aeronautical University, January 2024. . Available: <https://eaglepubs.erau.edu/introductiontoaerospaceflightvehicles/chapter/history-of-aircraft-and-aviation>
16. Конвенція про міжнародну цивільну авіацію. Додатки 1-18, Брюссель: Міжнародна організація цивільної авіації, 1990.
17. ICAO, Uniting Aviation a United Nations Specialized Agency, «DOC Series,» . Available: <https://www.icao.int/publications/Pages/doc-series.aspx>.
18. EASA Drones team, «U-space: a quick overview,» 2021. . Available: <https://www.easa.europa.eu/en/what-u-space>
19. Міністерство інфраструктури України, «Наказ № 1009 Про затвердження Класифікації повітряного простору обслуговування повітряного руху України,» 10 грудня 2013. . Available: <https://zakon.rada.gov.ua/laws/show/z2203-13#Text>.

Received (Надійшла) 10.12.2024

Accepted for publication (Прийнята до друку) 18.03.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Шмельова Тетяна Федорівна – доктор технічних наук, професорка, професор кафедри аеронавігаційних систем, Державне некомерційне підприємство «Державний університет «Київський авіаційний інститут», Київ, Україна;

Tetiana Shmelova – Doctor of Technical Sciences, Professor, Professor of Department Air Navigation, State Noncommercial Company "State University «Kyiv Aviation Institute», Kyiv, Ukraine;

e-mail: shmelova@ukr.net; ORCID Author ID: <https://orcid.org/0000-0002-9737-6906>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57202443723>.

Шаптала Сергій Олександрович – аспірант кафедри аеронавігаційних систем, Державне некомерційне підприємство «Державний університет «Київський авіаційний інститут», Київ, Україна;

Serhii Shaptala – PhD student, Department Air Navigation, State Noncommercial Company "State University «Kyiv Aviation Institute», Kyiv, Ukraine;

e-mail: shaptala128@ukr.net; ORCID Author ID: <https://orcid.org/0000-0002-0348-4050>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=58903596200>.

Research on the task of forming air traffic flows in integrated airspace

Serhii Shaptala, Tetiana Shmelova

Abstract. Relevance. Modern trends in the development of unmanned aviation, particularly the ability to operate beyond visual line of sight, require the integration of unmanned systems into general airspace. Existing air traffic management systems were designed primarily for manned aviation, which limits their efficiency under new conditions. **Object of the research:** the process of forming air traffic flows in integrated airspace. **Purpose of the article:** to identify directions of scientific research on adaptation, coordination, and decision-making efficiency in a changing air traffic environment; to define the main principles of airspace planning and organization; and to develop a mathematical model for optimizing air traffic flow in integrated airspace. **Research results.** The article outlines the main principles of airspace planning and organization, highlights the role of the U-space system, and proposes a mathematical model for setting the optimization problem of air traffic flows. Directions for scientific research on adaptation, coordination, and decision-making efficiency under dynamic air conditions are presented. **Conclusions.** The proposed approaches enable comprehensive automation of the processes of forming air traffic flows, monitoring, and ensuring navigational safety in integrated airspace. **Application area of the results:** air traffic management systems incorporating both manned and unmanned aerial vehicles, particularly for national and international aviation regulators.

Keywords: transport technologies, navigation, air traffic control, safety, airspace, airspace planning, airspace flows, adaptation, coordination, manned aviation, unmanned aviation, integration.

С. М. Авакін, С. О. Довгополий, О. В. Запорожець, І. О. Мощенко

Харківський національний університет радіоелектроніки, Харків, Україна

КОРИГУВАННЯ ФУНКЦІЇ ПЕРЕТВОРЕННЯ ЗАСОБІВ ВИМІРЮВАННЯ З ВИКОРИСТАННЯМ НЕЙРОМЕРЕЖЕВОГО ПІДХОДУ

Анотація. У статті розглянуто способи зменшення впливу нелінійності функції перетворення засобів вимірювання на точність результату вимірювання шляхом застосування додаткового пристрою-коректора, який реалізує зворотну по відношенню до функції перетворення залежність. Метою дослідження є вивчення можливостей застосування у якості такого коректора штучних нейронних мереж, а саме багатошарового перцептрона та радіально-базисної нейромережі. За допомогою імітаційного комп'ютерного моделювання досліджено працездатність запропонованих методів коригування функції перетворення, вплив виду нелінійності на якість такого коригування. Здійснено порівняльний аналіз з традиційними підходами, а саме коректором на базі поліноміального апроксиматора. Результати моделювання свідчать про те, що точність нейромережових коректорів не поступається точності поліноміального коректора, а у деяких випадках навіть вище. Це відкриває перспективи ширшого застосування у вимірювальній техніці таких сучасних методів обробки даних вимірювань, як штучні нейронні мережі.

Ключові слова: нелінійність, коригування, функція перетворення, штучна нейронна мережа, багатошаровий перцептрон, радіально-базисна нейромережа, навчання.

Вступ

Кілька останніх десятиліть ознаменувались стрімким розвитком інформаційних технологій, які радикально і докорінно змінили не тільки звичний уклад людського життя, а й цілі сектори економіки у більшості розвинених країн світу. Безумовним досягненням наукової думки стали успіхи в розробці штучного інтелекту. Дослідження в царині штучних нейронних мереж, глибокого навчання, фаззі-систем дозволили реалізувати досить складні методи математичної обробки та аналізу великих масивів даних та розв'язувати цілий ряд задач класифікації, оптимізації, управління, розпізнавання образів, ідентифікації, апроксимації складних нелінійних залежностей, діагностики тощо.

Галузь вимірювань та вимірювальної техніки теж не залишається осторонь і намагається активно використовувати більш ефективні методи обробки інформації, які дозволять підвищити точність і достовірність результатів вимірювань. Роль таких методів стає більш важливою з ускладненням вимірювань і підвищенням вимог до їх точності.

Нові напрямки прикладної математики, такі як теорія штучних нейронних мереж, інтервальный аналіз, робасна та непараметрична статистика, нечітка логіка, вейвлет-аналіз та ряд інших, дають математичний апарат для розв'язання тих задач, для яких класичні методи обробки даних є непридатними або малоефективними. До розробки нових методів обробки вимірювальної інформації спонукають потреби сучасної вимірювальної практики, для якої є характерним ускладнення вимірювальних задач і засобів вимірювання.

Розширення можливостей інтелектуальної обробки даних та результатів вимірювання тісно пов'язано з розвитком методології вимірювань, удосконаленням математичних методів, а також широ-

ким впровадженням засобів обчислювальної техніки у вимірювальний ланцюг, включаючи мікропроцесори і мікроконтролери.

Постановка задачі

Для засобу вимірювання одними з найбільш важливих є метрологічні характеристики, що впливають на результат вимірювання та точність цього результату. Однією з таких характеристик є номінальна статична функція перетворення засобу вимірювання (інші її назви – рівняння перетворення, градувальна характеристика). Вона встановлює залежність $y = F(x)$ інформативного параметра вихідного сигналу y від значення інформативного параметра вхідного сигналу x .

Зазвичай до функції перетворення висувається вимога лінійності у робочому діапазоні вимірювань. Однак, досить часто при вирішенні практичних вимірювальних задач доводиться мати справу з вимірювальними пристроями, які мають нелінійну функцію перетворення. Так, наприклад, аналогові електронні мегаомметри в режимі вимірювання великих опорів (одиниці і десятки мегаомів) мають зворотну нерівномірну шкалу. Чутливість приладу на різних ділянках такої шкали буде різною, що не зовсім зручно і при взятті відліку, і при оцінюванні похибки результату вимірювання. Окремі елементи вимірювального тракту, такі як напівпровідникові терморезистори або діоди, теж мають суттєво нелінійну робочу характеристику, що відбивається на функції перетворення приладу. Іноді наблизити функцію перетворення до лінійної вдається за допомогою схемотехнічних рішень, але не завжди. Зазвичай на практиці функцію перетворення апроксимують лінійною залежністю

$$\hat{y} = a + bx,$$

невідомі коефіцієнти якої a і b знаходять за допомогою методу найменших квадратів. Проте у випадку,

коли нелінійність функції перетворення є суттєвою, такий підхід не дає бажаного результату, оскільки внаслідок відхилення номінальної функції перетворення від реальної виникає велика методична похибка (рис. 1).

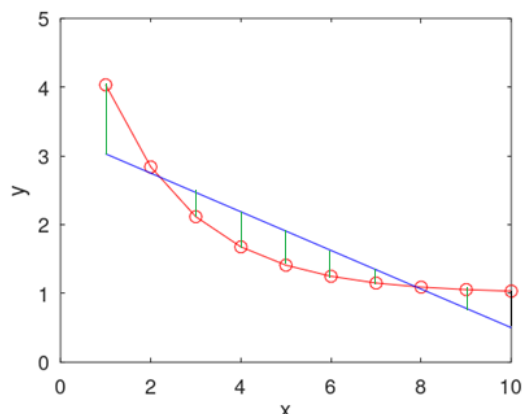


Рис. 1. Похибки лінеаризації функції перетворення

Для зменшення цієї похибки можна спробувати привести нелінійну функцію $y = F(x)$ до лінійного вигляду $\tilde{y} = a + b\tilde{x}$ шляхом заміни змінних

$$\tilde{y} = \varphi(y), \quad \tilde{x} = \psi(x)$$

з подальшим визначенням коефіцієнтів a і b лінійної функції за методом найменших квадратів [1]. Але у цьому випадку потрібно заздалегідь знати, як виглядає нелінійна функція перетворення, тобто мати апріорну інформацію про структуру математичної моделі засобу вимірювання. Обґрунтований вибір виду цієї нелінійної залежності є доволі складною задачею, яка погано піддається формалізації і здійснюється, виходячи з відомих фізичних законів або особистого досвіду фахівця, який займається розв'язанням задачі.

Одним з відомих підходів до зменшення впливу нелінійності функції перетворення на похибку результату вимірювання, який можна вважати універсальним, є алгоритмічна корекція функції перетворення. Послідовно із засобом вимірювання включається спеціальний пристрій-коректор, який здійснює зворотне по відношенню до його характеристики перетворення $\hat{x} = F^{-1}(y)$. В результаті такої корекції ми отримуємо оцінку $\hat{x}$ вхідного сигналу (вимірюваної величини) x , а результуюча функція

перетворення стає лінійною. Важливою додатковою умовою є інваріантність такого перетворювача до виду нелінійної функції, що коригується, тобто можливість адаптації до довільної функції перетворення.

Метою даної статті є дослідження адаптивної системи коригування функції перетворення засобу вимірювання, яка буде використовувати у якості пристрою корекції штучну нейронну мережу та забезпечуватиме лінійність функції перетворення в усьому діапазоні можливих значень вимірюваної вхідної величини.

Коригування функції перетворення за допомогою штучної нейронної мережі

Штучні нейронні мережі (ШНМ) – це обчислювальні структури, побудовані за принципом функціонування біологічних нейронних мереж, утворених клітинами мозку живих організмів. Відмінною особливістю ШНМ є здатність до навчання, що зробило їх пріоритетним напрямком досліджень в області штучного інтелекту. Теорія ШНМ стрімко розвивається в останні роки, що сприяє підвищенню зацікавленості до їх використання в різноманітних галузях науки, техніки, економіки, медицини, військової справи тощо. Завдяки таким властивостям, як висока надійність функціонування, завадостійкість, здатність до узагальнення, можливість реалізації складних багатовимірних відображень, нейромережі широко застосовуються для розпізнавання образів і класифікації, прийняття рішень і управління, оптимізації, апроксимації функцій, прогнозування, фільтрації, організації пам'яті [2-5].

Виходячи з вищевказаних особливостей ШНМ, зокрема здатності бути універсальним апроксиматором складних нелінійних залежностей, саме нейромережу найбільш доцільно використовувати у якості коригуючого пристрою, як було показано у ряді досліджень [6-10]. На рис. 2 наведено структуру такої адаптивної системи коригування функції перетворення засобу вимірювання. Дію внутрішніх та зовнішніх випадкових факторів на вимірювальний процес представлено адитивним шумом $\xi(t)$ на виході засобу вимірювання, де t – поточний час. Вихідний сигнал $y(t)$ подається на вхід нейромережі-коректора, яка здійснює зворотне перетворення та формує оцінку вхідного сигналу (вимірюваної величини) $\hat{x}(t)$.

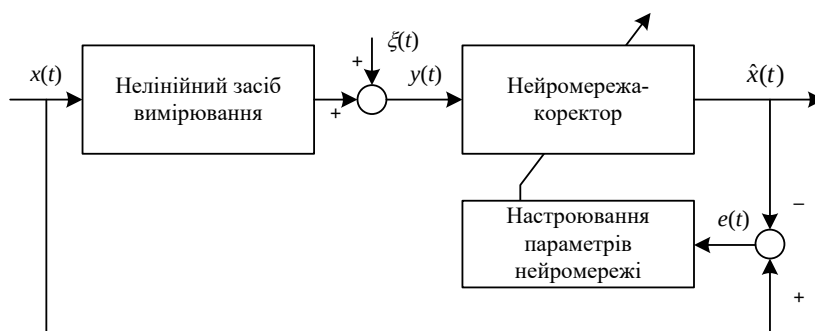


Рис. 2. Структура адаптивної системи корекції нелінійності функції перетворення засобу вимірювання

Алгоритм навчання здійснює настроювання синаптичних вагових коефіцієнтів нейромережі таким чином, щоб забезпечити найкраще наближення її вихідного сигналу $\hat{x}(t)$ до відомого значення вхідного сигналу $x(t)$. При цьому мінімізується середньоквадратичне значення помилки корекції $e(t) = x(t) - \hat{x}(t)$ за допомогою однієї з відомих градієнтних процедур оптимізації.

Пристрій корекції на базі багатошарового персептрона

Багатошарова нейронна мережа прямого поширення, або багатошаровий персептрон (Multilayer Perceptron, MLP) складається з декількох послідовно з'єднаних між собою шарів формальних нейронів: вхідного, прихованих та вихідного (рис. 3). Нейрони одного шару не зв'язані між собою, виходи нейронів n -го шару надходять на входи нейронів наступного $n+1$ -го шару. Вхідний векторний сигнал надходить на входи нейронів першого шару, а вихідні сигнали останнього шару формують вихідний векторний сигнал мережі. Конфігурація MLP визначається кількістю шарів, кількістю нейронів у кожному шарі та функціями активації нейронів. Навчання полягає у коригуванні коефіцієнтів зв'язків між нейронами таким чином, щоб забезпечити необхідні значення вихідних сигналів мережі. У процесі навчання багатошарова нейромережа здатна виявляти складні залежності між вхідними та вихідними сигналами та здійснювати узагальнення. При використанні парадигми навчання «з учителем» датасет, на якому навчається нейромережа, повинен бути розмічений або промаркований, тобто містити правильні відповіді (виходи нейромережі) на кожен вхідний зразок навчальної послідовності.

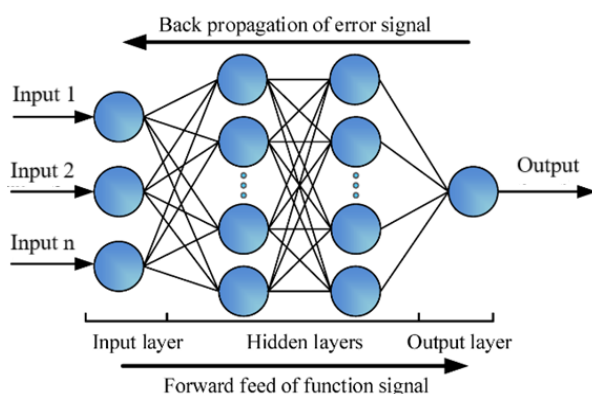


Рис. 3. Багатошаровий персептрон

Пропонується реалізувати пристрій для коригування нелінійності функції перетворення засобу вимірювання на базі тришарового персептрона, структуру якого наведено на рис. 4.

Вихідний шар персептрона утворено одним нейроном, який формує сигнал $\hat{x}$ як зважену суму вихідних сигналів нейронів прихованого шару

$$\hat{x} = \sum_{j=1}^n V_j O_j, \quad (1)$$

де O_j – вихідний сигнал j -го нейрона прихованого шару; V_j – синаптична вага j -го входу нейрона вихідного шару; n – кількість нейронів у прихованому шарі.

Прихований шар MLP утворено нейронами з сигмоїдальними функціями активації. Кожен нейрон цього шару описується такими рівняннями

$$O_j = \frac{1}{1 + e^{-S_j}},$$

$$S_j = \sum_{i=1}^m W_{ij} O_i,$$

де O_i – вихідний сигнал i -го нейрона вхідного шару; W_{ij} – синаптична вага i -го входу j -го нейрона прихованого шару; m – кількість нейронів у вхідному шарі.

Вхідний шар нейронів утворено самими вхідними сигналами нейромережі, якими у рамках задачі, що розглядається, є вихідний сигнал засобу вимірювання у та постійний сигнал, який дорівнює одиниці і вводиться для опису постійного зміщення.

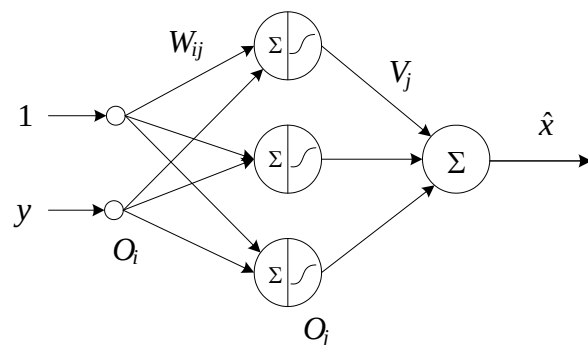


Рис. 4. MLP-коректор

Навчання MLP базується на мінімізації середньоквадратичної помилки

$$E = \frac{1}{2} \sum_{k=1}^N e^2(k) = \frac{1}{2} \sum_{k=1}^N (x(k) - \hat{x}(k))^2 \quad (2)$$

шляхом настроювання синаптичних вагових коефіцієнтів нейронів за допомогою градієнтних методів. Найбільш популярними є два алгоритми навчання:

– метод зворотного поширення помилки (back propagation);

– алгоритм Левенберга-Марквардта (Levenberg-Marquardt algorithm, LMA), який є комбінацією методу Гауса-Ньютона та методу градієнтного спуску.

У нашому випадку для навчання MLP доцільніше використовувати алгоритм Левенберга-Марквардта, оскільки він має вищу швидкість збіжності порівняно з іншими градієнтними методами оптимізації.

Датасет для процедури навчання нейромережі-коректора буде складатися з пар значень

$$\{ \langle y(1), x(1) \rangle, \langle y(2), x(2) \rangle \dots \langle y(N), x(N) \rangle \},$$

які отримують при калібруванні засобу вимірювання, подаючи на вхід зразковий сигнал $x(k)$ і отримуючи відповідне йому значення вихідного сигналу $y(k)$.

Пристрій корекції на базі радіально-базисної нейромережі

Ще одним універсальним апроксиматором складних функціональних залежностей є радіально-базисна нейронна мережа (Radial Basis Function Net, RBFN). Архітектура такої мережі складається з трьох шарів: вхідного, на який поступає вектор вхідних сигналів; прихованого, що складається з нейронів радіального типу; вихідного, який формує зважену лінійну комбінацію виходів нейронів прихованого шару (рис. 5).

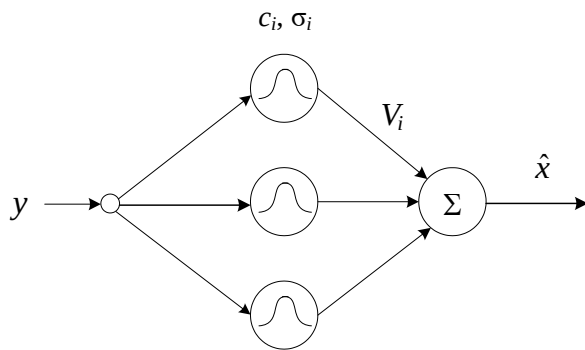


Рис. 5. RBFN-коректор

Якщо в якості функції активації нейронів прихованого шару використовувати функцію Гауса, то рівняння такого перетворювача матиме вигляд

$$\hat{x} = \sum_{i=1}^n V_i \exp\left(-\frac{(y - c_i)^2}{2\sigma_i^2}\right), \quad (3)$$

де c_i – центр i -ої базисної функції, а σ_i – її радіус (ширина). Навчання RBFN полягає у визначенні лінійних вагових коефіцієнтів V_i вихідного нейрона, центрів c_i та радіусів σ_i нейронів прихованого шару. При цьому можливі такі варіанти:

1) задаються фіксовані значення центрів та радіусів нейронів прихованого шару, а вагові коефіцієнти вихідного нейрона визначаються шляхом навчання;

2) центри і радіуси визначаються шляхом самонавчання (найчастіше за допомогою методів кластеризації), а потім корекцією ваг вихідного нейрона мінімізується цільова функція (2);

3) всі параметри мережі визначаються за допомогою навчання з учителем, наприклад методом зворотного поширення помилок.

Результати експериментальних досліджень

Для вивчення властивостей запропонованих нейромережевих систем коригування функції перетворення засобів вимірювання застосовувалось комп'ютерне моделювання в середовищі MATLAB з використанням бібліотеки Neural Network Toolbox.

У процесі моделювання досліджувався вплив виду нелінійності на якість коригування функції перетворення, при цьому для моделювання нелінійного засобу вимірювання використовувались такі види функцій [1]:

- 1) синусоїдальна $y = \sin(ax + b)$;
- 2) степенева $y = ax^b$;
- 3) гіперболічна $y = a + b/x$;
- 4) дробово-лінійна I $y = \frac{1}{a + bx}$;
- 5) дробово-лінійна II $y = \frac{cx}{a + bx}$;
- 6) показникова I $y = ae^{bx}$;
- 7) показникова II $y = ae^{b/x}$;
- 8) логарифмічна I $y = a + b \ln x$;
- 9) логарифмічна II $y = a + b/\ln x$.

Адитивний шум на виході засобу вимірювання моделювався гаусівським випадковим процесом з нульовим середнім і стандартним відхиленням 0,005. Якість коригування оцінювалась середньоквадратичним відхиленням оцінки сигналу після коригування від його дійсного значення

$$s = \sqrt{\frac{1}{N-1} \sum_{k=1}^N (x(k) - \hat{x}(k))^2}. \quad (4)$$

Для порівняння здійснювалось моделювання аналогічної системи коригування функції перетворення на базі поліноміального апроксиматора 5-го степеню. Результати моделювання наведено на рисунках 6-10 та в табл. 1.

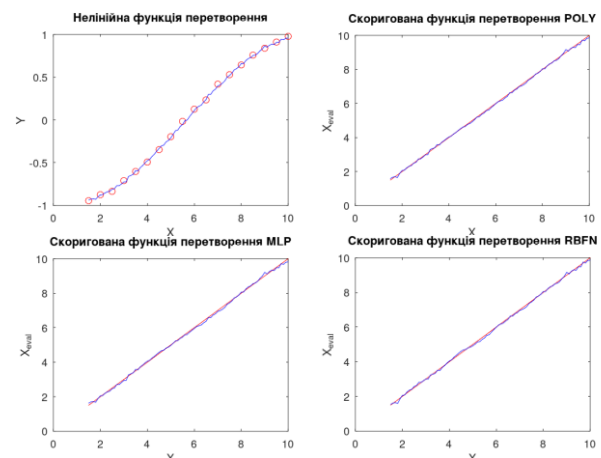


Рис. 6. Коригування синусоїдальної функції перетворення

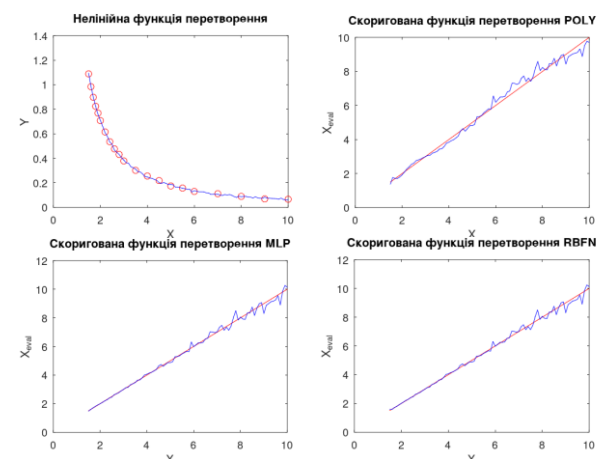


Рис. 7. Коригування степеневі функції перетворення

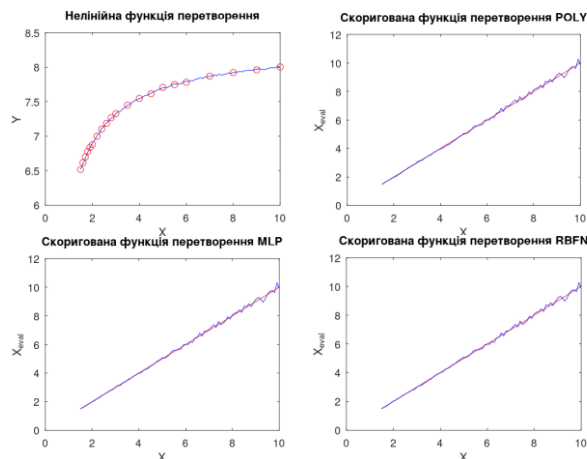


Рис. 8. Коригування дробово-лінійної II функції перетворення

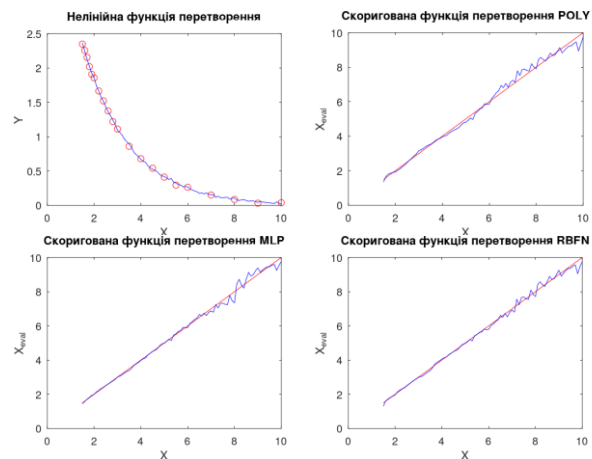


Рис. 9. Коригування показникової I функції перетворення

Таблиця 1 – Середньоквадратична похибка корекції для різних нелінійних функцій і коректорів

Функція	Формула	Середньоквадратична похибка корекції		
		POLY	MLP	RBFN
1. Синусоїдальна	$y = \sin(0,3x - 1,7)$	0,0615	0,0752	0,0700
2. Степенева	$y = 2 \cdot x^{-1,5}$	0,2621	0,1992	0,1983
3. Гіперболічна	$y = 5 + 3/x$	0,1360	0,1374	0,1414
4. Дробово-лінійна I	$y = \frac{1}{0,05 + 0,1x}$	0,0974	0,0985	0,1016
5. Дробово-лінійна II	$y = \frac{10x}{0,5 + 1,2x}$	0,1076	0,1061	0,1053
6. Показникова I	$y = 5 \cdot e^{-0,5 \cdot x}$	0,2160	0,1870	0,1611
7. Показникова II	$y = 5 \cdot e^{0,5/x}$	0,1780	0,1727	0,1590
8. Логарифмічна I	$y = 5 + 3 \cdot \ln x$	0,0704	0,0715	0,0911
9. Логарифмічна II	$y = 5 + 3/\ln x$	0,3099	0,2814	0,2849

Кружечками на графіку функції позначено точки навчальної вибірки.

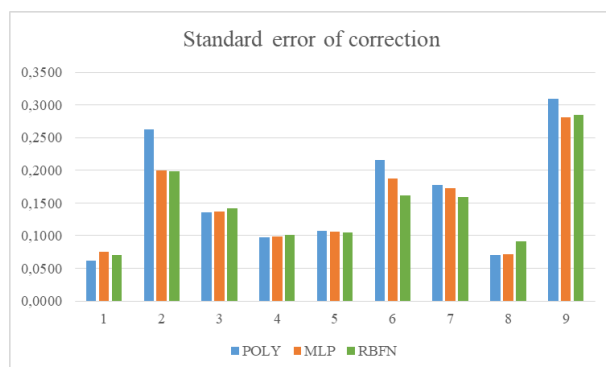


Рис. 10. Похибки коригування

Висновки

Аналіз результатів моделювання, наведених на рис. 6-10 та в табл.1, свідчить про те, що запропонована нейромережева система коригування нелінійності функції перетворення засобу вимірювання на базі MLP та RBFN є придатною для коригування

досить широкого класу нелінійних функцій перетворення, які зустрічаються у вимірювальних задачах. Порівняння характеристик нейромережевих коректорів і аналогічної системи на базі поліноміального апроксиматора показує, що середньоквадратична похибка корекції для деяких видів нелінійної функції перетворення у традиційного та нейромережевих коректорів відрізняється несуттєво, а для інших нейромережевий підхід показує кращі результати.

З цього можна зробити висновок, що результати моделювання повністю підтверджують працездатність запропонованої адаптивної системи коригування нелінійності функції перетворення на базі ШНМ і узгоджуються з теоретичними припущеннями. Перевагою запропонованого підходу є інваріантність нейромережевих коректорів до виду нелінійної характеристики засобу вимірювання та можливість синтезувати такі системи через навчання, без залучення складних методів проєктування. Це значно розширює можливості застосування таких систем у метрологічній практиці, зокрема в каналах інформаційно-вимірювальних систем, де досить просто

реалізувати методи цифрової обробки вимірювальної інформації. Використання запропонованого коректора на базі нейромережевого підходу дозволить суттєво зменшити систематичну похибку вимірювання, викликану невідповідністю номінальної та реальної функції перетворення засобу вимірювання.

СПИСОК ЛІТЕРАТУРИ

1. Грановский В.А., Сиряя Т.Н. Методы обработки экспериментальных данных при измерениях. Л.: Энергоатомиздат, 1990. 288 с.
2. Bishop C.M. Pattern Recognition and Machine Learning. Springer. 2006. 738 p.
3. Haykin S. Neural Networks and Learning Machines. 3 rd. ed. Pearson, 2009. 906 p.
4. Теорія та практика нейронних мереж : навч. посіб. /Л. М. Добровська, І. А. Добровська. К.: НТУУ «КПІ» Вид-во «Політехніка», 2015. 396 с.
5. Hornik K., Stinchcombe M., White H. Multilayer feedforward networks are universal approximators. *Neural Networks*. 1989. 2(5). pp. 359–366.
6. Запорожец О.В., Коротенко В.А., Овчарова Т.А. Компенсация нелинейности функции преобразования измерительных устройств с помощью искусственной нейронной сети. *Системы управления, навигации та зв'язку*. Вип. 4(16). 2010. С. 99–103.
7. Дегтярев А.В., Запорожец О.В., Овчарова Т.А. Адаптивная система компенсации нелинейности функции преобразования измерительных устройств на базе трехслойного персептрона. *Электротехнические и компьютерные системы*. № 06(82). 2012. С. 235–241.
8. Запорожец О.В., Овчарова Т.А. Компенсация нелинейности динамических измерительных преобразователей при помощи нейросетевых моделей. *Метрологія та прилади*. Вип. № 1/П/(45). 2014. С. 74–77.
9. Запорожец О.В., Овчарова Т.А., Руженцев И.В. Компенсация нелинейности полупроводникового терморезистора с помощью искусственной нейронной сети. *Системы обробки інформації*. 2015. №6 (131). С. 64–67.
10. Zaporozhets O.V., Shtefan N.V. Using Artificial Neural Network for Compensation of Semiconductor Thermistor Nonlinearity. *2019 IEEE 8th International Conference on Advanced Optoelectronics and Lasers (CAOL)*, Sozopol, Bulgaria, 6–8 Sept. 2019. PP. 703–706.

Received (Надійшла) 23.01.2025

Accepted for publication (Прийнята до друку) 09.04.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Авакін Сергій Миколайович – студент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Serhii Avakin – student, Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: serhii.avakin@nure.ua; ORCID Author ID: <https://orcid.org/0009-0004-4372-1913>.

Довгополий Сергій Олександрович – аспірант кафедри інформаційно-вимірювальних технологій, Харківський національний університет радіоелектроніки, Харків, Україна;

Serhii Dovhopolyi – PhD student, Department of Information and Measurement Technologies, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: serhii.dovhopolyi@nure.ua; ORCID Author ID: <https://orcid.org/0009-0008-8423-8812>.

Запорожець Олег Васильович – доктор філософії, доцент, доцент кафедри інформаційно-вимірювальних технологій, Харківський національний університет радіоелектроніки, Харків, Україна;

Oleg Zaporozhets – PhD, Associate Professor, Department of Information and Measurement Technologies, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: oleg.zaporozhets@nure.ua; ORCID Author ID: <https://orcid.org/0000-0002-7831-8479>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=15728942500>.

Мощенко Інна Олексіївна – доктор філософії, доцент, доцент кафедри інформаційно-вимірювальних технологій, Харківський національний університет радіоелектроніки, Харків, Україна;

Inna Moshchenko – PhD, Associate Professor, Department of Information and Measurement Technologies, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: inna.moshchenko@nure.ua; ORCID Author ID: <https://orcid.org/0000-0002-2738-0037>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57547207800>.

Correction of transformation function in measurement devices using neural network approach

S. Avakin, S. Dovhopolyi, O. Zaporozhets, I. Moshchenko

Abstract. The article considers ways to reduce the impact of the nonlinearity of the transformation function of measuring instruments on the measurement result accuracy by using an additional corrector device that implements the inverse dependence of the transformation function. The purpose is to study the possibilities of using artificial neural networks as such a corrector, namely a multilayer perceptron and a radial-basis neural network. Using computer simulation modeling, the performance of the proposed methods for correcting the transformation function and the influence of the type of nonlinearity on the quality of such correction were investigated. A comparative analysis with traditional approaches, namely a corrector based on a polynomial approximator was carried out. The modeling results indicate that the accuracy of neural network correctors is not inferior to the accuracy of the polynomial corrector, and in some cases even higher. This opens up prospects for wider application in measuring equipment of such modern methods of processing measurement data as artificial neural networks.

Keywords: nonlinearity, correction, transformation function, artificial neural network, multilayer perceptron, radial basis neural network, learning.

С. Я. Бовчалюк¹, Н. І. Бовчалюк², О. Ю. Дрозд¹

¹ Харківський національний університет радіоелектроніки, Харків, Україна

² Державний біотехнологічний університет, Харків, Україна

КОНЦЕПЦІЯ «МОДУЛЬНОЇ АРХІТЕКТУРИ» КЕРУЮЧИХ ПРИСТРОЇВ ПАРАЛЕЛЬНОЇ ДІЇ

Анотація. Актуальність. Публікації, що присвячені інформаційній технології паралельного логічного керування, або окремим її складовим частинам (керуючим пристроям з паралельною архітектурою, мові і технології їх програмування) показують, що вона є перспективною у плані реалізації систем керування переважно об'єктами критичної дії. У той же час спроби її застосування для об'єктів загального промислового сектору зіштовхуються з проблемами обмеженої функціональності, або скоріше з необхідністю розширення виконуваних функцій, тобто адаптації структури контролерів паралельної дії для розв'язання конкретних задач. Прикладами таких удосконалень можуть бути введення функціоналу нечіткого логічного висновку, внутрішніх таймерів, лічильників до структури класичного керуючого автомата паралельної дії. Таким чином реалізація «гнучкої» структури подібних автоматів дозволить значно спростити задачу побудови контролерів паралельної дії із заданими функціональними можливостями. **Метою даної роботи** є теоретичне обґрунтування можливості побудови «гнучкої» структури логічних керуючих автоматів паралельної дії, формулювання концепції «модульної архітектури», а також реалізація структури перспективного ЛКА ПД у відповідності до цієї концепції. **Висновок.** У результаті виконаних досліджень усунуто такий недолік керуючих пристроїв з паралельною архітектурою, як їх фактична побудова за принципом «жорсткої логіки», або недовикористання можливостей створення «гнучких» систем керування на базі сучасних кристалів ПЛІС. Запропоновано концепцію «модульної архітектури» керуючих пристроїв паралельної дії, що дозволяє створювати «гнучку» структуру відповідних логічних керуючих автоматів. На базі цієї концепції синтезовано структуру модульного керуючого автомата з можливістю реалізації заданої функціональності у залежності від конкретних задач системи керування, що будується на його основі.

Ключові слова: інформаційна технологія паралельного логічного керування, логічний керуючий автомат паралельної дії, програмований логічний контролер паралельної дії, ПЛІС-контролер, структура ЛКА ПД, модульний логічний керуючий автомат.

Вступ

З часу публікації структури сучасного керуючого пристрою паралельної дії, що вже можна назвати класичною [1-3], пройшло майже двадцять років. За цей час було виконано низку досліджень результатом яких стало формування інформаційної технології паралельного логічного керування (ІТПЛК) [4], як поєднання технічних пристроїв (логічних керуючих автоматів паралельної дії (ЛКА ПД), програмованих логічних контролерів паралельної дії (ПЛК ПД), ПЛІС-контролерів паралельної дії), мови (ЯППЛК, ЯПЛК-М) та технології їх програмування (Technological Visual Programming – TVP).

У низці останніх робіт, що стосуються ІТПЛК, показано основні тенденції та напрямки розвитку керуючих пристроїв з паралельною архітектурою – програмованих логічних контролерів паралельної дії або ПЛІС-контролерів паралельної дії. Так в [5] визначено три основні шляхи, або траєкторії, що пройшли складові ІТПЛК, які коротко можна сформулювати наступним чином:

- 1) реалізація ПЛК ПД на більш сучасній елементній базі [6-8];
- 2) вдосконалення мови і технології програмування ПЛК ПД [9-11];
- 3) розвиток і вдосконалення внутрішньої організації логічних керуючих автоматів паралельної дії [1, 3, 12-14].

Перша складова, на теперішній час, не має перспектив розвитку, оскільки відбувся остаточний перехід від побудови ПЛК ПД на дискретних ІМС

малого ступеня інтеграції, до їх реалізації на базі кристалів ПЛІС. На думку авторів, станом на зараз не існує мікроелектронних структур, що є більш перспективними для побудови керуючих пристроїв з паралельною архітектурою, ніж кристали сучасних ПЛІС. Залишається лише питання вибору виробника кристалу (AMD-Xilinx або Intel-Altera) та мови опису апаратних засобів (AHDL, VHDL, VerilogHDL).

Друга вказана складова є достатньо перспективною але, переважно, передбачає адаптацію технології TVP до змін, що відбуваються у структурній організації ПЛК ПД.

Прикладом можуть бути підтримка з боку TVP програмованих користувачем таймерів [14], або, у перспективі, лічильників, тощо. Тобто друга складова є повністю залежною від розвитку ІТПЛК за третьою складовою.

Таким чином саме третя вказана траєкторія є найбільш перспективною у плані розвитку технології ІТПЛК. Але саме цей напрямок, має усі ознаки стагнації, бо, як показано в [14]: «аналіз моделей і структур програмованих керуючих автоматів паралельної дії свідчить про їх «негнучкість», «закритість» або, фактично, реалізацію за допомогою концепції «жорсткої логіки». При цьому введення додаткового функціоналу ЛКА ПД у вигляді таймерів або, у перспективі, лічильників, елементів виконання арифметичних операцій, тощо, не ліквідує самої суті проблеми. Справа у тому, що сам підхід до створення ПЛК ПД повторює побудову ПЛК на дискретних елементах. Таким чином явно недовикористовується «гнучкість» ПЛІС-технологій, тобто

можливість створювати спеціалізовану архітектуру контролера під конкретну задачу, а не тільки пристосовувати вже існуючу «жорстку» структуру та створювати спеціалізоване програмне забезпечення для нього.

Перш ніж переходити безпосередньо до розв'язання вказаної вище проблеми, необхідно обов'язково зазначити дуже важливі особливості ЛКА ПД, які власне і визначають основні переваги керуючих пристроїв паралельної дії у порівнянні з класичними пристроями керування послідовної дії. Вони свого часу були сформульовані професором І. О. Фурманом: «принципова відмінність описаної структури від відомих структур мікропрограмних автоматів полягає в тому, що управління процесом відпрацювання програми відбувається не від імпульсів тактового генератора, як у мікропрограмних автоматах, побудованих за класичною схемою Вілкса і її удосконаленим варіантам, і не потоком команд, як в ПЛК послідовної дії, а потоком станів входів, тобто потоком, що формують в асинхронній послідовності інтервалів дискретного автоматного часу «дозволені» (такі що відповідають запрограмованим) комбінації фактичних станів керованого об'єкта та зовнішнього середовища. Крім того, паралельна реалізація операцій над булевими векторами в автоматі дозволяє оперувати машинними словами довільної довжини». [4, 15]. Таким чином, головною відмінністю ЛКА ПД, порівняно з автоматами послідовної дії, є можливість одночасного («за один такт дискретного автоматного часу») обслуговування практично необмеженого числа контрольованих входів і формування команд керування для практично необмеженої кількості виконавчих механізмів.

Метою роботи є теоретичне обґрунтування можливості побудови «гнучкої» структури логічних керуючих автоматів паралельної дії, формулювання концепції «модульної архітектури», а також реалізація структури перспективного ЛКА ПД у відповідності до цієї концепції.

Основна частина

Базовою, для побудови сучасних керуючих пристроїв паралельної дії, можна уважати структуру, що показана на рис. 1 [3]. Саме її покладено в основу промислових зразків ПЛІС-контролерів паралельної дії на кристалах компанії Intel-Altera [8]. Також результатом її розвитку є логічні керуючі автомати з розширеними функціональними можливостями [12-14].

Основу показаного керуючого автомату складають чотири блоки пам'яті: станів, переходів, команд і заборонених комбінацій – БПС, БПП, БПК і БПЗК відповідно. Базову логіку роботи ЛКА забезпечують схема порівняння (СП), лічильник адреси (ЛА) та блок логічного керування (БЛК). Ще три блоки – індикації (БІ), вибору операцій (БВО) та вихідний регістр (ВР) призначені для реалізації додаткових індикаторних функцій, розширення функціоналу логічного порівняння та фіксацію і (або) блокування) команд керування відповідно.

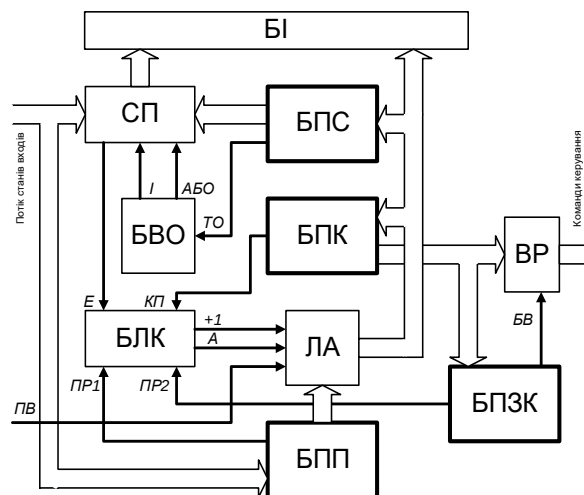


Рис. 1. Структура ЛКА паралельної дії

Принцип реалізації базових функцій логічного керування відбувається за рахунок перевірки істинності рівнянь (1) і (2), що зафіксовані у блоці логічного керування:

$$A = КП \vee ПР1 \vee ПР2, \quad (1)$$

$$+1 = E \wedge \overline{КП} \wedge \overline{ПР1} \wedge \overline{ПР2}. \quad (2)$$

де $КП$ – ознака кінця підпрограми; $ПР1$ – ознака переривання від БПП; $ПР2$ – ознака переривання від БПЗК; A – початкова адреса підпрограми; E – сигнал еквівалентності; $+1$ – сигнал переходу до наступного рядка.

Дуже стисло розглянемо принцип роботи ЛКА ПД, з яким детально можна ознайомитись у [3, 4].

Процес відпрацювання керуючої програми складається з двох етапів або частин:

1) аналіз комбінацій станів датчиків умов переходів (станів зовнішнього середовища) і формування початкової адреси підпрограми;

2) власне відпрацювання обраної підпрограми.

За вибір початкової адреси підпрограми відповідає блок пам'яті переходів, який, у разі появи на його вході однієї із запрограмованих комбінацій, встановлює ЛА у відповідний даний комбінації стан. Таким чином реалізується виконання рівняння (1), тобто істинною стає рівність $КП=A$ і БЛК формує сигнал «А» (Адреса), за яким лічильник адреси переадресовує БПС і БПК на перший рядок обраної підпрограми. В останньому рядку кожної підпрограми, а також у нульовому рядку програми записується тільки ознака кінця підпрограми «КП», що використовується як дозвіл переходу керуючого автомата до відпрацювання будь-якої із записаних у блоки пам'яті підпрограм.

Власне відпрацювання обраної підпрограми відбувається за рахунок реалізації рівняння (2). Якщо рівність виконується, то БЛК формує сигнал «+1», за яким ЛА адресує БПС і БПК на наступний $(i+1)$ рядок поточної підпрограми.

Умовою формування сигналу «+1» є поява сигналу «Е», який може бути сформований двома способами. Якщо на певному кроці керуючої програми необхідно порівнювати фактичний стан усіх датчи-

ків циклу з їх очікуваними значеннями, то до останнього стовпця i -го рядка, що записаний до блоку пам'яті станів, записується відповідна ознака і блок вибору операції формує сигнал « $I=I$ ». Цей сигнал перемикає схему порівняння на реалізацію логічної операції « I », тобто сигнал еквівалентності « E » на її виході з'явиться лише у разі збігу усіх фактичних станів датчиків циклу з їх очікуваними значеннями, записаними до i -го рядку БПС.

Якщо для переходу на наступний крок підпрограми достатньо спрацювання лише одного датчика, то БВО формує сигнал «АБО=1», що перемикає схему порівняння на реалізацію логічної операції «АБО».

Ознаки переривань «*ПР1*», «*ПР2*», сигнали «*ПВ*» і «*БВ*», не формують базову логіку роботи автомата, але є необхідними для практичної реалізації керуючого пристрою.

На рис. 2 і 3 показано варіанти ЛКА ПД з функціями реалізації функцій нечіткого логічного висновку – АПДН [12, 13].

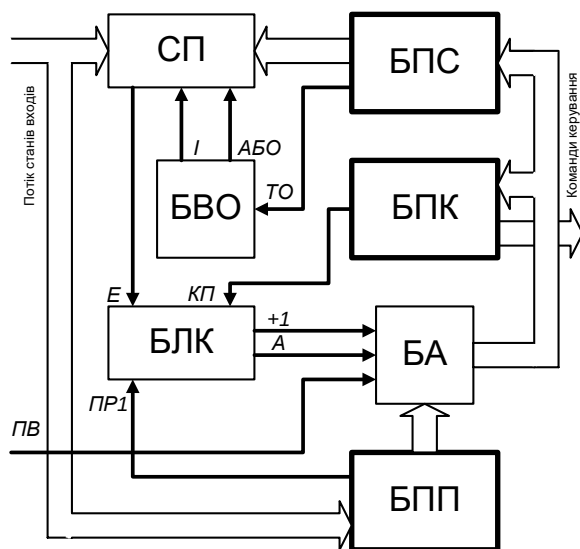


Рис. 2. Структура АПДН
для реалізації функцій збору даних

Переважна більшість блоків ЛКА ПД у структурі АПДН збереглась, хоча деякі з них частково або повністю змінили свою функціональність.

Розглянемо особливості та відмінності функціонування АПДН у порівнянні з ЛКА ПД для можливостей реалізації функцій збору інформації (рис. 2).

Так блок пам'яті переходів БПП, через відсутність необхідності формування алгоритмів з розгалуженнями, втрачає свою функціональність з формування адрес переходів, натомість разом з блоком БПК – команд, виконує функції зберігання таблиці виконання нечіткого логічного висновку. Оскільки саме процедура нечіткого логічного висновку є базовим функціоналом АПДН, то на ці два блоки пам'яті припадає основна частина програмного навантаження.

Блок пам'яті станів БПС так само виконує функції зі зберігання послідовності виконуваних команд, але у випадку відсутності такої послідовності

(тобто наявності лише ситуативного реагування на потік вхідних даних), цей блок може бути виключено зі структури АПДН.

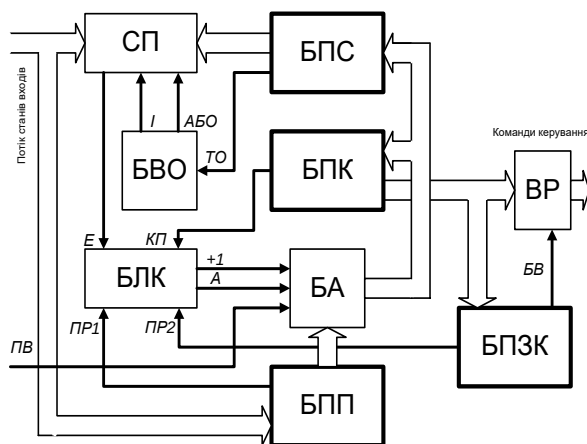


Рис. 3. Структура АПДН
для реалізації функцій керування

У структурі АПДН присутній новий блок – блок адресації (БА). Цей блок виконує функції подібні до лічильника адреси (ЛА) у структурі ЛКА ПД. Така зміна пов'язана з тим, що базовою функцією АПДН є формування нечіткого логічного висновку (зв'язка блоків БПП, БПК та БА), а не послідовне відпрацювання рядків програми (де основною є зв'язка блоків БПС та ЛА).

Окремого розгляду потребує той факт, що зі структури класичного ЛКА ПД повністю виключено елементи, що забезпечували неவிдачу заборонених (таких що можуть призвести до аварії на керованому об'єкті) комбінацій команд керування виконавчими механізмами – блок пам'яті заборонених комбінацій (БПЗК) і вихідний реєстр (ВР). Таке рішення обґрунтовано тим, що у випадку реалізації елементів збору даних, проблема видачі заборонених комбінацій команд керування відсутня. Це можна пояснити тим, що як таких керуючих механізмів може і не бути, а вихідними сигналами є інформаційний потік вихідних даних, для якого поняття «забороненої комбінації» відсутнє в принципі. У випадку побудови АПДН, що орієнтований на реалізацію систем контролю або керування, у його структурі також присутні відповідні блоки, які забезпечують неவிдачу заборонених комбінацій команд керування виконавчими механізмами (рис. 3).

Розвитком розглянутих структур, і на даний час останнім із відомих, є керуючий автомат паралельної дії із елементами реалізації внутрішніх програмованих таймерів [14] (рис. 4).

Базовою відмінністю структури є наявність додаткового блоку пам'яті внутрішніх змінних (БПВЗ), що призначений для зберігання чисельних значень внутрішніх таймерів.

Ще одним додатковим елементом є блок внутрішніх таймерів (БВТ), що інтегрує логіку роботи з внутрішніми таймерами до загальної логіки роботи ПЛК ПД. З більш докладним описом роботи ЛКА ПД з таймерами можна ознайомитись у [14].

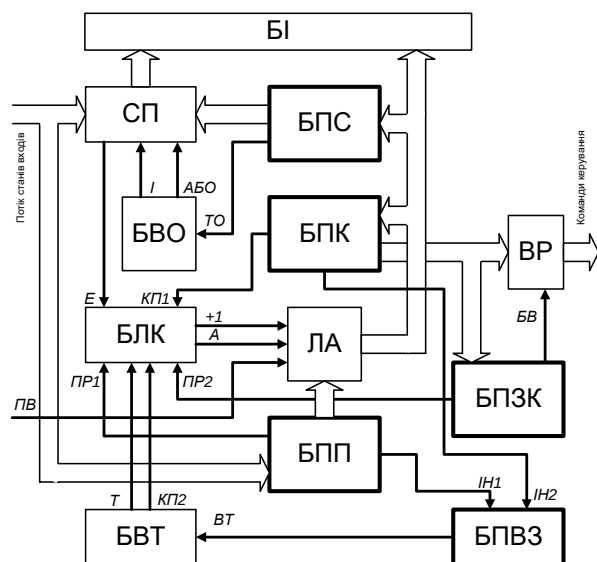


Рис. 4. Структура ЛКА ПД з елементами реалізації внутрішніх програмованих таймерів

На відміну від класичного ЛКА ПД, автомат з таймерами реалізує рівняння (3) і (4), згідно яких формуються сигнали «А» і «+1».

$$A = KP1 \vee KP2 \vee PR1 \vee PR2, \quad (3)$$

$$+1 = (E \vee T) \wedge \overline{KP1} \wedge \overline{KP2} \wedge \overline{PR1} \wedge \overline{PR2}. \quad (4)$$

Детальний аналіз розглянутих структур ЛКА ПД, а також згаданих у [5] перспективних напрямках їх розвитку у плані реалізації лічильників і елементів виконання арифметичних операцій, показує, що базова логіка роботи автомата зводиться до формування двох сигналів: «А» і «+1».

Тобто незалежно від того який саме блок ініціює їх формування (схема порівняння, БПП, БВТ, або у перспективі блок реалізації лічильників чи блок арифметичних операцій) у підсумку будуть лише трансформуватись (тобто ускладнюватись) рівняння від їх базового вигляду (1) і (2), до, наприклад, (3) і (4) відповідно.

Тобто до цих рівнянь буде входити все більше змінних і ознак переривань.

Таким чином можна сформулювати наступну думку, що становитиме суть концепції модульної архітектури керуючих пристроїв паралельної дії: «перспективна архітектура ЛКА ПД має передбачати можливості «гнучкої» зміни внутрішньої його структури за рахунок додавання або виключення блоків, що забезпечуватимуть необхідний рівень функціональності, без внесення змін до базових блоків, тобто його «ядра»».

Отже розглянемо основні положення, що формують пропоновану концепцію «модульної архітектури» ЛКА ПД:

1. ЛКА може складатись з необмеженої кількості блоків, що приймають участь у формуванні сигналів «А» і «+1» – назовемо їх блоками визначення наступного етапу (БВНЕ);

2. Усі блоки БВНЕ мають відповідати умові паралельного і незалежного один від одного функціонування;

3. Логіка роботи ЛКА ПД має забезпечувати реалізацію пріоритезації блоків БВНЕ;

4. Базова архітектура ЛКА має забезпечувати додавання або виключення блоків БВНЕ без внесення змін до усіх інших елементів структури.

Пропонована структура керуючого автомата, що реалізована у відповідності до вищеперерахованих положень, який отримав назву «модульний логічний керуючий автомат паралельної дії» (МЛКА ПД), представлено на рис. 5.

Детальному розгляду принципу та алгоритму функціонування, математичної та HDL-моделей МЛКА ПД буде присвячено окремі публікації. У рамках поточного дослідження розглянемо лише базові відмінності від класичного ЛКА ПД та призначення нових блоків і зв'язків.

Серед блоків визначення наступного етапу у структурі МЛКА ПД присутні чотири базових: БПС, БПП, БПК та БПЗК. Їх призначення і функціональність залишились незмінними і були розглянуті раніше для ЛКА ПД.

Новими блоками вибору наступного етапу є:

- БВТ – блок внутрішніх таймерів програмованих користувачем. Не дивлячись на присутність подібного блоку у структурі, що показана на рис. 4, логіка його взаємодії із іншими елементами МЛКА ПД є абсолютно іншою, хоча виконувані функції залишились такими ж самими, як це було показано в [14];

- БЛ – блок лічильників. Перспективний блок, що призначений для реалізації функцій лічби внутрішніх сигналів або сигналів із зовнішнього середовища;

- БРАО – блок реалізації арифметичних операцій. Також перспективний блок, що призначений для виконання арифметичних операцій з числами.

Як вже вказано до структури МЛКА ПД може бути додано необмежену кількість додаткових блоків БВНЕ, що, як і всі вищевказані, приєднуються до наступних внутрішніх шин:

- ШЕ – шина еквівалентностей. Шина призначена для передачі до БЛК сигналу про виконання у відповідному блоці умови, що є необхідною для формування сигналу переходу на наступний крок поточної підпрограми «+1». Прикладами подібних сигналів є «Е» та «Т» на структурі ЛКА ПД рис. 4;

- ШП – шина переривань. Шина призначена для передачі до БЛК сигналу про виконання у відповідному блоці умови, що є необхідною для формування адреси переходу до першого рядка іншої підпрограми «А». Прикладами таких сигналів у попередніх структурах є «ПР1» та «ПР2», тощо;

- ШЗ – шина внутрішніх змінних і констант. Уведена до структури МЛКА ПД для обміну даними між блоками, що їх використовують у процесі функціонування та блоком пам'яті внутрішніх змінних (БПВЗ), призначеним для їх постійного або тимчасового зберігання.

Слід зазначити, що БВНЕ можуть бути приєднані як до усіх трьох шин, так і до лише однієї з них. Прикладом може слугувати БПС, що приєднаний лише до ШЗ.

Саме за рахунок появи такого нового зв'язку значно розширюється функціональність формування умов переходу на наступний крок поточної

підпрограми за рахунок появи можливості оперативної (у процесі роботи) зміни умов, що записані до БПС.

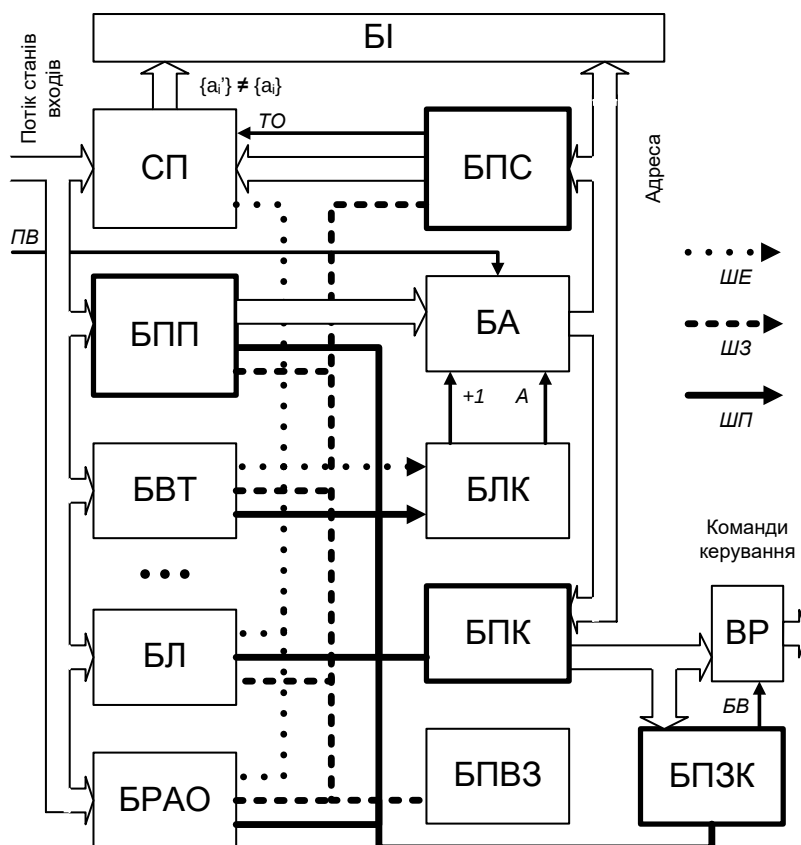


Рис. 5. Структура модульного логічного керуючого автомата паралельної дії

Зазначимо, що функціональність блоку логічного керування (БЛК), у порівнянні із структурою на рис. 1, значно розширилась. Із ключових моментів слід вказати наступне: у зв'язку з можливістю одночасної появи декількох активних сигналів на шині ШП, з'являється необхідність чіткого встановлення їх пріоритетів для визначення тієї підпрограми до якої має відбутись перехід на наступному етапі роботи обладнання.

Як можна побачити на рис. 5 сигнал «БВ» (блокування виходу), як і у більш ранніх структурах, формує блок пам'яті заборонених комбінацій, а не БЛК. Хоча за поточною логікою БПЗК є звичайним блоком визначення наступного етапу, що підключений до шини переривань і команду на заборону видачі команд керування виконавчими механізмами мав би формувати БЛК.

Таке рішення обумовлено тим, що тільки таким чином може бути забезпечена реалізація «безпечно» ПЛІС-контролера паралельної дії [4, 16, 17], тобто має бути забезпечене відокремлене розташування елементів БПЗК і ВР на кристалі ПЛІС, а в найкращому випадку для них має бути використано окремий кристал ПЛІС.

Також вкажемо, що найвищий пріоритет на шині ШП завжди має сигнал переривання саме від блоку БПЗК (у випадку його присутності у структурі конкретного МЛКА ПД).

Якщо говорити про функціональність блоку індикації, то на рис. 5, як і раніше, показано його підключення до лінії адреси поточного рядка програми і лінії індикації нееквівалентності схеми порівняння $\{a_i'\} \neq \{a_i\}$.

Таким чином може бути забезпечена функціональна діагностика роботи технологічного об'єкта [15].

Але враховуючи гнучкість пропонованої структури, БІ може виконувати значно ширше коло задач, тобто на етапі конфігурації кристала ПЛІС може бути підключений до будь якої з шин або сигналів МЛКА ПД (наприклад на етапі прототипування, налагодження або діагностики обладнання).

Окремо слід приділити увагу деяким аспектам реалізації блоку БРАО у перспективних проектах МЛКА ПД. В [5] обґрунтовано необхідну умову для його інтеграції до систем керування паралельної дії – паралельний принцип його роботи, тобто виконання арифметичних операцій за один такт дискретного автоматного часу. Це обумовлюється тим, що як було показано в [4, 11, 15], ПЛК ПД відрізняється від класичних ПЛК виконанням операцій за один такт дискретного автоматного часу (фізично тактів може бути більше через особливості структурної організації і функціонування кристала ПЛІС). В той же час арифметичні операції передбачають послідовний принцип свого виконання. Це пов'язано, що-

найменше (на прикладі операції додавання) з наявністю переносів з молодшого розряду у старший. Це принципове протиріччя розв'язати лише за рахунок реорганізації або простого вдосконалення структури ЛКА ПД неможливо.

Одним з можливих і, на думку авторів, найбільш перспективним підходом подолання цього протиріччя є використання системи залишкових класів (СЗК), або числової системи залишків (ЧСЗ) англ. Residue Number System [18-20]. Спроби інтеграції, або щонайменше, згадки застосування СЗК у контексті побудови систем керування паралельної дії наведено у [6].

Але, нажаль, до сих пір спроб реалізації реальних інженерних рішень або теоретичних структур керуючих автоматів паралельної дії, що мають у своєму складі обчислювальні елементи на базі СЗК не має.

Неважко показати, що пропонується структура МЛКА ПД дозволяє реалізувати керуючий пристрій паралельної дії з будь-якою (у межах задекларованих можливостей) функціональністю. Так у випадку використання лише чотирьох стандартних блоків БВНЕ (БПС, БПП, БПК і БПЗК), отримуємо структуру класичного ЛКА ПД, а додавши блоки БВТ і БПВЗ, отримуємо структуру, що за своєю функціональністю відповідає ЛКА, що показаний на рис 4. І, зрозуміло, що пропонується концепція «модульності» дозволить реалізувати контролери паралельної

дії із функціями, про які у рамках поточної публікації не йшлося і які будуть запропоновані лише у майбутньому.

На останок необхідно зазначити, що концепція «модульної архітектури» ЛКА ПД – це лише перший етап у розробці перспективних ПЛІС-контролерів паралельної дії.

На наступному етапі постає необхідність внесення змін до мови їх програмування ЯПЛК-М (а фактично створення нової мови, також за принципом «модульності»), а також реалізація підтримки з боку технології програмування TVP.

Висновки

У результаті виконаних досліджень усунуто такий недолік керуючих пристроїв з паралельною архітектурою, як їх фактична побудова за принципом «жорсткої логіки», або недовикористання можливостей створення «гнучких» систем керування на базі сучасних кристалів ПЛІС.

Запропоновано концепцію «модульної архітектури» керуючих пристроїв паралельної дії, що дозволяє створювати «гнучку» структуру відповідних логічних керуючих автоматів.

На базі цієї концепції синтезовано структуру модульного керуючого автомата з можливістю реалізації заданої функціональності у залежності від конкретних задач системи керування, що будується на його основі.

СПИСОК ЛІТЕРАТУРИ

1. Бовчалюк С. Я. Вдосконалення архітектури програмованих логічних контролерів паралельної дії / С. Я. Бовчалюк, І. О. Фурман, М. Л. Малиновський // Проблеми енергозабезпечення та енергозбереження в АПК України: Вісник ХНТУСГ імені Петра Василенка, вип. 37, том 2 – Харків, 2005. – С. 164–168.
2. Бовчалюк С. Я. Вдосконалення архітектури, мови програмування та технічної реалізації ПЛК паралельної дії // Міжнародна науково-технічна конференція «Інтегровані комп'ютерні технології в машинобудуванні ІКТМ-2005»: Тези доповідей. – Харків: Національний аерокосмічний університет «Харківський авіаційний інститут», 2005. – С. 290.
3. Бовчалюк С. Я. Вдосконалення математичної моделі та архітектури логічних керуючих автоматів паралельної дії / С. Я. Бовчалюк, І. О. Фурман // Інформаційно-керуючі системи на залізничному транспорті. – 2006. – №3(59). – С. 72–76.
4. Бовчалюк С. Я. Моделі, методи і засоби інформаційної технології паралельного логічного керування об'єктами залізничної автоматики: дис. ... канд. техн. наук: 05.13.06 / Бовчалюк Станіслав Ярославович. – Харків, 2008. – 203 с.
5. Бовчалюк С. Я. Визначення напрямків розвитку керуючих пристроїв з паралельною архітектурою на базі ПЛІС / С. Я. Бовчалюк, О. М. Піскарьов, С. С. Радченко, [та ін.] // Системи управління, навігації та зв'язку. Збірник наукових праць. – Полтава: ПНТУ, 2023. – Випуск 1 (71). – С. 69-72 DOI: 10.26906/SUNZ.2023.1.069.
6. Фурман І. О. Концепція, методи і засоби моделювання на ПЛІС контролерів і процесорів з паралельною архітектурою / І. О. Фурман, В. А. Краснобаєв, С. Я. Бовчалюк [та ін.] // Автомобільний транспорт: 36. наукових праць, вип. 16. – Харків, 2005. С. 338–341.
7. Бовчалюк С. Я. HDL-модель програмованого логічного керуючого автомата паралельної дії / С. Я. Бовчалюк, І. О. Фурман // Радіоелектронні і комп'ютерні системи. – 2007. – №6 (25). – С. 202–205.
8. Бовчалюк С. Я. Технічна реалізація промислового зразка ПЛІС-контролера паралельної дії / С. Я. Бовчалюк, І. О. Фурман, М. С. Деренько [та ін.] // Проблеми енергозабезпечення та енергозбереження в АПК України: Вісник ХНТУСГ імені Петра Василенка, вип. 87. – Харків, 2009. – С. 126–127.
9. Бовчалюк С. Я. Методологія побудови, синтаксис і семантика мови програмування ПЛІС-контролерів паралельної дії / С. Я. Бовчалюк, І. О. Фурман // Інформаційно-керуючі системи на залізничному транспорті. – 2007. – № 4 (66). – С. 38–44.
10. Бовчалюк С. Я. Концепція розробки технологічної мови для програмування ПЛІС-контролерів паралельної дії / С. Я. Бовчалюк, І. О. Фурман // Проблеми енергозабезпечення та енергозбереження в АПК України: Вісник ХНТУСГ імені Петра Василенка, вип. 57, том 2 – Харків, 2007. – С. 133–138.
11. Ilya Furman. Development and study of technological visual programming of logic control problems / Ilya Furman, Stanislav Bovchaliuk, Alexander Allashev, Aleksey Piskarev // Eastern-European Journal of Enterprise technologies, – 2017. – № 6/2 (90). – P. 23–31. DOI: 10.15587/1729-4061.2017.118833
12. Тимчук С. О. Реалізація функцій нечіткого логічного керування на базі автомата паралельної дії для інтелектуальних електричних мереж / С. О. Тимчук, С. Я. Бовчалюк, І. О. Фурман [та ін.] // Проблеми енергозабезпечення та енергозбереження в АПК України: Вісник ХНТУСГ імені Петра Василенка, вип. 196. – Харків, 2018. – С. 86–88.

13. Serhii Tymchuk. The Architecture of Fuzzy Logic Automat of Parallel Action for the Intelligent Smart Grid Networks / Serhii Tymchuk, Stanislav Bovchaliuk, Vira Shendryk, Sergii Shendryk New Technologies, Development and Application III. NT 2020. Lecture Notes in Networks and Systems, vol. 128. Springer, – 2020. – P. 462–468. DOI:10.1007/978-3-030-46817-0_53
14. Бовчалюк С. Я. Розвиток моделі та структури керуючих пристроїв з паралельною архітектурою / С. Я. Бовчалюк, Я. В. Гаращенко, Б. М. Коломоєць [та ін.] // Системи управління, навігації та зв'язку. Збірник наукових праць. – Полтава: ПНТУ, 2024. – Випуск 2 (76). – С. 64-66. DOI: 10.26906/SUNZ.2024.2.064
15. Фурман І. О. Науково-технічні основи створення і промислового застосування програмованих логічних контролерів на програмованих ВІС з матричною структурою: дис. ... докт. техн. наук: 05.13.05 / Фурман Ілля Олександрович. – К., 1989. – 197 с.
16. Бовчалюк С. Я. Вдосконалення алгоритму функціонування програмованого логічного контролера паралельної дії / С. Я. Бовчалюк, І. О. Фурман // Інформаційно-керуючі системи на залізничному транспорті. – 2007. – №2 (64). – С. 38–42.
17. Бовчалюк С. Я. Розвиток архітектури ПЛК паралельної дії: від абстрактної моделі паралельного автомата до інженерної реалізації безпечного ПЛІС-контролера / С. Я. Бовчалюк, І. О. Фурман, М. Л. Малиновський // Енергетика та комп'ютерно-інтегровані технології в АПК. 2016. – №. 2 (5). – С. 62-66.
18. Akushskiy I, Iuditskiy D. Mashinnaia arifmetika v ostatochnikh klassakh: M., Radio i Sviaz, 1968. 444 с.
19. Краснобаев В. А. Основные vlastivosti nepozitsionnoi sistemi chislennia u klasi liushkiv i ikh vpliv na strukturu ta printsipi realizatsii arifmeticheskikh operatsii komp'yuternoii sistemi / В. А. Краснобаев, С. О. Кошман, В. М. Курчанов [та ін.] // Системи управління, навігації та зв'язку, 2019. – Вип. 2(54) – С. 114-118.
20. Krasnobayev, V., Koshman, S., & Kovalchuk, D. (2022). THE CONCEPT OF PERFORMING THE ADDITION OPERATION IN THE SYSTEM OF RESIDUAL CLASSES. Advanced Information Systems, 6(1), 43–47. DOI: 10.20998/2522-9052.2022.1.07

Received (Надійшла) 10.02.2025

Accepted for publication (Прийнята до друку) 17.04.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Бовчалюк Станіслав Ярославович – кандидат технічних наук, доцент, доцент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Stanislav Bovchaliuk – PhD, Associate Professor, Associate Professor of Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: stanislav.bovchaliuk@nure.ua; ORCID Author ID: <http://orcid.org/0000-0002-3046-5904>.

Бовчалюк Наталія Іванівна – асистент кафедри автоматизації та комп'ютерно-інтегрованих технологій, Державний біотехнологічний університет, Харків, Україна;

Nataliya Bovchaliuk – assistant, Department of Automation and Computer-Integrated Technologies, State Biotechnological University, Kharkiv, Ukraine;

e-mail: natabov@gmail.com; ORCID Author ID: <https://orcid.org/0009-0004-5929-2981>.

Дрозд Олег Юрійович – студент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Oleh Drozd – student, Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: oleh.drozd@nure.ua; ORCID Author ID: <http://orcid.org/0009-0007-4285-4505>.

The concept of «module architecture» control devices of parallel action

Stanislav Bovchaliuk, Nataliya Bovchaliuk, Oleg Drozd

Abstract. Topicality. Publications devoted to the information technology of parallel logic control, or its individual components (control devices with parallel architecture, language and technology of their programming) show that it is promising in terms of implementing control systems mainly for critical objects. At the same time, attempts to use it for general industrial sector facilities face problems of limited functionality, or rather the need to expand the functions performed, i.e., to adapt the structure of parallel-action controllers to solve specific problems. Examples of such improvements include the introduction of fuzzy inference functionality, internal timers, and counters into the structure of a classical parallel-action control automaton. Thus, the implementation of a “flexible” structure of such automata will greatly simplify the task of building parallel-action controllers with specified functionalities. **The purpose of this paper** is to theoretically substantiate the possibility of building a “flexible” structure of logic control automata of parallel action, to formulate the concept of “modular architecture”, and to implement the structure of a promising LCA PA in accordance with this concept. **Conclusion.** As a result of the research, such a disadvantage of control devices with a parallel architecture as their actual construction on the principle of “hard logic” or underutilization of the possibilities of creating “flexible” control systems based on modern FPGA crystals has been eliminated. The concept of a “modular architecture” of parallel-acting control devices is proposed, which allows creating a “flexible” structure of the corresponding logic control automata. On the basis of this concept, the structure of a modular control automaton is synthesized with the possibility of implementing a given functionality depending on the specific tasks of the control system built on its basis.

Keywords: information technology of parallel logic control, parallel-action logic control automaton, programmable logic controller of parallel action, CPLD controller, structure of LCA PA, module logic control automaton.

Vladislava Harmash, Vladyslav Diachenko, Oleg Mikhal, Vasyl Znaidiuk

Kharkiv National University of Radio Electronics, Kharkiv, Ukraine

ENERGY-SAVING METHOD IN WIRELESS SENSOR NETWORKS

Abstract. Relevance. WSNs are characterized by many autonomous nodes that are expected to operate for extended periods without battery replacement or recharging. This poses a significant challenge for researchers and developers to find effective energy-saving methods that can substantially extend the autonomous operating time of the nodes and ensure the stable and reliable functioning of the entire network. **The object of research** is the processes of energy consumption and functioning of Wireless Sensor Networks. **Purpose of the article** is the development and investigation of an energy-saving method in Wireless Sensor Networks based on the use of machine learning algorithms, particularly artificial neural networks of the Kohonen map type and their modifications, in order to ensure the most efficient utilization of node energy resources, optimize data transmission and processing processes, and increase the duration of autonomous operation and the reliability of sensor network functioning under variable operating conditions. **Research results.** An energy-saving method based on modified Kohonen maps has been developed. The proposed approach involves multifactor clustering of nodes considering their energy parameters, adaptive selection of transmission routes, regulation of node activity modes, and online retraining of the map. **Conclusions.** The energy consumption of nodes depends not only on the hardware configuration but also on the method of data exchange organization, the chosen network topology, transmission frequency, and environmental conditions. The issue of energy depletion in individual nodes is critical, as it can lead to network fragmentation or complete network failure. Therefore, there is a need for dynamic, intelligent management that considers both local and global characteristics of the network.

Keywords: wireless sensor network, energy consumption, Kohonen map, machine learning, clustering, adaptive routing, node.

Introduction

Modern wireless sensor networks are one of the key infrastructure components across various fields of human activity, such as environmental monitoring, industrial process control, healthcare, smart homes and cities, agriculture, and military applications. The widespread adoption of such networks is driven by their capability for continuous data collection, transmission, and analysis from many spatially distributed sensor nodes. At the same time, the growing number of sensor devices, the increasing volume of processed data, and the expansion of geographic deployment areas significantly raise the demand for energy autonomy of these systems.

One of the most critical challenges faced by wireless sensor networks is ensuring a prolonged operational period of the nodes without the need to replace power sources or recharge batteries. A typical solution is to power the nodes using autonomous batteries, the replacement of which in large-scale or hard-to-reach networks is a complex and economically inefficient task. This is why energy efficiency becomes a crucial factor that directly impacts the viability and practical applicability of sensor networks.

The issue of node energy consumption is further complicated by the fact that a typical sensor device expends energy not only for data transmission but also for data reception, processing, and maintaining functionality in standby mode. Given the limited power supply resources, the need for a comprehensive approach to energy efficiency is evident. Therefore, the relevance of researching energy-saving methods lies in finding a balance between performance, reliability, and the duration of autonomous node operation.

This work presents a detailed analysis of modern methods for reducing energy consumption in wireless sensor networks. Particular attention is given to energy-efficient routing protocols, algorithms for managing node

and sensor operating modes, strategies for reducing the volume of data to be transmitted, methods of adaptive signal power control, and the potential for using renewable energy sources to ensure maximum autonomy of sensor devices. The results of this analysis will help formulate recommendations for the integrated use of various technologies and approaches, which will significantly enhance the efficiency of wireless sensor networks under real-world conditions.

The purpose of this work is the development and investigation of an energy-saving method in wireless sensor networks based on the use of machine learning algorithms, particularly artificial neural networks of the Kohonen map type and their modifications, to ensure the most efficient utilization of node energy resources, optimize data transmission and processing processes, and increase the duration of autonomous operation and the reliability of sensor network functioning under variable operating conditions.

Main part

The operation of sensor nodes in wireless sensor networks is determined by a combination of hardware, software, and environmental factors, each of which affects performance, stability, energy consumption, and the duration of autonomous operation. Key factors include: the type and sensitivity of the sensor (the choice of sensor depends on the type of physical quantity being measured. The higher the resolution and sensitivity, the more energy the sensor may consume); the power and architecture of the microcontroller (the processor's computing power determines the ability to locally process data before transmission. More powerful microcontrollers reduce the amount of data transmitted but may consume more energy themselves); the type and capacity of the power source (the battery or energy storage device limits the node's operation time. The type of power source and the possibility of energy harvesting are also important); energy

consumption modes (the availability of sleep, wake-up, standby, and active communication modes are critical for optimizing energy usage. Switching between modes must be controlled and adaptive); the communication protocol (the selected protocol affects transmission range, speed, and energy cost. Low-energy protocols are often chosen for long-term deployments); network topology and routing algorithms (a node may be an endpoint or a relay. Acting as a relay for others increases its energy consumption. Uneven load distribution among nodes leads to early depletion of certain elements in the network); data collection and transmission frequency (frequent measurements and transmissions increase energy usage. Optimizing this frequency is critical in designing energy-efficient algorithms); environmental conditions (temperature, humidity, interference levels, terrain, and other physical factors can affect sensor reliability and communication quality, causing additional energy expenditures); node software (optimized software, task schedulers, and adaptive algorithms can significantly reduce energy usage); local decision-making capability (nodes capable of analyzing collected data independently – e.g., classifying events or filtering noise – can reduce the number of transmitted packets and save energy).

In traditional wireless sensor network systems, decisions regarding routing, energy management, and node operation modes are made based on rigid rules or statistical models. This approach often fails to account for the complex dynamics of the network, changing environmental conditions, battery degradation, or unexpected overloads of individual nodes. Therefore, the application of machine learning algorithms, which allow adaptive optimization of network operation considering variable parameters and accumulated experience, is gaining increasing relevance.

One of the key tasks in the context of energy saving is node clustering, i.e., grouping nodes into clusters based on such features as battery level, geographic location, traffic density, and load. Machine learning enables this clustering to be performed dynamically rather than statically, with the structure of the clusters being updated as the network state changes. For instance, k-means methods, hierarchical clustering, or algorithms based on self-organizing maps can automatically determine the optimal number and configuration of clusters, helping reduce energy losses from routing and prevent overload of individual nodes.

Another important area is activity prediction of nodes and intelligent scheduling of sleep/wake cycles. Using historical activity data, typical sensing patterns, or external parameters (e.g., daily temperature variations, movement, or noise), learning algorithms can predict when a node is likely to be heavily loaded and adjust its operating mode accordingly. This approach maintains a balance between energy consumption and data transmission quality.

In addition, the use of classification and regression methods enables real-time assessment of the network state and decision-making regarding route changes or switching nodes to energy-saving modes.

Reinforcement learning-based approaches also deserve attention. In such models, nodes or node groups

learn to make energy-efficient decisions through interaction with the environment – by trial and error. For example, an agent can learn to deactivate secondary nodes or adjust transmission power in response to changing network conditions, receiving “rewards” for saving energy without compromising service quality.

Finally, recent research focuses on integrating deep learning to detect complex dependencies in the large volumes of data generated by the network. For example, convolutional neural networks enable the identification of recurring traffic patterns, while recurrent neural networks allow for forecasting future node loads. This creates the foundation for proactive energy and transmission management.

In article [1], a modified Kohonen map method is proposed to improve the performance of artificial neural networks through the clustering of sensor nodes—parameters such as hop count, energy levels, sensitivity, and delay are calculated. The author suggests using artificial neural network-based clustering of sensor data, which leads to enhanced energy saving in wireless sensor networks.

In article [2], the DL-GMA approach is proposed, which is a deep learning-based grouping model aimed at improving energy efficiency in wireless sensor networks. By utilizing advanced grouping and clustering techniques, DL-GMA optimizes energy consumption, enhances network stability and scalability, reduces congestion, and improves quality of service. However, the authors note certain limitations in their work, including network heterogeneity, mobile nodes, communication overhead, and the limited scale of the evaluation, which require further investigation.

In article [3], an integrated concept is proposed to enhance energy efficiency based on machine learning methods and intelligent data analysis. The author states that a unified model for sensor selection and event detection can learn from historical data and solve operational tasks such as energy efficiency, event detection accuracy, and service quality assurance in the case of sensor failures. However, the work has some shortcomings, as many functional and practical issues remain unresolved, particularly those related to network retraining.

The analysis of the sources confirms the relevance of using artificial neural networks of the Kohonen map type to improve energy efficiency in wireless sensor networks.

The essence of using Kohonen maps in energy-saving tasks lies in building an optimized model for distributing functions among network nodes. In typical wireless sensor networks, only a portion of the nodes perform critical tasks such as routing or data aggregation. The use of Kohonen maps enables the identification of nodes best suited to perform these functions based on a set of criteria: battery charge level, call frequency, signal quality, transmission intensity, load level, etc. Thus, node clusters are formed, within which cluster heads or relays are selected to minimize energy consumption across the entire system.

A key advantage of using Kohonen maps in this task is their ability for unsupervised learning, meaning such a neural network does not require pre-labeled data to

function. This is especially important in scenarios where reliable a priori information about the network status may not be available. During the learning process, the Kohonen map forms a topologically ordered structure where nodes with similar characteristics are grouped together, allowing for effective dynamic clustering in real time.

Moreover, this type of artificial neural network is well suited for adaptation to changes within the network. For example, if certain nodes fail or the environmental configuration changes, the Kohonen map can be retrained or adapted to the new conditions without requiring a complete system restart. This ensures the resilience and survivability of wireless sensor networks, which is critically important in remote or hard-to-reach environments.

Despite the effectiveness of classical self-organizing Kohonen maps for clustering tasks and reducing energy consumption in wireless sensor networks, for complex, dynamic, and large-scale environments, the standard Kohonen map architecture or its sequential learning process may prove insufficiently flexible [4]. This has led to the emergence of a few modifications that improve adaptability, scalability, and real-time accuracy. Such enhanced models open new horizons for implementing energy-efficient management in sensor networks, considering changes in network topology, load levels, power element status, and external influences.

There are also several architectures that are modifications of the classical Kohonen map. In hierarchical architecture, the network consists of multiple organizational levels, where each successive level refines the previous one. This approach is particularly effective for large sensor networks with heterogeneous node distribution. A hierarchical structure enables multilevel clustering – first dividing the network into regions based on their position on the grid and then forming energy-optimal clusters within each region. This significantly simplifies routing, reduces the number of inter-cluster transmissions, and allows for localized network adaptation.

Another promising direction involves adaptive maps that change their structure during the learning process. Unlike classical maps with a fixed number of neurons, adaptive maps can add or remove neurons in response to changes in data input. This is critically important for sensor networks where the number of active nodes can vary over time due to energy depletion, physical damage, or changes in the environmental configuration. Adaptive topology modification allows the system to maintain clustering accuracy and effective network management without the need for complete retraining. Another type of modification enables dynamic map expansion in the direction of highest data complexity. In sensor networks, this allows more attention to be automatically devoted to areas with high traffic or a high risk of overload, providing more detailed energy management in critical regions.

An alternative approach to improving the efficiency of Kohonen maps is the modification of the training process. The standard Kohonen map training algorithm is based on unsupervised learning with gradual reduction of the neighborhood radius and learning rate. Classical training requires a significant number of iterations and constant processing of input data, which may result in excessive energy consumption when implemented on real

nodes with limited resources. Therefore, several modifications have been proposed to improve energy efficiency and adaptation to sensor network conditions.

Training on data subsets suggests using random or statistically representative subsets instead of the entire input dataset. This reduces the amount of information transmitted to the main node or computing center, thereby lowering energy costs for communication. Additionally, local training on nodes or clusters with limited data reduces the number of iterations and improves scalability.

The classical Kohonen map algorithm requires pre-processing of the entire training set. Incremental learning, in contrast, allows the map to adapt to each new input vector individually. This approach accounts for changes in network state in real time without full retraining and also reduces the need for storing large data volumes on the nodes. Training with a dynamic neighborhood radius implies that the neighborhood radius is adapted dynamically – for example, depending on the variance of the input data or the energy resource level of the sensor nodes. This allows for the reduction of unnecessary updates and focuses the training process only on those nodes where changes are truly significant, thereby saving energy. In energy-aware learning, an additional parameter is introduced into the weight update function that accounts for the current battery charge level of a node. If a node has a low charge, its participation in the training process is reduced or excluded. In this way, the system automatically avoids overloading the resources of nodes that are close to depletion and ensures a more balanced load distribution across the network.

In priority-based learning, the weight coefficients are updated more frequently for inputs corresponding to high-priority events or clusters with high energy consumption. This allows computational resources to be focused on more energy-intensive areas and enables quicker adaptation to change that may threaten the network's viability. In the quantized Kohonen map, discretization of the input data space is introduced, which reduces the dimensionality of the training dataset and computational costs. In WSNs, this lowers the need for precise measurement of physical quantities and allows nodes to operate with coarser but more energy-efficient sensors or to transmit smaller-sized data packets.

The development of an effective energy-saving method for wireless sensor networks requires a combination of adaptive input data analysis, consideration of node energy constraints, and the network's ability to self-adjust [4]. In this context, modified Kohonen maps serve as a key component of the intelligent energy management model. A method is proposed based on dynamic node clustering using an adapted Kohonen map training algorithm [5], which considers the energy state of nodes, traffic density, position on the coordinate grid, and transmission history. The concept of the method involves building an energy-weighted Kohonen map that is updated in online mode and makes decisions regarding the formation of node clusters, selection of energy-optimal cluster centers, determination of efficient transmission routes with minimal energy loss, and adaptation of node operating modes (active, idle, sleep). The implementation stages of the method (Fig. 1) are as follows:

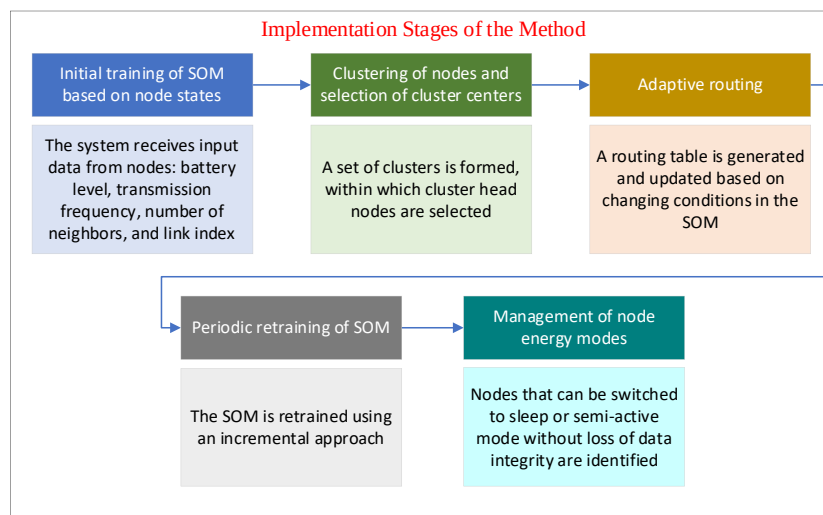


Fig. 1. Implementation Stages of the Method

1) Initial SOM Training Based on Node States. In the first stage, the system receives input data from the nodes: battery level, transmission frequency, number of neighbors, and link index. Based on these features, a Kohonen map is created with a dynamically adapted neighborhood radius and energy-weighted weight updates, where nodes with lower battery levels have reduced influence on the formation of the topology.

2) Node Clustering and Cluster Head Selection. After training, the map forms a set of clusters within which cluster head nodes are selected. The cluster head selection algorithm considers the maximum remaining energy, minimum distance to other nodes in the cluster, and load history. This ensures balanced distribution of balanced load and avoids premature exhaustion of individual nodes.

3) Adaptive Routing. Within each cluster, a routing table is formed and updated considering changing conditions such as energy depletion in a node or traffic increase. SOM is used not only for clustering but also for building a network-wide load-weighted map. This enables traffic redirection to avoid overloaded or low-energy nodes.

4) Periodic Retraining of the Kohonen Map. The method includes cyclic retraining of the map using an incremental online approach. This allows the system to remain adaptive even in dynamic conditions, such as changes in network topology or partial node failures. The learning radius and update coefficient are reduced depending on the stability of the cluster structure.

5) Management of Node Energy Modes. Based on the cluster structure and the state of the network, the algorithm identifies nodes that can be switched to sleep or semi-active mode without compromising data integrity. This implements a cyclic wake-up strategy for the nodes, which significantly reduces average energy consumption.

The proposed method represents a practical and promising approach to resource management in wireless sensor networks.

Conclusions

During this research, an analysis of the energy-saving problem in wireless sensor networks was conducted,

and the rationale for using machine learning methods – particularly modified Kohonen maps – as an effective tool for adaptive network resource management was substantiated. The energy consumption of nodes depends not only on the hardware configuration but also on the organization of data exchange, the selected network topology, transmission frequency, and environmental conditions. The problem of energy depletion in individual nodes is critical, as it can lead to network fragmentation or complete system failure. Therefore, there is a need for dynamic, intelligent management that considers both local and global characteristics of the network.

Clustering, prediction, and adaptive control models – especially those implemented through reinforcement learning or deep learning – can enhance both the energy resilience of the network and the overall quality of service.

As a result, machine learning not only provides energy savings but also creates the foundation for building self-learning, autonomous next-generation systems.

Thanks to their topological ordering and unsupervised learning capability, Kohonen maps enable the formation of energy-optimal clusters, the selection of the most suitable nodes to act as cluster heads, and the dynamic reconfiguration of network structure based on changes in energy states. This ensures balanced load distribution and reduces the number of transmissions within the network.

Machine learning algorithms offer qualitatively new opportunities for adaptive, self-learning, and energy-efficient functioning of wireless sensor networks. Their implementation results in reduced energy consumption, increased reliability, and minimized human intervention in the operation of such systems.

Modified Kohonen maps are an effective, scalable, and adaptive tool for managing energy consumption in wireless sensor networks. They not only enhance energy efficiency but also ensure fault tolerance, adaptability, and integration with other intelligent technologies.

An energy-saving method based on modified Kohonen maps has been developed.

The proposed approach involves multifactor clustering of nodes considering energy parameters, adaptive

selection of transmission routes, regulation of node activity modes, and online retraining of the map. Thus, modified Kohonen maps, integrated into the architecture of the

sensor network, represent an effective means for building energy-efficient, self-learning, and flexible sensor systems.

REFERENCES

1. V. Diachenko. Intellectual approaches to energy saving in wireless sensor computer networks. Control, Navigation and Communication Systems, vol.4, 2020. P.114-118. <https://doi.org/10.26906/SUNZ.2020.4.114>
2. Surenter, K.P. Sridhar, Michaelraj Kingston Roberts. Maximizing energy efficiency in wireless sensor networks for data transmission: A Deep Learning-Based Grouping Model approach. Alexandria Engineering Journal, 2023. P. 53-65. <https://doi.org/10.1016/j.aej.2023.10.016>
3. Mohammad Abdulaziz Alwadi. Energy efficient wireless sensor networks based on machine learning. University of Canberra, 2015. 123 p. <https://doi.org/10.26191/eenc-j011>
4. Diachenko, V., Liashenko, O., Mikhal, O., Ibrahim, BF., Koltun Y. Kohonen network with parallel training: Operation structure and algorithm. International Journal of Advanced Trends in Computer Science and Engineering 8 (1),2019. P. 35 – 38. <https://doi.org/10.30534/ijatcse/2019/0681.22019>
5. V. Diachenko, O. Mikhal. Perspectives of application of the advanced classical algorithm of Kohonen maps in distributed energy-critical sensor networks. Control, Navigation and Communication Systems, vol. 4, 2023. P.75-79. <https://doi.org/10.26906/SUNZ.2023.4.075>

Received (Надійшла) 12.03.2025

Accepted for publication (Прийнята до друку) 14.05.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Гармаш Владіслава Сергіївна – студентка кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Vladislava Harmash – student, Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: vladislava.harmash@nure.ua; ORCID Author ID: <https://orcid.org/0009-0009-4331-5508>.

Дяченко Владислав Олександрович – старший викладач, аспірант кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Vladyslav Diachenko – PhD student, Senior Lecturer Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: vladyslav.diachenko@nure.ua; ORCID Author ID: <http://orcid.org/0000-0003-2725-8784>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57207260441>.

Міхаль Олег Пилипович – доктор технічних наук, доцент, кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Oleg Mikhal – Doctor of Technical Sciences, Associate Professor, Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: oleg.mikhal@nure.ua; ORCID Author ID: <http://orcid.org/0000-0002-5977-3177>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=6506562747>.

Знайдюк Василь Григорович – кандидат технічних наук, старший викладач кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Vasyl Znaidiuk – Candidate of Technical Sciences, Senior Lecturer of Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: vasyl.znaidiuk@nure.ua; ORCID Author ID: <http://orcid.org/0000-0001-8590-8007>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57210340749>.

Метод енергозбереження в безпроводних сенсорних мережах

В. С. Гармаш, В. О. Дяченко, О. П. Міхаль, В. Г. Знайдюк

Анотація. Актуальність. Безпроводні сенсорні мережі характеризуються великою кількістю автономних вузлів, які повинні тривалий час працювати без заміни або підзарядки джерел живлення. Це ставить перед дослідниками і розробниками завдання пошуку ефективних методів енергозбереження, здатних суттєво продовжити період автономної роботи вузлів та забезпечити стабільну й надійну роботу всієї мережі. **Об'єкт дослідження:** процеси енергоспоживання та функціонування безпроводних сенсорних мереж. **Мета статті:** розробка методу багатоваріантного синтезу дискретно-казуальних моделей елементарних функцій операцій керованих інформацією який розширить можливості проектування мало ресурсних пристроїв реалізації СЕТ-операцій для побудови криптографічних систем з подвійним управлінням процесом шифрування. **Результати дослідження.** Розроблено метод енергозбереження, що базується на модифікованих картах Кохонена. Запропонований підхід передбачає багатофакторну кластеризацію вузлів з урахуванням енергетичних параметрів, адаптивний вибір маршрутів передачі, регулювання режимів активності вузлів і онлайнове перенавчання карти. **Висновки.** Енергоспоживання вузлів залежить не лише від апаратної конфігурації, а й від способу організації обміну даними, обраної топології мережі, частоти передачі та умов навколишнього середовища. Проблема енергетичного виснаження окремих вузлів є критичною, оскільки вона здатна спричинити фрагментацію або повний вихід мережі з ладу. Саме тому виникає потреба в динамічному, інтелектуальному управлінні, яке враховує як локальні, так і глобальні характеристики мережі.

Ключові слова: безпроводна сенсорна мережа, енергоспоживання, карта Кохонена, машинне навчання, кластеризація, адаптивна маршрутизація, вузол.

О. Г. Дрючко, О. В. Шефер, С. Г. Кислиця, В. М. Галай, О. В. Сухоребрий

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

РОЗРОБКА АДАПТИВНИХ СИСТЕМ УПРАВЛІННЯ ТЕРМІЧНОЮ ПОРИЗАЦІЄЮ ЛУЖНО-СИЛІКАТНИХ ТЕПЛОІЗОЛЯЦІЙНИХ ВИРОБІВ ТА З'ЯСУВАННЯ ОСОБЛИВОСТЕЙ РЕЖИМІВ ЇЇ ВІДТВОРЕННЯ

Анотація. Запропоноване дослідження спрямоване на розроблення й реалізацію адаптивної системи зонованого управління режимами термооброблення вихідної кремній-вмісної високо модульної сирцевої маси і з'ясування умов відтворення якісних водостійких теплоізоляційних виробів із жорсткою поризованою структурою на основі розчинного скла та мінеральних наповнювачів техногенного походження (відходів вугле-енергетичного комплексу). Інноваційність й актуальність запропонованих технічних рішень базуються на з'ясованих авторами властивостях, що виявляються композиціями; експлуатаційних особливостях та ефективності даного типу неорганічних утеплювачів з низькими значеннями умовної густини, теплопровідності здатними зберігати достатню міцність структури; легкій керованості процесів спінювання й твердіння у широких межах складу композицій. Запропонований варіант типу управління, відтворення умов технологічного перетворення, його режимів (послідовність, тривалість, стадійність, температурні межі, інтенсивність і швидкість) зумовлюють особливості поведінки переробляємих лужно-силікатних композиційних систем. Даний варіант відрізняється від відомих рішень складом, рецептурою вихідної сировинної суміші, послідовністю й режимами, стадійністю формування цільового продукту, технологічним оснащенням, розробленими і застосованими засобами керування виробництвом виробів, апаратними і програмними рішеннями їх управління. Виконана робота складається із трьох взаємопов'язаних пошукових емпіричних етапів, що базуються на практичній реалізації послідовної методології і результатів одержання цільових зразків за розробленими тепловими виробничими регламентами з використанням запропонованого комплексу апаратного і програмного забезпечення. Метрологічна якість отриманих результатів і їх достовірність на етапі з'ясування і виявлення цінних для виробництва функціональних властивостей і залежностей забезпечувались використанням програмованих прецизійних систем термоуправління власної розробки; у виробничих масштабах – вибором відповідних серійних засобів.

Ключові слова: адаптивні системи управління, лужно-силікатні композиційні матеріали, термічна поризація, теплофізичні властивості, відтворення регламентних умов, забезпечення метрологічної якості

Вступ

У зв'язку зі зростанням темпів й об'ємів сучасного промислового виробництва великого значення набувають питання, пов'язані зі створенням високопродуктивного переробного обладнання та вдосконаленням чинного. Серед численних промислових напрямів регламентів переробки є і методи виготовлення лужно-силікатних теплоізоляційних жорсткопоризованих матеріалів, отриманих шляхом термічного чи холодного спінювання водних розчинів силікатів лужних металів («рідкого скла») чи твердих лужно-силікатних гідрогелів [1–13]. Їх відносять до інноваційних ефективних неорганічних утеплювачів та вогнетривких теплозахисних матеріалів з широким колом практичних застосувань. Перспективу таких застосунків зумовлюють можливості досягнення низьких значень умовної густини й теплопровідності при збереженні достатньої міцності структури і легкій керованості процесів спінювання й твердіння у широких межах складу композицій. Їх застосування дає можливість хорошого захисту вразливих об'єктів, обладнання від високотемпературних впливів. Вказані переваги базуються на рівноважності і гомогенності основного компоненту сировинної суміші – «рідкого скла» і гідрогелів на його основі.

Композиційні лужно-силікатні пористі теплоізоляційні матеріали як гранульованого, так і блочного типу містять в своєму об'ємі значні кількості газової фази. Існують різні технологічні підходи в

отриманні таких матеріалів при газовиділенні безпосередньо в товщі сформованої композиції. Причому процес газовиділення при високих температурах може бути заснований як на реакціях спеціальних добавок, так і на виділенні парів кристалізаційної і хімічно зв'язаної води.

І у даній роботі приведені власні результати напрацювань авторів у вказаній області.

Постановка завдання. За розробленою рецептурою, вивченими властивостями поведінки сирцевої маси запропонувати режими й технічну реалізацію процесу відтворення якісних водостійких теплоізоляційних матеріалів із жорсткою поризованою структурою на основі «розчинного скла» та мінеральних наповнювачів техногенного походження – відходів вуглеенергетичного комплексу, яка відрізняється від аналогів складом вихідної сировинної суміші, послідовністю й режимами формування цільового продукту, технологічним оснащенням, апаратними і програмними рішеннями їх управління.

Експериментальна частина

Поставлена задача по виготовленню цільового поризованого теплоізоляційного матеріалу реалізується використанням технології гарячого спінювання сирцевої маси, схема одержання якого включає чотири головні стадії:

- 1) підготовку вихідних складових сировинної суміші і гомогенізацію останньої;
- 2) «затворення» композиційної системи «розчинним склом» й утворення стійкого гелю; фраг-

ментування затверділої сирцевої маси і поміщення грануляту у футеровані розбірні форми;

3) нагрівання і переведення речовини заготовок у піропластичний стан (110 – 115 °С);

4) подальше гаряче спінювання і відтворення регулярної поризованої макроструктури композиційних систем (130– 20°С) та формування цільових властивостей продукту перероблення (500–550°С).

Пороутворювачем у цьому випадку є вода (в основному силанольна чи молекулярна, міцно

зв'язана водневими зв'язками з немістовими атомами кисню), що виділяється при термообробленні композиційних систем.

У сировинній суміші використано промислове «розчинне скло», золу-винесення теплових електростанцій усередненого хімічного складу (табл. 1), лежаний портландцемент і додатково – загусник – попередньо підготовлене частково обезводнене затверділе «сухе скло».

Таблиця 1 – Хімічний склад золи-винесення теплових електростанцій, мас. %

SiO ₂	Al ₂ O ₃	Fe ₂ O ₃	MgO	CaO	Na ₂ O	K ₂ O	Mn ₃ O ₄	TiO ₂	SO ₃	P ₂ O ₅
51,68	16,75	14,47	0,88	4,38	0,35	2,58	0,04	0,86	4,24	0,49

У приготовлених зразках зола-винесення виявляє хороші армуючі властивості, високу термічну стабільність, достатню стійкість до агресивних середовищ, має малу насипну густину.

При формуванні складу сировинної суміші були враховані також результати покращення водостійкості лужно-силікатних композицій шляхом заміни водозміцнюючих складових двокальцієвого силікату на лежаний портландцемент.

Регулювання швидкості «схвачування» запропонованої сирцевої маси у ході формування ксерогелю кремнієвої кислоти (в залежності від цілей втілюваних завдань) здійснювали варіюванням властивостей загусника.

Оптимізована рецептура приготування сирцевої суміші дозволяє перероблення композицій різними способами з формуванням термоізоляційних матеріалів широкого призначення.

Важливою передумовою їх відтворення з необхідною системою властивостей є строге виконання регламентних вимог, установлених попередніми емпіричними дослідженнями.

Паралельно з розробкою рецептури опробувалась технологія виготовлення зразків.

Визначальним при цьому було виключення стадії гранулювання сировинної маси після термооброблення 110 – 115 °С, використання герметичних закритих форм при їх температурному відпалі та технічне запровадження відтворюваного темпу їх нагрівання. Усвідомлений вибір його оптимального режиму мотивують емпіричні дані із з'ясування особливостей термічного спінювання композиційних систем, одержаних методом диференційно-термічного аналізу (ДТА) на розробленому авторами аналітичному прецизійному комплексі [14] з термоперетворювачами типу К [15], та представлені на рис. 1.

Процес поризації складається із трьох основних етапів, тривалість і характер протікання яких залежить від виду і кількості води, що містить у собі сирцева маса:

– в інтервалі 100 – 110 °С вихідна затверділа композиційна система частково переходить у псевдопіропластичний стан і починає деформування зі збільшенням об'єму;

– в інтервалі 130 – 147 °С відбувається інтенсивне виділення вільної і адсорбованої води та інтенсивна поризація маси зразка;

– при значеннях вище 147 °С спостерігається видалення конституційної вологи, завершення структурної перебудови та фізико-хімічних перетворень композиційних систем.

З аналізу термографічних даних і макроструктури одержаних зразків, можна зробити висновок, що найбільший вклад у формуванні структури виробу з максимальною однорідністю вносить конституційна вода, тоді як видалення надлишку адсорбційної вологи на початкових стадіях приводить до утворення великих наскрізних пор і каналів у сирцевій масі. Тому вихідна рідко-скляна композиція повинна містити мінімальну кількість вільної й адсорбованої води.

Лужно-силікатні композиції у розчинах при нагріванні утворюють низку гідратованих асоціатів з відмінними властивостями (рис. 2). Це дозволяє здійснювати модифікування властивостей загусника сировинної маси – тертого «сухого скла» – частковим обезводненням купованого продукту у рідкому стані при різних значеннях температури, в умовах технологічного циклу формування цільового продукту, одночасно паралельно у тому ж оснащенні, без використання додаткового обладнання. А емпірично з'ясовані фізико-хімічна поведінка композиційних силікатних систем, особливості обезводнення і проходження в'язкопластичного стану, факт сильного прилипання проміжних продуктів перетворення до металів, кераміки, скла дозволяють запропонувати технологічні регламенти, стадійність, послідовність операцій у ході оброблення, розробку і вибір матеріалів оснащення, варіації способів одержання і використання поризованих цільових композитів.

Лабораторна практика свідчить, що виникає потреба у знаходженні ефективного способу регулювання швидкості гелеутворення способами зміщення рівноваги фізико-хімічних процесів обезводнювання дисперсних систем та вибором темпу термооброблення, а регулярність пороутворення в об'ємі зразків знаходженням «швидкості відповідності» термовпінювання у конкретному переробленні.

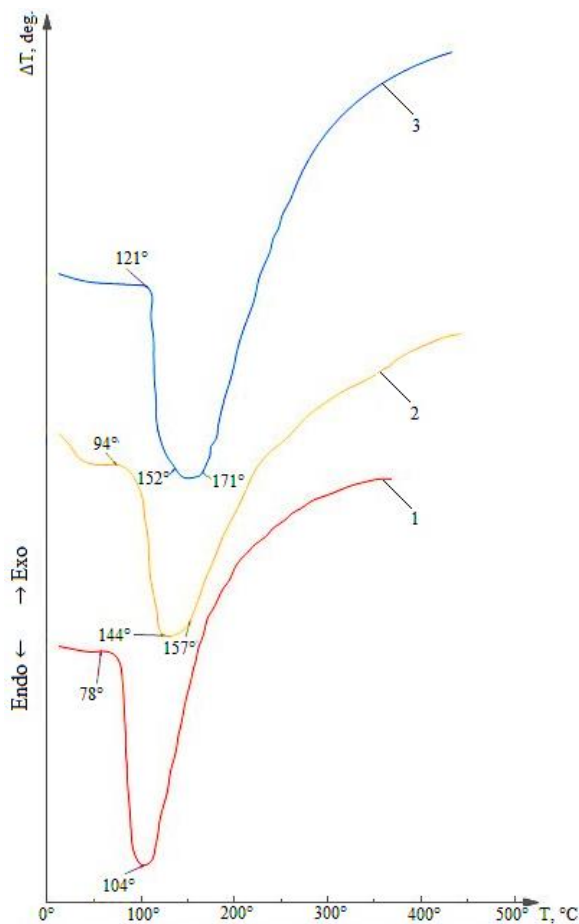


Рис. 1. Термограми ДТА ксерогелів натрієвих рідко-скляних композитів в координатах $\Delta T - T$, записані при нагріванні зразків в адекватних умовах з різною швидкістю: крива 1 – 4 град. / хв.; крива 2 – 7 град. / хв.; крива 3 – 20 град. / хв.

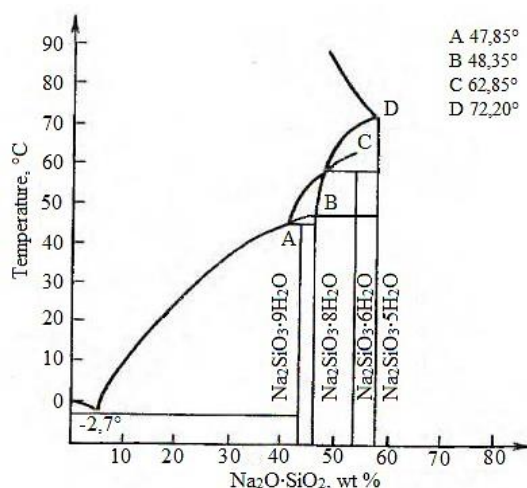


Рис. 2. Розчинність у системі $\text{Na}_2\text{O} - \text{SiO}_2 - \text{H}_2\text{O}$ [6]

Послідовні підпроцеси перероблення високомодульних силікатних композиційних систем. Розробка регламентів

Методологія послідовного перероблення включає послідовні підпроцеси:

1) підготовку вихідних складових сировинної суміші перетиранням, розмелюванням, просіюван-

ням через сито (80-315 мкм): золи-винесення ТЕС, лежалого портландцементу, частково обезводненого "сухого скла", висушеного при заданій, згідно із запропонованим регламентом перероблення композиційної системи, температурі (105, 190°C);

2) приготування наважок складових сировинних порошкових компонентів і дозованого об'єму розчинного скла - "затворювача", згідно з відпрацьованою оптимізованою рецептурою;

3) гомогенізацію порошкових складових сировинної суміші ретельним змішуванням;

4) "затворення" сировинної маси розчинним склом, перемішування до загущення композиції й утворення стійкого гелю за звичайних умов;

5) механічне фрагментування затверділого пластичного коржа на окремі елементи заданих розмірів, форми і конфігурації, з урахуванням коефіцієнта спучення (залежно від особливостей вирішуваних задач);

6) поміщення фрагментованих елементів у футеровану форму з необмеженим об'ємом на термостійкий полімерний піддон окремо (чи обмеженої конструкції розбірної форми);

7) нагрівання зі швидкістю 4-7 град. / хв. і переведення речовини гранульованих заготовок у піропластичний стан (110-115 °C, витримка 15 хв., можливе ущільнення) з наступним двостадійним термообробленням - гаряче спінювання і відтворення регулярної поризованої структури тіла заготовок виробів (130-220 °C, витримка 20 хв.), перенесення нелипких заготовок із футерованої форми на термостійкий піддон з подальшим нагріванням до 500-550 °C і їх витримкою (для надання їм задовільної водо- і паростійкості), охолодження й одержання готових виробів.

Одержані емпіричні відомості поведінки досліджуваних композицій дають відповідь про число способів і видів їх застосувань; стадійність, послідовність, температурні інтервали, режими і тривалість їх оброблення; з'ясовують вимоги до систем вимірювання і управління тепловим зонованим переробленням, до вибору оснащення і футерування його поверхонь.

Результати та їх обговорення

Оптимізована рецептура приготування сирцевої суміші дозволяє перероблення композицій різними способами з формуванням термоізоляційних матеріалів широкого призначення: гранульованих теплоізоляційних наповнювачів (рис. 3), матеріалів для термоізолювання у складних за формою конструкціях (рис. 4), плитного і оболонкоподібного видів термоізоляційних матеріалів (рис. 5). Така поставлена задача в залежності від цілей і особливостей вирішуваних завдань досягається можливістю проведення декількома відмінними способами завершальних стадій їх отримання. Використання двох стадійної схеми запропонованого перероблення в технології приготування поризованих теплоізоляційних матеріалів зумовлюють:

1) характер і особливості поведінки складових компонентів рідко-скляних композиційних систем у ході термооброблення, виявлення ними сильної ад-

гезії по відношенню до більшості конструкційних матеріалів;

2) необхідність розв'язання питання легкого видалення заготовок із оснастки формування;

3) вибір способу футерування внутрішніх поверхонь розбірних форм оснащення;

4) термофізичні і хімічні властивості використаного матеріалу футеровки.

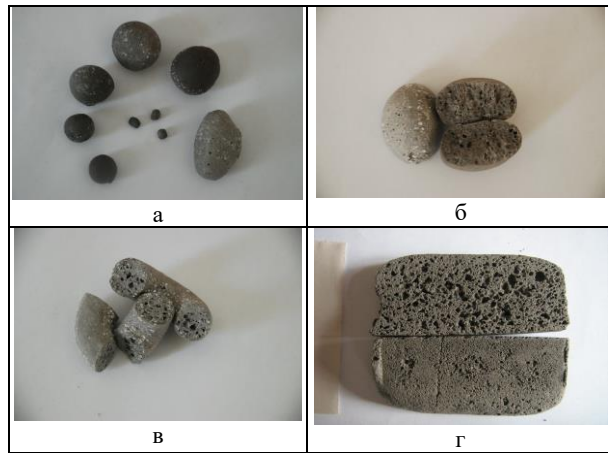


Рис. 3. Зображення зразків виробів гранульованих теплоізолюючих наповнювачів, одержаних у футерованих формах без обмеження об'єму формування:

- а, б – нарізанням ізорозмірних елементів;
в – висічкою елементів із пластичного затверділого сирцевого коржа заданої сформованої товщини;
г – із заготовок, сформованих в окремих розбірних формах



Рис. 4. Зображення фрагментів ділянок зон теплоізолювання у складних за формою конструкціях, проведених з заповненням робочої зони фрагментованими елементами і наступним термообробленням у розбірному оснащенні різної складності:

- а – без обмеження вільного об'єму формування;
б – з обмеженням простору формування

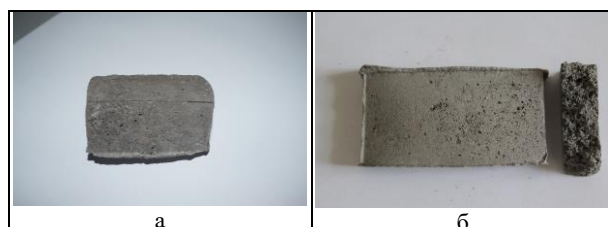


Рис. 5. Зображення фрагментів виробів, сформованих:
а – в оболонкоподібному виді;
б – у плитному виді

Особливістю технічного рішення є:

– легкість і доступність одержання складових і приготування сировинної суміші;

– формування сирцевої маси безпосередньо при «затворенні» її рідким склом за звичайних умов;

– спосіб термоізолювання є швидким;

– можливість легкого формування і фрагментування сирцевих заготовок, притаманних їм властивостей дає можливість рознести у часі і просторі окремі стадії проведення теплової ізоляції: стадії підготовки, формування грануляту (можливо, на спеціалізованій дільниці); зберігання; транспортування; технологічного укладання у робочій зоні з вимогами підвищеного опору теплопередачі (можливо, на будівельному майданчику);

– оброблення складних робочих ділянок – підбором товщини коржа сирцевого масиву, розмірів й форми вихідних фрагментованих елементів (залежно від цільового завдання та для можливості більш щільного їх упакування);

– універсальність способу термоізолювання (основується на прояві значної адгезійної здатності лужно-силікатних композиційних систем по відношенню до більшості конструкційних матеріалів – металів, кераміки, скла, дерева);

– мала усадка при запропонованій рецептурі сировинної суміші і способу оброблення;

– індиферентність до більшості складових компонентів і стабільність характеристик теплоізоляційного матеріалу, високі тепло- і хімічна стійкість, негорючість, здатність витримувати значні температури;

– поєднання набору цінних властивостей: низького коефіцієнта теплопровідності, термостійкості, негорючості, довговічності, невисокої собівартості.

Практична реалізація виробництва лужно-силікатних поризованих теплоізоляційних виробів у виробничих масштабах. Вимоги до роботи й функціонування адаптивних систем керування термодоброблення сирцевої маси

Відтворення досліджуваних теплоізоляційних матеріалів за розробленими рецептурою й технологічним регламентом у масових виробничих масштабах з набором цільових характеристик у великій мірі залежить від якісної і надійної роботи засобів формування теплового поля у зонах перероблення сирцевої маси і правильного вибору і ефективної роботи системи автоматизованого управління теплообміном у нагрівальному обладнанні (електричного опору чи газовому).

Виявлені особливості термодоброблення сирцевої маси – процеси складні, послідовні, у різних температурних інтервалах, потребують нагрівання з різним темпом (в залежності від складу, рецептури, маси, конструкційних особливостей), зі зміною фазового стану, поведінки і теплоємності проміжних форм, необхідною витримкою; потребують розбиття на стадії для можливості заміни характеристик матеріалу поверхні футеровки технологічного оснащення.

Для можливості практичного відтворення теплових регламентних умов підготовчих стадій і проведення термічної поризації високомодульних силікатних предформ різних варіантів теплоізоляційних

прикладних рішень пропонується вибрати вимірювач – ПІД-регулятор багатофункціональний двока-

нальний ОВЕН ТРМ212 (Інструкція з експлуатації АРАВ.421210.001-05ІЕ [16], табл. 2, табл. 3).

Таблиця 2 – Характеристики ПІД-регулятора ОВЕН ТРМ212

Найменування	Значення
Час опитування входу	Не більше 1 с
Тип інтерфейсу	RS-485
Швидкість передачі	2,4;4,8;9,6;14,4;19,6;28,8;38,4;57,6;115,2 кбіт/с
Тип кабелю	Екранована кручена пара
Тип протоколу передачі	ОВЕН, Modbus RTU (Slave), Modbus ASCII (Slave)

Таблиця 3 – Датчики та вхідні сигнали

Датчик або вхідний сигнал	Діапазон вимірювання	Значення одиниці молодшого розряду	Межа основної наведеної похибки
ТП с НСХ по ДСТУ EN60584-1			
ТХА (К)	-200...+1300°C	0,1°C	±0,5%

Пристрій призначений для вимірювання та автоматичного (а також ручного та дистанційного режимів) регулювання температури (при використанні як датчиків ТП), а також інших фізичних параметрів, значення яких може бути перетворене датчиками на уніфікований сигнал постійного струму або напруги.

Інформація про будь-який із вимірних фізичних параметрів відображається у цифровому вигляді на вбудованому чотирьохрозрядному ЦІ.

Пристрій дозволяє виконувати такі функції:

- вимірювання температури та інших фізичних величин (тиску, вологості, витрати, рівня тощо) у двох різних точках за допомогою датчиків;
- обробку вхідних сигналів: – цифрову фільтрацію та корекцію;
- масштабування уніфікованого сигналу для відображення на ЦІ фізичної величини;
- обчислення та індикацію квадратного кореня з вимірюваної величини;
- обчислення різниці, суми, відносини кореня з різниці двох вимірюваних величин.
- ПІД регулювання виміряної або обчисленої величини в системі із запірно-регулюючим або триходовим клапаном в автоматичному, ручному та дистанційному режимах;
- корекцію уставки регулятора залежно від зовнішнього параметра;
- автоналаштування ПІД-регулятора на об'єкті;
- дистанційне керування процесом регулювання (запуск/зупинка);
- реєстрацію даних на ПК та конфігурування приладу з комп'ютера через інтерфейс RS-485.

ТП до приладу слід підключати за допомогою спеціальних компенсаційних (термоелектродних) проводів, виготовлених з тих самих матеріалів, що і ТП (рис. 6).

Допускається використовувати дроти з металів з термоелектричними характеристиками, які в діапазоні температур від 0 до 100°C аналогічні характеристикам матеріалів електродів ТП. З'єднуючи компенсаційні дроти з ТП і приладом слід дотримувати

тись полярності. У разі порушень зазначених умов можуть виникати значні похибки при вимірюванні.

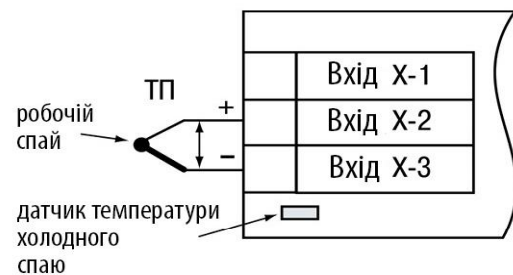


Рис. 6. Схема підключення термопари

Підключення за інтерфейсом RS-485. Інтерфейс зв'язку призначений для включення приладу до мережі, організованої за стандартом RS-485. Використання приладу в мережі RS-485 дозволяє:

- збирати дані про вимірювані величини та хід регулювання в системі диспетчеризації;
- встановити параметри приладу та дистанційно керувати за допомогою програми "Конфігуратор ТРМ101 ТРМ2хх".

Усі прилади в мережі з'єднуються у послідовну шину. Для якісної роботи приймачів та запобігання впливу завад на кінцях лінії зв'язку повинен бути узгоджувач резистор з опором 120 Ом. Резистор слід підключати безпосередньо до клем приладу.

Підключення приладу до ПК здійснюється через адаптер інтерфейсу RS-485.

Підключення навантаження до ОП типу С (рис. 7, 8). Ланцюги дискретних та аналогових вихідних елементів гальванічно ізолювані від схеми приладу. Оптосимістор включається до ланцюга керування силового симістора через обмежуючий резистор R1. Значення опору резистора визначає величина струму управління симістора. Оптосимістор може також керувати парою зустрічно-паралельно включених тиристорів VS1 та VS2. Для запобігання пробією тиристорів через високовольтні стрибки напруги в мережі до їх виводів рекомендується підключати RC-ланцюжок, що фільтрує (R2C1).

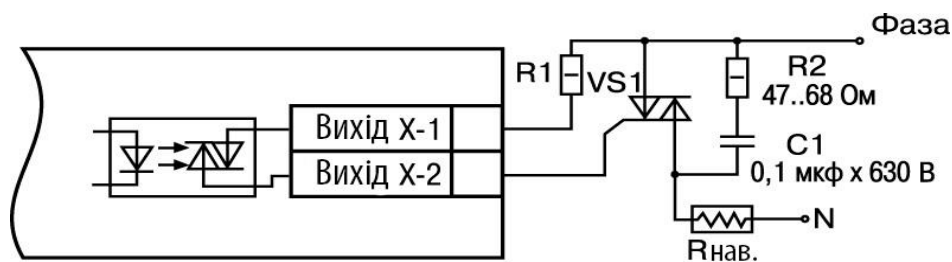


Рис. 7. Схема підключення силового симістора до ОП типу С

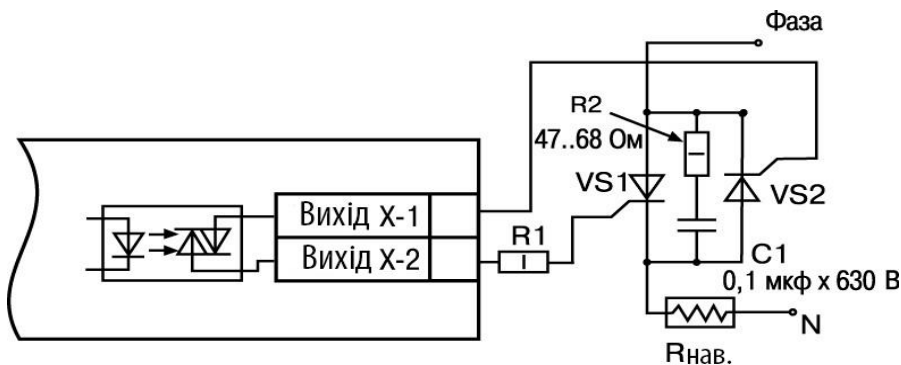


Рис. 8. Схема зустрічно-паралельного підключення двох тиристорів до ОП типу С

Висновки

Дане комплексне дослідження спрямоване на розроблення і реалізацію адаптивної системи зонованого управління режимами термооброблення кремній-вмісних високомодульних сирцевих композицій за умовами необхідними для відтворення якісних водостійких теплоізоляційних виробів із жорсткою поризованою структурою на основі розчинного скла та мінеральних наповнювачів техногенного походження (відходів вугле-енергетичного

комплексу). Запропонований проект відрізняється від відомих рішень складом, рецептурою вихідної сировинної суміші, послідовністю й режимами, стадійністю формування цільового продукту; технологічним оснащенням; розробленими і застосованими засобами керування виробництвом виробів; апаратними і програмними рішеннями їх управління; застосуванням отриманих достовірних знань про сумісну поведінку складових компонентів і дію впливаючих факторів в умовах технологічного перероблення.

СПИСОК ЛІТЕРАТУРИ

1. Рыжков И.В. Физико-химические основы формирования свойств смеси с жидким стеклом / И.В. Рыжков, В.С. Толстой. – Харьков: Вища школа, 1975. – 139 с.
2. Фиговский О.Л. Жидкое стекло и водные растворы силикатов, как перспективная основа технологических процессов получения новых наноконпозиционных материалов / О.Л. Фиговский, П.Г. Кудрявцев // Электронный научный журнал: Инженерный вестник. – 2014. – Т. 29. – № 2. – С. 55 – 97.
3. Figovsky O., Borisov Yu., Beilin D. Nanostructured Binder for Acid-Resisting Building Materials // J. Scientific Israel-Technological Advantages. 2012. Vol. 14. № 1. P. 7–12.
4. Патент 4124471 (США) Регулювання розміру частинок золю кремнезему / Лайб Д.Ф. Заявл. 22.08.77, опубл. 7.11.78 МКІ В 01 D 13/02.
5. Леонович С.Н. Особенности получения щелочно-силикатных теплоизоляционных материалов / С.Н. Леонович, Г.Л. Щукин, А.Л. Беланович и др. // Наука и техника. – 2012. – № 6. – С. 45 – 50
6. Fei SHI, Lijiu WANG, Jingxiao LIU, Miao ZENG. Effect of heat treatment on silica aerogels prepared via ambient drying // J. Mater. Sci. Technol. – 2007. – Vol.23. – No.3 – P. 402 – 406 (in Chinese).
7. Малявский Н.И. Аэрокстекло – новый щелочно-силикатный пеноматериал / Н.И. Малявский, Б.В. Покидько, Л.И. Шумаков // Кровельные и изоляционные материалы. – 2007. – № 1. – С. 28 – 30.
8. Малявский Н.И., Зверева В.В. Кальций–силикатные отвердители жидкого стекла для получения водостойких щелочно-силикатных утеплителей ISSN 1994-0351. Инженерный вестник. 2015. Вып. 2 (38).
9. Figovsky O., Beilin D. Improvement of Strength and Chemical Resistance of Silicate Polymer Concrete. International Journal of Concrete Structures and Materials. 2009, vol. 3, no. 2, pp. 97–101. doi: 10.4334/IJCSM.2009.3.2.097.
10. Figovsky O., Borisov Yu., Beilin D. Nanostructured Binder for Acid-Resisting Building Materials // J. Scientific Israel-Technological Advantages. 2012. Vol. 14. № 1. P. 7–12.
11. Kudryavtsev P.G. Alkoxides of chemical elements - promising class of chemical compounds which are raw materials for Hi-Tech industries; Journal "Scientific Israel - Technological Advantages", Vol.16, N2, 2014, p.147-170
12. Kudryavtsev P.G., Pilipenko V.G. Application of sol-gel processes for producing ultra-light weight composite materials, 6ht International Workshop on glass and ceramics from gels, Spain, 1991.
13. Kudryavtsev P.G., Kavalerova O.B., Kazakova I.L., Volkhin V.V., Koltakov A.I. Preparation and stabilization of colloidal solutions of metal oxides, 6ht International Workshop on glass and ceramics from gels, Spain, 1991.

14. Пат. 43549 Україна. МПК G 05 D 23/00. Спосіб програмного формування лінійного закону зміни температури нагрівника / О.Г. Дрючко, Д.О. Стороженко, Н.В. Бунякіна, І.О. Іваницька – у 2009 01783; Заявлено 02.03.2009; Опубл. 25.08.2009, Бюл. №16. – 10 с.
15. ITS-90 Table for Type K Thermocouple (Ref Junction 0°C) <http://reotemp.com>
16. Вимірювач – ПІД-регулятор багатофункціональний двоканальний ОВЕН ТРМ212. // Інструкція з експлуатації АРАВ.421210.001-05ІЕ.
17. Б. Р. Боряк, О. Г. Дрючко, Д. О. Ненич, О. В. Сухоребрий Автоматизована система лінійної регуляції керованої величини з використанням ПІД-регулятора. - 2023. - № 4. – С. 49-51. - <https://doi.org/10.26906/SUNZ.2023.4.049>

Received (Надійшла) 28.02.2025

Accepted for publication (Прийнята до друку) 30.04.2025

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, Y. 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>.

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>.

Кислиця Світлана Григорівна – кандидат технічних наук, доцент, доцент кафедри автоматики, електроніки та телекомунікацій, Національний університет «Полтавська політехніка імені Юрія Кондратюка», Полтава, Україна;

Svetlana Kyslytsia - 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.skyslytsia@nupp.edu.ua; ORCID Author ID: <https://orcid.org/0000-0002-2431-9900>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57671474500>.

Галай Василь Миколайович – кандидат технічних наук, доцент, доцент кафедри автоматики, електроніки та телекомунікацій, Національний університет «Полтавська політехніка імені Юрія Кондратюка», Полтава, Україна;

Vasyl Galai - 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>.

Сухоребрий Олександр Володимирович – магістр кафедри автоматики, електроніки та телекомунікацій, Національний університет «Полтавська політехніка імені Юрія Кондратюка», Полтава, Україна;

Oleksandr Suhorebryi - Master of Science, Department of Automation, Electronics and Telecommunications, Yuriy Kondratyuk Poltava Polytechnic National University, Poltava, Ukraine;

e-mail – ssanyok268@gmail.com; ORCID Author ID: <https://orcid.org/0009-0007-6847-1334>.

Development of adaptive systems for controlling thermal porization of alkali-silicate thermal insulation products and clarification of the features of its reproduction modes

O. Dryuchko, O. Shefer, S. Kyslytsia, V. Galai, O. Sukhorebryi

Abstract. The proposed research is aimed at developing and implementing an adaptive system for zoned control of heat treatment modes of the initial silicon-containing high-modulus raw mass and clarifying the conditions for reproducing high-quality waterproof thermal insulation products with a rigid porous structure based on soluble glass and mineral fillers of technogenic origin (waste of the coal-energy complex). The innovativeness and relevance of the proposed technical solutions are based on the properties revealed by the authors, which are manifested by the compositions; operational features and efficiency of this type of inorganic insulation with a low value of relative density, thermal conductivity capable of maintaining sufficient structural strength; easy controllability of the foaming and hardening processes within a wide range of composition compositions. The proposed variant of the type of control, the reproduction of the conditions of technological transformation, its modes (sequence, duration, stages, temperature limits, intensity and speed) determine the peculiarities of the behavior of the processed alkali-silicate composite systems. This variant differs from the known solutions in the composition, recipe of the initial raw material mixture, sequence and modes, stages of formation of the target product, technological equipment, developed and applied means of controlling the production of products, hardware and software solutions for their management. The work performed consists of three interconnected search empirical stages, based on the practical implementation of the sequential methodology and results of obtaining target samples according to the developed thermal production regulations using the proposed complex of hardware and software. The metrological quality of the obtained results and their reliability at the stage of clarification and identification of functional properties and dependencies valuable for production were ensured by the use of programmable precision thermal control systems of our own design; on a production scale – by selecting appropriate serial means.

Keywords: adaptive control systems, alkali-silicate composite materials, thermal porization, thermophysical properties, reproduction of regulatory conditions, metrological quality assurance.

П. О. Качанов, М. В. Тарасенко

Національний технічний університет «Харківський політехнічний інститут», Харків, Україна

СИСТЕМА АВТОМАТИЧНОГО РЕГУЛЮВАННЯ РОЗРІДЖЕННЯ В ТОПЦІ КОТЛА

Анотація. Автоматичне регулювання процесу горіння є ключовим фактором підвищення ефективності, безпеки та екологічності котельних установок. Відхилення від оптимальних параметрів горіння можуть призводити до неповного згорання палива, збільшення викидів шкідливих речовин або аварійних ситуацій. Автоматичні системи управління дозволяють підтримувати стабільний та економічний процес, контролюючи подачу палива і окисника в реальному часі. Одним із критичних аспектів регулювання є підтримка оптимального розрідження в топці. Воно забезпечує стабільність факела та сприяє повному згорянню палива. У роботі розглянуто вплив параметрів розрідження на ефективність процесу, а також методи регулювання димососами та дуттьовими вентиляторами. Застосування сучасних регуляторів, зокрема ПІД-регуляторів, дозволяє оперативно реагувати на зміну теплового навантаження, підтримуючи необхідний баланс між подачею повітря та відведенням продуктів згорання. Отримані результати свідчать про ефективність запропонованої системи автоматичного регулювання розрідження з показником перерегулювання у 15% та часом регулювання 5 секунд. Перспективними напрямками розвитку є застосування алгоритмів машинного навчання для прогнозування змін у процесі горіння та вдосконалення системи на основі аналізу хімічного складу димових газів.

Ключові слова: автоматичне регулювання, котельні установки, паровий котел, ПІД-регулятор, розрідження в топці.

Вступ

Автоматичне регулювання процесу горіння є важливим фактором забезпечення ефективності, безпеки та екологічності котельних установок. Горіння – це складний фізико-хімічний процес, що залежить від багатьох параметрів. Відхилення від оптимальних умов можуть призвести до неповного згорання палива, підвищеного утворення шкідливих речовин або навіть до аварійних ситуацій, таких як детонація або згасання полум'я. Автоматична система управління дозволяє постійно контролювати ці параметри і підтримувати їх на заданому рівні, забезпечуючи стабільний та економічний процес горіння. Крім підвищення ефективності, автоматичне регулювання мінімізує людський фактор, що знижує ймовірність помилок оператора і забезпечує швидку адаптацію до мінливих умов експлуатації. Система може в режимі реального часу коригувати подачу палива та окисника, оптимізуючи коефіцієнт надлишку повітря для максимального тепловиділення при мінімальних викидах. Це критично важливо не лише для зменшення витрат палива, а й для дотримання екологічних норм та продовження терміну служби котельного обладнання. Автоматичне регулювання процесу горіння включає регулювання подачі палива в топку в залежності від навантаження котла, підтримання оптимального співвідношення палива та повітря для економічного спалювання палива та підтримання необхідного стійкого розрідження в топці. Оптимальне розрідження в топці відіграє ключову роль у забезпеченні ефективності згорання. Слабке розрідження може призвести до неповного згорання палива через недостатню кількість повітря, а сильне розрідження — до надмірного підсмоктування повітря, збільшення теплових втрат та зниження температури в топці. Крім того, підвищене розрідження сприяє нестабільності факела, що негативно впливає на процес горіння. Різке підвищення розрідження (до 200 кПа) в топці може призвести до зриву факелу.

Під час згорання палива утворюються гарячі димові гази, густина яких є нижчою за густину

навколишнього повітря. Це призводить до виникнення природної тяги, що забезпечує видалення газів через димову трубу. Сила тяги визначається висотою димаря та температурою газів, а також витрачається на подолання опорів газового тракту, включаючи топку, димоходи, борів та шибири.

Таким чином, регулювання розрідження в топці котла є важливим завданням, що потребує врахування динамічних характеристик системи, зовнішніх впливів та технічних параметрів використовуваних димососів і дуттьових вентиляторів.

Огляд існуючих рішень

У статті [1] автори вказують, що розрідження в топці це важливий параметр, який вимірюється та контролюється в котельні. В рамках загального керування котлом, яке моделюється в цій роботі, розрідження регулюється через систему керування потоком димових газів з використанням ПІ-регулятора з негативним зворотнім зв'язком. Однак розрідження згадується як частина загальної схеми управління, але не є центральною темою роботи. Автори [2] розглядають систему управління подачею повітря в котел (forced draft system) і зазначають, що одним з її елементів є система примусової тяги, яка відповідає за підтримку розрідження в топці. Вони вказують, що розрідження в топці залежить від роботи тягового вентилятора, який регулює відведення відпрацьованих газів. Цей процес пов'язаний з балансом подачі первинного та вторинного повітря, а також із кількістю палива. У статті йдеться, що існуючі системи регулювання страждають від обурень та перешкод, що робить керування складним. Для вирішення цих проблем автори пропонують активне управління гасіння збурень (ADRC), яке здатне краще впоратися з нестабільностями та забезпечувати більш точну підтримку розрідження. Загалом стаття розглядає управління розрідженням як частину загального управління повітряним потоком у котлі, і автори пропонують удосконалені методи його регулювання. Стаття [3] присвячена методам виявлення та діагностики несправностей у системі розрідження топки котла. Автори підкреслюють, що стабільне розрідження не-

обхідне для безпечної та ефективної роботи котла, оскільки його відхилення можуть призвести до втрат тепла, нестабільного горіння та зниження ефективності установки. У роботі описується, що для контролю розрідження зазвичай регулюють швидкість димососів чи положення лопаток вентилятора. У статті запропоновано новий метод діагностики несправностей, що базується на аналізі канонічних змінних (CVA) та алгоритмі GA-XGBoost. CVA використовується для зниження надмірності даних та виявлення значних відхилень у динаміці процесу, а XGBoost, оптимізований за допомогою генетичного алгоритму, дозволяє з високою точністю класифікувати нормальні та аномальні стани системи. У роботі [4] вивчається вплив розрідження в топці на стабільність горіння в котлі. Автори розглядають основні фактори, що впливають на флуктуацію розрідження, включаючи роботу системи подачі повітря та димових газів, а також процес горіння. Було проведено спектральний аналіз тиску в топці, який дозволив виявити закономірності коливань за різних режимів роботи котла. Для оцінки стабільності горіння автори запропонували двопараметричну модель, засновану на частоті коливань розрідження та змінах навантаження. Аналіз показав, що у стійкому режимі різниця між позитивними та негативними флуктуаціями тиску мінімальна, тоді як за нестабільності ця різниця значно збільшується. Запропонований метод діагностики на основі даних про розрідження може бути використаний для оптимізації процесів горіння та запобігання можливим порушенням у роботі котла.

Метою даної статті є розробка та синтез системи автоматичного регулювання розрідження в топці котлоагрегата. Автоматичний регулятор повинен забезпечувати сталість показника розрідження в топці незалежно від навантаження котлоагрегата та інших збурюючих впливів, таких як зміна витрати палива або повітря.

Властивості котла, як об'єкта регулювання розрідження в топці

Для ефективного відведення продуктів згоряння використовується димова труба, однак за наявності теплоутилізаційних пристроїв (економайзерів, повітропідігрівачів, рекуператорів) опір газового тракту зростає. У таких випадках природної тяги стає недостатньо, і застосовується штучна тяга, що створюється дуттьовими вентиляторами або димососами. Димососи можуть бути прямої та непрямої дії, проте найбільш поширені пристрої прямої дії, що вирізняються високою економічністю. Конструкція димососа включає обертовий ротор з лопатками та нерухомий корпус, що виконує функцію направляючого апарата. Регулювання потоку димових газів здійснюється шляхом зміни кута повороту лопаток, що дозволяє керувати витратою та підтримувати задане розрідження.

Робочі умови димососів значно складніші, ніж у вентиляторів, оскільки вони працюють із газами високої температури. Для підвищення надійності конструкція передбачає термостійкі матеріали, посилені лопатки, а в деяких випадках – систему водяного охолодження підшипників.

Наявність невеликого (до 20-30 Па) постійного розрідження у верхній частині топки необхідно за умов нормального топкового режиму. Це перешкоджає вибиванню газів з топки, сприяє стійкості факела і служить непрямим показником матеріального балансу між повітрям, що подається в топку, і газами, що йдуть. Об'єкт регулювання по розрідженню – топкова камера з включеними послідовно з нею газоходами від поворотної камери до всмоктувальних патрубків димососів. Вхідним регулюючим впливом цієї ділянки служить витрата димових газів, що визначається подачею димососів.

До зовнішніх збурюючих впливів відносять зміну витрати повітря в залежності від теплового навантаження агрегату, до внутрішніх – порушення газоповітряного режиму, пов'язані з роботою систем пилоприготування, операціями з видалення шлаку і т.п. На рівень розрідження впливають атмосферний тиск, температура навколишнього повітря та витоки в газоходах. Узимку розрідження збільшується через більшу різницю температур між димовими газами та зовнішнім повітрям. Вітер також має вплив: за несприятливих умов у невеликих установках можливий ефект перекидання тяги, що може призвести до потрапляння газів у приміщення. Температура газів, що відсмоктуються, при штучній тязі повинна бути не нижче 100 °C і не вище 250 °C, так як при температурі нижче 100 °C відбувається іржавіння димососів, виділення конденсату в трубі, а температура вище 250 °C негативно впливає на їх міцність. Якщо температура відпрацьованих газів вище 250 °C, слід підмішувати холодне повітря.

Регулювання напрямним апаратом широко застосовується на теплових електростанціях через свою простоту, надійність, безпеку і можливість автоматизації процесу управління. Для підвищення ефективності регулювання направляючі апарати слід розташовувати у безпосередній близькості від входу до робочого колеса. Діапазон регулювання за допомогою цих пристроїв зазвичай не перевищує 30% від номінальної подачі, тому цей спосіб зазвичай застосовують спільно з дросельним регулюванням або іншими способами регулювання. Ділянка по розрідженню не має запізнення, має малу інерційність і значне самовирівнювання. Негативною властивістю ділянки є коливання регульованої величини близько середнього значення з амплітудою до 30-50 Па (3-5 мм вод. ст.) і частотою до декількох герц. Такі коливання (пульсації) залежить від великої кількості чинників, зокрема від пульсацій витрат палива і повітря. Вони ускладнюють роботу регулюючих приладів, особливо сильно вони впливають на релейні елементи, викликаючи їх занадто часті спрацьовування.

Для згладжування пульсацій перед первинними вимірювальними приладами встановлюють спеціальні демпфуючі пристрої: дросельні трубки та шайби, імпульсні труби підвищеного діаметра або проміжні балони (ємності). Для цього використовують також електричний демпфер, що є в електричних схемах вимірювальних блоків регулюючих приладів.

Регулювання розрідження зазвичай здійснюють за допомогою зміни кількості газів, що відсмоктують-

ся димососами. Найбільшого поширення набула схема регулювання розрідження з одноімпульсним ПІД-регулятором, що реалізує принцип регулювання по відхиленню. Під час роботи котла в регульовальному режимі часто відбуваються зміни теплового навантаження і, отже, зміни витрати повітря. Робота регулятора повітря призводить до тимчасового порушення матеріального балансу між повітрям, що надходить, і газами, що йдуть. Для попередження цього порушення та збільшення швидкості регулятора розрідження рекомендують ввести на його вхід додаткову дію від регулятора повітря через пристрій динамічного зв'язку. Як пристрій динамічного зв'язку використовують аперіодичну ланку, вихідний сигнал якої надходить на вхід регулятора розрідження лише в моменти переміщення виконавчого механізму регулятора повітря.

Синтез САР

Топка котла може бути описана, як об'єкт регулювання, що має високий рівень самовирівнювання та низький час запізнення. Експериментальним методом отримано криву розгону по розрідженню в топці котла вона зображена на рис. 1.

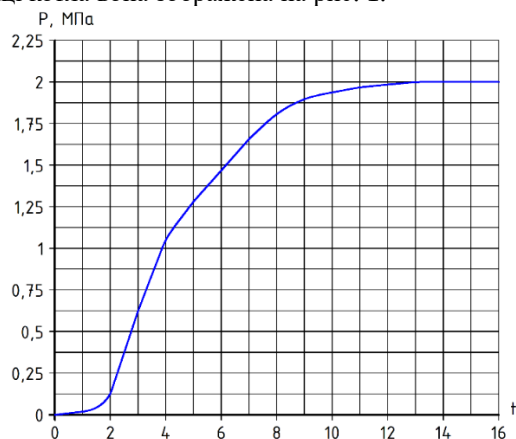


Рис. 1. Крива розгону топки по розрідженню

Приймемо постійну часу об'єкта регулювання $T = 4,3\text{с}$, а час запізнення $\tau_{зан} = 1,4\text{с}$. Коефіцієнт передачі $K_{об} = 0,2\text{кгс/м}^2 \cdot \Delta\text{У}^{-1}$. Топка котла, як об'єкт регулювання розрідження може бути представлена у вигляді інерційної ланки першого порядку з запізненням. Передавальна функція ланки:

$$W_{об}(p) = \frac{K_{об} e^{-p\tau_{зан}}}{T_{об} p + 1} = \frac{0,2e^{-1,4p}}{4,3 p + 1}.$$

Повна структурна схема системи регулювання розрідження в топці приведена на рис. 2.

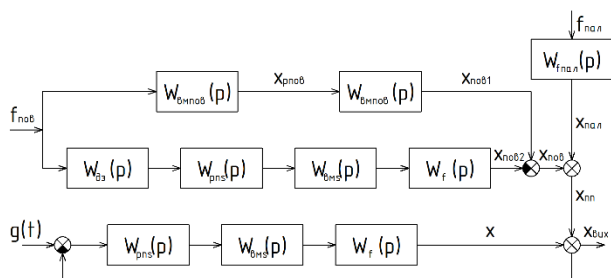


Рис. 2. Повна структурна схема САР

Передавальні функції об'єкта по каналам збурення виглядають наступним чином:

При збуренні повітрям: $W_{f_{нов}} = \frac{x_{нов}(p)}{F_{нов}(p)}$, де $x_{нов}$ – витрата повітря; $F_{нов}$ – збурювальний вплив повітрям.

При збуренні паливом $W_{f_{пал}} = \frac{x_{пал}(p)}{F_{пал}(p)}$, де $x_{пал}$ – витрата палива; $F_{пал}$ – збурювальний вплив паливом.

Оскільки витрата палива змінюється одночасно з витратою повітря, то можемо виключити величину збурення $f_{нов}$ з подальшого розгляду в схемі САР. Для подальших розрахунків введемо фільтр $W_\phi(p)$, вхідною величиною якого є збурення $f_{нов}$, а вихідною — зміна регулюючого параметру $x_{нов}$. Отже,

$$W_\phi(p) = \frac{x_{нов}}{F_{нов}(p)} = \frac{W_{вмпов}(p) * W_{f_{нов}}(p)}{W_{дз}(p) * W_{пнс}(p) * W_{вмс}(p) * W_f(p)}.$$

Підібравши відповідну передавальну функцію пристрою динамічного зв'язку можна досягти максимальної фільтруючої здатності фільтру, тобто досягти значення $W_\phi(p) = 0$. В такому випадку:

$$W_{дз}(p) = \frac{W_{вмпов}(p) * W_{f_{нов}}(p)}{W_{пнс}(p) * W_{вмс}(p) * W_f(p)} = \frac{1}{T_{вмпов} p} * \frac{K_{f_{нов}}}{T_{f_{нов}} p + 1} e^{-p\tau_{зан}} = \frac{1}{T_{вмс} p} * \frac{K_f * e^{-p\tau_{зан}}}{T_f p + 1} * K_{пнс} * (Tp + 1)$$

Зазвичай динамічні характеристики виконавчих механізмів по повітря (дугтеві вентилятори) та димовим газам (димососи) є майже однаковими, а постійні часу та запізнення об'єкта при збурення витратою повітря та тягою приблизно рівні. Звідси отримуємо:

$$W_{дз}(p) = \frac{K_{f_{нов}}}{K_f * K_{пнс} * (Tp + 1)} = \frac{K_\delta}{T_\delta p + 1}.$$

де K_δ – коефіцієнт підсилення пристрою динамічного зв'язку; T_δ – постійна часу пристрою динамічного зворотного зв'язку.

Отже, пристрій динамічного зв'язку має бути виконаний у вигляді інерційної ланки першого порядку з настроюваними параметрами K_δ і T_δ . Підібравши параметри динамічної настройки можна досягти компенсації зовнішніх збурень, що з'являються під впливом регулятора витрати повітря. Після здійснення компенсації зовнішніх збурень, розрахунок параметрів настройки регулюючого пристрою САР розрідження проводиться виходячи з величини керуючого впливу, згідно до структурної схеми. Отже, наведену вище структурну схему (рис. 2) можна спростити до схеми, що зображена на рис. 3.

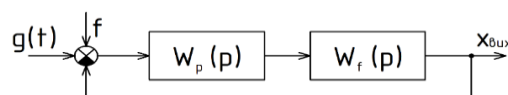


Рис. 3. Спрощена структурна схема САР

Передавальна функція САР розрідження по керуючому впливу буде виглядати наступним чином:

$$W(p) = \frac{x_{вих}(p)}{G(p)} = \frac{W_f(p) * W_p(p)}{1 + W_f(p) * W_p(p)},$$

де $W_p(p)$ – передавальна функція регулятора.

Моделювання перехідного процесу було проведено у середовищі Matlab. Отримані такі показники якості перехідного процесу: час регулювання – 5 секунд; перерегулювання – 15 %.

Висновки

Отримана САР з показником перерегулювання у 15% та часом регулювання 5 секунд. Одержані результати задовольняють вимогам до САР розрідження в топці котла та забезпечують необхідну

швидкістю перехідного процесу та точність регулювання. Серед майбутніх напрямків досліджень у сфері САР слід виділити технології машинного навчання. Сучасні котли оснащені датчиками, які записують інформацію про численні робочі параметри, включаючи температуру, тиск і витрати. Після цього ці дані можуть бути відправлені на платформу обробки та аналізу в реальному часі. Оператори можуть передбачати такі проблеми та вживати запобіжних заходів, використовуючи алгоритми машинного навчання для виявлення тенденцій і відхилень у даних. Для систем регулювання процесів горіння та зокрема САР розрідження в топці актуальним є регулювання за хімічним складом димових газів. Дана тема потребує подальших досліджень та вдосконалення вже існуючих методів.

СПИСОК ЛІТЕРАТУРИ

1. Wang J., Ji C., Cao L., Jin Q. "Design and realization of automatic control system for boiler based on model free adaptive control," Proc. of 2011 CCDC, Mianyang, China, 2011, pp. 1881-1886, DOI: 10.1109/CCDC.2011.5968507.
2. Wang Q., Xu H., Pan L., Sun L. "Active disturbance rejection control of boiler forced draft system: a data-driven practice." *Sustainability* 12(10): 4171, 2020. DOI:10.3390/su12104171
3. Ling D., Li C., Wang Y., Zhang P. "Fault detection and identification of furnace negative pressure system with CVA and GA-XGBoost." *Energies* 15(17): 6355, 2022. DOI:10.3390/en15176355
4. Chen Z., Niu Y., Xiao C., Li E. "Analysis of Combustion Stability Based on the Negative Pressure Characteristic of Furnace." *Journal of Physics*. Vol. 1087. No. 4. IOP Publishing, 2018. DOI: 10.1088/1742-6596/1087/4/042082
5. Попович М. Г., Ковальчук О. В. Теорія автоматичного керування – К.: Либідь. 2007. – 655 с.
6. Zak S. H. *Systems and Control*. Oxford University Press, 2003. – 770 p. ISBN: 978-019-51-5011-7
7. Gilman G.F. (Jerry), Gilman J. "Boiler control systems engineering." – ISA Publ., 2010. – 198 p. ISBN: 978-193-60-0720-2
8. Kristinsson H. Lang S. "Boiler control improving efficiency of boiler systems." CODEN: LUTEDX/TEIE (2011). URL: <http://lup.lub.lu.se/student-papers/record/3632662>
9. Александров С. С., Голуб О. П., Костенко Ю. Т. Теорія автоматичного керування. Т1. – Харків: ХДПУ, 2000. – 154 с.
10. Александров С. С., Козлов С. П., Кузнецов Б. І. Автоматичне керування рухомими об'єктами і технологічними процесами: Підручник у 3-х томах. Т.1 Теорія автоматичного керування. – Харків: НТУ „ХПІ”, 2002. – 490 с.

Received (Надійшла) 23.01.2025

Accepted for publication (Прийнята до друку) 16.04.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Качанов Петро Олексійович – доктор технічних наук, професор, професор кафедри автоматики та управління в технічних системах, Національний технічний університет «Харківський політехнічний інститут», Харків, Україна;
Petro Kachanov – Doctor of Technical Sciences, Professor, Professor of Department of automation and control in technical systems, National Technical University «Kharkiv Polytechnic Institute», Kharkiv, Ukraine;
 e-mail: petro.kachanov@khp.edu.ua; ORCID Author ID: <http://orcid.org/0000-0002-7532-5913>;
 Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57190405644>.

Тарасенко Микола Валерійович – аспірант кафедри автоматики та управління в технічних системах, Національний технічний університет «Харківський політехнічний інститут», Харків, Україна;
Mykola Tarasenko – PhD student, Department of automation and control in technical systems, National Technical University «Kharkiv Polytechnic Institute», Kharkiv, Ukraine;
 e-mail: porolah2008@gmail.com; ORCID Author ID: <http://orcid.org/0000-0001-6877-0173>.

Automatic draft control system in a boiler furnace

Petro Kachanov, Mykola Tarasenko

Abstract. Automatic combustion control is a key factor in improving the efficiency, safety, and environmental sustainability of boiler installations. Deviations from optimal combustion parameters can lead to incomplete fuel combustion, increased emissions of harmful substances, or even emergency situations. Automatic control systems enable a stable and economical process by regulating the fuel and oxidizer supply in real time. One of the critical aspects of regulation is maintaining optimal draft pressure in the furnace, which ensures flame stability and promotes complete fuel combustion. This study examines the impact of draft pressure parameters on combustion efficiency, as well as methods for regulating draft using induced draft fans and forced draft fans. The application of modern controllers, particularly PID controllers, allows for rapid adaptation to changes in thermal load while maintaining the necessary balance between air supply and flue gas removal. The results confirm the effectiveness of the proposed automatic draft control system, achieving a 15% overshoot and a regulation time of 5 seconds. Promising areas for further development include the use of machine learning algorithms to predict combustion process changes and the enhancement of control systems based on flue gas composition analysis.

Keywords: automatic control, boiler installations, steam boiler, PID controller, furnace draft.

А. С. Янко, О. В. Михайліченко, О. І. Крук

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

ОПТИМАЛЬНІ АЛГОРИТМИ АВТОМАТИЧНОГО КЕРУВАННЯ РОБОТОТЕХНІЧНОЮ ПЛАТФОРМОЮ, ЯКІ ЗАБЕЗПЕЧУЮТЬ ЇЇ ВИСОКУ ПРОДУКТИВНІСТЬ ТА НАДІЙНІСТЬ ФУНКЦІОНУВАННЯ

Анотація. У статті представлено дослідження сучасних оптимальних алгоритмів автоматичного керування роботизованими платформами, що гарантують їх високу продуктивність, адаптивність та надійність. Проаналізовано основні способи руху мобільних роботів, такі як колісні, гусеничні та багатоланкові конструкції. Описані переваги та недоліки кожного підходу з точки зору підвищення енергоефективності та маневреності в складних умовах експлуатації. Особливу увагу приділено розробці адаптивних алгоритмів керування, які дозволяють платформі змінювати параметри руху у відповідь на умови навколишнього середовища. Для підвищення адаптивності системи пропонується використовувати лінійні квадратурні регулятори, а також нечіткі регулятори для забезпечення точного балансування та стабілізації рухів. Представлено алгоритмічні рішення для автоматичного керування параметрами руху платформи в умовах пересіченої місцевості та можливість динамічної реконфігурації маршрутів руху. Дослідження показує, що використання адаптивних алгоритмів дозволяє значно підвищити продуктивність роботизованих платформ, зменшити ризик перекидання та гарантувати стабільну роботу навіть в умовах високої мінливості навколишнього середовища. Запропонований підхід може бути використаний для створення нового покоління автономних систем, здатних ефективно вирішувати задачі навігації та руху в реальному часі.

Ключові слова: автоматичне керування, адаптивний алгоритм, колісні мобільні роботи, лінійний квадратичний регулятор, нечітка логіка, робототехнічна платформа.

Вступ

Розвиток сучасної робототехніки вимагає передових методів автоматичного керування для забезпечення надійності та ефективності мобільних робототехнічних платформ [1-3]. Саме автономні колісні мобільні роботи є дуже цікавим предметом як для наукових досліджень, так і для практичного застосування.

Однією з найважливіших проблем є адаптація алгоритмів керування до складних умов експлуатації, таких як пересічена місцевість, нестабільні навантаження та динамічні перешкоди.

Відсутність ефективних механізмів адаптації призводить до низької продуктивності робототехнічних платформ, підвищеного енергоспоживання та ризику виходу з ладу внаслідок механічних впливів.

Ця проблема особливо актуальна при розробці автономних систем, що використовуються в складних умовах, таких як пошуково-рятувальні операції та інженерні завдання у важкодоступних районах.

Традиційні методи керування, засновані на алгоритмах руху жорсткого тіла, часто виявляються неефективними через непередбачувані зміни зовнішніх факторів.

Розглядаються сучасні адаптивні підходи до автоматичного керування, які забезпечують гнучкість платформи та підвищують стійкість до зовнішніх впливів.

Запропоновано використання алгоритмів оптимізації, що поєднують класичні методи керування, такі як лінійні квадратичні регулятори, з сучасними підходами штучного інтелекту, такими як нечітка логіка та машинне навчання. Такий підхід має на меті покращити маневреність, зменшити ризик перекидання та ефективно використовувати наявні енергетичні ресурси.

Оптимальні алгоритми керування робототехнічною платформою

Для того, щоб визначити ефективний алгоритм автоматичного керування, необхідно спочатку визначити тип приводу. Це пов'язано з тим, що пересування мобільних платформ по землі і в різних середовищах базується на різних принципах. Найпоширенішими типами наземної мобільності є ті, що використовують спеціальні зовнішні пристрої, такі як колеса, ноги або гусениці. Рух багатотілесних систем також може базуватися на періодичній зміні конфігурації мобільної системи.

Змієподібні рухи завжди викликали інтерес у біомеханіці та робототехніці. Розглянемо багатоланкову систему (рис. 1), що лежить у горизонтальній площині і складається з прямих жорстких ланок, з'єднаних послідовно циліндричними шарнірами, в яких розміщені виконавчі механізми. Позначимо через l_i довжину i -ї ланки, через m_i – масу i -го шарніра, а через M_i – крутний момент, що створюється i -м приводом. Для простоти ми нехтуємо масами ланок порівняно з масами шарнірів [4]. Система знаходиться в умовах сухого тертя, що підкоряється закону Кулона:

$$F_i = \frac{km_i g v_i}{|v_i|}, \text{ якщо } v_i \neq 0, \quad (1)$$

$$F_i \leq km_i g, \text{ якщо } v_i = 0.$$

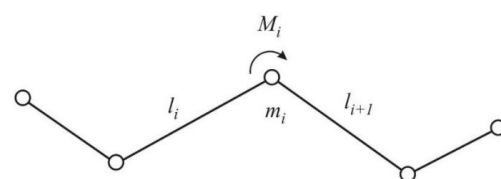


Рис. 1. Багатоланкова система

Багатоланкова система за умови (1), що складається з декількох ланок, може рухатися вздовж площини, чергуючи повільні та швидкі фази руху. Використовуючи повільну і швидку фази, багатоланкова система може виконувати поздовжні і поперечні обертальні рухи, а також може переміщатися з початкового положення і конфігурації в будь-яке положення і конфігурацію на площині. У повільній фазі рухаються лише кінцеві ланки, тоді як інші ланки залишаються нерухомими. Однак зі збільшенням ваги платформи такий тип руху не є надійним, оскільки збільшує навантаження на кріплення рухомих частин і навіть без урахування низької зносостійкості, ефективність знижена через тип матеріалу і вагу окремих ланок [4].

Перший виклик, з яким стикаються розробники автономних систем, – це автоматизована маневреність. Маневреність є найважливішою характеристикою в екстремальних ситуаціях, таких як дослідження далеких планет. Адже застрягання або перекидання позначатиме повний провал дорогої місії. Існує багато хороших прикладів механічних рішень, які поєднують в собі мобільність. Одним з найуспішніших прикладів є концепція MER explorer. Колеса – не єдине рішення, їх можна замінити рейками для кращої продуктивності або, як у випадку з Big-dog, ногами. Незважаючи на те, що існує широкий спектр механічних рішень для платформ, що розробляються, колеса залишаються найдешевшим і найпоширенішим елементом рушія [5].

Найпоширеніший спосіб побудови мобільних робототехнічних платформ – це використання двокілісного приводу з диференціальним керуванням і

заднім стабілізатором; завдяки незалежному керуванню двома двигунами платформа може рухатися «голим». Такі робототехнічні платформи добре працюють у приміщеннях на рівних поверхнях і в місцях, доступних для людей на інвалідних візках. Існує багато комерційних платформ, заснованих на цьому механізмі пересування. Такі платформи дозволяють зосередитися на двох важливих аспектах інтелектуальних автономних платформ: зондуванні та прийнятті рішень. Однак у реальних умовах (як на вулиці, так і в приміщенні) доводиться мати справу з нерівностями поверхонь. Використання традиційних колісних робототехнічних платформ обмежене, оскільки платформа повинен підніматися і спускатися по сходах і долати перешкоди. Крім того, при проектуванні платформ, які повинні працювати в таких умовах, необхідно враховувати більш жорсткі обмеження, такі як вага, міцність і розмір, порівняно з пересуванням по рівній поверхні. Проектування робототехнічних платформ, здатних вирішувати складні завдання в тривимірному світі, зміщує акцент на кінематичні можливості автономних машин, які є такими ж важливими, як сприйняття і прийняття рішень. Схема апаратної побудови подібної системи наведена на рис. 2.

Окрім підвищення маневреності робототехнічних платформ на різних ділянках місцевості, необхідно забезпечити автоматичний захист від перекидання, що є однією з найбільших проблем для всіх наземних транспортних засобів. Вирішення цієї проблеми лежить в площині модифікації конструкції шасі та автоматичних алгоритмів контролю за рухом.

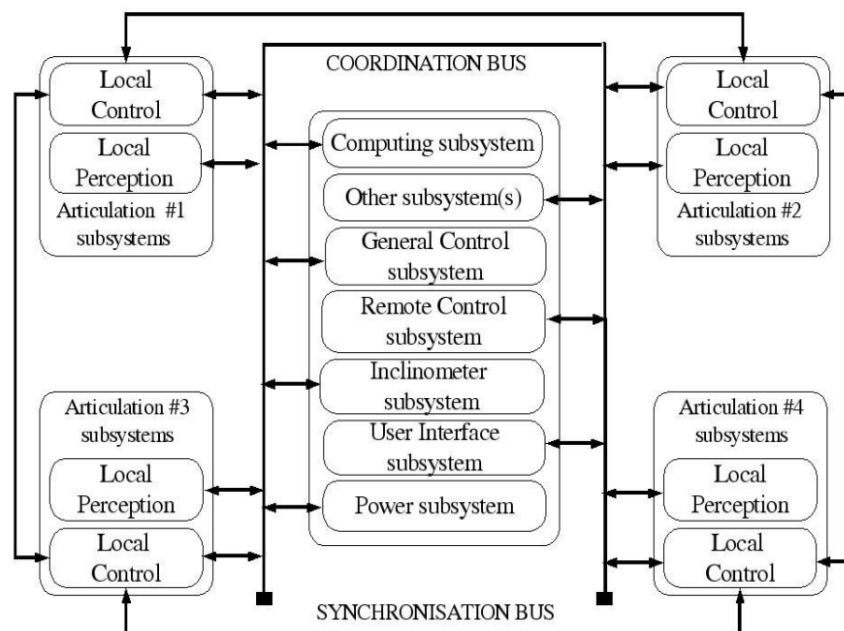


Рис. 2. Схема системи керування [5]

Для вирішення проблеми нестійкості колісно-ножних роботів на складних ділянках місцевості пропонується платформа з підпружиненими колесами і ногами та адаптивним механізмом втягування колісно-ножного механізму, який утримує тіло па-

ралельно землі для забезпечення плавного подолання перешкод.

Основними режимами роботи робототехнічної платформи є режим роботи на колесах і режим роботи на колесах і ногах, які перемикаються обертан-

ням тазостегнового двигуна. Енергоефективність підвищується завдяки використанню колісного режиму на рівній місцевості та колісного режиму на перешкодах і пересіченій місцевості.

Коли платформа зберігає положення стоячи, механізм з'єднання ніг утримується нерухомим за рахунок блокування двигуна тазостегнового суглоба. В цей час вона може бути еквівалентною двоколісному самобалансуючому роботу.

Центр мас знаходиться над віссю коліс робота, а положення робота у світовій системі координат – $[x_b, y_b, z_b, \alpha]^T$, де координата положення середньої точки осі ведучих коліс робота – $[x_b, y_b, z_b + r]$, α – курсовий кут робота, відстань між центрами двох коліс – b , а радіус колеса – r , кути повороту лівого

та правого коліс – θ_L, θ_R , а переміщення лівого та правого коліс – X_L, X_R відповідно. Припустимо, що маса тіла спрощеного робота дорівнює M , довжина тіла дорівнює L , момент інерції тіла навколо осі Y дорівнює I_b , а координата положення центру мас тіла дорівнює $[x, 0, z]$.

Момент інерції кожного шатуну в центрі мас дорівнює I_{cl} , який отримано за допомогою програмного забезпечення для моделювання [6]. Кут нахилу та довжина тіла еквівалентної моделі зображені на рис. 3.

Розрахунок куту проведений за формулою:

$$\phi = \arctan\left(\frac{x}{z-r}\right), \quad L = \sqrt{x^2 + (z-r)^2}. \quad (2)$$

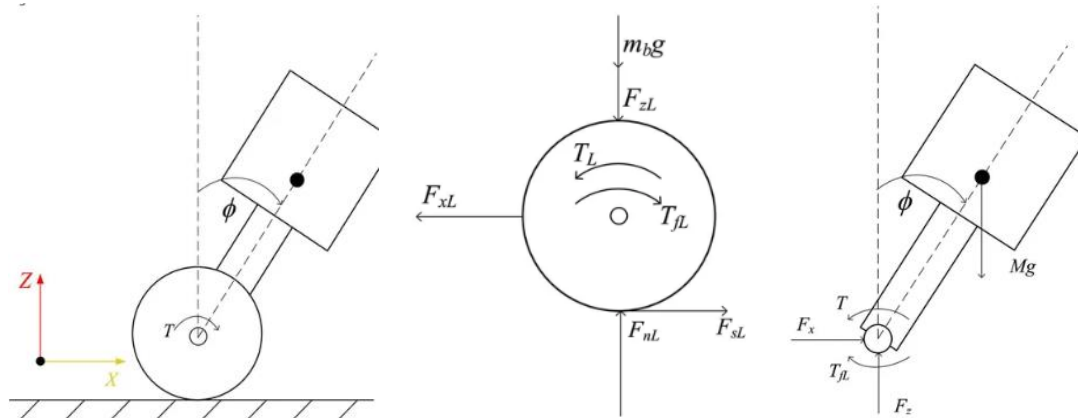


Рис. 3. Схематичні зображення механізмів [6]

При аналізі стрибків робота через перешкоди, завдяки симетричній конструкції робота, центр мас в процесі руху може утримуватися тільки в зміщенні в напрямку X і Z у світовій системі координат, а процес стрибка аналізується відповідно до методу аналізу плоских роботів. Балансування та керування швидкістю робота реалізовано методом лінійного квадратичного регулятора (Linear quadratic regulator, LQR).

Хоча розроблене механічне рішення забезпечує гнучкість і мобільність, що дуже важливо в умовах пересіченої місцевості, воно також створює певні проблеми, пов'язані з автономною роботою платформи. В автономному режимі контролер платформи покладається виключно на власні дані з датчиків і

правила керування та повинен мати можливість коригувати їх на ходу. Проблеми управління можна пояснити двома типовими ситуаціями. Найпростіша ситуація зображена на рис. 4, коли обидві секції коліс мають однакову перешкоду однакової форми. Значення $\vec{V}$ на рис. 4 і 5 вказують тип і напрямлення активної швидкості колеса [7].

Більш складна ситуація зображена на рис. 5, коли колісні секції (wheel section) наштовхуються на перешкоди різної форми, створюючи таким чином різні навантаження на відповідні колісні пари на кожній стороні платформи ($\vec{V}_{RSection}, \vec{V}_{LSection}$ – індикатори швидкості правої та лівої колісної секції).

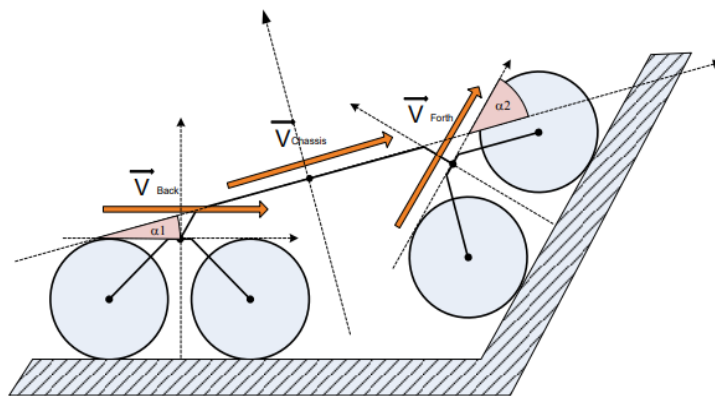


Рис. 4. Однонаправлене переміщення

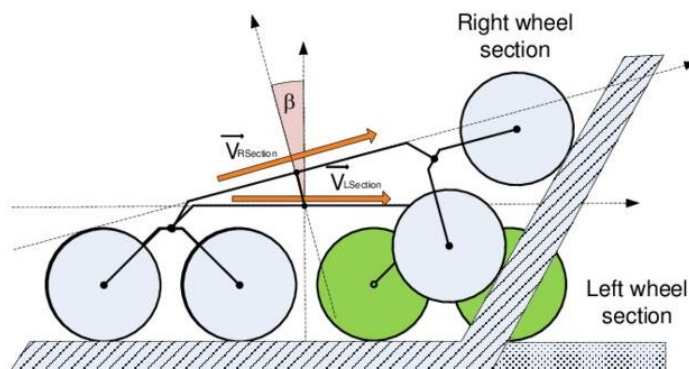


Рис. 5. Двонаправлене переміщення

Алгоритм для вирішення двох проблем розділяється на етапи, в яких:

1. Фактичні значення коліс отримуються за допомогою датчиків обертання двигуна.
2. Положення колісних пар і секцій коліс зчитуються з вбудованих датчиків положення повороту потенціометрів;
3. Фактичні швидкості колісних пар і секцій коліс розраховуються за формулами відносно рисунків 4 і 5 відповідно:

$$\vec{V}_{new} = \frac{\vec{V}_{old}}{\cos(\alpha_1 - \alpha_2)}, \quad \vec{V}_{RSection} = \frac{\vec{V}_{RSection_old}}{\sin(\beta)}. \quad (3)$$

4. На цьому кроці застосовується нечіткий регулятор для того, щоб відрегулювати задані швидкості секцій коліс відповідно до розрахованої помилки секцій відповідно до розрахованих значень.

5. Визначає, чи зіткнулася платформа з винятком. У всьому циклі управління платформи управління платформою існує дисбаланс між необхідністю плавного руху, щоб уникнути фізичних пошкоджень, і швидкою реакцією, яка необхідна для реагування на потенційно небезпечні ситуації. Наприклад, потенційно небезпечні ситуації можуть виникнути при швидкій зміні положення четвертої колісної пари високої швидкості платформи. Кожне з можливих виключень описується правилами виключень правилами, які визначаються експертами або емпіричним способом на тестах [7]. Покрокова структура такого алгоритму розглядається як відправна точка для розробки алгоритмів, які пропонують найвищу автоматичну стійкість до помилок і перешкод з можливістю додавання нових правил, тобто модульність алгоритму. Розуміння поведінки автоматичного керування робототехнічною платформою визначається шляхом порівняння ручного та автоматичного керування.

Як показано на рис. 6, чотириколісна керована робототехнічна платформа складається з модуля робота і модуля дистанційного керування. Модуль робота складається з чотирьох рухових серводвигунів, приводного двигуна постійного струму, електричного регулятора швидкості (Electric speed controller, ESC), мікроконтролера (Microcontroller) і модуля ZigBee. Модуль телеуправління складається з двох джойстика, мікроконтролера і модуля ZigBee. Модуль робота оснащений двома диференціальними редукторами, що з'єднують колеса, по одному в центрі кожної передньої і задньої осі [8].

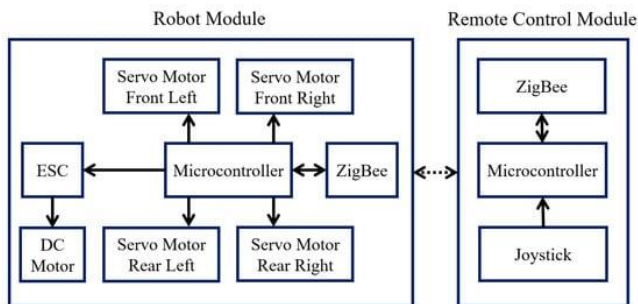


Рис. 6. Схема модулів контролеру

Ці диференціальні передачі передають потужність від приводних двигунів постійного струму через приводний вал на кожне колесо, дозволяючи колесам обертатися з різною швидкістю під час повороту робота. Використання чотирьох серводвигунів для горизонтальної орієнтації дозволяє розміщувати додаткові датчики і контролери, такі як LiDAR, камери і вбудовані контролери, на верхній частині рами робота. ZigBee в модулі робота використовується для отримання команд управління від модуля дистанційного керування і передачі інформації про стан робота, наприклад, про рівень заряду батареї, поточний кут повороту і швидкість робота [8]. Приклад стандартного розміщення вказаний на рис. 7.

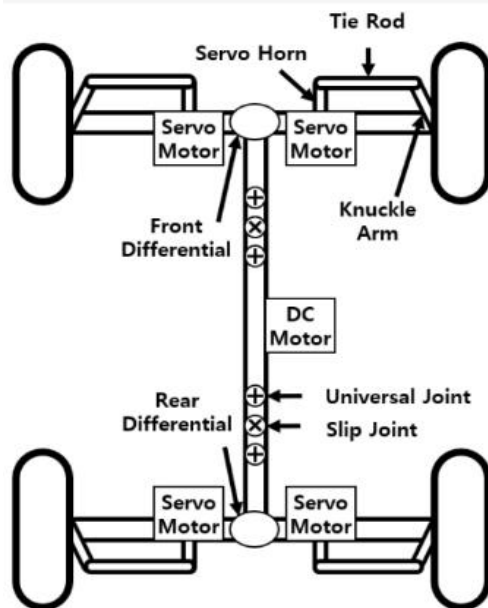


Рис. 7. Схема-базис розміщення модулів

Для кращого керування колісними роботами можна використовувати кілька методів. Пропорційно-інтегрально-диференціальний (ПІД) регулятор є найпростішим в реалізації і може забезпечити розумне керування, якщо правильно налаштувати параметри. Однак його продуктивність недостатня, і він неефективний для систем високого порядку і дробових систем. В [9] використовується за основну ПІД-регулятор дробового порядку, що має переваги ПІД-регулятора, а також виправлені недоліки.

Дробовий ПІД-коректор реалізується в звичайних системах керування зі зворотним зв'язком, як показано на рис. 8, де $R(s)$ – керуючий сигнал, $E(s)$ – відхилення, що виникає внаслідок різниці між заданим значенням $R(s)$ і керованою змінною $Y(s)$, $C(s)$ –

передавальна функція дробового коректора, $G(s)$ – передавальна функція системи [9].

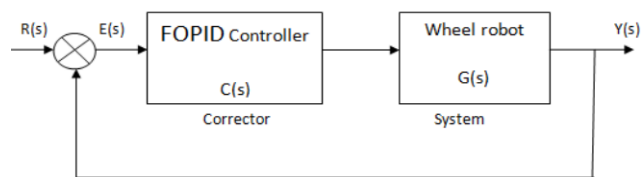


Рис. 8. Схема ПІД-коректору

Враховуючи порівняння існуючих методів і алгоритмів побудови виведемо базову структурну схему системи автоматичного руху в обхід перешкод (рис. 9).

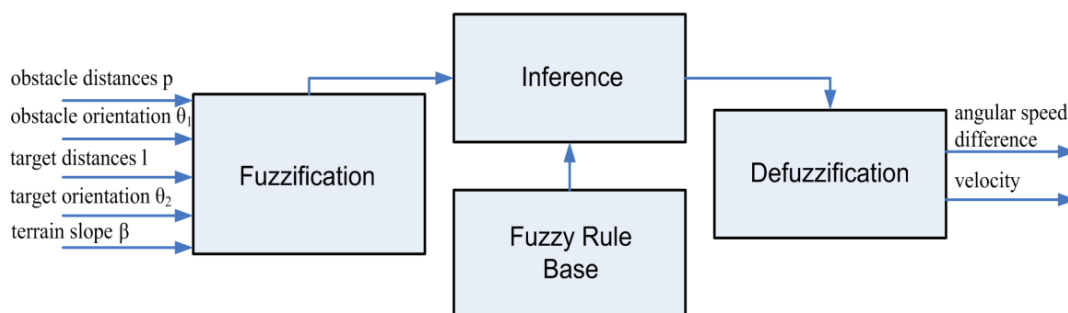


Рис. 9. Схема-базис алгоритму уникнення перешкод

Схема представлена на рис. 9 передбачає нечітке керування рухом автономної колісної мобільної роботизованої платформи в неструктурованому середовищі з перешкодами [10]. Результати моделювання показують ефективність та обґрунтованість поведінки уникнення перешкод у неструктурованому середовищі та керування швидкістю руху колісної мобільної роботизованої платформи запропонованої стратегії нечіткого керування. Пропонується для реалізації нечіткого регулятора використовувати метод дефазифікації за центром області. Ефективність запропонованого підходу підтверджується результатами моделювання, які демонструють високу точність та швидкість реакції системи на зміни у навколишньому середовищі. Застосування нечіткого керування дозволяє робототехнічній платформі плавно та безпечно обходити перешкоди, уникаючи різких змін траєкторії та забезпечуючи стабільність руху. Це особливо важливо для роботи в умовах обмеженого простору та високої щільності перешкод.

Подальші дослідження будуть спрямовані на розробку більш складних алгоритмів нечіткого керування, які враховують динамічні зміни в навколишньому середовищі та дозволяють платформі адаптуватися до непередбачуваних ситуацій. Також планується провести експериментальні дослідження на реальній робототехнічній платформі для оцінки ефективності запропонованого підходу в реальних умовах.

Висновки

Автоматичне керування робототехнічними платформами є важливим фактором підвищення їх

продуктивності та ефективності.

Дослідження показали, що вибір найкращого алгоритму керування залежить від конкретного завдання, вимог до продуктивності та складності системи. Кожен з розглянутих алгоритмів має переваги і недоліки, а їх ефективність може змінюватися в залежності від умов експлуатації. ПІД-регулятори широко використовуються в різних системах управління, в тому числі робототехнічних платформах, завдяки своїй простоті і універсальності.

Адаптивне керування ефективне для систем з невизначеними або мінливими параметрами, де алгоритм керування може автоматично підлаштовувати параметри у відповідь на зміну зовнішніх умов.

Нечітке управління та нейромережеве управління є потужними інструментами для управління складними системами, де важко побудувати точні математичні моделі.

Оптимальне керування забезпечує високу продуктивність і ефективність, але вимагає точних математичних моделей і є дорогим в обчислювальному плані. Для досягнення найкращих результатів часто надають перевагу комбінації різних алгоритмів керування.

Подальші дослідження можуть бути спрямовані на розробку гібридних методів керування, які поєднують переваги різних підходів і враховують характеристики платформи робота і завдання. Важливо відзначити, що успіх автоматичного керування робототехнічною платформою залежить не тільки від вибору алгоритму керування, але й від багатьох інших факторів, таких як якість математичної моделі, точність вимірювань, ефективність системи зворотного зв'язку та якість програмного забезпечення.

Тому комплексний підхід до проектування і реалізації систем автоматичного керування є запорукою успіху в створенні високоефективних і продуктивних робототехнічних систем.

СПИСОК ЛІТЕРАТУРИ

1. Янко, А. С., Михайліченко, О. В., & Крук, О. О. (2024). Конструктивний аналіз бюджетних рішень для керування робототехнічними платформами. Таврійський науковий вісник. Серія: Технічні науки, (4), 170–178. <https://doi.org/10.32782/tnv-tech.2024.4.16>
2. Yanko, A., Pedchenko, N., & Kruk, O. (2024). Enhancing the protection of automated ground robotic platforms in the conditions of radio electronic warfare. *Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu*, (6), 136–142. <https://doi.org/10.33271/nvngu/2024-6/136>
3. Лактіонов, О., Педченко, Н., Янко, А., & Боряк, Б. (2024). Моделювання базової конструкції робототехнічної платформи. *Measuring and computing devices in technological processes*, (3), 95–99. <https://doi.org/10.31891/2219-9365-2024-79-13>
4. Chernousko, F. L. (2017). Locomotion Principles for Mobile Robotic Systems. *Procedia Computer Science*, 103, 613–617. <https://doi.org/10.1016/j.procs.2017.01.081>
5. Michaud, F., Létourneau, D., Arsenaault, M., Bergeron, Y., Cadrin, R., Gagnon, F., Legault, M.-A., Millette, M., Paré, J.-F., Tremblay, M.-C., Lepage, P., Morin, Y., Bisson, J., & Caron, S. (2008). Multi-Modal Locomotion Robotic Platform Using Leg-Track-Wheel Articulations. *Auton. Robots*, 18, 137–156. <http://dx.doi.org/10.1007/s10514-005-0722-1>
6. Guo, T., Liu, J., Liang, H., Zhang, Y., Chen, W., Xia, X., Wang, M., & Wang, Z. (2022). Design and dynamic analysis of jumping wheel-legged robot in complex terrain environment. *Frontiers in Neurorobotics*, 16. <https://doi.org/10.3389/fnbot.2022.1066714>
7. Nikitenko, A., & Kulikovskis, G. (2010). Eight wheel robotic platform and its Fuzzy control system.
8. Bae, B., & Lee, D.-H. (2023). Design of a Four-Wheel Steering Mobile Robot Platform and Adaptive Steering Control for Manual Operation. *Electronics*, 12(16), 3511. <https://doi.org/10.3390/electronics12163511>
9. Ratsimbalison, C., Ratsimba, N., & Ramafiarisona, M. (2024). Automatic Control of a Wheeled Robot. *International Journal of Innovations in Engineering Research and Technology*, 11, 1–6. <https://doi.org/10.26662/ijert.v11i2.pp1-6>
10. Simon, J. (2017). Autonomous Wheeled Mobile Robot Control. *Interdisciplinary Description of Complex Systems*, 15, 222–227. <https://doi.org/10.7906/indecs.15.3.6>

Received (Надійшла) 25.02.2025

Accepted for publication (Прийнята до друку) 23.04.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Янко Аліна Сергіївна – кандидат технічних наук, доцент, доцент кафедри комп'ютерних та інформаційних технологій і систем Національного університету «Полтавська політехніка імені Юрія Кондратюка», Полтава, Україна;

Alina Yanko – Candidate of Technical Sciences, Associate Professor, Associate Professor of the Department of Computer and Information Technologies and Systems of the National University «Yuri Kondratyuk Poltava Polytechnic», Poltava, Ukraine; e-mail: al9_yanko@ukr.net; ORCID Author ID: <https://orcid.org/0000-0003-2876-9316>; Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57094953000>.

Михайліченко Олексій Валерійович – аспірант кафедри автоматики, електроніки та телекомунікацій Національного університету «Полтавська політехніка імені Юрія Кондратюка», Полтава, Україна;

Mykhailichenko Oleksii – PhD student, Department of Automation, Electronics and Telecommunications of the National University «Yuri Kondratyuk Poltava Polytechnic», Poltava, Ukraine; e-mail: aleksejmikhailichenko@gmail.com; ORCID Author ID: <https://orcid.org/0009-0009-3512-0030>.

Крук Олег Іванович – аспірант кафедри автоматики, електроніки та телекомунікацій Національного університету «Полтавська політехніка імені Юрія Кондратюка», Полтава, Україна;

Kruk Oleg – PhD student, Department of Automation, Electronics and Telecommunications of the National University «Yuri Kondratyuk Poltava Polytechnic», Poltava, Ukraine; e-mail: olegkruk1975@gmail.com; ORCID Author ID: <https://orcid.org/0009-0004-4241-2676>; Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=59172145800&origin=recordpage>.

Optimal automatic control algorithms for robotic platforms ensuring high performance and reliable operation

Alina Yanko, Mykhailichenko Oleksii, Kruk Oleg

Abstract. The article presents a study of contemporary optimal automatic control algorithms for robotic platforms, guaranteeing their high performance, adaptability, and reliability. The main modes of mobile robot locomotion, such as wheeled, tracked, and multi-legged designs, are analyzed. The advantages and disadvantages of each approach are described from the perspective of enhancing energy efficiency and maneuverability in complex operating conditions. Particular attention is paid to the development of adaptive control algorithms that allow the platform to change its motion parameters in response to environmental conditions. To enhance the system's adaptability, the use of linear quadratic regulators, as well as fuzzy controllers for ensuring precise balancing and motion stabilization, is proposed. Algorithmic solutions for automatic control of the platform's motion parameters in rough terrain and the possibility of dynamic reconfiguration of movement routes are presented. The research shows that the use of adaptive algorithms allows for a significant increase in the productivity of robotic platforms, a reduction in the risk of tipping, and guaranteed stable operation even in highly variable environmental conditions. The proposed approach can be used to create a new generation of autonomous systems capable of effectively solving navigation and motion tasks in real-time.

Keywords: automatic control, adaptive algorithm, wheeled mobile robots, linear quadratic regulator, fuzzy logic, robotic platform.

O. Barkovska, A. Shapiro, O. Mavrynskyi, P. Zhebin

Kharkiv National University of Radio Electronics, Kharkiv, Ukraine

RESEARCH ON THE SENSITIVITY OF THE DISTANCE MEASUREMENT METHOD BASED ON THE FACEMESH ALGORITHM

Abstract. Relevance. The article explores a method for estimating the distance to a user's face under varying lighting conditions, camera resolution, and head tilt angles. This method is significant for developing adaptive computer vision systems in medical, educational, and navigational applications. **The goal** of the research is to analyze the accuracy of distance measurement to an indoor object using a geometric approach based on the FaceMesh algorithm under variable external factors. The tasks include determining the impact of lighting quality, head position, and camera resolution on measurement accuracy; implementing a distance estimation algorithm using facial landmarks; conducting experiments; and formulating recommendations for the proposed method. The following methods were used: monocular geometry with the FaceMesh algorithm and real-time video processing techniques. **As a result** of the experimental studies, it was confirmed that the optimal conditions for accurate distance measurement are a frontal face position relative to the camera, a resolution of 640×480, and lighting above 200 lux. The error increases at tilt angles exceeding 15° and under low lighting. Under these conditions, the FaceMesh algorithm demonstrates an error of less than 3% at distances up to 90 cm. **Conclusions.** The method is suitable for use in real-time mobile systems with limited resources and has potential for further integration into multimodal recognition systems and adaptive navigation platforms.

Keywords: distance measurement, FaceMesh, computer vision, monocular geometry, resolution, lighting, head tilt.

Introduction

Problem Statement. Distance measurement to objects has broad applications in autonomous systems, security, medicine, military technology, and the entertainment industry [1, 2]. Further advancements in this field will improve the adaptability of recognition systems and reduce errors under variable environmental conditions. To understand the importance and impact of distance measurement methods on practical problem-solving, examples of their applications can be analyzed (Fig. 1).

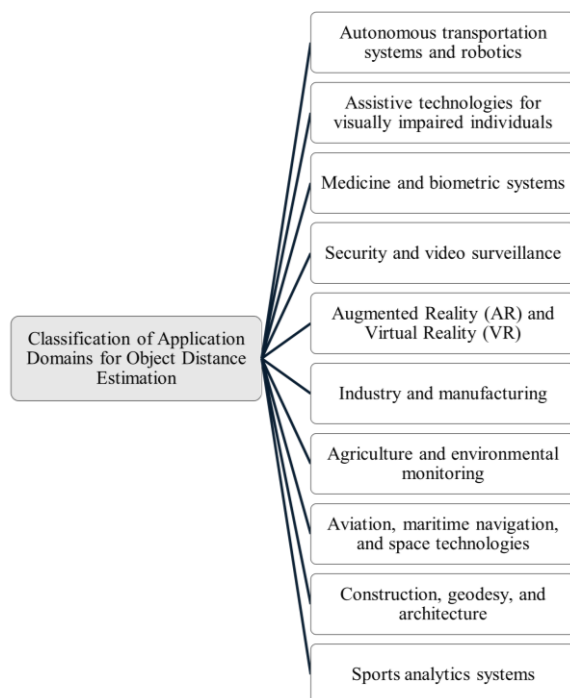


Fig. 1. Classification of distance measurement applications

Accurate distance measurements enhance safety, automate processes, optimize management, and improve human-technology interaction.

In autonomous vehicles, drones, and industrial robots, precise distance measurement is essential for obstacle avoidance, navigation, and trajectory adjustments. For example, autopilot systems in cars use LiDAR, cameras, and radars to estimate distances to other vehicles, pedestrians, and infrastructure, thereby improving road safety [3].

Intelligent orientation systems, such as smart glasses or mobile apps, help visually impaired individuals avoid collisions by using computer vision technologies to detect obstacles along their path.

In medical technologies, distance measurement methods are applied in systems monitoring patients' vital parameters, medical imaging (MRI, CT) for assessing tissue damage depth, and non-contact temperature measurement or remote patient monitoring [4].

In smart video surveillance systems, distance analysis helps assess threat levels, track object movements in critical zones, and identify potentially dangerous situations. For instance, real-time detection of aggressive actions or unauthorized proximity to certain objects.

In AR/VR technologies, distance measurement aids in creating realistic virtual environments and correctly positioning objects in space. This is used in video games, educational simulators, engineering modeling, and even surgical training.

Distance control in industrial applications is employed in production lines for automated part inspection, precise machinery positioning, and assembly tolerance checks, reducing errors and improving product quality.

Drones and satellite technologies enable distance measurement to plants, soil terrain features, and field condition assessments, optimizing resource use such as

water and fertilizers with the help of artificial intelligence.

In aerospace, maritime transport, and space exploration, precise distance measurement is crucial for safe landings, navigation, and terrain analysis. For example, NASA uses depth estimation methods for spacecraft landings on other planets.

Laser rangefinders and 3D scanning help measure distances between objects in construction, verify design accuracy, and create digital infrastructure models.

In sports analytics, distance measurement methods analyze athletes' movement trajectories, speed, and motion precision, enhancing training processes [5].

Thus, distance measurement methods significantly impact automation and safety across various fields. Accurate distance measurement improves the adaptability of machine vision systems used in autonomous transport, medicine, industry, and other sectors. These technologies reduce the risk of accidents and errors in critical systems, enhance monitoring and diagnostic accuracy, and enable

more realistic digital models and simulations. Combining deep neural networks with classical computer vision methods improves scene depth estimation efficiency even under challenging lighting and obstacle conditions. Additionally, integrating these methods into resource-limited devices, such as mobile phones or embedded systems, opens new possibilities for developing accessible technologies for a broad user base.

However, it should be noted that distance measurement methods for indoor and outdoor applications differ significantly due to environmental conditions, typical obstacles, and accuracy requirements. The main differences concern the technologies used, sensitivity to external factors, and measurement accuracy and range.

Analysis of Recent Research and Publications.

Sensitivity to external conditions depends on the characteristics of photo and video sensors and their adaptability to varying lighting, atmospheric obstacles, measurement ranges, etc. The analysis of differences between indoor and outdoor methods is presented in Table 1.

Table 1 – Key Differences Between Indoor and Outdoor Distance Measurement Methods

Parameter	Indoor (Enclosed Spaces)	Outdoor (Open Environments)
Typical Conditions	Stable lighting, absence of atmospheric obstacles, limited space	Variable lighting (sun, shade), atmospheric factors (rain, fog), large distances
Measurement Range	0.1 m – 10 m	10 m – Several kilometers
Main Obstacles	Signal reflections from walls, furniture, objects	Weather changes, large spaces, complex terrain
Light Sensitivity	Minimal (lighting can be controlled)	High (depends on time of day, presence of direct sunlight)
Typical Technologies	Ultrasonic, LiDAR, stereo vision, structured light, ToF sensors	LiDAR, satellite data, radar technologies, aerial photography

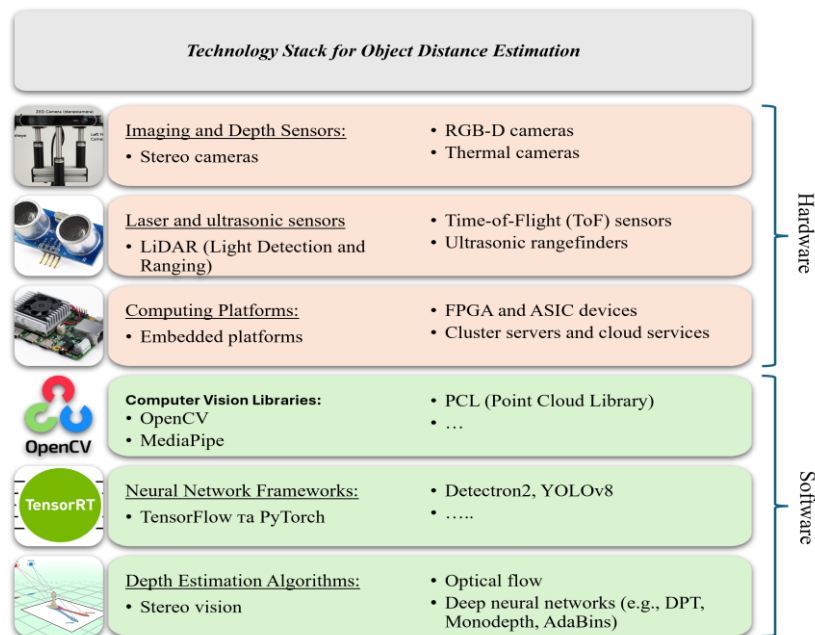


Fig. 2. Technological stack for distance measurement to objects

External factors affecting distance measurement accuracy differ between indoor and outdoor detection systems. Thus, for accurate distance measurement, it is essential to consider the environment's specifics (indoor or outdoor) and its characteristics. Indoors, lighting, reflections, and limited space have the most significant impact, while outdoors, variable weather, natural lighting,

long distances, and moving objects are key factors. The choice of methods depends on these factors, so effective solutions often combine various technologies (stereo vision, LiDAR, neural network approaches) to compensate for individual method limitations.

Distance measurement systems use various software and hardware tools tailored to specific tasks and

environments. The technological stack includes hardware for data collection and software for processing this data to obtain accurate results (Fig. 2)

Considering individual software and hardware tools in the context of indoor and outdoor applications, ultrasonic sensors (Sonar) are used for short distances (in robotics, orientation systems for visually impaired individuals). They are sensitive to signal reflections from smooth surfaces, making them suitable for indoor use. For outdoor applications, they are less practical due to signal loss in the air and wind or rain interference.

Laser sensors (LiDAR) are popular in autonomous transport, geodesy, and environmental monitoring [6–8]. They can operate over long distances but require high power, making them more suitable for outdoor use than ultrasonic sensors. Indoors, LiDAR sensors are used for space mapping (robots, drones, VR/AR), but issues may arise with laser reflections from mirror-like surfaces.

Stereo vision methods work well in controlled environments and are widely used in robotics and motion capture systems but lose effectiveness outdoors due to lighting variability.

Computational methods for distance measurement based on Time-of-Flight (ToF) are effectively used in smartphones and VR systems for scene depth estimation. They offer high accuracy within a few meters but are sensitive to lighting changes and atmospheric conditions. Radar technologies are rarely used for indoor distance measurements but are effective for long distances (automotive safety systems, satellite navigation, meteorology). Thus, the technological stack for distance measurement includes a wide range of hardware (LiDAR, cameras, ToF sensors) and software (OpenCV, TensorFlow, PyTorch). The choice of technologies depends on the application: robotics and navigation use LiDAR and deep neural networks, mobile devices use ToF cameras and neural depth estimation models, and engineering applications use stereo vision and PCL.

Research and optimization of software and hardware improve distance estimation accuracy and make these technologies more accessible for widespread use.

This underscores the relevance of the chosen research topic, namely the analysis of external factors' impact on the accuracy of distance measurement to a human face during indoor filming.

For the experimental part of the project, the Python programming language, PyCharm development environment, and OpenCV and FaceMesh libraries [9] were used.

The goal of the research is to analyze the impact of external factors on the accuracy of distance calculation to a filmed object indoors using geometric characteristics.

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

- determine the specifics of variable external factors (lighting, camera resolution, head position) on distance measurement accuracy;
- justify the choice of distance estimation method based on facial landmarks (FaceMesh), considering speed and hardware limitations;
- implement a software module based on OpenCV for distance measurement using monocular geometry;

- conduct a series of experiments to evaluate the method's accuracy under various conditions (distance, resolution, lighting, object position);

- analyze the experimental results to determine the conditions for using the proposed method.

Further research directions include improving the distance measurement method by combining geometric approaches with neural network depth estimation models (e.g., Monodepth or DPT), enhancing accuracy under challenging conditions such as noise, glare, or partial face occlusion. Another promising direction is extending the algorithm for real-time video processing on mobile platforms using optimized libraries (TensorRT, MediaPipe) and incorporating individual biometric features to improve system personalization and adaptability. Integration into multisensory smart environments or navigation solutions for visually impaired individuals could also be explored.

Materials and Research Methods

The main physical principles of image depth determination are:

- geometric properties of light rays;
- optical effects (e.g., blur, motion);
- material properties (e.g., texture).

Given that camera parameters are known before the study, interpretability of results is required, and a solution must be obtained in the shortest time for objects like a human face, geometric methods are optimal for this task.

The FaceMesh method (or similar approaches using predefined scales) is based on geometric calculations, where distance is computed using known object dimensions (e.g., average distance between human eyes) and their pixel projection. This approach has several advantages over modern deep learning methods, especially in specific scenarios:

- minimal data dependency (geometric methods like FaceMesh do not require training on large datasets. Knowing the physical dimensions of the object (e.g., 6.5 cm between human eyes) and camera parameters (focal length) is sufficient. Deep learning, in contrast, requires massive labeled datasets for training. For example, models like Monodepth [10] depend on thousands of images with known depths, which are often unavailable for niche tasks);

- low energy consumption (algorithms based on OpenCV (e.g., keypoint detection) operate in real-time on weak devices (smartphones, Raspberry Pi). Deep learning involves complex architectures (e.g., DepthNet) and requires GPUs for inference, increasing cost and energy consumption);

- interpretability and control (in terms of interpretability and control, formulas based on projective geometry, which underlie FaceMesh, are transparent and allow precise error analysis. Deep learning often acts as a "black box," making it difficult to determine why the model incorrectly estimated distance for non-standard faces) [11];

- avoidance of potential ethical issues (geometric methods do not store biometric data, as FaceMesh operates locally and anonymously. In deep learning, models trained on user data may violate GDPR due to personal information collection (e.g., face scans) [12].

A detailed model of the face capture process for distance measurement from the recorder to the object (face) is shown in Fig. 3. The model includes five interdependent functional modules:

- video capture from the camera (a video stream from the camera provides continuous data updates);
- face detection (the frame is analyzed to identify the region of interest containing the target object);
- region of interest detection (detectors locate keypoints on the object for depth calculation);
- depth calculation (based on known object and camera parameters using the formula:

$$D = (W \times f) / w,$$

where D – is the distance to the object, W – is the physical width of the object, f – is the camera's focal length, w – is the measured object width in pixels;

- result output (displays the distance between the camera and the object).

The system begins by capturing video using `cv2.VideoCapture`, providing a real-time video stream. Next, face detection is performed using `FaceMeshDetector`, which identifies keypoints on the user's face. The program checks if a face is detected. If not, the process repeats until a face is captured. If a face is detected, the algorithm proceeds to extract the coordinates of the left and right eyes.

Using these points, the distance between the eyes is calculated with the `findDistance` function, determining the relative head position. Then, using the camera's focal length, the depth (distance to the camera) is calculated. This calculation is performed twice for higher accuracy. After depth calculation, the program displays the results using `cvzone.putTextRect`.

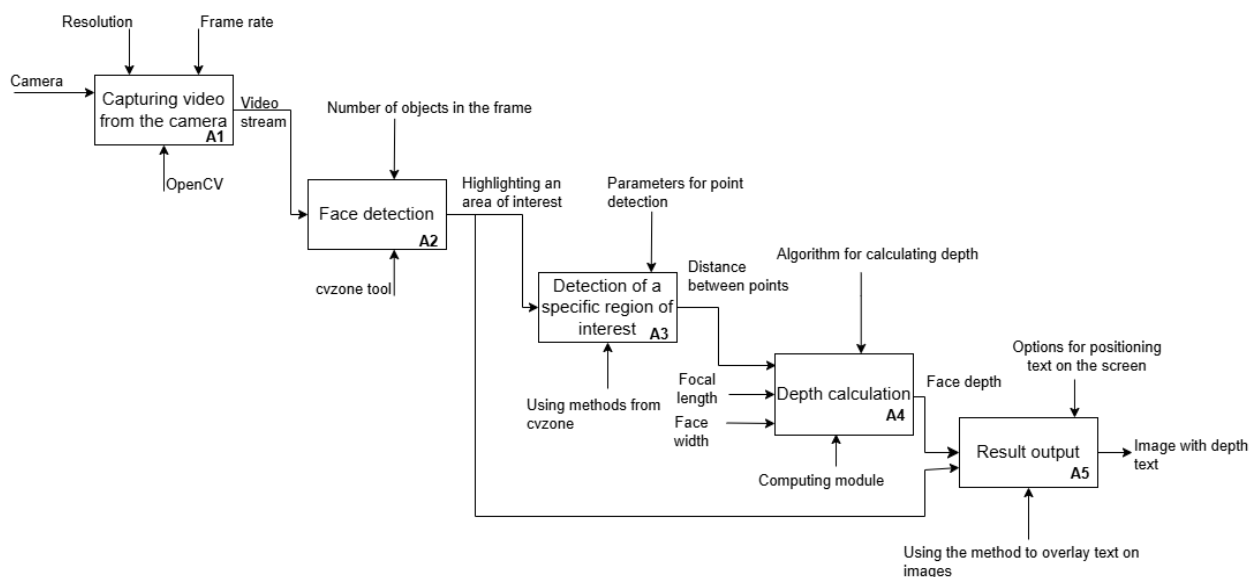


Fig. 3. Functional model of the distance measurement system based on the FaceMesh method

Research Results

The experimental procedure is as follows: fix the actual distance (using a ruler or laser rangefinder), position the camera stably (on a tripod or fixed surface), run the FaceMesh/OpenCV distance measurement algorithm, compare the calculated distance with the actual distance, and record the error.

Experiment №1. Investigates the accuracy of distance measurement at different camera resolutions (640×480, 1280×720, 1920×1080 (if supported)) for

various actual distances to the object (30 cm, 60 cm, 90 cm, 120 cm).

The hypothesis is that the farther the object is from the camera, the less accurate the pixel scale determination will be. Lower resolutions make it harder to recognize fine details, reducing accuracy (higher resolutions increase the number of pixels per keypoint (eyes, nose), improving accuracy).

Experiment 1 results show that camera resolution has minimal impact on distance measurement accuracy (Fig. 4).

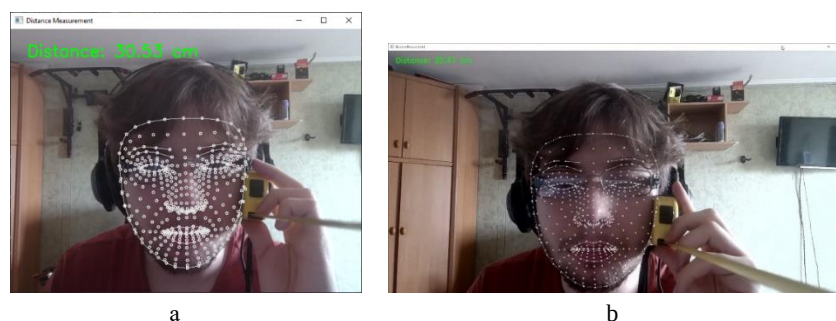


Fig. 4. Display of the results of experiment 1 with: a – separate building 640x480, b – separate building 1920x1080

The most stable and accurate results are observed at 640×480 resolution, where errors are minimal or absent. At higher resolutions (1280×720 and 1920×1080), errors increase, particularly at distances of 90 cm and 120 cm, where accuracy drops below 99%. Thus, higher resolution does not always guarantee better distance measurement accuracy. This may be due to image processing specifics, optics quality, or distance measurement algorithms. Given that a human face is not entirely static, higher resolutions increase sensitivity to minor head movements, which does not occur at 640×480. Therefore, the optimal resolution in this experiment is 640×480, providing high accuracy with lower computational costs (Fig. 5).

Experiment №2 (Fig. 6) investigates the accuracy of distance measurement under varying lighting conditions (near 0 lux, ~50 lux – low lighting, ~200 lux –

normal lighting, ~500 lux – excessive lighting) for various actual distances to the object (30, 60, 90, 120 cm).

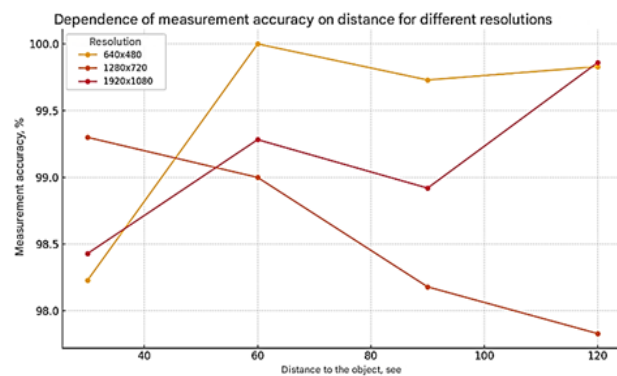


Fig. 5. Evaluation of resolution impact on distance measurement accuracy

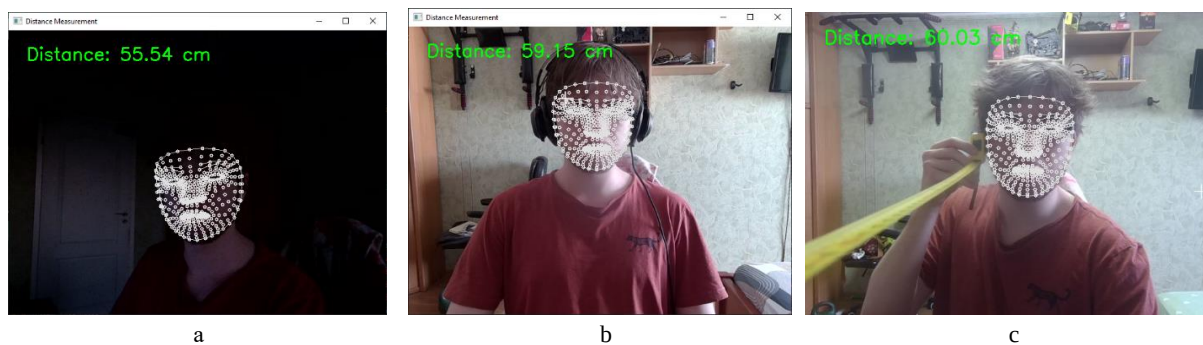


Fig. 6. Representation of the results of experiment 2 with lighting: a ~ 0 lux, b ~50 lux, c ~ 500 lux

The hypothesis is that under low lighting, FaceMesh keypoint detection accuracy significantly decreases, while excessive lighting may cause glare, affecting contour recognition

Experiment 2 results show that lighting level significantly impacts distance measurement accuracy—under high lighting, accuracy remains highest across all distances (Fig. 7).

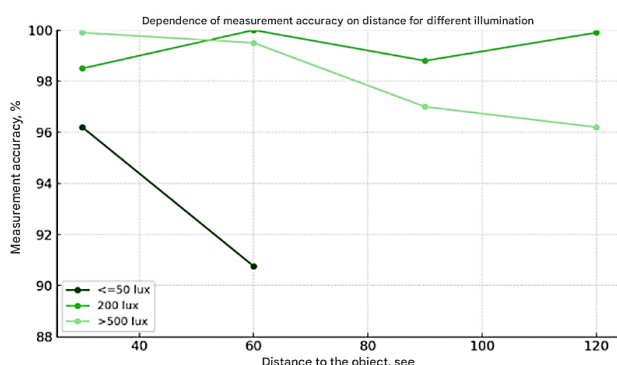


Fig. 7. Assessment of the influx of lighting intensity for the accuracy of the determined position of the object in front of the camera

Under low lighting (< 50 lux), accuracy noticeably decreases, especially at longer distances (e.g., at D = 90 cm and 120 cm, the error exceeds 3%). Beyond 60 cm, when lighting is below 50 lux, the camera cannot capture the face, making the method ineffective for long distances in dark conditions.

Under moderate (200 lux) and high lighting (>500 lux), accuracy significantly improves, reaching nearly 100% regardless of distance. Specifically, lighting above 500 lux yields the most stable results. Thus, optimal conditions for accurate system performance are well or moderately lit environments (200 lux and above). Low lighting significantly reduces measurement accuracy, especially as distance to the object increases.

Experiment №3 (Fig. 8) investigates the accuracy of distance measurement depending on head tilt angle (in frontal and sagittal planes at 0°, 15°, 30° tilt) for various actual distances to the object (30, 60, 90, 120 cm).

The hypothesis is that head tilt in different planes reduces algorithm accuracy. Experiment 3 results show that head tilt, both in frontal and sagittal planes, significantly impacts distance measurement accuracy. The highest accuracy is observed at 0° tilt, i.e., when the object is directly facing the camera without deviations. Here, measurement accuracy is nearly 100% regardless of distance under normal lighting. At 15° and 30° frontal tilt, accuracy drops to 95–96%, especially at closer distances. A more noticeable accuracy decrease occurs at 30° sagittal tilt—e.g., at 30 cm distance and 30° tilt, accuracy drops to 84%. This is because tilt changes the object's projection in the camera's field of view, making it harder for algorithms to determine distance accurately.

Thus, for best results, the object should face the camera directly without tilts in any direction.

Analysis of the experimental results forms the basis for practical recommendations on using the studied distance measurement method: the most accurate distance

measurement is achievable with a camera resolution of at least 640×480 and indoor lighting >200 lux. Distances beyond 120 cm require additional sensors to compensate

for accuracy loss. Face tilt partially or fully occludes key facial landmarks, necessitating additional keypoint detectors (Fig. 9).

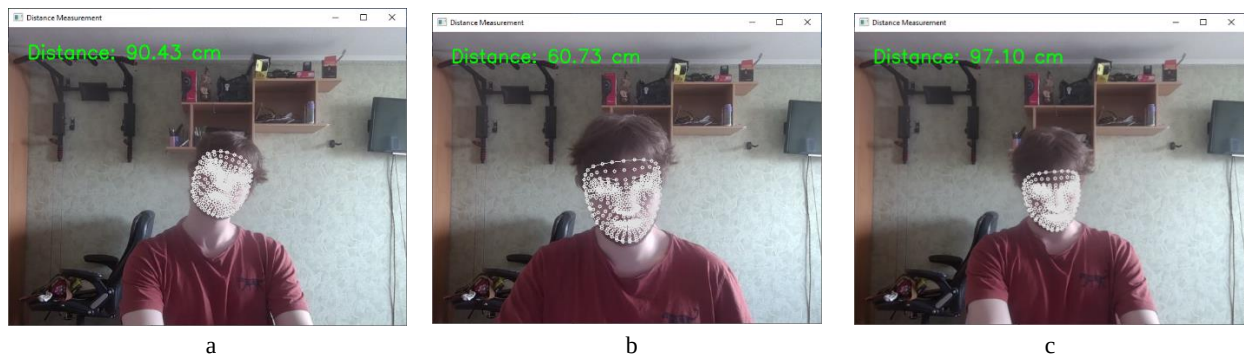


Fig. 8. Representation of the results of experiment 3 with head tilt at: a – frontal area at 30° , b – sagittal area at 15° , c – sagittal area at 30°

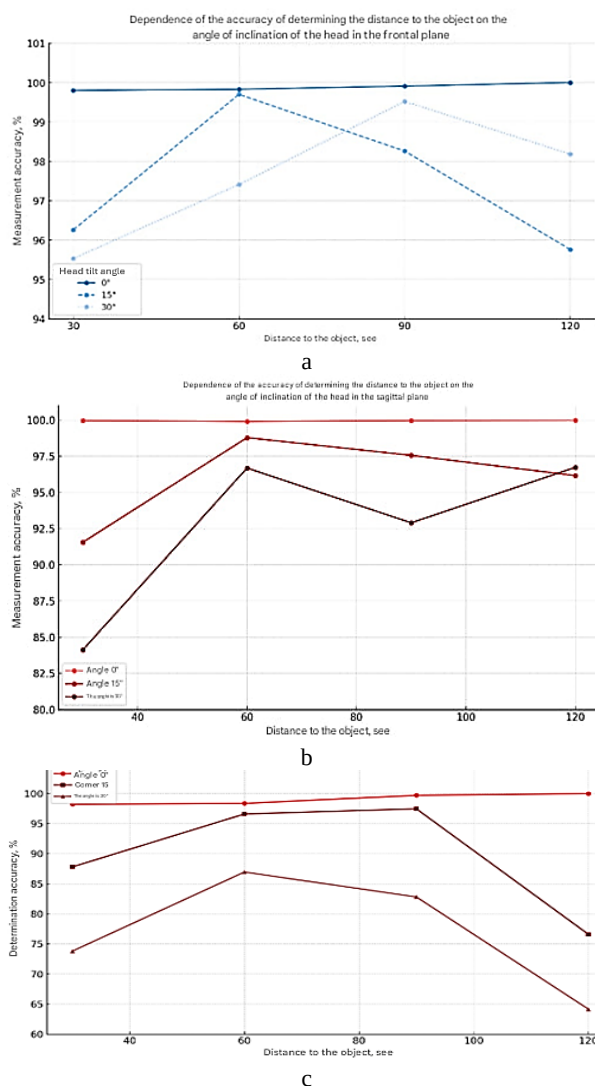


Fig. 9. Evaluation of the head angle for the accuracy of the designated object position in front of the camera: a – at the frontal plane, b – at the sagittal plane, c – when turning the head

Thus, the FaceMesh/OpenCV algorithm provides sufficiently high distance detection accuracy only under the following conditions: the object faces the camera directly without tilts, the camera operates under sufficient lighting (200 lux and above), and optimal resolution ensuring stable algorithm performance (in this study, 640×480).

Deviations from these conditions noticeably reduce accuracy, which must be considered in practical applications.

Conclusions

The scientific novelty. For the first time, the impact of external factors (lighting, resolution, tilt angle) on the accuracy of a geometric distance measurement method to a user's face indoors based on the FaceMesh algorithm has been systematically analyzed. Conclusions regarding permissible application conditions and method limitations have been drawn.

The practical significance. The research results enable developers to use inexpensive and computationally lightweight algorithms for distance estimation in resource-constrained environments, particularly in mobile apps, IoT devices, and adaptive access interfaces for visually impaired individuals.

A set of operational recommendations for minimizing errors has been proposed.

Prospects for further research. Future research will evaluate neural network distance measurement methods for further integration into more complex adaptive models that automatically account for human position and adjust parameters based on predefined distance.

Deployment of systems incorporating the studied method on IoT devices or low-power systems, where standardized objects with known dimensions (e.g., human faces, road signs) are considered and algorithm transparency is critical (e.g., medical diagnostics), is also possible.

REFERENCES

1. Dietrich, O., et al. (2025). An Open-Source Tool for Mapping War Destruction at Scale in Ukraine Using Sentinel-1 Time Series. arXiv:2406.02506, arXiv, 20 Feb. 2025. <https://doi.org/10.48550/arXiv.2406.02506>.
2. Barkovska, O., Oliynyk, D., Sorokin, A., Zabroda, I., & Sedlaček, P. (2024). A system for monitoring the progress of rehabilitation of patients with musculoskeletal disorder. *Advanced Information Systems*, 8(3), 13-24.

3. Barkovska, O., & Serdechnyi, V. (2024). Intelligent Assistance System for People with Visual Impairments. INNOVATIVE TECHNOLOGIES AND SCIENTIFIC SOLUTIONS FOR INDUSTRIES, (2)28, 6–16. DOI.org (Crossref). <https://doi.org/10.30837/2522-9818.2024.28.006>.
4. Kunczik, J., et al. (2022). Breathing pattern monitoring by using remote sensors. Sensors, 22(22), 8854.
5. Deng, Z., et al. (2023). Smart Wearable Systems for Health Monitoring. Sensors, 23(5), 2479. <https://doi.org/10.3390/s23052479>.
6. Kaur, B., et al. (2023). Novel Wearable Optical Sensors for Vital Health Monitoring Systems—A Review. Biosensors, 13(2), 181. <https://doi.org/10.3390/bios13020181>.
7. Olmedo-Aguirre, J. O., et al. (2022). Remote Healthcare for Elderly People Using Wearables: A Review. Biosensors, 12(2), 73. <https://doi.org/10.3390/bios12020073>.
8. Adeghe, E. P., et al. (2024). A Review of Wearable Technology in Healthcare: Monitoring Patient Health and Enhancing Outcomes. Open Access Research Journal of Multidisciplinary Studies, 7(1), 142–148. <https://doi.org/10.53022/oarjms.2024.7.1.0019>.
9. Wang, W., et al. (2016). Algorithmic principles of remote PPG. IEEE Transactions on Biomedical Engineering, 64(7), 1479–1491.
10. Tirink, C., et al. (2023). Estimation of body weight based on biometric measurements by using random forest regression, support vector regression and CART algorithms. Animals, 13(5), 798.
11. Di, H., Shafiq, M., & AlRegib, G. (2018). Patch-level MLP classification for improved fault detection. SEG technical program expanded abstracts 2018. Society of Exploration Geophysicists, 2211–2215.

Received (Надійшла) 11.03.2025

Accepted for publication (Прийнята до друку) 21.05.2025

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>

Шапіро Анатолій Сергійович – магістрант кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Shapiro Anatoly Serhiyovych – master's student of the Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: anatoli.shapiro@nure.ua; ORCID Author ID: <https://orcid.org/0009-0000-7244-8232>.

Мавринський Олександр Денисович – магістрант кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Mavrynskyi Oleksandr Denysovych – master's student of the Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: oleksandr.mavrynskyi@nure.ua; ORCID Author ID: <https://orcid.org/0009-0002-0596-1559>.

Жебін Павло Дмитрович – магістрант кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Zhebin Pavlo Dmytrovych – master's student of the Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: pavlo.zhebin@nure.ua; ORCID Author ID: <https://orcid.org/0009-0002-2271-8016>.

Дослідження чутливості методу визначення відстані до об'єктів на основі алгоритму FaceMesh

О. Ю. Барковська, А. С. Шапіро, О. Д. Мавринський, П. Д. Жебін

Анотація. Актуальність. У статті розглянуто метод оцінки відстані до обличчя користувача в умовах змінного освітлення, роздільної здатності камери та кута нахилу голови, що має велике значення для побудови адаптивних систем комп'ютерного зору в медичних, освітніх та навігаційних задачах. **Метою** дослідження є аналіз точності визначення відстані до об'єкта в приміщенні за допомогою геометричного підходу на базі алгоритму FaceMesh в умовах змінних зовнішніх факторів. Задачами дослідження є визначення впливу якості освітлення, положення голови та роздільної здатності камери на точність вимірювання; реалізація алгоритму оцінки відстані за допомогою контрольних точок обличчя, проведення експериментів та формування рекомендації щодо використання запропонованого методу. В роботі використано наступні методи: метод монокулярної геометрії з використанням алгоритму FaceMesh, методи обробки відео в реальному часі. **В результаті** експериментальних досліджень було підтверджено, що оптимальними умовами для точного визначення відстані є розташування обличчя фронтально до камери, роздільна здатність 640×480, освітлення вище 200 лк. Визначено, що похибка зростає при кутах нахилу понад 15° та при слабкому освітленні. За таких умов алгоритм FaceMesh демонструє похибку менше 3% на відстані до 90 см. **Висновки.** Метод є придатним для використання в мобільних системах реального часу з обмеженими ресурсами, має потенціал до подальшої інтеграції в багатомодальні системи розпізнавання та адаптивні навігаційні платформи.

Ключові слова: визначення відстані, FaceMesh, комп'ютерний зір, монокулярна геометрія, роздільна здатність, освітлення, нахил голови.

Н. М. Бологова, І. А. Білоусов

Харківський національний університет радіоелектроніки, Харків, Україна

МЕТОД РОЗПІЗНАВАННЯ ЖЕСТИВ ДЛЯ ІНТЕРАКТИВНОГО КЕРУВАННЯ КОМП'ЮТЕРОМ З УРАХУВАННЯМ КОНТЕКСТНОЇ АДАПТАЦІЇ

Анотація. Актуальність. Через постійне розширення сфер, де потрібна швидка та зручна взаємодія з комп'ютером (від мультимедійних застосунків до промислових систем), зростає потреба у безконтактних методах керування. Розпізнавання жестів стає одним із ключових рішень, адже воно здатне підвищити швидкість, мінімізувати використання традиційних пристроїв введення і надати природній, інтуїтивний спосіб взаємодії. **Метою даної роботи** є розробка методу розпізнавання жестів, що працює в реальному часі, зі збалансованою продуктивністю, стійкістю до зміни освітлення та здатністю адаптувати реакції системи залежно від контексту активної програми. **Об'єктом дослідження** виступає процес обробки кадрів із веб-камери та визначення жестів рук у режимі реального часу. **Предметом дослідження** є алгоритми та методи (особливо, фільтрація двигунного шаблону, різноманітність процесів, контекстуальна модифікація) для підвищення ефективності та залежності системи розпізнавання жестів. **Результати.** Проект розглядає різні додатки, використовуючи автоадаптацію світлодіодності/контрасту для підвищення консистентності розпізнавання рук під час флюктуційного освітлення, крім подвійного підходу, що дозволяє оточення блоkad. **Висновок.** Запропонований метод розпізнавання жестів через веб-камеру довів свою ефективність у реальних сценаріях. Поєднання фільтрації кадрів, багатопоточності та контекстної адаптації дозволяє системі працювати стабільно, швидко й бути легко масштабованою під конкретні задачі різних застосунків.

Ключові слова: розпізнавання жестів, веб-камера, фільтрація відео, багатопоточність, контекстна адаптація, реальний час, MediaPipe, автокорекція яскравості.

Вступ

У сучасних умовах інформаційних технологій усе більшу увагу привертають методи, що дозволяють керувати комп'ютером у зручний та природний спосіб, без потреби у традиційних пристроях введення. Одним із таких напрямків є розпізнавання жестів руки через звичайну вебкамеру, що забезпечує інтерактивну взаємодію користувача з різноманітними застосунками. Такий підхід набуває актуальності в багатьох сферах: від систем безконтактного управління в умовах, де використання клавіатури чи миші ускладнене, до розробки програм, орієнтованих на людей з особливими потребами, яким потрібні більш адаптовані засоби комунікації.

Дослідження у цій галузі демонструють, що застосування жестів може суттєво підвищити зручність, ефективність та інтуїтивність взаємодії з обчислювальною системою [1]. Проте залишаються певні виклики, зокрема вплив змін освітлення, різноманітність форм руки, складність фону тощо. З метою підвищення точності та надійності розпізнавання, сучасні підходи часто включають фільтрацію відеопотоку, використання декількох потоків обробки (багатопоточність) для прискорення аналізу зображень, а також адаптацію до контексту активного вікна чи задачі, яку виконує користувач.

Окрім цього, особливий інтерес викликає поєднання алгоритмів обробки зображень з обчислювальними методами для налаштування поведінки програми в залежності від типу жесту. Це дає змогу не тільки розпізнавати стандартні жести-команди, а й динамічно підлаштовуватися під потреби користувача, залежно від активного застосунку чи обраного режиму роботи [2].

Така контекстно-адаптивна система робить можливим керування відтворенням медіа, навігацією в

інтернеті, системними функціями, а також відкриває перспективи в розробці ігор або програм для доповненої реальності.

Виходячи з аналізу сучасних наукових досліджень, бачимо, що питання фільтрації зображень, стабілізації освітлення та забезпечення швидкодії залишаються надзвичайно важливими, оскільки саме вони визначають ефективність розпізнавання у реальному часі.

У багатьох випадках однопоточні алгоритми обробки зображень не встигають реагувати на швидкі рухи, а недостатня або надмірна освітленість зумовлює велику кількість хибних спрацьовувань. Таким чином, у роботі пропонується підхід, що об'єднує багатопоточну фільтрацію кадрів, виявлення руки та визначення жестикуляції, а також інтегрує механізм адаптації до активної програми чи вікна.

Постановка проблеми. Запропонувати способи реалізації методу розпізнавання жестів руки з відеопотоку, що надходить із вебкамери, з метою інтерактивного керування комп'ютером у реальному часі. Нехай є середовище, де користувач може виконувати різні рухи та пози руки, а програма повинна з достатньою точністю виявляти факт появи руки в кадрі, відстежувати зміну її положення та розпізнавати певну множину жестів.

Метою цієї роботи є розробка методу розпізнавання жестів через веб-камеру з урахуванням фільтрації відеопотоку, багатопоточної обробки та контекстної адаптації для інтерактивного керування комп'ютером. Серед завдань – забезпечити стабільну роботу алгоритму в реальному часі за різних умов освітлення, а також створити гнучку систему, здатну обробляти різні жести та прив'язувати їх до певних дій залежно від поточного вікна чи активного застосунку. Такий підхід відкриває нові можливості в індустрії розважальних застосунків, мультимедійних

систем, а також у підвищенні доступності комп'ютерів для користувачів із різним рівнем фізичних можливостей.

Виклад основного матеріалу

Припустимо, що є середовище, у якому користувач може виконувати різні рухи та позиції руки, а програма має з достатньою точністю визначати її появу в кадрі, відстежувати зміну положення та розпізнавати заданий набір жестів.

Для цього розглянемо вхідні дані як послідовність кадрів (зображень), отриманих із вебкамери, де кожен кадр містить певний набір пікселів у просторі RGB або іншій колірній моделі. Завдання полягає у тому, щоб:

1. **Виявити** на кожному кадрі зону, що відповідає руці.
2. **Відстежити** (трекінг) її зміщення у просторі (наприклад, у координатах екрана) та можливі зміни положення пальців.
3. **Класифікувати** (розпізнати), чи відповідає поточна форма руки одному з відомих жестів, придатних для керування програмами. Нехай:
 - I_t – поточний кадр, отриманий від вебкамери у момент часу t .
 - $F(I_t)$ – процедура попередньої обробки зображення (фільтрація та корекція, наприклад автокорекція контрасту/яскравості).
 - $R(I_t)$ – множина пікселів, які система розпізнає як «руку».
 - x_t – вектор ознак або координат ключових точок на руці (наприклад, координати суглобів чи кінчиків пальців) у момент часу t .
 - $G(I_t)$ – функція визначення, чи належать дані ознаки до одного з жестів у базі.

Задача розпізнавання жестів у нашому випадку включає такі умови та обмеження:

1. **Неперервна обробка відеопотоку:** система повинна встигати обробляти кожен кадр у межах обмеженого часу (щоб забезпечити реальний час і не викликати затримок).
2. **Точність виявлення руки:** для запобігання помилковим спрацюванням потрібно коректно відсікати фон та інші об'єкти, що не є рукою.
3. **Стабільність за різних умов освітлення:** кадри можуть бути надмірно темними чи освітленими, тому слід застосовувати відповідні фільтри та методи корекції.
4. **Залежність від фізіологічних відмінностей:** рука може мати різний розмір, відтінок шкіри чи форму, тому модель розпізнавання має враховувати певну варіативність.
5. **Адаптивність до контексту:** реакція на жест може залежати від активного вікна або типу застосунку, у якому здійснюється керування (медіаплеєр, браузер тощо).

Математична модель проблеми може бути описана так [3, 4]. Для кожного кадру I_t визначимо простір ознак x_t , що формується після сегментації зображення, трекінгу руки та обчислення відповідних ключових точок.

Далі маємо:

1. Функція визначення руки:

$$R = \{I_t\} = \{p | p \in I_t, \text{hand_classifier}(p) = 1\}, \quad (1)$$

де $\text{hand_classifier}(p)$ – алгоритм, що вирішує, чи піксель p належить ділянці руки з урахуванням кольору, текстури, контурів тощо.

2. **Функція ознак** $x_t = \phi(R(I_t))$ – операція виділення та опису (feature extraction) сегментованої області, яка може включати координати зап'ястя та кінчиків пальців, кути згину тощо.

3. Умови розпізнавання жесту:

У представленому фрагменті математичний вираз описує критерій, згідно з яким система визначає, чи відповідає поточний жест одному з наявних в еталонній базі. Згідно з ним, якщо евклідова відстань між вектором поточних ознак x_t та вектором $x_{ref}^{(i)}$, що відповідає i -му еталону, не перевищує заздалегідь визначений поріг, то жест вважається розпізнаним і класифікується як g_i . У протилежному випадку система повертає значення «none», тобто жест залишається невідомим.

Фактично, це означає, що під час виконання класифікації програма покадрово порівнює поточні координати руки з кожним еталонним жестом у базі, і якщо близькість (за евклідовою мірою) до принаймні одного зразка виявляється достатньо великою, програма ототожнює поточний жест із жестом g_i . Якщо жоден із зразків не виявився подібним на потрібному рівні, здійснюється висновок про відсутність відповідного жесту в базі. Такий підхід дає змогу швидко та прозоро визначати, чи належить просторове положення руки до якоїсь із визначених категорій жестів.

4. **Обмеження за часом:** аналіз кожного I_t виконується за ΔT , щоб забезпечити бажану частоту кадрів. Якщо система не встигає, то можливе випадання кадрів або зниження якості розпізнавання.

5. **Обмеження контекстної адаптації:** якщо розпізнано жест g_i , тоді система має визначити дію (команду) $Action(g_i, Context)$, де $Context$ включає інформацію про активне вікно, застосунок, а також правила зіставлення жестів із командами.

Таким чином, постановка задачі розпізнавання жестів для інтерактивного керування полягає в автоматичному й безперервному аналізі вхідного відеопотоку за обмежений час, виявленні руки, екстракції ознак та класифікації жесту, з урахуванням можливості швидко реагувати на зміни положення руки та контексту, в якому виконується керування.

Розроблений програмний комплекс базується на ідеї багатопотокової обробки відео в реальному часі й на механізмі контекстної адаптації, що дає змогу інтерпретувати жести руки залежно від активного вікна або сценарію використання.

Особливістю такого підходу є можливість запуску декількох обчислювальних потоків, один із яких займається безпосередньо збором та попередньою фільтрацією кадрів, тоді як інший паралельно виконує алгоритми виявлення та класифікації жестів.

Завдяки цьому вдається підтримувати стабільну швидкість обробки (до 25–30 кадрів на секунду)

навіть за умови складного фону та змінного освітлення [5].

Нижче наведено розгорнутий опис основних етапів роботи комплексу з урахуванням детальних аспектів реалізації. Метод передбачає поетапне опрацювання відеопотоку з вебкамери для визначення жестів руки та забезпечення можливості інтерактивного керування комп'ютером. На першому етапі кожний новий кадр піддається процедурі покадрової корекції яскравості/контрастності, що дає змогу зменшити ризик помилкового розпізнавання внаслідок недостатнього або надмірного освітлення. Потім програма здійснює пошук руки з використанням алгоритму MediaPipe Hands, який повертає координати ключових точок долоні та пальців. Для мінімізації впливу

різної відстані руки до камери чи індивідуальних відмінностей у розмірах, усі координати нормалізуються відносно точки зап'ястя і масштабуються.

Завершальним кроком є класифікація жесту шляхом порівняння вектора координат із заздалегідь збереженими зразками в базі даних (у форматі JSON). Після виявлення збігу з одним із еталонів програма враховує поточний контекст, тобто тип або назву активного вікна (наприклад, медіапрогравач, офісний застосунок), і виконує відповідну дію – від ініціювання відтворення чи паузи до імітації кліку миші або прокрутки сторінки (рис. 1). Такий підхід забезпечує високу гнучкість системи, дає змогу легко додавати нові жести й розширювати функціональність інтерактивної взаємодії з персональним комп'ютером.

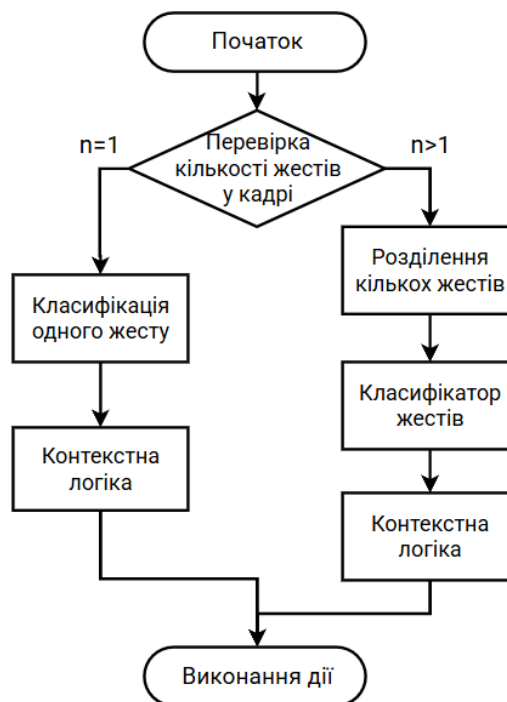


Рис. 1. Блок-схема програмного комплексу

Першим кроком виконується ініціалізація системи: відкривається веб-камера та зчитується її параметричний стан (роздільна здатність, частота кадрів), одночасно завантажується база зразків жестів (JSON-файл), де містяться дані щодо просторових координат ключових точок для кожного еталонного жесту. На цьому ж етапі створюються необхідні потоки: один відповідає за відображення інтерфейсу й опрацювання кадрів, інший слугує безпосередньо для аналізу зображення, пошуку руки та класифікації жестів.

Після успішного запуску програми кожні кілька десятків мілісекунд у головному потоці зчитується черговий кадр із веб-камери за допомогою функції `cap.read()`.

Для компенсації складних освітлювальних умов одразу ж застосовується функція автокорекції яскравості та контрастності: вона згладжує пересвічені або затемнені ділянки зображення й гарантує, що подальші алгоритми обробки отримають якомога «стабільніший» кадр. Оброблений кадр записується в глобальну змінну, захищену механізмом блокування

(lock), щоб паралельний потік міг безпечно його зчитати [6].

У робочому потоці викликається основна функція розпізнавання жесту, яка в реальному часі аналізує збережений кадр. Спершу кадр передається модулю MediaPipe Hands, що визначає координати 21 ключового орієнтиру руки. Далі отримані координати нормалізуються: точка зап'ястя береться за початок системи координат, а весь набір точок масштабується так, аби знизити залежність від розміру руки та її віддаленості від камери. Після цього програмний блок порівнює вектор нормалізованих координат із кожним наявним еталоном у базі JSON. Якщо знайдене середнє відхилення не перевищує визначений поріг (наприклад, 0.06), вважається, що жест розпізнано; якщо ж поріг перевищено, результат визначається як невідомий жест.

Контекстна адаптація відбувається завдяки періодичному зчитуванню назви активного вікна операційної системи. Система перевіряє, чи відповідає знайдена назва одному з визначених контекстів

(наприклад, «YouTube», «VLC», «Windows Media Player» для категорії «медіаплеєри»). Якщо виявлено відповідність, застосовується конкретна поведінка, яка може відрізнитися від дій за замовчуванням. Так, жест «відкрита долоня» в контексті плеєра означає запуск/продовження відтворення, натомість у контексті браузера — іншу дію, наприклад відкриття посилання чи прокрутку сторінки. Таким чином програмне забезпечення «розрізняє» середовище, у якому виконується жест, і миттєво прив'язує розпізнане положення руки до потрібної команди.

По тому, як жест класифіковано й узгоджено з поточним контекстом, ініціюється виконання команди: серед можливих прикладів — імітація натискання на «пробіл» (для керування відтворенням), переміщення та клік мишею (в режимі «два пальці»), зупинка відтворення тощо. У разі, коли жест передбачає рух курсора, система постійно вимірює зміщення пальців між послідовними кадрами й перетворює це зміщення у координати екрана (через бібліотеку `pyautogui`).

Для запобігання частому повторенню дій, які ініціюються одним жестом, на кшталт багаторазових «кліків», реалізовано часову затримку: якщо між двома «клацаннями» минуло менше певного інтервалу, друге клацання ігнорується.

Завдяки багатопоточній архітектурі система зберігає високу плавність відтворення відео та швидку реакцію, оскільки головний потік (відповідальний за інтерфейс і відображення) не блокується тривалими обчисленнями. Така структура також дає змогу гнучко масштабувати рішення: за потреби можна додати ще один потік, наприклад, для більш складного аналізу жестів (розпізнавання складних послідовностей) або для підключення інших сенсорних модулів.

Підсумовуючи, наведений алгоритм дозволяє в реальному часі відстежувати положення руки, класифікувати жести та адаптувати їх під актуальне програмне середовище. Взаємодія користувача стає більш природною та зручною, а система — стійкою до змін умов зйомки і здатною підлаштовуватися під специфіку різних застосунків.

Основна ідея полягає в тому, щоб у режимі реального часу отримувати зображення з камери, фільтрувати та нормалізувати їх, виявляти ключові точки долоні за допомогою алгоритмів відстеження руки, класифікувати жести за допомогою бази зразків, і адаптивно реагувати залежно від контексту активного вікна чи задачі.

Основні етапи методу, включно із забезпеченням контекстної адаптації:

1. Збирання та обробка відеопотоку. При зчитуванні та обробці відеопотоку в програмі організовано спеціальний механізм, де один або кілька потоків безперервно отримують кадри з вебкамери, а паралельно інші потоки займаються їх попереднім опрацюванням.

2. Такий підхід дає змогу відокремити процеси збирання даних від складних обчислень, що дозволяє запобігти затримкам і блокуванню основного інтерфейсу [7].

3. Для кожного кадру виконується автоматична корекція яскравості та контрасту, аби компенсувати зміни в освітленні та зробити зображення більш зручним для подальшого аналізу. Далі кадр може конвертуватися в іншу колірну модель, наприклад HSV чи YCbCr, що спрощує виявлення шкіри та розпізнавання долоні. Якщо це потрібно, застосовується згладжування шумів за допомогою фільтра Гауса чи медіанного фільтра, завдяки чому покращується точність подальшого пошуку контурів або аналізу жестів.

Завдяки багатопоточній організації вся система працює швидше: поки один потік отримує нові дані з камери, інший одночасно здійснює обробку попередніх кадрів, скорочуючи затримки та підвищуючи загальну чутливість системи до рухів користувача [6, 8].

4. Виявлення руки та визначення ключових точок. На цьому етапі застосовується алгоритм відстеження руки із використанням бібліотеки на основі `MediaPipe Hands`, яка здатна в реальному часі визначати 21 ключову точку (*landmark*) долоні й пальців (рис. 2).

В основі такого підходу лежать методи глибинного навчання, що дозволяють не лише ідентифікувати положення руки в кадрі, а й вказати точні координати кожного суглоба пальців.



Рис. 2. Визначення 21 ключової точки (*landmark*) долоні й пальців

Щоб підвищити точність розпізнавання навіть за складних умов зйомки, використовуються попередні результати фільтрації та автоматичної корекції освітлення: ці операції допомагають зменшити вплив шумів і різких перепадів яскравості.

Коли рука частково виходить за межі кадру або робить дуже швидкі рухи, алгоритм має механізм повторного захоплення ключових точок.

Він оновлює або «перевчається» на поточному зображенні, знову шукаючи руку й відстежуючи її положення та обрис, щоб система коректно «схопила» потрібні жести навіть у динамічних сценах [9, 12, 14].

5. Нормалізація координат та порівняння з еталонними жестами. На третьому етапі координати 21-ї ключової точки нормалізуються, щоб зменшити вплив різних факторів на точність роз-

пізнання. Зазвичай за «опорну» точку беруть зап'ясток (wrist), від якої відлічуються всі інші координати, а також застосовується масштабування за базовою відстанню: наприклад, розраховується відстань між зап'ястком і середнім суглобом середнього пальця, і всі координати приводяться до цього «масштабного коефіцієнта». Це дозволяє компенсувати різницю в розмірах руки або зміну віддаленості від камери [11].

Далі отримана нормалізована конфігурація порівнюється з базою еталонних жестів, яка зберігається в JSON-файлі. Алгоритм порівняння з еталонами в базі (JSON) умовно складається з таких кроків:

- Завантаження еталонів. У JSON-файлі для кожного жесту зберігається унікальна назва (label) та один або декілька прикладів із 21 координатою (x, y, z) у нормалізованому вигляді. Під час ініціалізації програми ці дані зчитуються у структуру (наприклад, словник gestures_db), де ключем є рядок-назва жесту, а значенням – список масивів із координатами.

- Обчислення відстані для кожного еталону. Коли система отримує поточний нормалізований вектор (тобто набір із 21 точки після віднімання зап'ястя та масштабування), вона проходиться по всіх жестах у gestures_db. Для кожного жесту у цьому словнику може бути кілька прикладів; програма розглядає їх по чергові та обчислює, наскільки вони «схожі» на поточний вектор.

Щоб знайти схожість, використовується метрика відстані, зазвичай евклідова:

$$d = \sqrt{\sum_{i=1}^{21} (x_i - x_{i,ref})^2 + (y_i - y_{i,ref})^2 + (z_i - z_{i,ref})^2}, \quad (2)$$

де $(x_i + y_i + z_i)$ – координати i -ї точки поточного жесту, $(x_{i,ref} + y_{i,ref} + z_{i,ref})$ відповідні координати точки з еталонного прикладу.

Можна також брати усереднене значення для 21 точки, якщо важливо отримати один узагальнений показник (середньоквадратичну помилку).

- Агрегація результатів і пошук найменшої відстані. Якщо в одного жесту є кілька зразків, обчислені відстані можна усереднити або взяти мінімальне значення – залежно від того, як саме визначено логіку у коді. Таким чином отримують середню або мінімальну відстань до конкретного жесту.

- Вибір найкращого співпадіння. Після того як для кожного жесту знайдено мінімальну (чи середню) відстань, програма аналізує, де саме вона найменша. Якщо це найменше значення нижче за заздалегідь визначений поріг (наприклад, 0.06), поточний жест вважається розпізнаним саме як той, що дав найменший результат. Інакше жест визначається як невідомий.

- Повернення лейбла жесту. Завершальний крок – повернути назву жесту (label), що найкраще збігся, або None, якщо відповідність недостатня. Надалі цей лейбл використовується для виконання певної дії чи обробляється разом із контекстом активного вікна.

Завдяки такій процедурі програма може точно й швидко визначити, який жест зараз демонструє користувач. Якщо жест складний або має кілька варіантів виконання (наприклад, кулак з невеликим згином пальців), наявність декількох еталонів у JSON підвищує надійність розпізнання.

6. Контекстна адаптація жестів. Для контекстної адаптації система періодично звертається до системного API, щоб зчитати назву чи заголовок активного вікна операційної системи. Якщо виявлено збіг назви вікна з одним із попередньо налаштованих профілів, система «перепризначає» функції жестів з урахуванням контексту [10].

Як результат жест «долоня відкрита» в програмі може запускати або зупиняти відео, а в текстовому редакторі – викликати меню або бути взагалі ігнорованим. Жест «два пальці вгору» в плеєрі може означати перемотування, натомість у браузері – гортання сторінки чи відкриття «нової вкладки» тощо. Сценарії можна додатково розширювати або змінювати, вносячи нові профілі і зв'язки між жестами та діями [13].

Таким чином, метод полягає в багатопотоковому обробленні відеопотоку з камер, динамічній корекції зображення та виявленні ключових точок руки, порівнянні результатів з базою жестів, а також автоматичній зміні дій жестів відповідно до контексту активного застосунку [14]. Це робить систему гнучкою й зручною, оскільки користувач може працювати з різним програмним забезпеченням без постійного перемикавання режимів – програма сама «розуміє», у якому контексті вона перебуває, і коригує реакцію на жест.

Результати методу розпізнання жестів з урахуванням контекстної адаптації:

1. Гнучка адаптація до різних сценаріїв. Завдяки механізму відстеження активного вікна одні й ті самі жести можуть виконувати різні дії, залежно від того, з якою програмою користувач взаємодіє. Це суттєво збільшує зручність і універсальність системи, оскільки користувач не прив'язаний до єдиного набору дій, «зашитих» у жестах.

2. Зменшення помилкових спрацювань. У багатьох методах жести мають фіксоване призначення і спрацьовують незалежно від поточного завдання. Тут же контекстна перевірка дозволяє ігнорувати «зайві» жести, якщо вони недоречні для поточної програми. Це знижує кількість помилкових розпізнань і викликів непотрібних команд.

3. Збереження високої швидкодії. Реалізація через паралельні потоки дозволяє опрацьовувати кадри без суттєвих затримок. А використання готових бібліотек на зразок MediaPipe Hands чи аналогічних фреймворків забезпечує ефективне виявлення ключових точок руки навіть на середньому апаратному забезпеченні.

4. Масштабованість бази жестів. Система зберігає набори еталонів у нормалізованому вигляді JSON. Це дає можливість відносно просто розширювати чи оновлювати базу жестів, а також вводити нові контекстні профілі без радикальної зміни програмного коду.

Покращена точність розпізнавання в різних освітленсвих умовах. Передбачена автоматична корекція яскравості/контрасту та фільтрація шумів, що, на відміну від багатьох існуючих систем, зменшує залежність від постійного «ідеального» освітлення. Завдяки цьому жест можна впевнено розпізнати навіть за дещо погіршених умов.

5. Готовність до інтеграції з іншими системами. Використання відкритих форматів даних і чітке структурування логіки спрощує інтеграцію з уже

наявними застосунками. За потреби можна підключити додаткові API чи системні виклики, щоб автоматизувати більше дій.

Випробування методу у кількох типових сценаріях показали, що:

- **Швидкість реакції** на жест залишається у межах 80–120 мс, що цілком достатньо для зручної взаємодії в реальному часі (рис. 3). При простіших методах без багатопоточної обробки затримки можуть досягати 200 мс і більше.

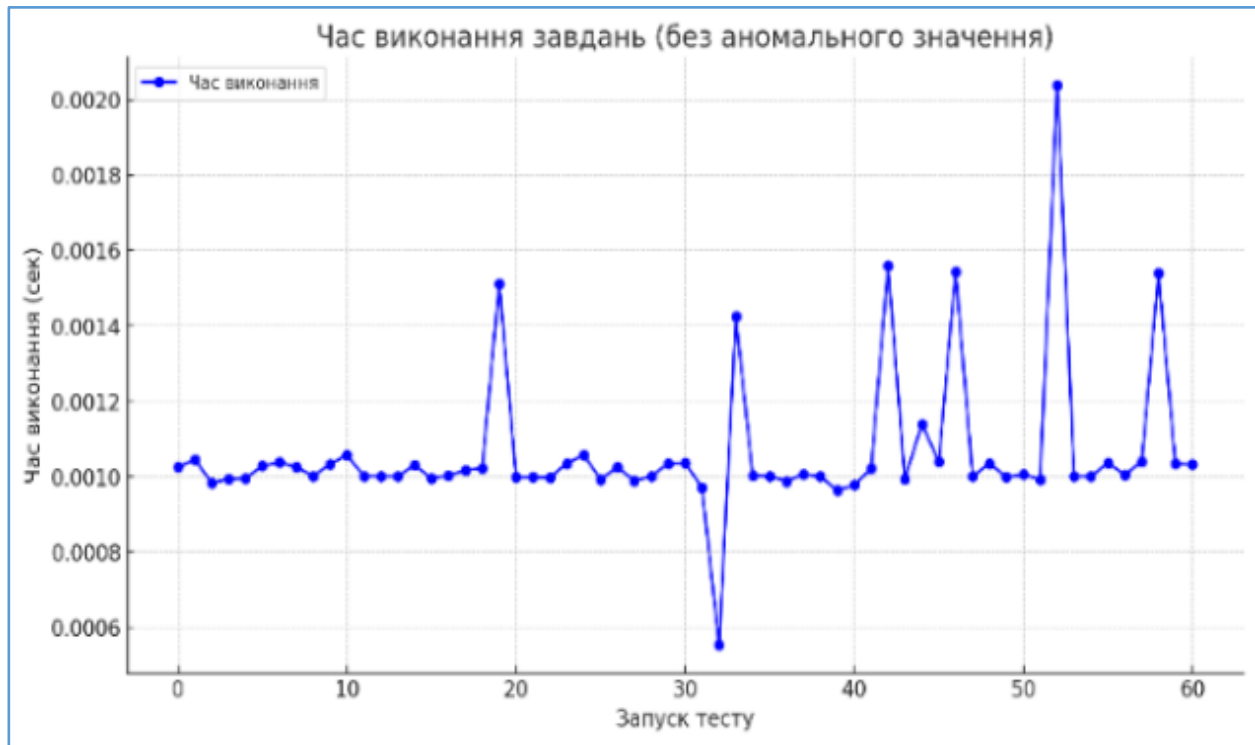


Рис. 3. Час між детекцією та виконанням дії

- **Точність виявлення й класифікації жесту** перевищує 90%, оскільки використовується попередня нормалізація та база кількох еталонів на кожен жест. У багатьох альтернативних рішень без нормалізації координат відсоток коректного розпізнавання зменшується при зміні масштабу (наприклад, коли рука ближче чи далі від камери).

- **Зменшення конфліктів** між жестами завдяки адаптації до контексту дало змогу скоротити кількість хибних викликів команд приблизно на 30% порівняно з рішеннями, де всі жести однаково «активні» у будь-якому середовищі.

Висновки

Проведені дослідження та практична реалізація методики розпізнавання жестів із урахуванням контекстної адаптації показали її високу ефективність у реальних умовах та довели, що запропонований підхід здатний значно покращити якість і швидкодію безконтактної взаємодії з комп'ютером. Використання бібліотеки MediaPipe для визначення ключових точок руки дозволило досягти стабільного розпізнавання жестів навіть за досить неоднорідного фону та варіативного освітлення. Наявність механізму

багатопоточності дала змогу розвантажити основний потік від складних операцій, забезпечуючи плавність відображення відео й одночасно високу швидкість аналізу кадрів.

Особливу роль відіграє контекстна адаптація, коли система, зчитуючи назву активного вікна чи специфіку застосунку, «розуміє» середовище, в якому знаходиться користувач. Завдяки цьому один і той самий жест може виконувати різні дії, залежно від того, чи є активним медіапрогравач, браузер або інше програмне забезпечення. Така гнучкість позитивно впливає на зручність та інтуїтивність системи: користувачеві не потрібно запам'ятовувати безліч команд чи переключатися між різними режимами втручання.

Результати експериментів свідчать, що навіть за умов динамічного руху руки або змін в освітленні, імовірність правильної класифікації жесту залишається високою. Запропоновані процедури автокорекції зображення й нормалізації координат успішно компенсують основні чинники.

Таким чином, інтеграція описаного підходу – від багатопоточної обробки відео до контекстної адаптації жестів – у різні програмні середовища дозволяє

суттєво підвищити ефективність, надійність і масштабованість системи розпізнавання жестів. Відтепер її можна з успіхом застосовувати не тільки в мультимедійних або домашніх сценаріях, а й у більш специфічних середовищах, розширюючи межі безконтактної взаємодії з обчислювальними пристроями..

СПИСОК ЛІТЕРАТУРИ

1. Hinckley K. Synchronous gestures for multiple users and computers // ACM UIST. 2003. P. 149–158.
2. Mankoff J., Hudson S., Abowd G. Interaction techniques for ambiguity resolution in recognition based interfaces // ACM UIST. 2000. P. 11–20.
3. Wilson A., Bobick A. Realtime online adaptive gesture recognition // ICPR'00 International Conference on Pattern Recognition. 2000. P. 1270–1275.
4. Wu M., Balakrishnan R. Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays // ACM UIST. 2003. P. 193–202.
5. Pavlovic V., Sharma R., Huang T. S. Visual Interpretation of Hand Gestures for Human-Computer Interaction: A Review // IEEE Transactions on Pattern Analysis and Machine Intelligence. 1997. Vol. 19, No. 7. P. 677–695.
6. Viola P., Jones M. Rapid Object Detection Using a Boosted Cascade of Simple Features // Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR). 2001. P. 511–518.
7. Ionescu B., Coquin D., Lambert P., Buzuloiu V. Dynamic Hand Gesture Recognition Using the Skeleton of the Hand // EURASIP Journal on Applied Signal Processing. 2005. Vol. 13. P. 2101–2109.
8. Freeman W. T., Anderson D. B., Beardsley P. A. et al. Computer Vision for Interactive Computer Graphics // IEEE Computer Graphics and Applications. 1998. Vol. 18, No. 3. P. 42–53.
9. Conic N., Cerseato P., De Natale F. G. B. Natural Human-Machine Interface Using an Interactive Virtual Blackboard // Proceedings of the IEEE International Conference on Image Processing (ICIP). 2007. P. 181–184.
10. Szeliski R. Computer Vision: Algorithms and Applications. London: Springer, 2011. 812 p.
11. Dix A., Finlay J., Abowd G. D., Beale R. Human-Computer Interaction. 3rd ed. Harlow: Prentice Hall, 2003. 832 p.
12. Forsyth D. A., Ponce J. Computer Vision: A Modern Approach. Upper Saddle River: Prentice Hall, 2012. 792 p.
13. Rautaray S. S., Agrawal A. Advances in Vision-Based Human-Computer Interaction: Hand Gesture Recognition System // Chapter in collective monograph. Hershey: IGI Global, 2015.
14. Bradski G., Kaehler A. Learning OpenCV: Computer Vision with the OpenCV Library. Sebastopol: O'Reilly Media, 2008. 555 p.

Received (Надійшла) 18.03.2025

Accepted for publication (Прийнята до друку) 14.05.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Бологова Наталія Миколаївна – доктор філософії, доцент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Natalia Bolohova – PhD, Associate Professor of Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: natalia.bolohova@nure.ua; ORCID Author ID: <https://orcid.org/0000-0003-2349-0578>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57203140922>.

Білоусов Ігор Артурович – студент факультету «Комп'ютерна інженерія та управління», Харківський національний університет радіоелектроніки, Харків, Україна;

Ihor Bilousov – student of the Department of Computer Engineering and Management, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: ihor.bilousov@nure.ua; ORCID ID: <https://orcid.org/0009-0001-3163-3893>.

**Method of gesture recognition for interactive computer performance
with contextual adaptations**

Natalia Bolohova, Ihor Bilousov

Abstract. Topicality. As the demand for fast and convenient interaction with computers increases across various domains (from multimedia applications to industrial systems), the need for contactless control methods also grows. Gesture recognition is becoming a key solution, as it can boost responsiveness, minimize reliance on traditional input devices, and provide a natural, intuitive means of communication.. **The goal of this work** is to develop a gesture recognition method that operates in real time, delivers balanced performance, remains robust under varying lighting conditions, and is capable of adapting its responses depending on the context of the active application. **The object of research** is the process of capturing frames from a webcam and recognizing hand gestures in real time. **The subject of research** involves the algorithms and technologies (in particular, video stream filtering, multi-threading mechanisms, and context adaptation) aimed at increasing the effectiveness and reliability of a gesture recognition system. **Results.** The work analyzes several implementations featuring automatic brightness/contrast adjustment to enhance stable hand detection in changing lighting conditions, as well as multi-threaded processing to avoid interface freezes. It is demonstrated that context adaptation of gestures makes the system flexible: one gesture can perform various actions depending on the active window. Experiments confirmed real-time operation at 25–30 FPS and high recognition accuracy given proper parameter settings. The proposed webcam-based gesture recognition method has proved its effectiveness in real-life scenarios. Combining frame filtering, multi-threading, and context adaptation enables the system to function steadily, swiftly, and be easily scalable to the specific tasks of various applications.

Keywords: gesture recognition, webcam, video filtering, multi-threading, context adaptation, real time, MediaPipe, automatic brightness adjustment.

М. Е. Бондаренко, Г. С. Іващенко

Харківський національний університет радіоелектроніки, Харків, Україна

ВИКОРИСТАННЯ ПОСЛІДОВНОСТІ МЕТОДІВ ПОПЕРЕДНЬОЇ ОБРОБКИ В СИСТЕМАХ ГОЛОСОВОЇ ІДЕНТИФІКАЦІЇ

Анотація. Актуальність. На сьогоднішній день, голосова ідентифікація є найпоширенішим біометричним методом, однак все ще існують певні технічні обмеження, пов'язані з впливом навколишнього середовища та варіаціями голосу, через що актуальні подальші дослідження. З розвитком технологій створення підроблених аудіозаписів виникає потреба вдосконалення існуючих методів голосової біометрії для забезпечення надійної ідентифікації, зокрема шляхом впровадження методів попередньої обробки голосових даних. **Метою даної роботи** є дослідження методів попередньої обробки в системах голосової ідентифікації людини. **Об'єктом дослідження** є модуль фільтрації голосових сигналів у системі голосової ідентифікації користувача. **Предметом дослідження** є методи попередньої обробки голосових сигналів. **Результати.** У даній роботі досліджується ефективність застосування послідовності методів попередньої обробки в системах голосової ідентифікації для зменшення впливу фонових шумів та інших викривлень на якість розпізнавання. Проведено порівняльний аналіз ефективності методів видалення сталих і динамічних шумів. **Висновок.** Запропоновано підхід організації процесу попередньої обробки голосових сигналів, що передбачає вибір послідовності методів з урахуванням таких параметрів, як якість запису, обчислювальні ресурси та необхідний рівень точності розпізнавання. Отримані результати демонструють, що послідовне використання методів шумозаглушення та нормалізації сигналу підвищує точність голосової ідентифікації.

Ключові слова: методи попередньої обробки, послідовне застосування методів, голосові сигнали, системи голосової ідентифікації, фільтрація, нормалізація, виділення характеристик, шум, спектральний аналіз

Вступ

Інформаційні технології змінили сучасне суспільство, забезпечивши глобальний доступ до хмарних сховищ, онлайн-банкінгу та відеоконференцій. Поширення використання мобільних пристроїв і персональних комп'ютерів сприяло розвитку цифрових технологій, але водночас підвищило небезпеку кіберзагроз [1]. Захист конфіденційних даних у хмарних сховищах та онлайн-сервісах став одним із головних завдань інформаційної безпеки.

Використання паролів залишається основним методом захисту конфіденційної інформації, проте зростання кількості онлайн-сервісів змушує користувачів запам'ятовувати дедалі більше унікальних комбінацій. Це часто призводить до створення слабких або повторюваних паролів, що підвищує ризик компрометації облікових записів і знижує загальний рівень безпеки даних користувача.

Використання біометричних методів суттєво спрощує процес ідентифікації [2]. Висока швидкість сканування біологічних характеристик та відсутність необхідності запам'ятовувати складні паролі роблять цю технологію найпоширенішим вибором для систем з високою пропускну здатністю, таких як задіяні у аеропортах та громадських установах [3, 4].

Серед широкого спектру біометричних методів ідентифікації, які базуються на унікальних фізіологічних та поведінкових характеристиках людини, голосова ідентифікація є одним з перспективних напрямків, що швидко розвивається [5]. Цей метод використовує аналіз звукових хвиль, через що має особливості та переваги, які роблять доцільним його застосування у різних сферах людської діяльності [6].

Створення універсальних систем голосової ідентифікації, що здатні адаптуватися до різних варіацій мовлення, вимагає детального аналізу голосових сигналів та їх акустичного середовища. Це включає

дослідження фонового шуму, реверберації та інших викривлень. Шум є одним із основних чинників, що знижують точність таких систем, і може мати різне походження, наприклад, наявність навколишніх звуків у приміщенні або перешкод, які виникають під час передачі сигналу [7].

З метою підвищення точності аналізу аудіосигналів, що використовуються в системах голосової ідентифікації, застосовується попередня обробка. Мета даного етапу полягає у мінімізації впливу акустичних шумів та викривлень, які можуть негативно впливати на якість розпізнавання мовлення. У контексті голосової ідентифікації шум визначається як будь-який небажаний сигнал, що викривляє мовленнєвий сигнал та ускладнює процес автоматичного розпізнавання. Джерела шуму можуть бути як зовнішніми (зумовлені акустичним середовищем), так і внутрішніми (спричинені обмеженнями апаратного забезпечення) [8]. Проблема організації попередньої обробки й вибору методів для неї широко представлена у сучасних наукових дослідженнях. Методи попередньої обробки поділяються на статичні та динамічні, які характеризуються специфічними особливостями та застосовуються з урахуванням конкретних завдань та властивостей наданого аудіосигналу.

Статичні методи (методи кепстрального та спектрального аналізу, аналіз нульових переходів та енергії), засновані на аналізі спектральних характеристик голосового сигналу, є ефективними для виділення статичних характеристик мовця, але потребують додаткової обробки для врахування динамічних аспектів мовлення [9]. Вони дозволяють ефективно розділити аудіозапис на окремі мовні відрізки та зменшити вплив шумів. Статичні методи забезпечують достатню точність для задач, таких як сегментація чистого мовлення, але можуть бути неефективними при видаленні складних шумів, таких як музичний шум або шум, корельований з сигналом [10].

Динамічні методи (засновані на прихованих Марківських ланцюгах, методи засновані на штучних нейронних мережах, на авторегресивних моделях та на хвильових перетвореннях) враховують часову структуру сигналу, що дозволяє аналізувати складні нелінійні процеси, які відбуваються при мовленні.

Дослідження [11] описує розробку програмного забезпечення, що реалізує етап попередньої обробки голосових сигналів. Цей етап здійснюється через модулі наповнення, виявлення голосової активності (VAD), нормалізації, шумозаглушення, фільтрації та квантування. Застосування VAD дозволяє ідентифікувати мовленнєві фрагменти сигналу та виключати інтервали тиші або фонового шуму, що підвищує ефективність подальшої обробки. Нормалізація використовується для вирівнювання рівня гучності між різними записами та їх окремими частинами. Фільтрація здійснюється для виділення частотних компонентів, що містять найбільш значущу інформацію для розпізнавання мовлення. Методи шумозаглушення базуються на застосуванні Швидкого Перетворення Фур'є (FFT). Водночас, для досягнення оптимальних результатів часто необхідна комбінація FFT з іншими підходами до аналізу та обробки голосових сигналів, а також ретельне налаштування параметрів системи [12].

У дослідженні [13] описується застосування статистичних критеріїв для голосової ідентифікації та сегментації мовних сигналів з метою виділення мовних та немовних фрагментів. Рішення базується на методі оцінки часових характеристик (STE) – це метод у обробці мовлення, спрямований на визначення часових характеристик звукових сигналів, таких як початок і кінець звуків, тривалість пауз тощо. Метод дозволяє не тільки знайти мовні фрагменти мовленнєвого запису, а й нормалізувати швидкість записаного мовлення до заданого значення та розбивати запис на окремі фонемі.

У дослідженні [14] для зменшення переривчатості мовного сигналу на початку і в кінці кожного кадру пропонується використовувати віконний метод обробки кожного кадру сигналу для збільшення кореляції мел-частотних кепстральних коефіцієнтів (MFCC).

В ході аналізу результатів існуючих досліджень, зроблено висновок, що результативність описаних підходів наразі є недостатньою, та існують ще не задіяні напрямки вдосконалення поширених методів попередньої обробки, що потребують окремого детального експериментального дослідження.

Метою цієї роботи є дослідження послідовного застосування методів попередньої обробки мовних записів для підвищення точності роботи систем голосової ідентифікації. Вибір методів попередньої обробки потребує аналізу впливу різних типів шумових завад та інших факторів викривлення на якість записів та на їх подальшу обробку.

Постановка задачі

Голосовий сигнал можна представити як часовий ряд, де кожне окреме значення відповідає амплітуді сигналу в певний момент часу (рис. 1), $x(t) =$

$= [x_1, x_2, \dots, x_N]$, де $x(t)$ – вектор, що представляє голосовий сигнал, x_i – значення амплітуди сигналу в момент часу i , N – загальна кількість відліків сигналу.

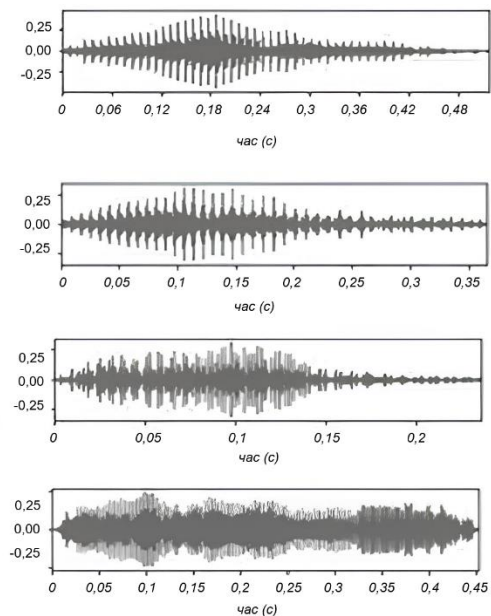


Рис. 1. Спектральний вигляд аудіозаписів чотирьох мовців, котрі промовляють однаковий текст з додаванням шумових ефектів

Дослідження передбачає використання методів попередньої обробки, таких як видалення шуму, нормалізація та фільтрація. Необхідно проаналізувати ефективність різних методів попередньої обробки та їх поєднання у системах голосової ідентифікації.

З метою дослідження стійкості системи голосової ідентифікації до різних типів викривлень було сформовано спеціалізовану базу даних, яка включала 4900 аудіозаписів [15, 16]. Дана база містить записи мовлення 10 різних дикторів, які вимовляли одне й те саме слово з варіаціями швидкості мовлення в діапазоні від 0,5 до 2,0.

Крім того, аудіозаписи були виконані в умовах різних рівнів акустичних завад для моделювання впливу шумового середовища.

Для оцінки ефективності системи в умовах відсутності мовленнєвого сигналу або під час обробки мовлення, що належить іншому диктору, до бази даних також були додані відповідні типи записів. Такий підхід дозволив всебічно проаналізувати стійкість системи до різних сценаріїв реального використання та оцінити її здатність коректно функціонувати за умов змінної акустичної обстановки.

Рішення задачі попередньої обробки голосових сигналів

У процесі попередньої обробки голосового сигналу застосовуються методи зниження шуму, нормалізації та фільтрації. Зокрема, використовуються підходи спектрального віднімання та фільтра Вінера, а також динамічні методи, такі як LMS-алгоритм, Гаар-вейвлет і Морлет-вейвлет, що дозволяють ефективно покращити якість сигналу перед його подальшим аналізом.

У залежності від характеру шумів та викривлень зазначені методи можуть застосовуватися як окремо, так і в комбінації, забезпечуючи багатоступеневу обробку сигналу. Мультиmodalний підхід, що передбачає комбінацію різних методів попередньої обробки, дозволяє адаптувати систему до широкого спектру умов запису та підвищити її стійкість до різноманітних типів шумів (рис. 2).

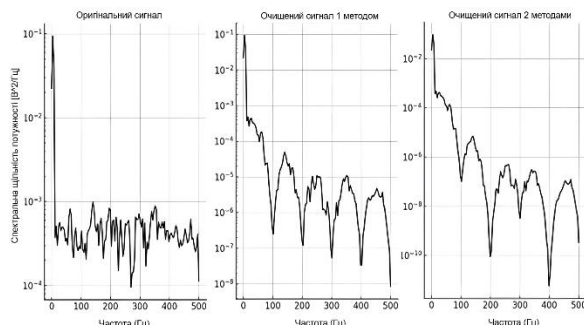


Рис. 2. Спектральний вигляд голосового фрагменту до обробки й після обробки одним та двома методами

Першим етапом є пригнічення сталих шумів за допомогою спектрального віднімання та фільтру Вінера, що забезпечує очищення сигналу від небажаних компонентів.

Спектральне віднімання базується на припущенні, що шум має постійний спектр, який можна оцінити під час пауз у мовленні:

$$S(f) = X(f) - N(f), \quad (1)$$

де $S(f)$ – спектр очищеного сигналу, $X(f)$ – спектр вихідного сигналу, $N(f)$ – оцінка спектра шуму.

Фільтр Вінера здійснює фільтрацію на основі статистичних властивостей як корисного сигналу, так і шуму. Його робота вимагає оцінки відповідних характеристик, зокрема розрахунку автокореляційних функцій та спектрів потужності. Використовуючи відомі статистичні характеристики, розраховується оптимальний фільтр, який мінімізує середньоквадратичну похибку. Формула для фільтра Вінера в частотній області має вигляд:

$$H(f) = \frac{S_{xx}(f)}{S_{xx}(f) + S_{nn}(f)}, \quad (2)$$

де $H(f)$ – передаточна функція фільтра Вінера, $S_{xx}(f)$ – спектр потужності корисного сигналу, $S_{nn}(f)$ – спектр потужності шуму.

Наступним етапом є нормалізація, яка вирівнює амплітудні та частотні характеристики сигналу. Амплітудна нормалізація здійснюється шляхом ділення значень сигналу на його максимальну амплітуду, що дозволяє уникнути перевищення динамічного діапазону. Частотна нормалізація виконується з використанням технік фільтрації, таких як високочастотні та низькочастотні фільтри, що дозволяє видалити небажані частоти.

Фільтрація, що є наступним кроком, видаляє специфічні небажані частотні компоненти сигналу. На цьому етапі застосовуються смугові фільтри, які дозволяють виділити найбільш інформативні частотні компоненти для задачі розпізнавання:

$$H(f) = \begin{cases} 1, & f_1 \leq f \leq f_2, \\ 0, & \text{інакше} \end{cases}, \quad (3)$$

де f_1 та f_2 – межі смуги пропускання.

Фреймування є необхідним етапом для підготовки сигналу до спектрального аналізу. Сигнал розбивається на короткі сегменти, або фрейми, тривалістю від 20 до 40 мс, що дозволяє врахувати зміну властивостей сигналу у часі. Застосування віконних функцій, таких як функція Хеннінга або Хеммінга, допомагає згладити краї фреймів, що мінімізує спектральні витікання. Віконна функція $\omega(n)$ може бути описана як:

$$\omega(n) = 0,54 - 0,46 \cos\left(\frac{2\pi n}{N-1}\right), \quad (4)$$

де N – довжина вікна, коефіцієнт 0,54 обирається таким чином, щоб забезпечити оптимальне співвідношення згладжування та збереження корисного сигналу (відповідає за зсув середнього значення амплітуди, забезпечуючи підйом центральної частини вікна), коефіцієнт 0,46 використовується для корекції і згладжування країв вікна (забезпечує зменшення амплітуди косинусної функції, що сприяє мінімізації спектрального витікання).

Перетворення сигналу з часової області в частотну (рис. 3) здійснюється за допомогою швидкого перетворення Фур'є (FFT). Перетворення Фур'є є одним із найпоширеніших методів спектрального аналізу, що дозволяє розкласти складний мовний сигнал на простіші частотні компоненти, забезпечуючи детальне представлення його спектральної структури. Аналізуючи ці компоненти, можна виявити індивідуальні особливості мовлення, такі як інтонація, ритм, тембр та артикуляційні характеристики, які відіграють ключову роль у розпізнаванні мови та ідентифікації мовця. Таке перетворення дозволяє отримати спектральне представлення сигналу, яке є більш зручним для подальшого аналізу. Дискретне перетворення Фур'є виконується згідно:

$$X(k) = \sum_{n=0}^{N-1} x(n)e^{-j2\pi nk/N}, \quad (5)$$

де $X(k)$ – спектральні компоненти, $X(n)$ – часові відліки сигналу, N – кількість відліків.

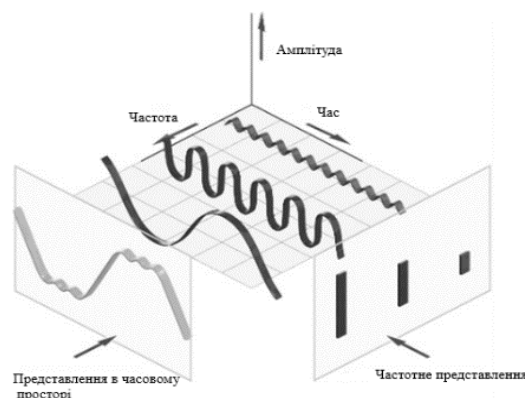


Рис. 3. Перехід від часової області до частотної області сигналу

Подальша обробка включає перетворення спектра в мел-шкалу з використанням мел-фільтробанку.

Мел-шкала більш точно відображає особливості сприйняття частот людським слуховим апаратом. Спектральні компоненти агрегуються в мел-частотні ке́пстральні коефіцієнти (MFCC) за допомогою дискретного косинусного перетворення (DCT). Розрахунок MFCC виконується відповідно до:

$$c_j = \sum_{k=1}^K \log(S(k)) \cos\left(\frac{\pi i}{K}(k - 0,5)\right), \quad (6)$$

де c_i – ке́пстральний коефіцієнт, $S(k)$ – відповідний енергетичний спектр, K – кількість фільтрів у мел-фільтробанку.

Логарифмування енергетичного спектру зменшує динамічний діапазон сигналу, що забезпечує меншу чутливість до варіацій у гучності. Це досягається шляхом застосування логарифмічної функції $\log(E)$, де E – енергія спектрального компонента.

Для завершального етапу (декореляції ознак значень вхідних даних) виконується дискретне косинусне перетворення (DCT), що дозволяє зменшити взаємозв'язок між ознаками та робить ознаки (ке́пстральні, спектральні коефіцієнти та форманти) більш незалежними одна від одної.

Декореляція спрощує подальший аналіз та обробку даних і дозволяє відфільтрувати високочастотні компоненти, які в більшості випадках містять в собі шум. Отримано мел-ке́пстральні коефіцієнти:

$$C_n = \sum_{m=0}^{M-1} \log M_m \cdot \cos\left(\frac{\pi n(2m+1)}{2M}\right), \quad (7)$$

де M – кількість мел-каналів.

Після отримання мел-ке́пстральних коефіцієнтів можна застосувати алгоритми пригнічення динамічного шуму, які потребують точної оцінки шумового порогу. Для цього використовуються алгоритми адаптивних фільтрів, таких як LMS (Least Mean Squares) та RLS (Recursive Least Squares), які передбачають постійну корекцію ваг фільтра на основі поточної помилки, що дозволяє їм адаптуватися до змінних умов шляхом корегування свої ваги $\omega(n)$ на кожному кроці на основі поточної помилки $e(n)$:

$$\omega(n+1) = \omega(n) + \mu e(n)x(n), \quad (8)$$

де $\omega(n)$ – вектор ваг фільтра на n -му кроці, μ – крок навчання, що визначає швидкість адаптації, $e(n) = d(n) - y(n)$ – поточна помилка, де $d(n)$ – бажаний сигнал, а $y(n) = \omega(n)^T x(n)$ – вихід адаптивного фільтра, $x(n)$ – вхідний сигнал на n -му кроці. Алгоритм RLS використовує вдосконалену рекурсивну схему для мінімізації середньоквадратичної помилки:

$$\omega(n) = \omega(n-1) + k(n)[d(n) - \omega(n-1)^T x(n)], \quad (9)$$

де $k(n)$ – коефіцієнт підсилення.

Також можуть бути використані методи вейвлет-фільтрації та маскування спектрального середовища.

Вейвлет-фільтрація заснована на розкладі сигналу на різні частотні підполоси за допомогою вейвлет-перетворення. Це дозволяє аналізувати сигнал на різних масштабах і виявляти компоненти, пов'язані з шумом.

$$W(j, k) = \sum_{n=0}^{N-1} x[n] \cdot \varphi_{j,k}[n], \quad (10)$$

де $W(j, k)$ – коефіцієнти вейвлет-перетворення.

Далі, виконується порогова модифікація коефіцієнтів вейвлет-перетворення, при якій коефіцієнти, що відповідають шуму, зменшуються або обнуляються, тоді як значення, що відповідають корисному сигналу, залишаються незмінними:

$$\hat{W}(j, k) = \begin{cases} W(j, k) - \lambda, & \text{якщо } W(j, k) > \lambda, \\ W(j, k) + \lambda, & \text{якщо } W(j, k) < -\lambda, \\ 0, & \text{якщо } |W(j, k)| \leq \lambda, \end{cases} \quad (11)$$

Метод маскування спектрального середовища заснований на оцінці спектрального відношення сигнал-шум (SNR) в кожній частотній смузі. На основі отриманого SNR будується маска, яка використовується для ослаблення або посилення різних частотних компонентів сигналу:

$$H(k) = \begin{cases} 0, & \text{якщо } SNR(k) \leq \gamma, \\ 1, & \text{якщо } SNR(k) > \gamma, \end{cases} \quad (12)$$

де $H(k)$ – функція маскування на k -й частоті, SNR – відношення сигнал/шум на k -й частоті, γ – порогове значення для SNR.

Результати порівняльного аналізу

У ході дослідження створено проект на основі засобів Pandas та Teras (бібліотеки для обробки аудіо, яка розроблена для спрощення аналізу та перетворення аудіо сигналів), підготовлені тестові дані в форматі WAV-файлів (Waveform Audio File Format) з записами промови чисел різними мовцями.

В процесі експериментів порівняно результати застосування різних методів пригнічення шумів у системі. Проведено окремі експерименти з фільтрацією лише статичних шумів: за допомогою спектрального віднімання та фільтра Вінера та досліджено ефективність фільтрації лише динамічних шумів: за допомогою алгоритму LMS та вейвлет-фільтрації – Гааровий вейвлет та Морлет вейвлет.

В експериментах аналізувалися голосові сигнали, зокрема оригінальний та зашумлений, представлені у спектральному вигляді. Для оцінки змін у спектрі було обрано п'ять контрольних точок, у яких визначалися відповідні коефіцієнти згідно з графічними даними. Подальша обробка сигналів здійснювалася із застосуванням різних методів шумопригнічення, зокрема спектрального віднімання, фільтра Вінера, LMS-фільтра та вейвлет-фільтрації. Після фільтрації для кожного методу було повторно визначено спектральні коефіцієнти у тих самих контрольних точках.

Основним критерієм оцінки якості фільтрації є ступінь наближення коефіцієнтів фільтрованого сигналу до коефіцієнтів оригіналу. Перевищення цих значень після обробки вказує на залишкові шуми, тоді як занижені коефіцієнти свідчать про втрату корисної інформації, що погіршує відновлення сигналу.

Таким чином, мета експерименту полягала у визначенні ефективності різних методів шумопригнічення шляхом аналізу впливу змін коефіцієнтів у спектрі після обробки сигналів.

Отримані результати дають можливість оцінити ступінь збереження корисної інформації при застосуванні кожного з розглянутих методів фільтрації та обробки (табл. 1).

Таблиця 1 – Результати обробки сигналу обраними методами фільтрації шумів

Вид сигналу	Вхідні сигнали				
	1	2	3	4	5
Оригінальний сигнал	0,50	0,32	0,26	0,68	0,44
Зашумлений сигнал	0,71	0,70	0,40	0,89	0,87
Результати обробки оригінального сигналу					
Метод обробки	1	2	3	4	5
Спектральне віднімання	0,45	0,48	0,35	0,62	0,31
Фільтр Вінера	0,52	0,33	0,32	0,63	0,39
LMS	0,53	0,32	0,31	0,62	0,46
Гааровий вейвлет	0,49	0,32	0,29	0,65	0,43
Морлет вейвлет	0,46	0,34	0,25	0,64	0,41
Результати обробки зашумленого сигналу					
Зашумлений сигнал	0,71	0,70	0,40	0,89	0,87
Спектральне віднімання	0,64	0,39	0,42	0,54	0,43
Фільтр Вінера	0,58	0,34	0,32	0,60	0,46
LMS	0,51	0,38	0,25	0,66	0,52
Гааровий вейвлет	0,48	0,37	0,24	0,67	0,47
Морлет вейвлет	0,43	0,36	0,23	0,69	0,44

Повне видалення шуму є складною задачею у умовах динамічного рівня шуму та може бути вирішено шляхом організації двоступеневої обробки голосового сигналу, яка передбачає поєднання методів попередньої обробки та глибшого аналізу акустичних характеристик. Такий підхід дозволяє поєднати переваги різних методів, забезпечуючи ефективне відфільтровування шумів, нормалізацію сигналів і покращення якості голосового сигналу перед подальшим розпізнаванням. Вибір методів фільтрації для двоступеневої обробки голосового сигналу визначається характеристиками шуму, співвідношенням сигнал-шум та бажаною якістю обробленого сигналу. Ефективність фільтрації оцінюється шляхом експериментального дослідження використання різних методів, таких як лінійні, нелінійні фільтри та нейронні мережі. Для порівняння ефективності статичних методів (спектрального віднімання (1) та фільтра Вінера (2)), а також динамічних підходів (LMS-алгоритму (3), Гаар-вейвлета (4) і Морлет-вейвлета (5)), було реалізовано двоступеневу систему пригнічення шумів, яка передбачає послідовне застосування фільтрації як статичних, так і динамічних шумів. Для цього поєднано: спектральне віднімання та LMS алгоритм (6), Фільтр Вінера та Гаар вейвлет (7), Фільтр Вінера та Морлет вейвлет (8), а також реалізовано повторну обробку мовного запису за допомогою статичного методу – спектрального віднімання (9).

Для оцінки точності алгоритмів фільтрації шуму шляхом порівняння очищеного сигналу з оригінальним сигналом використовується середня абсолютна помилка, яка розраховується згідно:

$$MAE = \frac{1}{N} \sum_{i=1}^N |x_i - \hat{x}_i|, \quad (13)$$

де N – кількість елементів у сигналу, x_i – значення оригінального (еталонного) сигналу в момент часу, $\hat{x}_i$ – значення очищеного сигналу в момент часу.

Результати, наведені у табл. 2, демонструють, що більшість застосованих методів фільтрації знижують рівень шуму в сигналах, що підтверджується

зменшенням середньої абсолютної помилки (рис. 4). Комбіновані методи фільтрації, такі як поєднання фільтра Вінера та вейвлет-фільтрації (7, 8) є більш стійкими до різнотипових шумів та викривлень.

Таблиця 2 – Результати обробки сигналів методами фільтрації та успішність ідентифікації користувача

№ методу	Оригінальний сигнал		Зашумлений сигнал	
	Успішність ідентифікації (%)	MAE	Успішність ідентифікації (%)	MAE
Методи фільтрації статичних шумів				
1	57,23	9,34	48,32	11,44
2	53,61	10,53	54,68	12,04
Методи фільтрації динамічних шумів				
3	89,12	3,24	82,49	6,48
4	85,81	4,92	83,26	8,52
5	83,62	7,12	83,73	7,57
Двоступенева система пригнічення шумів				
6	91,12	2,13	87,34	5,36
7	94,23	1,58	92,14	2,16
8	98,85	0,22	97,39	1,14
9	47,32	11,64	44,99	12,11

Організація етапу фільтрації дозволяє покращити якість голосового сигналу, що, у свою чергу, позитивно впливає на успішність ідентифікації користувача. Згідно з отриманими експериментальними результатами, комбінація фільтра Вінера та Морлет вейвлета (8) продемонструвала найвищу ефективність серед усіх розглянутих методів, забезпечивши 97,39% успішності ідентифікації навіть у зашумлених умовах, що свідчить про здатність цього підходу ефективно адаптуватися до динамічних шумів (рис. 5). Поєднання фільтра Вінера та Гаар вейвлета (7) також показало значне покращення результатів порівняно з окремим використанням цих методів.

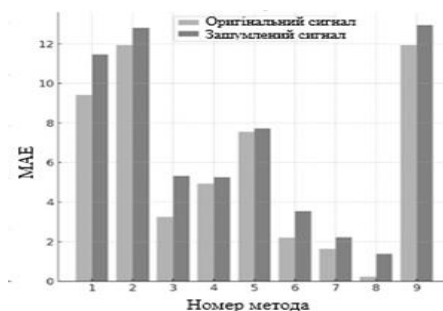


Рис. 4. Діаграма порівняння показника MAE досліджуваних методів фільтрації шумів

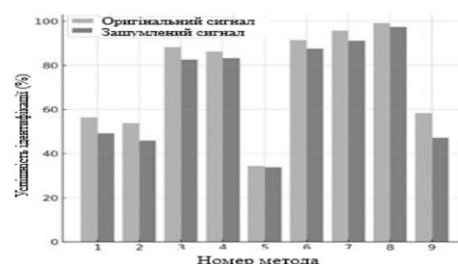


Рис. 5. Діаграма порівняння показника успішності ідентифікації користувача після в залежності від метода попередньої обробки

Застосування повторної обробки мовного запису із використанням статичного методу спектрального віднімання [9] виявило тенденцію до зниження загальної якості обробленого сигналу, що зумовлено перетворенням частини корисної мовленнєвої інформації на шумові артефакти. Аналіз отриманих результатів показує, що при повторному застосуванні алгоритму спостерігається його схильність до некоректного інтерпретування залишкових низькоенергетичних компонентів мовного сигналу як небажаних шумів, що призводить до їхнього видалення. Це може знижувати розбірливість корисної інформаційної складової мовленнєвого сигналу та призводити до втрати важливих акустичних характеристик, критичних для точного розпізнавання та ідентифікації мовця. Цей ефект особливо помітний при низькій інтенсивності вхідного сигналу або значній варіативності амплітудного спектра. Повторне спектральне віднімання без адаптивного налаштування може негативно впливати на якість фільтрованого сигналу.

У табл. 3 представлено середній час обробки голосового сигналу для різних методів фільтрації шумів, що дозволяє здійснити порівняльний аналіз їхньої обчислювальної складності.

Таблиця 3 – Час виконання методів очищення сигналу

№ методу	Час виконання, с	
	Оригінальний сигнал	Зашумлений сигнал
1	41,1	50,3
2	70,1	73,5
3	22,3	26,5
4	73,3	72,9
5	145,1	144,1
6	67,5	71,5
7	124,3	132,6
8	227,7	241,4
9	91,2	98,4

Отримані результати свідчать про те, що послідовне застосування двох методів попередньої обробки очікувано призводить до збільшення загального часу виконання, що зумовлено накопиченням обчислювальних витрат на кожному етапі фільтрації. Дане явище пояснюється необхідністю проведення додаткових операцій аналізу, фреймування, оцінки характеристик сигналу та їхньої корекції, що, у свою чергу, підвищує вимоги до обчислювальних ресурсів.

Зростання обчислювальної складності особливо актуальне в системах, що працюють у режимі реального часу, де збільшення часу обробки може суттєво

впливати на продуктивність та швидкість реагування. Вибір методів шумозаглушення має враховувати не лише покращення якості сигналу, а й обчислювальні витрати, важливі для застосування в умовах обмежених ресурсів.

Результати проведених досліджень підтверджують високу ефективність двоступеневої системи шумозаглушення, яка забезпечує значне покращення якості обробки голосового сигналу шляхом комбінованої фільтрації статичних і динамічних шумів. Комплексний підхід до пригнічення акустичних перешкод демонструє суттєві переваги порівняно з методами, що застосовують лише однокомпонентну фільтрацію, оскільки дозволяє більш точно виділяти корисну мовленнєву інформацію. Крім того, результати експериментальних досліджень вказують на особливу важливість попередньої обробки голосових записів для підвищення точності ідентифікації користувача, що особливо актуально в умовах наявності високого рівня фонових шумів або змін у характеристиках мовлення протягом запису.

Висновки

В роботі досліджено методи попередньої обробки голосових сигналів та їх вплив на кінцеву точність роботи системи голосової ідентифікації. Проведено порівняння отриманих результатів методів пригнічення шумів, що дозволило визначити найбільш ефективні підходи для покращення якості голосового сигналу та підвищення точності ідентифікації користувача. Запропонований двоступеневий підхід пригнічення шуму (сталого та динамічного) є сучасним рішенням поліпшення якості мовних сигналів в системах голосової ідентифікації та потребує подальших досліджень, спрямованих на розробку більш ефективних і робастних методів.

При послідовному застосуванні двох методів кожен з них може вносити власні артефакти або залишкові шуми, відсутні на початковому етапі обробки. Несумісність методів шумозаглушення може призводити до зміни структури шуму першим методом, що знижує ефективність другого на етапі його усунення. Важливим етапом конфігурації двоступеневих методів є точне налаштування параметрів кожного методу та оцінка їх сумісності. У середньому похибка у двоступеневих системах шумозаглушення на 4-6% менша, ніж у динамічних методів, і на 8-9% нижча порівняно зі статичними.

Аналіз отриманих даних свідчить про те, що послідовне застосування методів для видалення статичних та динамічних шумів дозволяє досягти значно кращих результатів у відновленні чистого сигналу.

СПИСОК ЛІТЕРАТУРИ

1. Mykhailichenko I., Ivashchenko H., Barkovska O., Liashenko O., "Application of Deep Neural Network for Real-Time Voice Command Recognition", IEEE 3rd KhPI Week on Advanced Technology (KhPIWeek), Kharkiv, Ukraine, pp. 1-4, doi: <https://doi.org/10.1109/KhPIWeek57572.2022.9916473>
2. Barnwal S.K., Gupta P. (2022), "Evaluation of AI System's Voice Recognition Performance in Social Conversation", 5th International Conference on Contemporary Computing and Informatics (IC3I), Uttar Pradesh, India, pp. 804-808, doi: <https://doi.org/10.1109/IC3I56241.2022.10073242>
3. Cornacchia M., Papa F., Sapio B. (2020), "User acceptance of voice biometrics in managing the physical access to a secure area of an international airport", Technology Analysis & Strategic Management, Vol. 32(10), pp. 1236-1250, doi: <https://doi.org/10.1080/09537325.2020.1758655>

4. Li Z.L., Sajat M.S., Yusof Y., Fazea Y., Santana Purba H. (2021), "The Design and User Acceptance of IoT-based Access and Entrance Control System Using Voice Recognition", Knowledge Management International Conference (KMICe), pp. 315-321, doi: <https://doi.org/10.13140/RG.2.2.12611.73762>
5. Kinkiri S., Keates S. (2020), "Speaker Identification: Variations of a Human voice", 2020 International Conference on Advances in Computing and Communication Engineering (ICACCE), Las Vegas, NV, USA, 2020, pp. 1-4, doi: <https://doi.org/10.1109/ICACCE49060.2020.9154998>
6. Kambampati P., Rane S., Shoeb A., Dhannawat R. (2024), "PAYV - Payment Voice: A Platform using Voice Recognition to Enable Payment Transactions", Asia Pacific Conference on Innovation in Technology (APCIT), MYSORE, India, 2024, pp. 1-6, doi: <https://doi.org/10.1109/APCIT62007.2024.10673442>
7. Papadopoulos P., Tsiartas A., Gibson J., Narayanan S. (2014), "A supervised signal-to-noise ratio estimation of speech signals", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Florence, Italy, pp. 8237-8241, doi: <https://doi.org/10.1109/ICASSP.2014.6855207>
8. Mishra P., Singh S., Singh S.K., Dixit A. (2018), "Pre-Processing and Partition of Voice for Semi-Voice Authentication", Fourth International Conference on Advances in Computing, Communication & Automation (ICACCA), Subang Jaya, Malaysia, pp. 1-5, doi: <https://doi.org/10.1109/ICACCAF.2018.8776849>
9. Awan S.N., Giovinco A., Owens J. (2012), "Effects of Vocal Intensity and Vowel Type on Cepstral Analysis of Voice", Journal of voice, Vol. 26(5), pp. 15-20, doi: <https://doi.org/10.1016/j.jvoice.2011.12.001>
10. Lecomte I., Lever M., Boudy J., Tassy A. (1989), "Car noise processing for speech input", International Conference on Acoustics, Speech, and Signal Processing, Glasgow, vol. 1, pp. 512-515, doi: <https://doi.org/10.1109/ICASSP.1989.266476>
11. Mamatov N., Niyozmatova N., Samijonov A. (2021), "Software for preprocessing voice signals", Tashkent University of information technologies, Vol. 18, pp. 1-8, doi: [https://doi.org/10.6703/IJASE.202103_18\(1\).006](https://doi.org/10.6703/IJASE.202103_18(1).006)
12. Slimane A.B.; Zaid A.O. (2021), "Real-Time Fast Fourier Transform-Based Notch Filter for Single-Frequency Noise Cancellation Application to Electrocardiogram Signal Denoising", Journal of Medical Signals & Sensors, Vol. 11(1), pp. 52-61, doi: https://doi.org/10.4103/jmss.JMSS_3_20
13. Zaw T.H., War N. (2017), "The combination of spectral entropy, zero crossing rate, short time energy and linear prediction error for voice activity detection", 20th International Conference of Computer and Information Technology (ICCIT), Dhaka, Bangladesh, pp. 1-5, doi: <https://doi.org/10.1109/ICCITECHN.2017.8281794>
14. Singh S., Rajan E. (2011), "MFCC VQ based speaker identification and its accuracy affecting factors", International Journal of Computer Applications, Vol. 21(6), pp. 1-6, doi: <https://doi.org/10.5120/2519-3423>
15. Ahmed A., Khondkar M.J.A., Herrick A., Schuckers S., Imtiaz M.H. (2024), "Descriptor: voice pre-processing and quality assessment dataset", IEEE Data Descriptions, Vol. 1, pp. 146-153, DOI: <https://doi.org/10.1109/IEEEDATA.2024.3493798>
16. Nagrani A., Chung J.S., Ziccerman A. (2017), "VoxCeleb: a large-scale speaker identification dataset", Sound, p. 6, doi: <https://doi.org/10.21437/Interspeech.2017-950>

Received (Надійшла) 15.01.2025

Accepted for publication (Прийнята до друку) 26.03.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Івашченко 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>.

Using a sequence of preprocessing methods in voice identification systems

Maksym Bondarenko, Heorhii Ivashchenko

Abstract. Topicality. To date, voice identification is the most common method, but some technical limitations associated with environmental influences and voice variations require further research. With the development of technologies for creating fake audio recordings, there is a need for additional improvements to ensure reliable identification using existing voice biometrics methods, mainly by introducing voice data pre-processing methods. **The goal of this work** is to study the methods of pre-processing in human voice identification systems. **The object of research** is the voice signal filtering module in the user's voice identification system. **The subject of research** is the methods of pre-processing voice signals. **Results.** This article studies the effectiveness of applying a sequence of pre-processing methods in voice identification systems to reduce the impact of background noise and other distortions on recognition quality. A comparative analysis of the effectiveness of removing steady and dynamic noise was performed. **Conclusion.** The proposed approach to organizing the procedure of pre-processing voice signals involves the selection of an optimal sequence of methods, considering such parameters as recording quality, computing resources, and the required level of accuracy. The obtained results demonstrate that the sequential use of noise reduction and signal normalization methods increases the accuracy of voice identification.

Keywords: preprocessing methods, sequential application of methods, voice signals, voice identification systems, filtering, normalization, feature extraction, noise, spectral analysis.

Gennady Golovko, Ivan Taranenko, Oleksiy Kushch

National University «Yuri Kondratyuk Poltava Polytechnic», Poltava, Ukraine

ZERO TRUST: WHY TRADITIONAL SECURITY MODELS NO LONGER WORK

Abstract. In the context of escalating cyber threats and the rapid evolution of digital work environments, traditional security models based on perimeter defense are increasingly proving inadequate. The Zero Trust model introduces a fundamental paradigm shift by eliminating the assumption of trust within corporate networks and instead adopting a continuous verification approach. This article explores the core principles and architectural components of Zero Trust, highlights its operational and strategic advantages over legacy systems, and examines the practical challenges organizations face during implementation. Particular attention is paid to how Zero Trust addresses modern cybersecurity risks such as remote work, cloud infrastructure, insider threats, and the proliferation of mobile and IoT devices. The study also analyzes real-world applications, including Google's BeyondCorp framework, and provides practical recommendations for organizations of various sizes seeking to enhance their security posture. By offering a comparative overview and actionable guidance, this paper contributes to the development of adaptive, scalable, and resilient cybersecurity strategies for enterprises operating in a rapidly changing digital landscape.

Keywords: Zero Trust, cybersecurity, perimeter defense, network security, cloud security, access control, identity verification, digital transformation, remote work, cyber threats.

Introduction

In recent years, the cybersecurity landscape has changed dramatically due to the rapid adoption of cloud services, the rise of remote work, and the increasing sophistication of cyber threats. Traditional security models, which rely heavily on perimeter-based defenses, have become insufficient in protecting modern IT environments. These legacy models operate on the assumption that everything inside the network can be trusted, which has proven to be a critical vulnerability in the age of advanced persistent threats and insider risks.

The concept of Zero Trust, first introduced by analyst John Kindervag in 2010 while at Forrester Research, challenges this outdated assumption by adopting a “never trust, always verify” approach to security. Instead of trusting users or devices by default, Zero Trust continuously evaluates access requests based on identity, context, and risk level, regardless of whether the request originates from inside or outside the network [1].

Organizations such as Google, with its BeyondCorp framework, have successfully implemented Zero Trust principles to protect their infrastructure without relying on traditional VPNs or firewalls. Furthermore, according to the National Institute of Standards and Technology (NIST), Zero Trust is now considered a recommended strategy for reducing the attack surface and increasing organizational resilience.

This article explores the limitations of traditional security models, outlines the core tenets of Zero Trust, and examines its role as a necessary evolution in the face of today's cyber risks.

Analysis of Existing Research. The growing interest in Zero Trust security architecture is reflected in numerous recent publications, which examine its core concepts, implementation strategies, and operational benefits. Numerous studies have been dedicated to defining the principles of Zero Trust, its implementation

models, and the benefits it brings over traditional perimeter-based security frameworks. Notably, NIST's Special Publication 800-207 provides a comprehensive architectural model and guidance for Zero Trust implementation across different organizational contexts. Similarly, industry reports such as CrowdStrike's Global Threat Report and IBM's Zero Trust Maturity Model examine the evolving threat landscape and advocate for Zero Trust as a strategic response [2].

Many of these works focus on describing Zero Trust's theoretical foundation or reporting on case studies from large enterprises such as Google's BeyondCorp framework. They effectively highlight the vulnerabilities of legacy systems and outline high-level recommendations for Zero Trust adoption. Furthermore, several studies have explored the integration of artificial intelligence (AI) and behavioral analytics to enhance Zero Trust decision-making processes, particularly for anomaly detection and adaptive access control.

However, a critical limitation of these studies is the lack of practical, comparative analyses detailing the specific operational challenges organizations encounter when transitioning from traditional to Zero Trust architectures. Additionally, existing literature often emphasizes enterprise-scale implementations, leaving a research gap in addressing how small and medium-sized enterprises (SMEs) can feasibly adopt Zero Trust principles within resource-constrained environments.

This article addresses these gaps by offering a balanced examination of Zero Trust benefits and real-world implementation challenges, contextualized not only for large corporations but also for medium and small organizations. It provides a practical framework for incremental adoption strategies, tool integration, and policy development, thereby contributing to the operationalization of Zero Trust security models across diverse business environments.

Recent academic works have also expanded the understanding of Zero Trust in specialized domains. Kim et al. (2024) explored the implementation of Zero

Trust principles in defense cybersecurity environments, emphasizing dynamic access control and continuous monitoring [3].

Gambo and Almulhem (2025) conducted a systematic literature review covering the development of Zero Trust architectures from 2016 to 2025, identifying key technological trends and challenges [4].

Additionally, Zanasi et al. (2024) proposed a flexible Zero Trust model tailored for Industrial IoT infrastructures, highlighting the importance of microsegmentation and adaptive access policies in highly heterogeneous environments [5].

Purpose and Objectives of the Study. Based on the analysis of existing research on Zero Trust security models, this study aims to examine the practical implementation of Zero Trust principles in modern IT environments and identify strategies to overcome operational and organizational challenges that arise during this transition. While many prior works provide theoretical foundations or large-scale enterprise case studies, there remains a lack of applied guidance for medium and small organizations adopting Zero Trust under limited resources and legacy infrastructure constraints.

The primary objectives of this study are as follows:

- To analyze the weaknesses of traditional perimeter-based security models in the context of contemporary cyber threats and digital transformation.
- To systematize the core principles and framework components of the Zero Trust model, highlighting their relevance to various organizational environments.
- To investigate real-world Zero Trust implementations, with a focus on both large-scale and resource-constrained organizations.
- To identify the key benefits and operational challenges associated with Zero Trust adoption.
- To propose a practical, incremental implementation approach suitable for different organizational scales, emphasizing operational feasibility and risk mitigation.

By addressing these objectives, the study contributes to closing the existing research gap between

theoretical Zero Trust concepts and their real-world operationalization, offering applicable insights for a broader range of enterprises.

Main part

1. The Weaknesses of Traditional Security Models. Traditional cybersecurity strategies are based on perimeter defense, which assumes that everything inside the network is trustworthy, while everything outside is not. This “trust but verify” mindset made sense in an era when organizations operated entirely within on-premise environments. However, in the age of cloud computing and remote work, this model no longer provides sufficient protection.

Modern organizations face a range of threats that traditional models struggle to defend against:

- Remote work has become standard across industries, weakening the internal network concept. VPNs are often overloaded and outdated, failing to provide secure access control for distributed teams (Cisco, 2022);
- Cloud security concerns arise when data is stored and processed across multiple third-party platforms, making traditional firewalls ineffective;
- Insider threats remain a major concern. Once inside the network, attackers (or malicious employees) can move laterally without restriction;
- Digital transformation has introduced more endpoints, devices, and integrations, complicating traditional network security monitoring.

2. The Zero Trust Model: Principles and Framework. Zero Trust is a cybersecurity architecture designed to address the shortcomings of perimeter-based defense. It is built on the idea of “never trust, always verify.” Instead of assuming that internal systems and users are safe, it treats every access request — regardless of origin — as potentially malicious until proven otherwise.

The core pillars of Zero Trust are shown in Table 1.

This shift toward **identity verification**, **adaptive access control**, and **visibility** over users and devices allows Zero Trust to secure environments across cloud, hybrid, and on-prem systems [6].

Table 1 – The core principle of Zero Trust

Principle	Description
Continuous verification	No user or device is trusted automatically. Identity, device, and context are always checked.
Least privilege access	Users only get access to what they absolutely need — reducing risk exposure.
Micro-segmentation	Networks are broken into zones to minimize damage in case of a breach.
Assume breach	The system is built with the assumption that intrusions will happen — the goal is to contain them.

3. Real-World Application: Google's BeyondCorp. A leading real-world example of Zero Trust in action is Google's BeyondCorp framework. Launched after a major cyber attack in 2009, BeyondCorp aims to

allow employees to work securely from any location without a traditional VPN. Every access request is evaluated based on multiple factors: device posture, user identity, location, time of request, etc (Fig. 1).

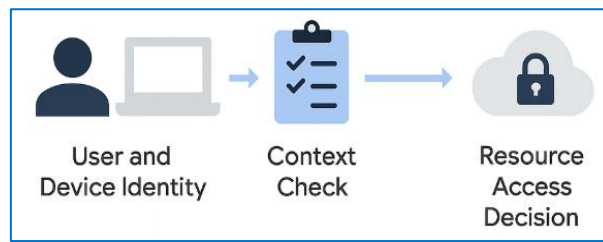


Fig 1. BeyondCorp Access Flow

This model enhances network security by making internal systems less reliant on perimeter defenses. Instead of assuming trust, Zero Trust verifies identity and context every time access is requested [7].

4. Benefits of Zero Trust. The transition to a Zero Trust architecture provides organizations with a robust response to today's dynamic cyber threats. By removing the implicit trust that traditional systems relied upon, Zero Trust significantly reduces the attack surface and increases network security resilience (Table 2) [8].

Key Benefits:

- **Minimized impact of breaches:** Since access is segmented and tightly controlled, attackers cannot easily move laterally through the network;

- **Improved visibility and control:** Zero Trust solutions require continuous monitoring of users, devices, and traffic — enhancing situational awareness;

- **Stronger cloud security:** By applying access policies based on context rather than network location, organizations can more securely adopt cloud services;

- **Enhanced compliance:** Regulatory frameworks increasingly recommend or mandate Zero Trust elements, especially for protecting identity and access management;

- **Support for remote work:** Employees can access resources securely from any location or device, without relying on VPNs or centralized infrastructure.

Table 2 – Key Differences Between Traditional Security and Zero Trust

Aspect	Traditional Security	Zero Trust Architecture
Trust model	Trust inside, block outside	Trust no one, verify always
Access control	Static, location-based	Dynamic, identity and context-aware
Attack surface	Wide (flat network)	Minimized through segmentation
Cloud and remote support	Limited or dependent on VPN	Built-in and adaptive
Breach response	Reactive, post-incident	Proactive, continuous monitoring
Compliance readiness	Often manual, inconsistent	Auditable, policy-driven

According to recent industry research, organizations implementing Zero Trust experience fewer incidents and lower breach-related costs than those using perimeter-based models. [9]

Improvements after Zero Trust implementation:

- 50% reduction in data breach incidents;
- 2x increase in endpoint visibility;
- 40% improvement in compliance audit readiness;

- Faster detection and isolation of anomalies.

These outcomes make Zero Trust an essential part of digital transformation initiatives, particularly in sectors like finance, healthcare, and government where access control and data integrity are critical.

5. Implementation Challenges. Despite its advantages, implementing Zero Trust is not without obstacles. Organizations often face several technical, cultural, and operational barriers during the transition [10].

Major challenges include:

Legacy infrastructure: Older systems may not support modern identity and policy enforcement mechanisms;

- **Organizational resistance:** Shifting from a perimeter-focused to an identity-centric mindset can meet opposition from stakeholders used to traditional models;

- **Tool and vendor complexity:** Integrating different tools for identity verification, device posture, policy engines, and monitoring requires careful planning;

- **Cost and resource allocation:** The initial investment — both financial and in terms of skilled personnel — may be substantial.

These issues can delay full adoption or lead to partial implementations that fall short of delivering the full benefits of Zero Trust. However, many organizations adopt an incremental approach — starting

with high-risk assets or user groups and expanding gradually.

Tip: Start with an identity and access audit, followed by segmentation and policy definition, to build the foundation of Zero Trust with minimal disruption.

6. Comparative Overview: Traditional vs. Zero Trust Security. As the digital landscape continues to evolve, Zero Trust is expected to become not only a best practice but a fundamental requirement for cybersecurity architectures.

The shift toward decentralized networks, increased remote work, and the proliferation of cloud-based and edge computing systems require a security model that is dynamic, adaptable, and identity-centric.

One of the most promising developments is the integration of artificial intelligence (AI) and machine learning (ML) into Zero Trust frameworks. These technologies enable real-time behavioral analytics, anomaly detection, and automated responses to potential threats. For example, if a user logs in from an unusual location or attempts to access resources outside their normal scope, the system can automatically trigger step-up authentication or deny access entirely.

Moreover, the Internet of Things (IoT) presents new security challenges due to the sheer number and heterogeneity of connected devices. Traditional network perimeter controls are insufficient to secure these endpoints. Zero Trust allows for granular policies based on device health, firmware version, and usage patterns, providing enhanced visibility and control over IoT assets [11].

Cloud-native Zero Trust architectures are also becoming a critical focus. As organizations migrate to hybrid and multi-cloud environments, Zero Trust principles ensure that data and applications remain protected across platforms. This includes continuous validation of access requests, encryption of data in motion and at rest, and policy enforcement at every access point.

In the coming years, we are likely to see Zero Trust extended beyond corporate IT to operational technology (OT) and critical infrastructure, where traditional systems were not designed with cybersecurity in mind. Governments and regulatory bodies are already pushing for Zero Trust adoption in these sectors as part of national security strategies.

Furthermore, the concept of Zero Trust-as-a-Service (ZTaaS) is emerging, where security vendors provide managed ZT solutions, reducing the burden on in-house teams. This trend is particularly beneficial for small and medium enterprises (SMEs), which may lack the resources to build and maintain complex security infrastructures [12].

In summary, the future of Zero Trust lies in deeper automation, broader adoption across sectors, tighter integration with AI and IoT technologies, and increased regulatory alignment. As organizations strive for resilience in the face of complex cyber threats, Zero Trust will be central to ensuring sustainable and scalable security strategies.

This comparison highlights why Zero Trust is increasingly viewed not just as a cybersecurity

improvement, but as a necessity in modern IT environments.

Discussion of Research Results

The analysis of existing scientific publications on Zero Trust models revealed that while previous studies have extensively described the theoretical foundations, core principles, and large-scale enterprise implementations of Zero Trust, they have not sufficiently addressed certain operational aspects and practical challenges organizations face during the transition from traditional perimeter-based models. This study successfully contributes to filling these gaps by:

- Providing a comparative analysis of traditional and Zero Trust security models within the context of modern cyber threats, digital transformation, and remote work environments.
- Systematically structuring the core principles of Zero Trust and relating them to real-world operational needs.
- Highlighting the practical outcomes of a well-documented enterprise-scale implementation (Google's BeyondCorp), offering readers a concrete example of Zero Trust in action.
- Detailing the key benefits organizations can achieve through Zero Trust adoption, including improved visibility, compliance readiness, breach mitigation, and support for distributed workforces.
- Identifying and systematizing the primary technical, cultural, and operational challenges encountered during Zero Trust implementation, which were either omitted or only briefly mentioned in prior works.

Unlike earlier studies that focused mainly on conceptual frameworks or enterprise-level deployments, this paper provides a balanced overview of both strategic advantages and practical obstacles relevant to organizations considering Zero Trust adoption today. By doing so, it expands the applied understanding of Zero Trust and supports informed decision-making for cybersecurity professionals and IT leaders [13].

Conclusion

In an era defined by digital transformation, remote work, and increasingly sophisticated cyber threats, traditional perimeter-based network security approaches are no longer sufficient.

The Zero Trust model emerges as a robust and modern solution that redefines how organizations protect their data, systems, and users. This study not only confirmed the critical weaknesses of traditional security models but also systematized the core principles of Zero Trust, providing a comparative overview of their practical application.

A significant contribution of this research lies in identifying specific operational and organizational challenges associated with Zero Trust implementation — a topic insufficiently addressed in prior studies. By analyzing real-world deployment examples and outlining key benefits, this paper closes existing research gaps and offers applicable insights for organizations of various sizes considering Zero Trust

adoption. Although implementing Zero Trust poses certain difficulties — particularly for organizations with legacy infrastructure — its long-term advantages in terms of resilience, adaptability, and compliance significantly outweigh the transitional challenges. This study emphasizes that Zero Trust should be viewed not merely as a technical solution but as a strategic

philosophy aligned with the realities of modern IT environments.

By addressing both conceptual and operational aspects, this research supports informed decision-making for cybersecurity practitioners and provides a foundation for future studies exploring scalable, resource-efficient Zero Trust implementations.

REFERENCES

1. Forrester Research. The State of Zero Trust Adoption. URL: <https://www.forrester.com/zero-trust/>
2. CrowdStrike. Global Threat Report. URL: <https://www.crowdstrike.com/en-us/resources/reports/crowdstrike-2023-global-threat-report/>
3. Kim Y., Sohn S.-G., Kim K.T., et al. (2024). Exploring Effective Zero Trust Architecture for Defense Cybersecurity: A Study. KSII Transactions on Internet and Information Systems, 18(9). DOI: <https://doi.org/10.3837/tiis.2024.09.011>
4. Gambo M.L., Almulhem A. (2025). Zero Trust Architecture: A Systematic Literature Review. arXiv preprint. DOI: <https://doi.org/10.48550/arXiv.2503.11659>
5. Zanasi C., Russo S., Colajanni M. (2024). Flexible Zero Trust Architecture for the Cybersecurity of Industrial IoT Infrastructures. Ad Hoc Networks, 157. DOI: <https://doi.org/10.1016/j.adhoc.2024.103414>
6. NIST. Special Publication 800-207: Zero Trust Architecture. URL: <https://doi.org/10.6028/NIST.SP.800-207>
7. Google Cloud. BeyondCorp Enterprise Whitepaper. URL: <https://cloud.google.com/beyondcorp>
8. Cisco. Hybrid Work and VPN Limitations Report. URL: https://www.cisco.com/c/m/en_us/solutions/global-hybrid-work-study.html
9. Verizon. Data Breach Investigations Report. URL: <https://www.verizon.com/business/resources/reports/dbir/>
10. IBM Security. Zero Trust Maturity Model: A Strategic Guide. URL: <https://www.ibm.com/zero-trust>
11. Gartner. Future of Work Trends: Remote Work and Cybersecurity. URL: <https://www.gartner.com/en/information-technology/insights/remote-work-technology>
12. Microsoft. Zero Trust Deployment Guidance. URL: <https://learn.microsoft.com/en-us/security/zero-trust/>
13. Okta. Identity-First Security in a Zero Trust World. URL: <https://www.okta.com/zero-trust/>

Received (Надійшла) 11.03.2025

Accepted for publication (Прийнята до друку) 21.05.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Головко Геннадій В'ячеславович – доцент кафедри комп'ютерних та інформаційних технологій і систем, Національний університет «Полтавська політехніка ім. Ю. Кондратюка», Полтава, Україна;

Gennady Golovko – Associate Professor, Department of Computer and Information Technologies and Systems, National University "Yuri Kondratyuk Poltava Polytechnic", Poltava, Ukraine;

e-mail: genvgolovko@ukr.net; ORCID Author ID: <http://orcid.org/0000-0002-1745-1321>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57195419168>.

Тараненко Іван Іванович – студент кафедри комп'ютерних та інформаційних технологій і систем, Національний університет «Полтавська політехніка ім. Ю. Кондратюка», Полтава, Україна;

Ivan Taranenko – student of the Department of Computer and Information Technologies and Systems, National University "Yuri Kondratyuk Poltava Polytechnic", Poltava, Ukraine;

e-mail: vanyo.taranenko@gmail.com; ORCID Author ID: <http://orcid.org/0009-000-4606-6943>.

Куш Олексій Олексійович – аспірант кафедри автоматизації, електроніки та телекомунікацій, Національний університет «Полтавська політехніка ім. Ю. Кондратюка», Полтава, Україна;

Kushch Oleksiy – PhD student at the Department of Automation, Electronics and Telecommunications, National University "Yuri Kondratyuk Poltava Polytechnic", Poltava, Ukraine;

e-mail: alechkusch881@gmail.com; ORCID Author ID: <https://orcid.org/0000-0002-7630-6298>.

Zero Trust: чому традиційні моделі безпеки більше не працюють

Г. В. Головко, І. І. Тараненко, О. О. Куш

Анотація. На тлі зростання кіберзагроз і стрімкої трансформації цифрового робочого середовища традиційні моделі безпеки, що базуються на периметровому захисті, дедалі частіше виявляються неефективними. Модель Zero Trust пропонує принципово новий підхід, усуваючи припущення про довіру всередині корпоративних мереж і запроваджуючи концепцію безперервної верифікації. У статті розглянуто основні принципи та архітектурні компоненти Zero Trust, окреслено її стратегічні й операційні переваги порівняно із застарілими системами, а також проаналізовано практичні виклики, що виникають у процесі впровадження. Особливу увагу приділено подоланню сучасних загроз кібербезпеки, таких як віддалена робота, хмарні сервіси, внутрішні загрози та поширення мобільних і IoT-пристроїв. Дослідження містить аналіз реальних прикладів застосування, зокрема фреймворку BeyondCorp від Google, і надає практичні рекомендації для організацій різного масштабу щодо підвищення їхньої кібербезпеки. Запропонований порівняльний огляд і прикладні рекомендації сприяють розробці адаптивних, масштабованих і стійких стратегій кіберзахисту в умовах динамічного цифрового середовища.

Ключові слова: Zero Trust, кібербезпека, периметровий захист, мережна безпека, хмарна безпека, контроль доступу, верифікація особистості, цифрова трансформація, віддалена робота, кіберзагрози.

Mykhailo Hulevych, Oleksii Kolomiitsev

National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine

AUTOMATED TEST GENERATION TECHNIQUES FOR C++ SOFTWARE

Abstract. Automation of test script generation is critically important in modern software quality assurance, particularly for complex languages such as C++, where manual testing becomes resource-intensive and error-prone. Automated approaches significantly reduce the testing effort, improve test effectiveness, and enhance overall software reliability. The CIDER tool, introduced by the author, offers a promising automated solution by generating test scenarios based on recorded program executions using harmony search for inputs optimization. Despite its benefits, tool faces next limitations: incomplete coverage of complex branching logic, inefficient exploration of large input spaces, and difficulties in handling semantically rich or contextually complex input data. The overview article aims to systematically explore, compare, and evaluate various automated test generation techniques such as symbolic execution, concolic testing, evolutionary algorithms, reinforcement learning, and model-based testing. The objectives are: to classify these methods based on specified criteria and identify suitable approaches that could enhance the tool's capability for automated test generation in terms of coverage efficiency.

Keywords: test automation, C++, regression testing, artificial bee colony, ant colony optimization, firefly algorithm, cuckoo search, q-learning.

Introduction

C++ is a foundational language for systems level and performance-critical software. It's widely used in domains such as embedded systems, operating systems, high-performance computing, financial, and real-time applications. However, the language's inherent complexity including manual memory management, template metaprogramming, and pointer arithmetic, introduces significant challenges for effective and reliable testing, particularly due to associated security risks. In response to the increasing demand for scalable and high-assurance software, the automation of test suite generation has become an essential focus in recent software engineering research. Ensurance the software quality requires additional attention to low-level constructs, tightly coupled modules, subtle language semantics, complex C++ memory model.

Thus, the study and advancement of automated test suite generation techniques is of paramount importance. The software systems are developed rapidly in Agile and DevOps environments and automated testing becomes a crucial enabler of continuous integration and regression safety. In the C++ software even minor changes can introduce subtle and difficult-to-detect bugs, automated test generation provides a reliable safety assurance for maintaining code quality and system reliability.

Recent publications. Research efforts continue to underline the importance of test automation, particularly for C++, that exhibits substantial structural and semantic complexity. Recently, Zhang et al. [1] introduced CITYWALK framework. The framework leverages large language models (LLMs) and language-specific aspects of knowledge to automate unit test generation, particularly for C++ codebases. It illustrates the growing role of artificial intelligence in addressing language-specific challenges. In addition, Rho et al. [2] presented the Coyote testing framework. It demonstrates the practical viability of concolic execution-based automation in industrial environments that achieves significant improvements in test coverage without manual intervention. Beyond individual tools and

frameworks, broader empirical studies reinforce the trend. A quantitative analysis [3] showed that higher levels of test automation maturity are directly associated with improved software quality across industrial projects. These findings are echoed by recent industry reports, which highlight automation as a core element of scalable and sustainable quality assurance practices in contemporary software engineering. As project complexity and delivery speed increase, automated testing becomes not only a technical asset but also an imperative for ensuring robustness and quality [4].

At the same time, several critical challenges of automated test suite generation persist. These include the absence of reflective capabilities, the presence of non-determinism in standard libraries, and the lack of standardized mechanisms for specifying expected behavior across software components. These issues have motivated ongoing work in areas such as method call sequence mining, context-sensitive test generation, and mutation-based test evaluation [5]. Collectively, these research directions underscored both the feasibility and necessity of continued investment in automated test generation for C++, particularly as software complexity and safety requirements continue to grow.

Research objectives. The CIDER [6] tool is designed to address these challenges. It exemplifies a fully automated black-box testing approach tailored specifically for C++ libraries. It leverages recorded program executions to generate automated test scenarios, subsequently applying heuristic optimization techniques, such as harmony search, to improve test coverage. By providing compatibility with industry-standard libraries and facilitating flexible script-based testing approaches, the tool enables effective and realistic testing scenarios for complex software systems and demonstrates several advantages: effective initial scenario generation based on actual program execution, the use of harmony search to enhance test coverage, compatibility with industry-standard libraries and tools via bindings and scripts. However, the tool has incomplete coverage of complex branching logic due to limited systematic path exploration, motivating the investigation of additional enhancement techniques.

The research objectives are:

RO1: To explore and systematically compare various automated test generation techniques, including concolic and symbolic execution, evolutionary algorithms, reinforcement learning, model-based and search-based testing.

RO2: To identify methodologies and best practices that are most effective in overcoming CIDER's identified limitations.

In the next sections an overview of most popular automation techniques will be provided, along with their classification with respect to CIDER improvement relevance.

Classification Criteria

The criteria for classification include:

- complexity of control-flow
- complexity of data-flow
- semantical-rich input
- automation potential
- regression testing potential
- black-box testing suitability

For each criterion the description is provided in Table 1.

Table 1 – Classification criteria description

	Criteria	Description
Source code characteristics	Control-flow complexity	Measures the ability of a testing technique to effectively handle intricate branching logic and multiple execution paths within the source code.
	Data-flow complexity	Indicates how effectively a testing technique can track and analyze data dependencies, variable usage, and transformations throughout the software.
	Semantical rich input	Assesses the capability of the technique to generate meaningful, contextually valid input data, particularly for structured or domain-specific scenarios (e.g., JSON, XML, API payloads).
Tool operational characteristics	Automation potential	Evaluates the degree to which a testing technique can operate autonomously without manual intervention after the initial setup.
	Regression testing potential	Reflects the suitability of a technique to repeatedly validate software correctness over time, particularly as the software evolves or changes.
	Black-box testing suitability	Indicates whether the testing technique can be effectively applied without access to internal source code details, relying exclusively on external interfaces or behavioral analysis.

The first three criteria relate to source code under test characteristics which indirectly include coding language specific aspects, project requirements, domain constraints, and typical use cases for the tested software.

▪ The control-flow complexity

It is measured using different metrics, one of them is cyclomatic complexity, which is defined as number of linearly independent execution paths of the module under test, corresponding to number of conditional statements in the source code (e.g., branches). The control-flow graph (CFG) approach [7] using CFG that consists of nodes (N) as conditional statements and edges (E) as execution paths between them. Thus, cyclomatic complexity is calculated as $C(G) = E - N + 2$. A higher value indicates that more complex testing for the module is required.

▪ The data-flow complexity

It has no universally accepted method of measurement and is instead evaluated as a combination of analytical approaches.

The definition-usage approach analyses pairs of definition and usage of variables through the source code and corresponding variable interactions [8]. The complexity grows with number of definitions, re-definitions and interactions of variables as shown, Fig 1.

```

1. int x = 10;           // definition
2. if (y > 0) {
3.   x = y + 5;         // re-definition
4. }
5. printf("%d", x);     // usage

```

Fig. 1. Def-Usage data-flow example

The data-flow coverage metrics analyze the test coverage output to check if every variable definition executed at least once (is reachable).

The fan-in and fan-out metrics analyze and compare the number of dependencies entering and leaving the module respectively, giving information about coupling and data exchange complexity.

▪ The semantical-rich input

The semantically legal generation ability of the testing tool is crucial for the libraries from specific domains, in which not correctly generated input leads to low coverage.

Techniques

▪ **Symbolic execution** exploration mechanism of control paths treats input variables as symbolic not as concrete values. This method systematically enumerates possible execution paths and generates symbolic constraints to represent these flows. As for **concolic testing**, it complements symbolic execution by executing programs with concrete inputs while simultaneously generating symbolic constraints. Such approach mitigates certain scalability issues associated with purely symbolic approaches. One of the classic technique usage example is KLEE [9]. It was used to achieve high coverage in testing of GNU COREUTILS and BUSYBOX that significantly improved fault detection by systematically exploring multiple execution paths. The results show precise fault localization and improved software reliability.

▪ **Search-based testing** utilizes natural selection concepts to iteratively evolve test inputs toward optimal

solutions defined by a fitness function. The fitness function commonly measures objectives such as code coverage, fault detection rates, or execution performance. The EvoSuite [10] automatically generates effective unit test suites for Java applications, improving code coverage and fault detection capability through the genetic evolution of test inputs.

▪ **Reinforcement learning (RL)** dynamically generates and refines test inputs through iterative interactions with the software system. The process is guided by reward function that reflects coverage metrics or fault detection metrics. The strategy of RL algorithms is to adjust input generation ability over time to maximize long-term rewards. The DRIFT tool [11] applies reinforcement learning to dynamically generate test cases. It significantly improves coverage and fault discovery efficiency, particularly in complex software systems. Tests within Windows 10 were performed to justify the efficiency of the method.

▪ **Model-based testing** generates test cases systematically from formal or semi-formal models of the software behavior, which capture specifications or functional descriptions. In addition, **property-based testing** complements such results using technique of validation software behavior against predefined properties or invariants. It automatically generates input scenarios to falsify these properties. The SpecExplorer [12] generated exhaustive test cases for verifying protocol implementations, which ensures compliance with detailed system specifications. The QuickCheck [13] systematically validates software behavior against specified properties, while Randoop [14] generates valid method call sequences that effectively uncover functional faults in software APIs, significantly enhancing black-box testing effectiveness.

▪ **Mutation Testing** assesses the quality and effectiveness of test suites using intentionally injecting small artificial faults (mutations) into the source code and observing whether existing test cases can detect these introduced faults. The PIT [15] has been used to assess and improve test suites in Java applications by injecting faults, leading to a significant increase in the robustness and quality of test cases. The Table 2 presents a comparison of the techniques based on their coverage effectiveness with respect to the characteristics of the tested source code.

Table 2 – Comparison of techniques based on source code characteristics criteria

Techniques	Control Flow Complexity	Data Flow Complexity	Semantical Rich Input
Symbolic execution	High	High	Low
Search-Based Testing	High	Medium	Medium
Reinforcement Learning	High	Medium	Medium
Model-Based Testing	High	High	High
Mutation Testing	Medium	Medium	Low

These capabilities are rated as High, Medium, Low respectively. The model-based testing technique receives a high rank in all three criteria. However, the comparison doesn't account for its automation potential, which is constrained by the need for manual settings and model creation limitations. To address this issue three operational criteria classification has been done. The comparison on operational properties shown on Fig. 2.

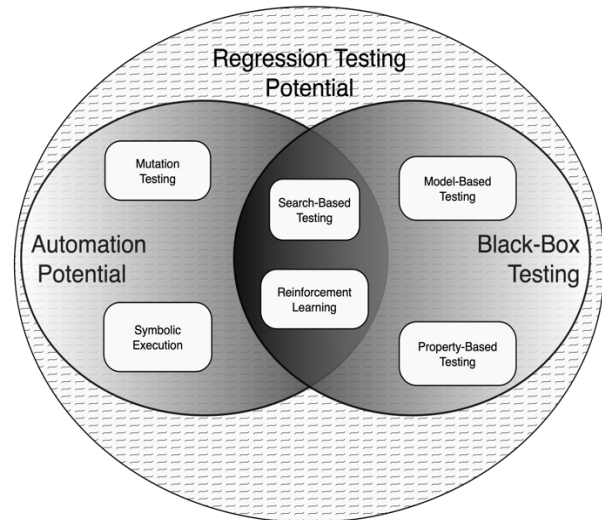


Fig. 2/ The techniques' classification depending on testing tool operational characteristics

Among all considered techniques, search-based testing emerges as the most appropriate candidate for enhancing CIDER. It aligns perfectly with CIDER's black-box nature, fully supports automation, and can replace current heuristic strategies with more adaptive and scalable optimizations. RL is also highly compatible, offering dynamic learning of effective test scenarios through interaction, ideal for evolving systems. In contrast, techniques like symbolic execution and mutation testing require source code access and are thus incompatible with CIDER's concept.

Investigation

Problem 1: Script input optimization.

Formulation. Let S denote a fixed test script consisting of a predefined sequence of function calls from a finite set of operations $\mathcal{F}$. Each invocation within S requires a vector of input parameters $x = (x_1, x_2, \dots, x_m)$, $m \in \mathbb{N}$. Executing the script S with a given input vector x yields a dynamic execution trace, from which a coverage metric $C(S(x)) \in \mathbb{R}$ is computed. The coverage function quantifies the extent to which the input-triggered execution explores the program under test, based on structural criteria such as line, branch, or function coverage.

Objective. The objective is to identify the input vector x' that maximizes the achieved code coverage.

Investigation. The possible directions of search-based testing algorithms usage are:

- ♦ Artificial Bee Colony (ABC)
- ♦ Firefly algorithm (FA)
- ♦ Cuckoo Search (CS)

A comparison of the algorithms based on their advantages and drawbacks is shown in Table 3.

Table 3 – Comparison of search-based metaheuristic algorithms

	Advantages	Disadvantages
ABC	Provides a well-balanced exploration and exploitation mechanism.	Exhibit slower convergence on complex landscapes.
FA	Rapid convergence in continuous domains. Highly effective for unimodal problems.	Low performance in discrete spaces. Risk of being trapped in local optima.
CS	Highly effective for multimodal landscapes. Requires minimal parameter tuning.	May converge prematurely; less interpretable in terms of search behavior.
ACO	Highly effective for sequential and graph-based optimization. Adapts well to problem structure.	Slower convergence and requires more elaborate parameter tuning.

In addition, common algorithm of metaheuristic search optimization is shown in Fig. 3, with basic input parameters as fitness function C , number of iterations T .

Metaheuristic Search Optimization	
1:	Initialize $x = (x_1, x_2, \dots, x_m)$
2:	Evaluate $C(S(x_i))$ for each x_i
3:	$x_{best} = \text{argmax}(C(S(x_i)))$
4:	for $t = 1$ to T do
5:	for each solution x_i do
6:	Generate $x'_i \leftarrow \text{Gen}(x_i, x_{best}, t, \dots)$
7:	Evaluate $C(S(x'_i))$
8:	if Acceptable(x'_i) then
9:	$x_i \leftarrow x'_i$
10:	if $C(S(x_i)) > C(S(x_{best}))$ then
11:	$x_{best} \leftarrow x_i$
12:	end if
13:	end if
14:	end for
15:	Optionally Diversify(x)
16:	end for
17:	return x_{best}

Fig. 3 Generic representation of metaheuristic search optimization algorithm

▪ **Artificial Bee Colony.** The algorithm was introduced by D. Karaboga [16] and inspired by behavior of honeybee colonies. Three kinds of branches of logic are included: employed bees, onlooker bees and scouts. The employed bee searches on m iteration the solution vector x'_{mi} according to formula:

$$x'_{mi} = x_{mi} + \varphi_{mi}(x_{mi} - x_{ki}), \quad (1)$$

where x_{mi} – current solution, $\varphi_{mi} \in [-1, 1]$ – the random number, x_{ki} – randomly chosen solution in memory.

The employed bees share their food source information with onlooker bees which choose a food source depending on the probability values, calculated as:

$$P_m = \frac{C_m(x_{mi})}{\sum_i C_m(x_{mi})}. \quad (2)$$

After the x_{mi} is chosen it's modified via eq. (1). The employed bees whose solutions failed to improve for limit times become the scouts. The scout bee initializes their solution randomly.

▪ **Firefly algorithm.** The algorithm was firstly introduced in 2009 [17] and is inspired by fireflies' behavioral attraction to bright light. The search iteration at m can be described as:

$$x'_{mi} = x_{mi} + \beta(x_{mi} - x_{ki}) + \alpha\varphi_{mi}, \quad (3)$$

where x_{mi} – firefly's current solution, $\varphi_{mi} \in [-1, 1]$ is the random number, α – scale factor, x_{ki} – randomly chosen solution in memory.

The attraction β from eq. (3) is as follows:

$$\beta = \beta_0 e^{-\gamma r_{mi}^2}, \quad (4)$$

where γ – absorption coefficient, r_{mi}^2 – distance between fireflies, β_0 – the attractiveness at distance 0.

The distance can be calculated as Cartesian or any other interpretation.

▪ **Cuckoo search.** The generation of new solutions behaves as cuckoo which lays own eggs in neighbor's nest [18]. The solution search formula incorporates Lévy distribution:

$$x'_{mi} = x_{mi} + \text{Lévy}(\lambda). \quad (5)$$

The Lévy distribution is as follows:

$$\text{Lévy} \sim u = t^{-\lambda}, \quad (6)$$

where $\lambda \in (1; 3]$. The best solutions are kept for the next iteration and worse are dropped by the nest owner with probability P .

Problem 2: Call sequence optimization.

Let $\mathcal{F} = \{F_1, F_2, \dots, F_n\}$ denote the finite set of available function calls provided by a software under test interface. A test script is defined as an ordered sequence of function invocations, denoted $S = \{F_{i1}, F_{i2}, \dots, F_{ik}\}$, where each $F_{ik} \in \mathcal{F}$, and the sequence length k may be fixed or subject to constraints reflecting practical limitations on test depth or execution time.

The execution of a script S produces a runtime behavior of the program under test, resulting in a dynamic trace from which a structural coverage metric $C(S(x)) \in \mathbb{R}$ is computed. This metric may correspond to branch coverage, path coverage, state coverage, or other relevant measures of behavioral exploration.

Objective. The objective is to determine the script $S' \in \mathcal{S}$ that maximizes the achieved coverage, where $\mathcal{S}$ is the space of syntactically and semantically valid call sequences.

Investigation. The possible directions of search-based testing algorithms usage are:

▪ **Ant Colony Optimization.** The algorithm is inspired by foraging behavior of ants [19]. The ants use pheromones to mark good paths between nest and food sources.

▪ Over time, shorter paths receive more pheromone, reinforcing them based on probability.

The problem can be modeled as a directed fully connected graph $G = (V, E)$, where $V \subseteq \mathcal{F}$ is the node representing method call, $E \subseteq V \times V$ are edges representing transitions between method calls - a path of

the graph represent searched call sequence S . Each edge has associated pheromone value τ_{ij} and a heuristic value η_{ij} representing desirability of the edge. Then each ant k in the state i (current method call) choses the next method j based on:

$$P_{ij}^{(k)} = (\tau_{ij})^\alpha (\eta_{ij})^\beta / \sum_{l \in \mathcal{N}_i^k} (\tau_{il})^\alpha (\eta_{il})^\beta, \quad (7)$$

where $\mathcal{N}_i^k$ - unvisited neighbors of node i for ant k , and α, β - balancing coefficients.

After construction of S the reward R_k is computed and pheromone matrix is updated using scaling factor Q and call sequence length L_k :

$$\tau'_{ij} = Q \cdot R_k / L_k. \quad (8)$$

The algorithm is stopped when best sequence is found, or time budget is over.

The next papers [20, 21] have deeply dived into experimental application of ACO optimization for method call sequences generation in test scenarios. Their results showed that ACO generated sequences achieved higher code coverage and robustness compared to random or greedy baselines, particularly in preserving object usage patterns and handling class constructors correctly.

In addition to graph-based metaheuristic techniques such as ACO, RL provides a learning-based approach to generating call sequences, particularly Q-learning.

▪ **Q-learning.** The algorithm was introduced by C. Watkins [22] and is a form of model-free RL.

The generic algorithm of learning phase is shown on Fig. 4, where:

s – state, representing partial sequence of function call,

a – action selecting next valid and capable method from $\mathcal{F}$,

r – reward, a scalar value based on $\mathcal{C}(s)$, and the return represents policy,

$\pi(s)$ – which maximizes the cumulative reward for doing action a in state s .

The update of Q-Table phase is done using formula:

$$Q'(s, a) \leftarrow Q(s, a) + \Delta Q(s', a'),$$

$$\Delta Q(s', a') = \alpha \cdot [r + \gamma \cdot \max_{a'} Q(s', a') - Q(s, a)],$$

where $\alpha \in (0; 1]$ – learning rate, γ – discount factor.

Q-learning Algorithm

```

1: Initialize Q-Table  $Q(s, a) \leftarrow 0$ 
2: for each episode do
3:   Initialize State ( $s$ )
4:   for  $t = 1$  to  $T$  do
5:     if Exploration ( $\epsilon$ ) then
6:        $a = \text{Random}(A)$ 
7:     else Exploitation
8:        $a = \text{argmax}_a Q(s, a)$ 
9:     end if
10:     $s' = \text{Step}(s, a)$ 
11:     $r = \text{GetReward}(s')$ 
12:    Update Q-Table ( $s', s, r$ )
13:    if TerminalState ( $s'$ )
14:      break
15:     $s \leftarrow s'$ 
16:  end for
17: end for each
18: return  $\pi(s) = \text{argmax}_a Q(s, a)$ 

```

Fig. 4. Generic representation of learning section of Q-learning algorithm

Conclusions

This paper presents a comprehensive overview of automated test generation techniques applicable to C++ software, with a particular focus on enhancing the capabilities of the CIDER tool.

The search-based and reinforcement learning techniques offer adaptive, scalable, and fully automated mechanisms for high-coverage test scenarios generation. In particular, search-based approaches such as Artificial Bee Colony, Firefly Algorithm, and Cuckoo Search show strong potential in test inputs optimization. On the other hand, Ant Colony Optimization has effectiveness in method call sequence generation. Reinforcement learning complements the approach via dynamic learning from test execution feedback, thereby improving test effectiveness over time. Symbolic and concolic execution, despite their high precision, are constrained by CIDER's lack of source code access, making them less suitable in this context. The mutation testing, although useful for evaluating test suite robustness, offers limited applicability for full process automation. In conclusion, integration of adaptive metaheuristics and learning-based optimization techniques into CIDER's architecture shows promising results in improvement of test coverage and full automation potential.

REFERENCES

1. Zhang, Y., Lu, Q., Liu, K., Dou, W., Zhu, J., Qian, L., Zhang, C., Lin, Z., & Wei, J. (2025). CITYWALK: Enhancing LLM-Based C++ Unit Test Generation via Project-Dependency Awareness and Language-Specific Knowledge. *arXiv preprint arXiv:2501.16155*. <https://doi.org/10.48550/arXiv.2501.16155>
2. Rho, S., Martens, P., Shin, S., Kim, Y., Heo, H., & Oh, S. (2023). Coyote C++: An Industrial-Strength Fully Automated Unit Testing Tool. *arXiv preprint arXiv:2310.14500*. <https://doi.org/10.1016/j.jss.2022.111259>
3. Wang, Y., Mäntylä, M. V., Liu, Z., & Markkula, J. (2022). Test automation maturity improves product quality - Quantitative study of open source projects using continuous integration. *Journal of Systems and Software*, 188, 111259. <https://doi.org/10.1016/j.jss.2022.111259>
4. Feloney, S. (2023). *State of Test Automation: Trends and Priorities for 2023*. DevOps Digest. <https://www.devopsdigest.com/state-of-test-automation-2023>
5. Papadakis, M., Kintis, M., Zhang, J., Jia, Y., Le Traon, Y., & Harman, M. (2018). Mutation testing advances: An analysis and survey. *Advances in Computers*, 112, 275–378. <https://doi.org/10.1016/bs.adcom.2018.03.015>
6. Hulevych, M. (2024). CIDER: Assisted automation tool for C++ libraries testing. *Control, Navigation and Communication Systems*, 2(76), 74–77. <https://doi.org/10.26906/sunz.2024.2.074>
7. McCabe, T. J. (1976). A complexity measure. *IEEE Transactions on Software Engineering*, 2(4), 308–320. <https://doi.org/10.1109/TSE.1976.233837>

8. Rapps, S., & Weyuker, E. J. (1985). Selecting software test data using data flow information. *IEEE Transactions on Software Engineering*, 11(4), 367–375. <https://doi.org/10.1109/TSE.1985.232226>
9. Cadar, C., Dunbar, D., & Engler, D. (2008). KLEE: Unassisted and automatic generation of high-coverage tests for complex systems programs. *Proceedings of the 8th USENIX Symposium on Operating Systems Design and Implementation (OSDI 2008)*, San Diego, CA, USA. https://www.usenix.org/legacy/event/osdi08/tech/full_papers/cadar/cadar.pdf
10. Fraser, G., & Arcuri, A. (2011). EvoSuite: Automatic test suite generation for object-oriented software. *Proceedings of the 19th ACM SIGSOFT Symposium and the 13th European Conference on Foundations of Software Engineering*, Szeged, Hungary, 416–419. <https://doi.org/10.1145/2025113.2025179>
11. Harries, L., Storan Clarke, R., Chapman, T., Nallamalli, S. V. P. L. N., Ozgur, L., Jain, S., Leung, A., Lim, S., Dietrich, A., Hernández-Lobato, J. M., Ellis, T., Zhang, C., & Ciosek, K. (2020). Drift: Deep reinforcement learning for functional software testing. *arXiv preprint arXiv:2007.08220*. <https://doi.org/10.48550/arXiv.2007.08220>
12. Veanes, M., Campbell, C., Grieskamp, W., Schulte, W., Tillmann, N., & Nachmanson, L. (2008). Model-based testing of object-oriented reactive systems with Spec Explorer. In R. M. Hierons, J. P. Bowen, & M. Harman (Eds.), *Formal methods and testing* (pp. 39–76). Springer. https://doi.org/10.1007/978-3-540-78917-8_2
13. Claessen, K., & Hughes, J. (2011). QuickCheck: A lightweight tool for random testing of Haskell programs. *SIGPLAN Notices*, 46(4), 53–64. <https://doi.org/10.1145/1988042.1988046>
14. Pacheco, C., Lahiri, S. K., Ernst, M. D., & Ball, T. (2007). Feedback-directed random test generation. *Proc. 29th Int. Conf. on Software Engineering (ICSE '07)*, Minneapolis, MN, USA, 75–84. <https://doi.org/10.1109/ICSE.2007.37>
15. Coles, H., Laurent, T., Henard, C., Papadakis, M., & Ventresque, A. (2016). PIT: A practical mutation testing tool for Java (Demo). *Proceedings of the 25th International Symposium on Software Testing and Analysis (ISSTA 2016)*, Saarbrücken, Germany, 449–452. <https://doi.org/10.1145/2931037.2948707>
16. Karaboga, D. (2010). Artificial bee colony algorithm. *Scholarpedia*, 5(3), 6915. <https://doi.org/10.4249/scholarpedia.6915>
17. Yang, X.-S. (2009). Firefly algorithms for multimodal optimization. In O. Watanabe & T. Zeugmann (Eds.), *Stochastic algorithms: Foundations and applications. SAGA 2009, Sapporo, Japan, October 26–28, 2009* (pp. 169–178). Lecture Notes in Computer Science (Vol. 5792). Springer. https://doi.org/10.1007/978-3-642-04944-6_14
18. Yang, X.-S., & Deb, S. (2009). Cuckoo search via Lévy flights. In *Proceedings of the World Congress on Nature & Biologically Inspired Computing (NaBIC 2009)*, 210–214. IEEE. <https://doi.org/10.1109/NABIC.2009.5393690>
19. Dorigo, M., Maniezzo, V., & Colnari, A. (1996). The Ant System: Optimization by a colony of cooperating agents. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 26(1), 29–41. <https://doi.org/10.1109/3477.484436>
20. Nayak, G., & Ray, M. (2022). Model-based test sequence generation and prioritization using ant colony optimization. *Journal of Information Technology Research*, 15(1), 1–17. <https://doi.org/10.4018/JITR.299946>
21. Palak, P., & Gulia, P. (2018). Ant Colony Optimization Based Test Case Selection for Component Based Software. *International Journal of Engineering and Technology*, 7(4), 2743–2745. <https://doi.org/10.14419/ijet.v7i4.17565>
22. Watkins, C. J. C. H., & Dayan, P. (1992). Q-learning. *Machine Learning*, 8(3–4), 279–292. <https://doi.org/10.1007/BF00992698>

Received (Надійшла) 18.03.2025

Accepted for publication (Прийнята до друку) 14.05.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Гулевич Михайло Володимирович – аспірант кафедри комп'ютерної інженерії та програмування, Національний технічний університет "Харківський політехнічний інститут", Харків, Україна;

Mykhailo Hulevych – PhD student, Computer Engineering and Programming Department, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;
e-mail: gulevich30misha@gmail.com; ORCID Author ID: <https://orcid.org/0009-0003-8622-3271>.

Коломійцев Олексій Володимирович – доктор технічних наук, професор, професор кафедри комп'ютерної інженерії та програмування, Національний технічний університет "Харківський політехнічний інститут", Харків, Україна;

Oleksii Kolomiitsev – Doctor of Technical Sciences, Professor, Professor of the Computer Engineering and Programming Department, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;
e-mail: Oleksii.Kolomiitsev@khp.edu.ua; ORCID Author ID: <https://orcid.org/0000-0001-8228-8404>;
Scopus ID: <https://www.scopus.com/authid/detail.uri?authorId=57211278112>.

Методики автоматизації генерації тестів для програмного забезпечення C++

М. В. Гулевич, О. В. Коломійцев

Анотація. Автоматизація тестування є критично важливою для забезпечення якості програмного забезпечення. Автоматизовані підходи значно скорочують зусилля на тестування, покращують ефективність тестування та підвищують загальну надійність програмного забезпечення. Інструмент CIDER, представлений автором, пропонує багатообіцяюче автоматизоване рішення шляхом генерації тестових сценаріїв на основі запису виконання програм і застосування евристичної пошукової оптимізації, зокрема гармонічного пошуку. Незважаючи на свої переваги, CIDER стикається з обмеженнями, включаючи неповне охоплення складної логіки розгалуження, неефективне дослідження великих просторів вхідних даних, труднощі в обробці семантично насичених або контекстуально складних вхідних даних. Ця оглядова стаття спрямована на систематичне дослідження, порівняння та оцінку різних автоматизованих методик генерації тестів, таких як символічне виконання, конколічне тестування, еволюційні алгоритми, навчання з підкріпленням і тестування на основі моделі. **Мета** полягає в тому, щоб класифікувати ці методи на основі їхніх характеристик, оцінити їхню ефективність у вирішенні наявних недоліків CIDER та визначити відповідні підходи, які могли б розширити можливості інструменту для автоматизованої генерації тестів з точки зору ефективності покриття коду.

Ключові слова: автоматизація тестування, C++, регресійне тестування, бджолиний алгоритм, алгоритм мурашиної оптимізації, алгоритм світлячків, пошук зозулі, навчання за q-функцією.

Maksym Demydenko, Anton Fesenko

National University "Yuri Kondratyuk Poltava Polytechnic", Poltava, Ukraine

DEVELOPMENT OF A WEB API FOR AN ONLINE CAR PARTS STORE USING N-TIER ARCHITECTURE

Abstract. Modern business environment challenges, market globalization, and increasing consumer demands have significantly influenced the development of e-commerce. Digitalization of buying and selling processes has become especially important, encompassing all industries, including the automotive sector. Despite significant progress in web technology development, the issue of ensuring an efficient software structure for online stores remains relevant. Most web applications are built on a monolithic architecture, which complicates their modernization, scaling, and integration with new services. This limits business growth and reduces its efficiency. The aim of this work is to develop a Web API for an online car parts store using an N-Tier architecture. This approach ensures a clear separation of the project into logical layers, simplifying development, testing, and maintenance while enhancing the system's flexibility and scalability.

Keywords: E-commerce, web development, multi-tier architecture.

Introduction

Modern business environment challenges, market globalization, and increasing consumer demands have significantly influenced the development of e-commerce. The digitalization of buying and selling processes has become especially important, covering all industries, including the automotive sector. Despite significant progress in web technology development, the challenge of ensuring an efficient software structure for online stores remains relevant.

Relevance of the Topic. The development of e-commerce, particularly in the field of car parts sales, requires the implementation of new approaches to managing online stores. To ensure the efficient operation of such platforms, it is necessary to use flexible and scalable information systems capable of processing large volumes of data, supporting integration with various services, and adapting to rapidly changing market demands.

Many modern e-commerce platforms struggle with effective scalability and rapid adaptation to new technological trends due to the limitations of traditional architectural approaches.

A multi-tier architecture enables a clear separation of functional system components, enhancing its flexibility and scalability. Therefore, the use of such architectural approaches is relevant for creating high-performance online stores.

Analysis of Recent Research. Most publications dedicated to multi-tier (N-Tier) architecture examine its application in various software development domains, including e-commerce and enterprise systems. For instance, in study [1], the fundamental principles of N-Tier architecture, its components, and key operational aspects are discussed. The study explains how a multi-tier approach enhances scalability, simplifies system maintenance, and facilitates integration with various services.

In study [2] on the Stackify website, it is noted that N-Tier architecture differs from the MVC framework by incorporating an additional middle (business logic) layer, which streamlines communication between different system layers. The study also compares N-Tier

architecture with other architectural approaches and provides real-world application examples.

An article [3] on Medium.com presents an approach to organizing multi-tier solutions to improve application performance while ensuring flexibility and adaptability to changes.

The objective of this work is to develop a Web API for an online car parts store using an N-Tier architecture.

Features of Multi-Tier Architecture in E-Commerce Application Development

The characteristics of information systems are determined not only by their functional capabilities and software tools but also by their architecture. An alternative to monolithic systems is the multi-tier (N-Tier) architecture, which divides the system into separate logical layers. This approach provides several advantages, including:

1. **Modularity** – Each layer is an independent project (e.g., a DLL or JAR file) that can be reused in other projects.

2. **Isolated Scalability** – Expanding or modifying one layer does not affect the operation of other layers.

3. **Development Efficiency** – Different development teams can work independently on separate layers, improving productivity.

4. **Testability** – Each layer can be tested individually, making it easier to detect errors and simplifying the testing of new features.

Limitations of Using Multi-Tier Architecture:

1. **Cost** – Multi-tier architecture requires more memory and resources to manage communication between layers.

2. **Security Concerns** – Strong security measures must be implemented to ensure that higher layers do not directly access lower layers but interact only through the business logic layer.

3. **Reduced Performance** – Additional delays in data exchange between layers may impact the overall system performance [2].

The project implementation diagram is shown in Fig. 1. The overall project consists of a website (front-end) and a server-side component (back-end), which

functions as a REST API utilizing a multi-tier architecture. It processes client requests (from the website or direct API calls) over the HTTP(HTTPS) protocol in JSON format and returns responses in the same format.

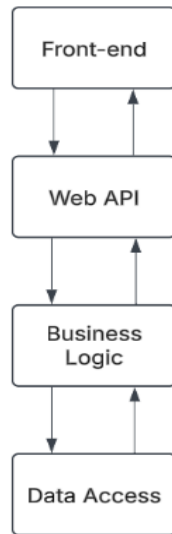


Fig. 1. Project Architecture

The system receives a request from the website and processes it down the schema to the API layer. From there, the request moves to the Business Logic (BL) layer, which checks the user's permissions to execute the request, transforms the data from the Data Transfer Object (DTO) into a fully-fledged object that represents an entity from the database, and performs any necessary calculations for required fields (for example, calculating the final price of a product based on the initial price and applying a discount).

After processing, the data is passed to the Data Access Layer (DAL), where it is stored, modified, deleted, or retrieved from the database [5].

After passing through all the layers, the request returns from the DAL layer to the BL layer, where it is transformed into the object required by the client. The processed data is then passed to the API layer, which returns the data in JSON format to the client.

The project includes the following list of data access services, as shown in Fig 2:

1. IAuthDAL – responsible for searching for users in the database by email, identifier, and recovery token, as well as providing functionality for creating and updating user records.

2. ICartItemDAL – responsible for storing items in the shopping cart in the database, as well as deleting and updating them.

3. IOnlinePaymentDAL – stores information about users' online payments and provides functionality for searching and updating records.

4. IOrderDAL – responsible for storing and updating orders.

5. IProductDAL – stores and updates products in the database. The deletion method is absent because products are not directly deleted from the database; instead, a "soft delete" technique is used. The product is marked as inactive for orders rather than being deleted.

6. IProductDetailsDAL – stores product details in the database. It serves as a base interface for services like IBrandDAL (product brands), ICategoryDAL (product categories), and ISpecialTagDAL (special tags), as they share similar logic, differing only in types.

7. IRefreshTokenDAL – stores and updates refresh tokens in the database.

8. IShoppingCartDAL – responsible for storing and updating user shopping carts.

The general database schema with which this level interacts is shown in Fig. 3. It represents the structure and relationships between various entities within the system, including users, products, orders, payments, shopping carts, and other relevant data components.

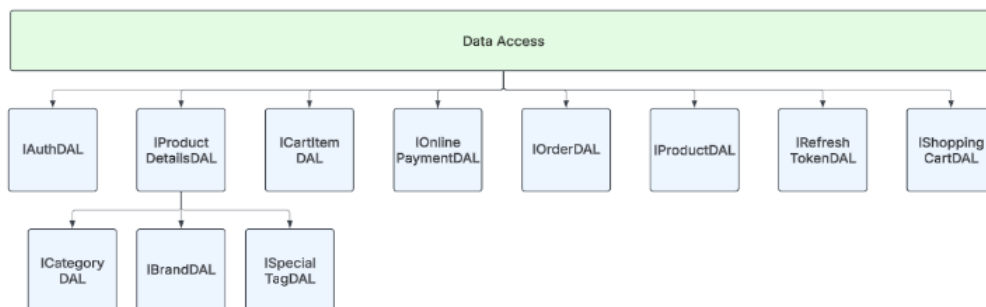


Fig. 2. Data Access Level



Fig. 3. Database Schema

The Business Logic (BL) layer consists of 13 services, which inherit from the interfaces shown in Fig. 4. These services are divided into primary and

auxiliary categories. Primary services are designed for use by the API layer, while auxiliary services are intended for internal use within the business logic layer.

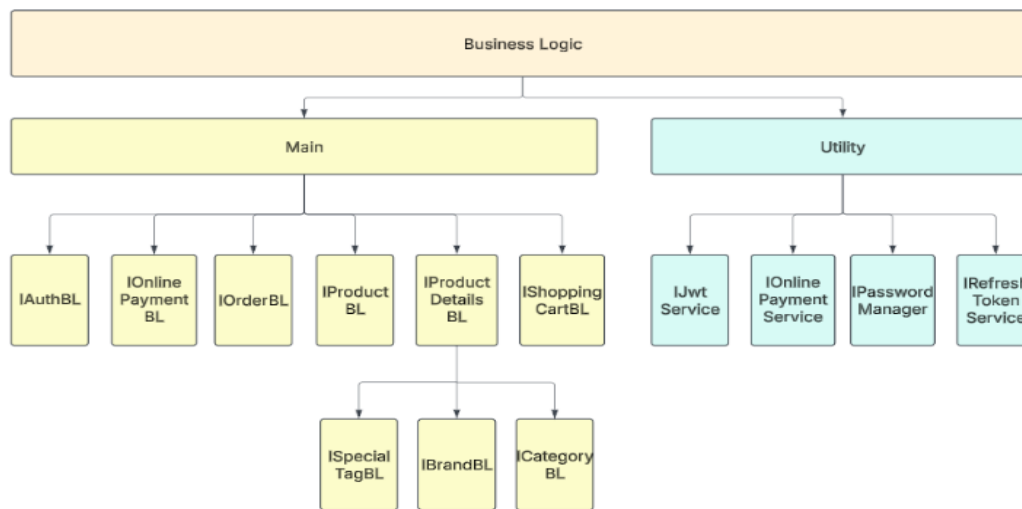


Fig. 4. Business Logic Level

The primary services include:

1. IAuthBL – responsible for user authentication and registration, role changes, and the generation of refresh tokens and JWT access tokens.
2. IOOnlinePaymentBL – manages online order payments (initializing payments and canceling them in case of order cancellation).
3. IOrderBL – responsible for creating and processing orders.
4. IProductBL – responsible for creating and editing products.
5. IProductDetailsBL – Responsible for creating and editing product details. Similar to the business logic layer, it includes three more services under it: ISpecialTagBL, IBrandBL, and ICategoryBL.

IShoppingCartBL – responsible for adding and removing items from the user's shopping cart.

The auxiliary services include the following:

1. IJwtService – a service for generating access tokens in JWT format. It is called by the authentication service to generate access tokens when users log in. The current implementation of this service sets the token's expiration time to 15 minutes.
2. IOOnlinePaymentService – a service for initializing online payments and canceling them. It is invoked by the main OnlinePaymentBL service. The difference between the two is that IOOnlinePaymentService serves as a contract for a specific online payment service, such as Stripe, Portmone, or Privat24. OnlinePaymentBL is unaware of the specific implementation and simply interacts with the interface.
3. IPasswordManager – a password manager responsible for hashing passwords and comparing them during authentication, password changes, or user registration.
4. IRefreshTokenService – responsible for generating refresh tokens. The current implementation specifies that the refresh token's lifespan is 2 weeks.

When a user logs into the system, two tokens are issued: a JWT token with a lifespan of 15 minutes and an access token, which is stored in an HTTP-only cookie and has a lifespan of 2 weeks.

JSON Web Token (JWT) is a compact, secure format for transmitting information between parties in the form of a JSON object. It is widely used for authentication and authorization in web applications. The token consists of three parts: the header, the payload, and the signature [4]. The header defines the token type and the signing algorithm, the payload contains the data (claims), and the signature guarantees the integrity of the token, preventing it from being altered without authorization. JWT can be signed using either a symmetric or asymmetric key, making it reliable for verifying the authenticity of the data. Due to its compact nature, the token can easily be transmitted through HTTP headers or URLs. The primary advantage of JWT is its ability to work securely in stateless systems, where the server does not store session information because all the necessary information is encoded directly within the token.

The refresh token allows for the automatic renewal of the access token when its expiration time is reached, without requiring the user to log in again. This token has a longer lifespan and is stored in an HTTP-only cookie, providing added security and convenience by eliminating the need to repeatedly enter the password.

The refresh token is a unique string of random characters, 32 bytes in length. It is stored in the database, which also associates the token with a specific user and defines its expiration time. After being stored in the database, the token is sent to the API via the Business Logic (BL) layer, which then adds the token to the response in the HTTP-only cookie. This process ensures that the user's session remains valid without requiring constant re-authentication.

HTTP-only cookies are accessible only by the server and cannot be read by JavaScript, which reduces

the risk of XSS (Cross-Site Scripting) attacks. This feature enhances security by preventing malicious scripts from accessing sensitive information stored in the cookies. Additionally, it allows tokens to be automatically sent with each request without the need for manually adding them to the headers, simplifying authentication while maintaining security.

The API layer handles the reception and sending of processed requests to the client. It consists of 8 controllers in total:

1. AuthController – responsible for user registration, authentication, role changes, password changes, and access token renewal. The overall workflow and dependencies of this functionality are shown in Fig. 5.

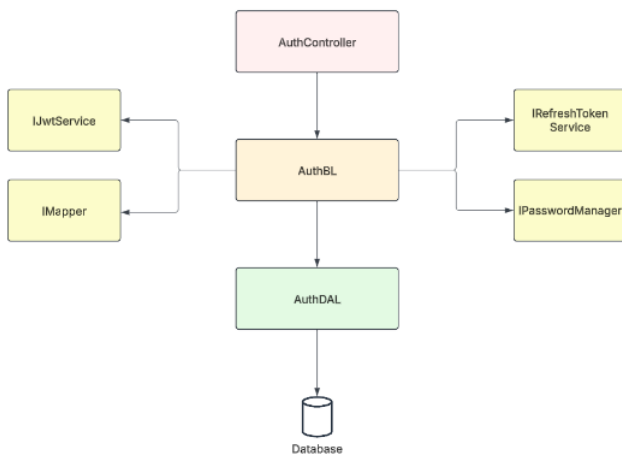


Fig. 5. Auth Component Functionality

2. BrandController – responsible for retrieving, creating, and deleting product brands. The functionality diagram is shown in Fig. 6.

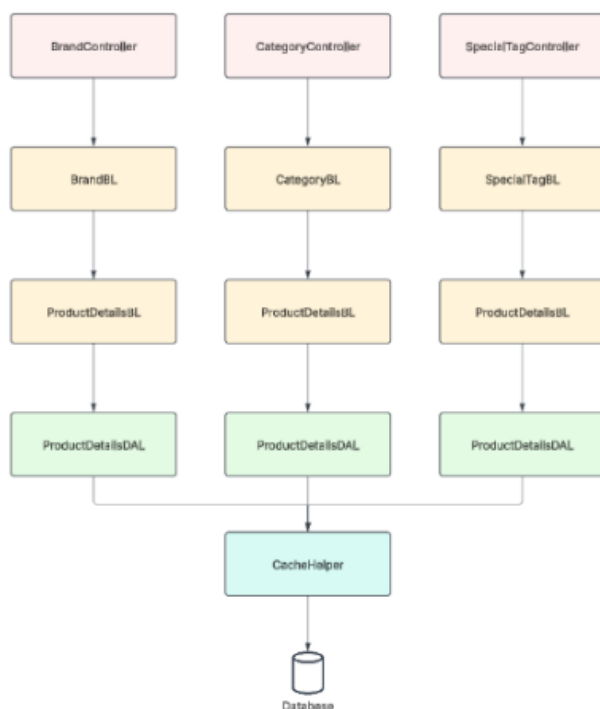


Fig. 6. Functionality of the Brand, Category and Special Tag Components

3. CategoryController – responsible for retrieving, creating, and deleting product categories. The functionality diagram is shown in Fig. 6.

4. SpecialTagController – responsible for accessing, creating, and deleting special tags for products. The functionality diagram is shown in Fig. 6.

5. OnlinePaymentController – responsible for retrieving, initializing, and canceling online payments. The functionality diagram is shown in Fig. 7.

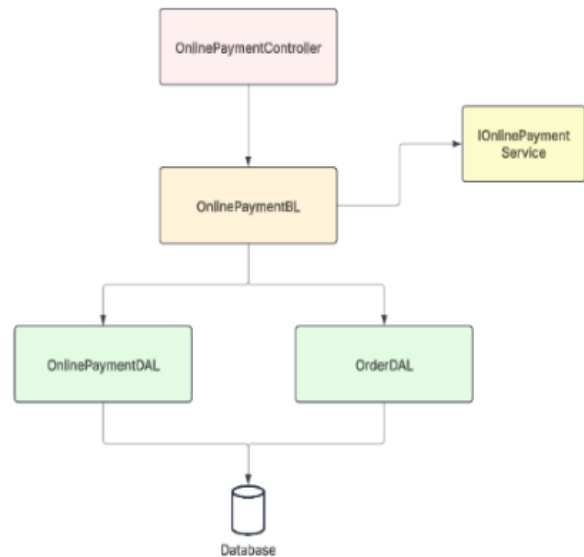


Fig. 7. Functionality of the Online Payment Component

6. OrderController – responsible for creating and updating orders. The functionality diagram is shown in Fig. 8.

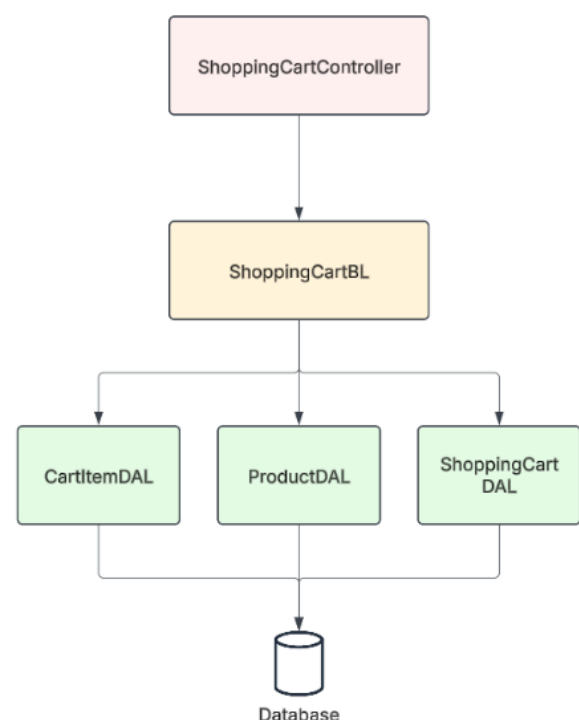


Fig. 8. Component for Order Processing

7. ProductsController – responsible for retrieving, creating, and updating products. The component is shown in Fig. 9.

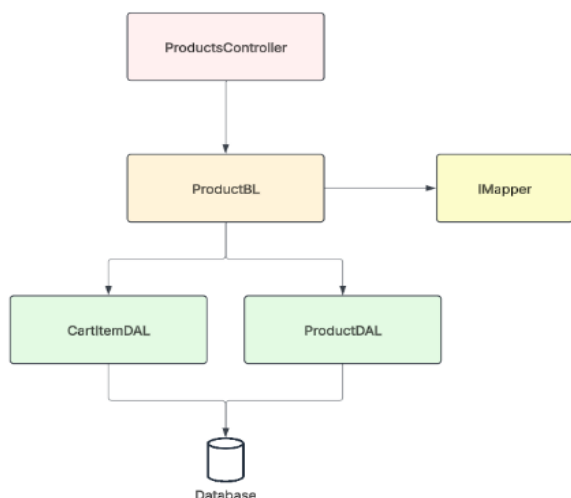


Fig. 9. Functionality of the Products Component

8. ShoppingCartController – responsible for accessing and updating the shopping cart. The component is shown in Fig. 10.

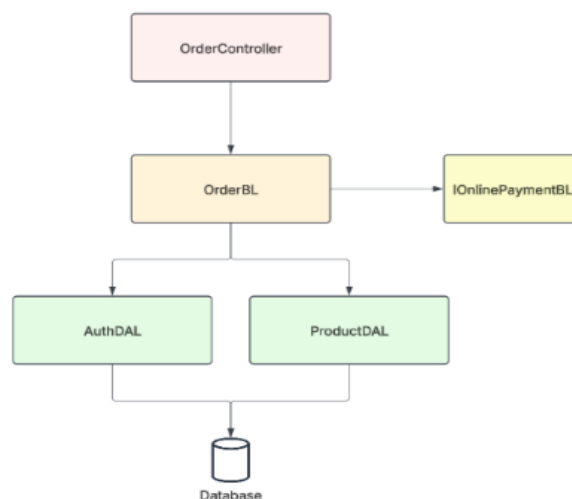


Fig. 10. Functionality of the Shopping Cart Component

Conclusion

The analysis of recent studies revealed that most publications on multi-tier (N-tier) architecture focus on its use in enterprise applications and business process management systems. Based on the findings, a prototype Web API for online stores selling auto parts has been developed using multi-tier architecture.

REFERENCES

1. Бойко Н. І. Еволюція побудови архітектур інформаційних систем. Перспективи розвитку «хмарної» архітектури». Вісник Національного університету «Львівська політехніка», Серія: Інформаційні системи та мережі. 2015. Вип. 832. С. 348–367. URL: <https://science.lpnu.ua/sites/default/files/journal-paper/2018/jun/12945/24-348-367.pdf>
2. What is N-Tier Architecture? How It Works, Examples, Tutorials, and More. URL: <https://stackify.com/n-tier-architecture>
3. Understanding N-Tier Architecture: Building Robust and Scalable Applications URL: <https://medium.com/@segekaratas/understanding-n-tier-architecture-building-robust-and-scalable-applications-62db30a40b5>
4. What is JWT? URL: <https://blog.postman.com/what-is-jwt/>
5. Understanding Layers, Tiers, and N-Tier Architecture in Application Development: URL: <https://dev.to/3bdelrahman/understanding-layers-tiers-and-n-tier-architecture-in-application-development-1hlb>

Received (Надійшла) 11.03.2025

Accepted for publication (Прийнята до друку) 07.05.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Демиденко Максим Ігорович – старший викладач кафедри комп'ютерних та інформаційних технологій і систем, Національний університет «Полтавська політехніка імені Юрія Кондратюка», Полтава, Україна;

Maksym Demydenko - Senior Lecturer, Department of Computer and Information Technologies and Systems, National University «Yuri Kondratyuk Poltava Polytechnic», Poltava, Ukraine;

e-mail: maxx74@ukr.net; ORCID Author ID: <https://orcid.org/0009-0001-1374-7379>

Фесенко Антон Богданович – студент кафедри комп'ютерних та інформаційних технологій і систем, Національний університет «Полтавська політехніка імені Юрія Кондратюка», Полтава, Україна;

Anton Fesenko - student of the Department of Computer and Information Technologies and Systems, National University «Yuri Kondratyuk Poltava Polytechnic», Poltava, Ukraine;

e-mail: tohafesenkolovep@gmail.com

Створення WEB API для онлайн-магазину автотоварів з використанням архітектури N-Tier

М. І. Демиденко, А. Б. Фесенко

Анотація. Сучасні виклики бізнес-середовища, глобалізація ринків та зростання потреб споживачів суттєво вплинули на розвиток електронної комерції. Особливо важливою стала цифровізація процесів купівлі-продажу, що охопила всі галузі, зокрема й автомобільну індустрію. Незважаючи на значний прогрес у розвитку веб-технологій, проблема забезпечення ефективної структури програмних рішень для онлайн-магазинів залишається актуальною. Більшість веб-додатків побудовані на монолітній архітектурі, що ускладнює їх модернізацію, масштабування та інтеграцію з новими сервісами. Це обмежує розвиток бізнесу та знижує його продуктивність. Метою даної роботи є створення Web API для онлайн-магазину автотоварів із використанням архітектури N-Tier. Такий підхід забезпечує чітке розділення проєкту на логічні рівні, що спрощує розробку, тестування й подальшу підтримку, а також підвищує гнучкість і масштабованість системи.

Ключові слова: електронна комерція, веб-розробка, багаторівнева архітектура.

Tetyana Derkach¹, Andrii Dmytrenko¹, Lina Klochko²

¹ National University «Yuri Kondratyuk Poltava Polytechnic», Poltava, Ukraine

² Geological Bureau BEG SA, Switzerland

MODELING AND IMPLEMENTATION OF A DISTANCE LEARNING SYSTEM: STAGES, REQUIREMENTS, FUNCTIONALITY

Abstract. The article examines the essential aspects of organizing and implementing a distance learning system in the context of modern educational development. The study analyzes the principal stages of distance education deployment, ranging from initial planning to the full-scale operation of the system. A comprehensive specification of requirements for the system's functionality is outlined, including technical, functional, and organizational dimensions. The article defines the core user groups of the distance learning system, with a particular focus on identifying their roles and specific needs within the learning environment. Special emphasis is placed on the design of the user interface, which must ensure intuitiveness, accessibility, and adaptability to diverse user expectations. Additionally, the article addresses the requirements for the database structure intended to manage and store information related to users, courses, and learning outcomes. The core system modules are identified and described, including modules for registration, administration, educational content management, assessment, and data analytics. The proposed approach facilitates the development of an efficient and user-centered distance learning environment designed to meet the evolving demands of contemporary education.

Keywords: distance learning, computer technologies, modular software.

Introduction

In the modern context of the global use of computer technologies in every sphere of human activity, there arises a necessity for the implementation of both elements of e-learning and distance learning in general within educational institutions.

Distance learning is a combination of modern technologies that ensure the delivery of information interactively through the use of ICT (information and communication technologies) from those who teach (teachers, prominent figures in specific scientific fields, politicians) to those who learn (students or users).

The main goal of distance learning (DL) is to provide all citizens with the opportunity to acquire quality knowledge and the corresponding skills and abilities at their place of residence through the use of information and communication technologies and appropriate software [4].

An important factor in ensuring a sufficient quality level of DL during the organization of the educational process is the well-reasoned and balanced choice of specialized software for managing distance learning. The characteristic features of this form include flexibility, modularity, parallelism, cost-effectiveness, technological advancement, social equality, and a new role for the teacher.

Thus, it is possible to effectively allocate users' time and enable them to determine the quality of their own preparation. Distance education requires a high level of professional training for teachers. For the implementation of this form of learning into the educational process, it is necessary to determine the distance learning system.

Distance learning systems (Learning Management Systems) are applied software products for managing the educational process. Today, a wide range of distance learning systems is used in educational institutions in Ukraine. Almost all modern LMSs offered by developers meet the specified requirements, but not all

of them are accessible to most educational institutions due to their high cost and complexity of use [8].

The absence of any standards and requirements for distance learning systems in the educational system compels the vast majority of educational institutions to make their choice among existing software based on the following main criteria:

- reliability in maintenance and security;
- compatibility;
- ease of use and administration;
- modularity;
- accessibility;
- cost of software, maintenance, and hardware.

Based on the above, the task was set to develop an accessible distance learning system that would meet the needs of any educational institution.

The aim of this article is to present a scheme of a distance learning course and to propose elements of a distance learning system.

Stages of Organizing and Implementing the Learning Process

The work of organizing and implementing the learning process is conventionally conducted in two stages:

0. Preparatory Stage.

At this stage, the instructor creates and populates the electronic course with logically-structured educational information, taking into account the specifics of the subject. The instructor also plans for student consultations and schedules the educational activities for all participants (Fig. 1).

1. The Main Stage.

This stage involves the organization of users' educational activities according to the plan, as well as adjustments to the educational material and the structure of the electronic course.

These stages are based on the working programs of academic disciplines, which define the subject, goals, and objectives of the academic discipline, as well as the

program of the academic discipline with a detailed course structure.

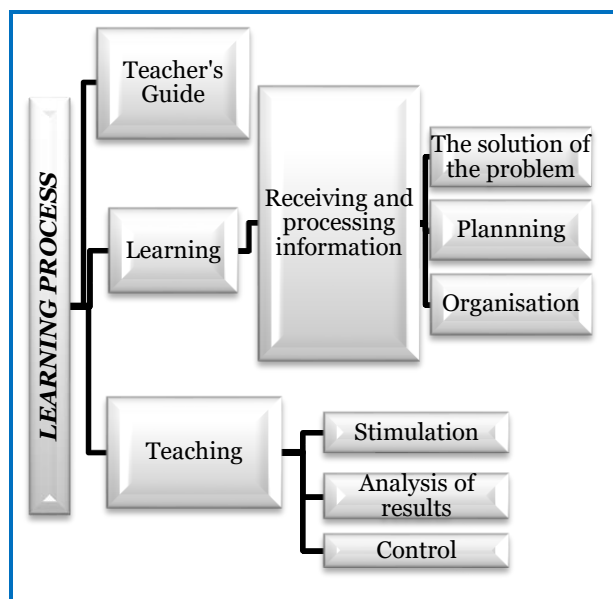


Fig. 1. The general schema of the learning process.

In the creation of a distance learning course at Poltava National Technical University named after Yuri Kondratyuk, a scheme was implemented as shown in Fig. 2.

The scheme was developed using the example of the discipline "Informatics," aimed at providing high-level computer training for future specialists. The course builds on the knowledge of computer technology acquired in school.

The main challenge in developing an electronic course is the unpreparedness of instructors to work with the software products used to create electronic versions of lectures and supplementary materials for the practical part of the course.

An important stage in the development and implementation of a distance learning course is the preparatory stage, which includes the development of the future course structure.

It is necessary to understand that, in modern conditions, there is still an issue with the level of computer science education among high school graduates, resulting in students at higher educational institutions having varying levels of preparation. Therefore, some initial topics overlap with the school curriculum. However, our goal is to provide more in-depth knowledge on these topics.

The subject matter was developed considering the updates in modern computer hardware and software (Fig. 3).

To implement a distance learning format in the educational institution, the creation of a modular software package was proposed. A work plan was developed to achieve this goal (Fig. 4).

2. Software Requirements Specification

The Software Requirements Specification (SRS) is a complete description of the behavior of the distance learning system to be developed.

Specification contents:

1. Purpose of creation.
2. User characteristics.
3. Product functions.
4. Assumptions and dependencies.
5. General limitations.
6. Operations.
7. Interface requirements.
8. Technological tools.
9. Memory constraints.
10. Maintenance.
11. Database requirements.
12. System architecture.

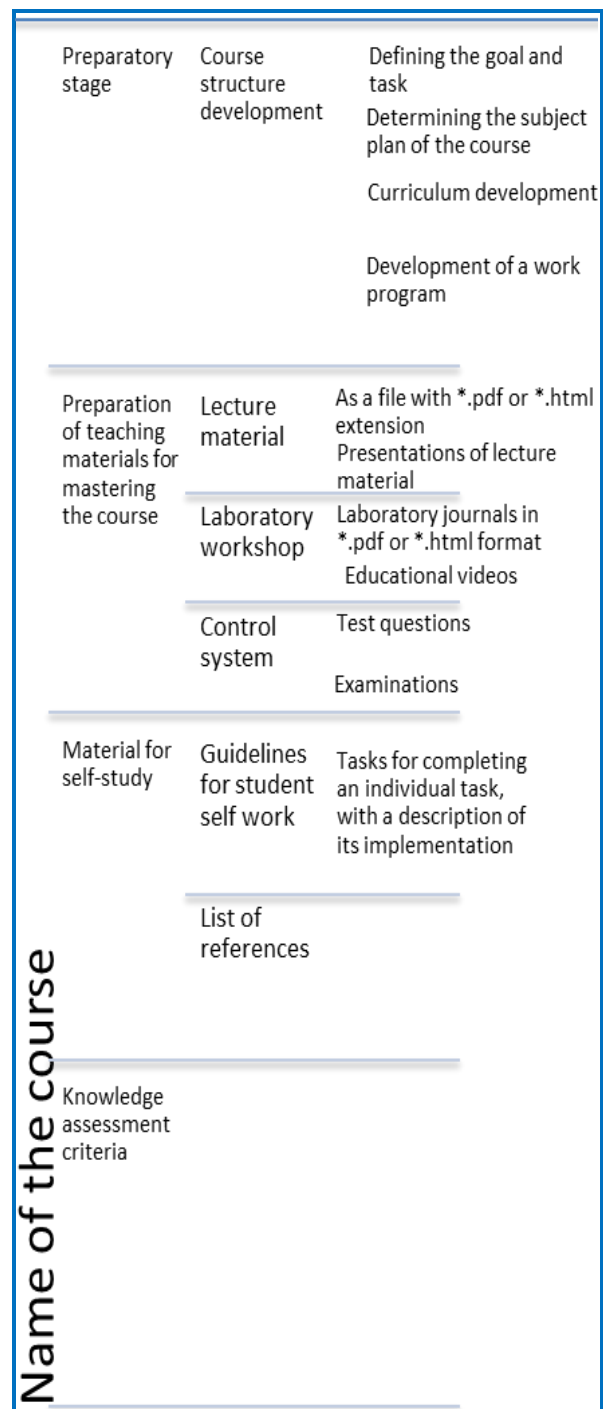


Fig. 2. Scheme of the educational process organization

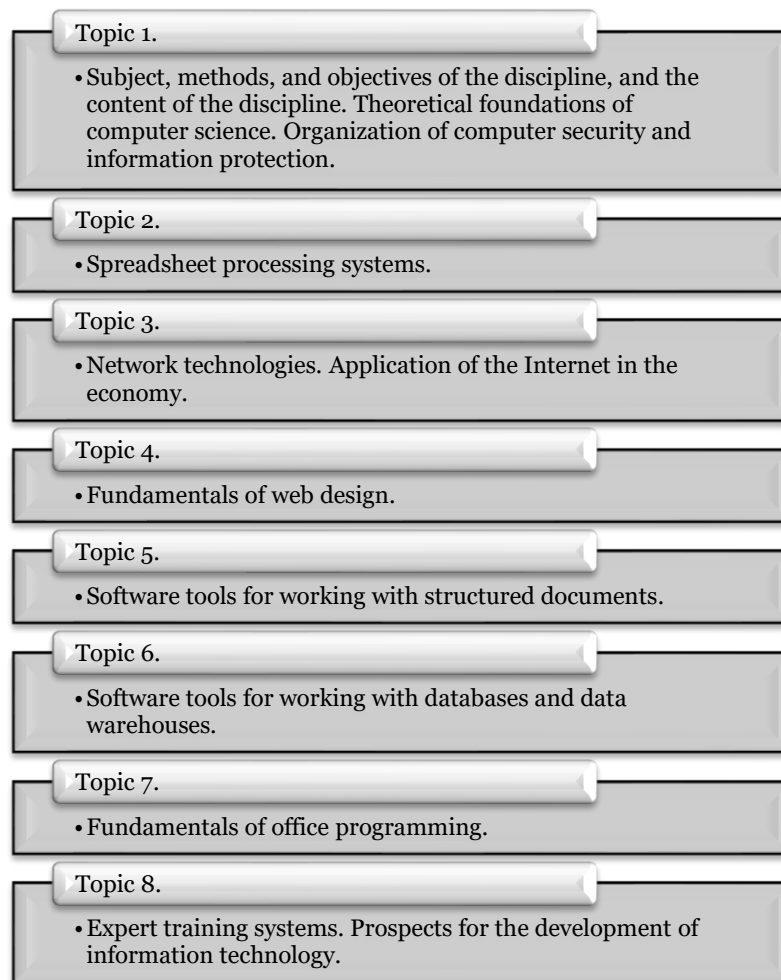


Fig. 3. Distance course topics

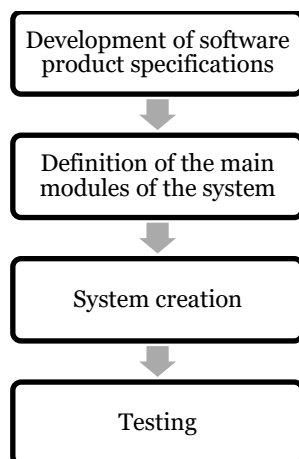


Fig. 4. Planning the implementation of a software development task

2.1. Purpose of creation

The implementation of the latest technologies in the educational process to ensure a distance learning format.

2.2. User characteristics

- Managers of various levels.
- Heads of regional management bodies.
- Individuals seeking a second education or retraining.

- Army officers facing downsizing.
- Individuals wishing to improve their qualifications in a specific field of knowledge.
- Citizens with limited mobility.
- Persons with disabilities.
- Citizens who wish to receive education at a convenient time and are located remotely from the educational institution.
- Highly capable students who already possess substantial knowledge and wish to complete an educational program in a shortened timeframe.
- Users who want to combine study with professional activities.
- Users wishing to complete special educational programs consisting of courses provided by various educational institutions, including those from different countries.
- Users geographically isolated from educational resources.
- Individuals who did not complete their education in their youth.

2.3. Product functions

- Accounting for learners, personalization, and access rights management to educational materials.
- Management of the learning process, including tracking learning and testing outcomes.
- Preparation of operational and analytical reports.

- Management and integration with electronic communication mechanisms.

- Integration with external information systems.

2.4. Assumptions and dependencies

- All users must have the technical and software capabilities necessary for quality learning (e.g., working with video, audio, multimedia).

- The presence of appropriate technical and software support, access to information, and the ability to use distance learning tools are required. The user must have a personal computer and internet access.

2.5. General limitations

- The system has limitations in knowledge verification.

- There is no direct face-to-face interaction between students and instructors. The presentation of material lacks emotional expression, making it difficult to create a creative atmosphere among students.

- High demands are placed on the organization of the learning process and its administration.

- A key challenge is the issue of user authentication during knowledge assessments.

- A set of individual psychological conditions is mandatory. The effectiveness of distance learning depends on the student's independence, awareness, and strict self-discipline.

- There is a lack of constant supervision over the learners.

- There is a noticeable lack of practical work.

2.6. Operations

- Authoring tools for developing educational content.

- Learning Management System (LMS).

- System for information exchange between participants in the educational process.

- Educational content delivery system (usually a website).

2.7. Interface requirements

- Since the quality of the user's interactive experience with the system (speed, convenience, low fatigue) is related to psychological characteristics such as short-term and medium-term memory, reaction time, and visual information perception capabilities, it is important to remember that the interface is the most crucial part of a Decision Support System (DSS) from both an advertising perspective and the direct user's perspective, who may work with it for several hours at a time.

- The interface influences the nature of decisions; it can accelerate decision-making time and improve or degrade their quality.

- The specific type of interface that can be created with the chosen tools and the fundamental possibilities offered by the tool system should be considered.

- The interface design was carried out based on the developed data flow diagram (Fig. 5).

Here is the English translation of the provided Russian technical text:

According to the international classification, the following interface requirements must be met:

- ✓ Alignment with user tasks.
- ✓ Ease of use.

- ✓ Manageability.
- ✓ Conformity to user expectations.
- ✓ Error tolerance.
- ✓ Adaptation/Personalization.
- ✓ Ease of learning.

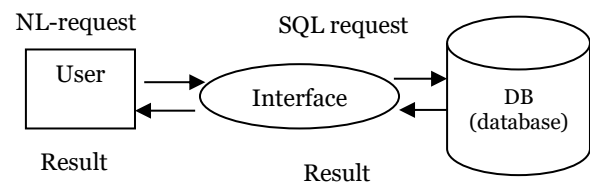


Fig 5. General diagram of data flows of a natural language (NL) interface

2.8. Technological Tools

- Internet technologies.

- World Wide Web: a system for organizing information on the Internet based on hypertext.

- FTP (File Transfer Protocol) for file transfers.

- E-mail.

- Synchronous, real-time online teleconferences.- Video conferencing in computer networks holds a special place.

- Given the needs for communication via global network channels and considering current trends in channel utilization methods, it is advisable to focus on the TCP/IP protocol stack. Accordingly, there are two main possibilities for implementing transport services: streaming mode, which is provided by the TCP protocol, and datagrams based on the UDP protocol [8].

2.9. Memory Constraints

- There are no memory resource limitations, reaction speed to system requests, etc.

2.10. Maintenance

- The software product is maintained by a team of specialists, including course instructors-designers and technology experts (television and video recording, computer graphics, network technologies, etc.).

2.11. Database Requirements

The main task of the database (DB) is to ensure the guaranteed preservation of large volumes of information and provide user access to it. Thus, the DB consists of two parts: stored information and a management system. In this project, the MySQL relational database management system was used. MySQL is an open-source database management system (DBMS).

MySQL server capabilities:

- Easy installation and use.

- Supports an unlimited number of users simultaneously working with the DB.

- The number of rows in tables can reach 50 million.

- High command execution speed.

- Simple and effective security system.

The main requirements for databases are:

- Structured: it should be organized according to a unified principle: by organizations, employees, industries.

- Ease of use.

- Maximum completeness of information.

2.12. Definition of Main System Modules

- Administration Module: includes role distribution, definition of restrictions, and general system settings.

- Learning Process Organization and Support Module: allows the introduction of materials for theoretical and practical training in any format into the system.

- Test Development and Support Module For technical universities, there is a need to implement both closed and open type test tasks. Therefore, when developing the testing module, the task of evaluating the assignment completion arose.

- Educational Materials Development and Presentation Module: should ensure the inclusion of theoretical material and methodological recommendations for completing the practical part in the course. In modern teaching conditions, the presentation of information changes, so it is necessary to provide students with video and multimedia materials. This means anticipating the inclusion of various formats and their compatibility with the system. Interactive Interaction Module for Course Users:

lecturer-student, student-student, student-lecturer interactions. The most commonly used tools are forums and chats.

- User Activity Registry Module: should provide complete information about the work of students in the system and allow the output of reporting information.

3. Graphical Modeling of the Information System

During the preliminary analysis of the task, graphical modeling of the information system was performed using a Data Flow Diagram (DFD) (Figure 6).

This methodology of graphical structural analysis describes external sources and destinations of data relative to the system, logical functions, data flows, and data stores that are accessed.

Information sources (external entities) generate information flows (data flows), carrying information to subsystems or processes.

These, in turn, transform the information and generate new flows that carry information to other processes or subsystems, data stores, or external entities - information consumers [2].

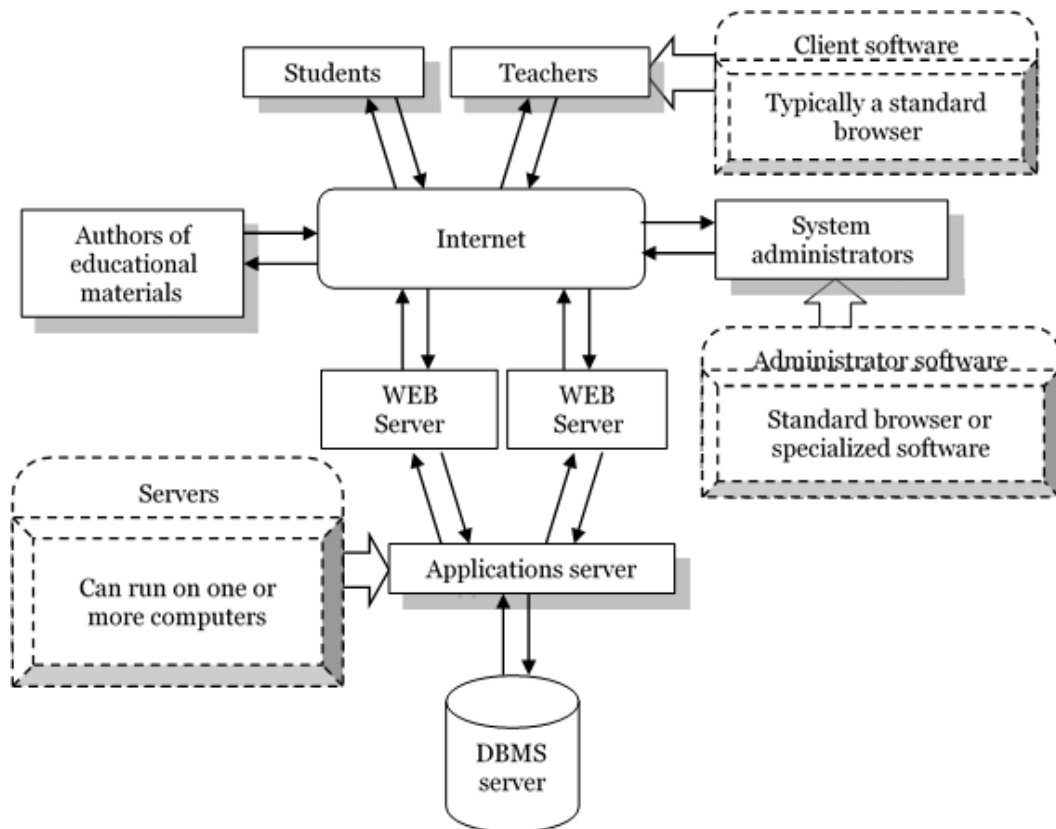


Fig 6. DFD Data Flow Diagram

Conclusions

The conducted analysis highlights that the effective organization and implementation of distance learning systems requires a comprehensive approach that integrates technical, functional, and organizational components. Clearly defined user roles, a user-friendly and adaptive interface, and a structured database for managing user, course, and learning outcome

information are key factors for ensuring the system's efficiency and sustainability.

The identification and integration of essential system modules – including registration, administration, content management, assessment, and analytics – form the foundation for creating a flexible and scalable educational environment. The proposed approach not only addresses the current needs of learners and educators but also contributes to the development of a

modern, accessible, and high-quality distance learning ecosystem.

The results of this research may serve as a methodological basis for further improvement of

distance education platforms and their integration into the broader educational process, facilitating the transition toward more personalized, data-driven, and technology-enhanced learning experiences.

REFERENCES

1. Bissell C. C. The Open University of the United Kingdom. Leadership in Science and Technology: A Reference Handbook. 2011. Vol. 2. 24–32.
2. A.A. Havrylyshyn, T.A. Dmytrenko, T.M. Derkach, A.O. Dmytrenko. Development of informational system for the electronic educational and methodical complex of the discipline. Системи управління, навігації та зв'язку. Полтава: ПолтНТУ, 2019. Т. 4 (56). 35–39.
3. Kalyanov G.N. Consulting in Enterprise Automation: Approaches, Methods, Tools. URL: http://www.business-process.ru/designing/methodology/dfd/dfd_theory_dfd.html
4. ISO/IEC 25000:2014 Systems and software engineering [Electronic source] – Available at: <https://www.iso.org/obp/ui#iso:std:iso-iec:25000>
5. Katernyak I. Guide to the preparation and organization of electronic education. K.: PLC "Farbovaniy lyst", 2016. - 48 p. (in Ukrainian)
6. Kukharenyk V.M., Rybalko O.V., Syrotenko N.H. "Distance Learning. Conditions of Application. Distance Course." Edited by Kukharenyk V.M. Kharkiv: Torsing, 2001. 320 p.
7. Lomanchenko A., Khloponin O., Derkach T., Dmytrenko T. Technological characteristics of the e-learning systems development. Системи управління, навігації та зв'язку. 2021. Т. 2 (64). 84–88.
8. Velychko V.Yu., Kamyshyn V.V., Stryzhak O.Ye. "Information Technologies for the Formation of Modern Knowledge Systems as the Basis for Innovative Development of Education" // Materials of the Interdisciplinary Scientific and Practical Conference "Innovative Teaching Technologies for Gifted Youth," December 08–09, 2010, in Kyiv. IOD. 2010. 168 p.
9. Деркач Т. М. Дистанційне навчання в сучасних умовах освітнього процесу. Наука та освіта. ХНУ, 2025. 46–49.
10. Дмитренко Т.А., Деркач Т.М., Ладатко Д.С. та ін. Модуль «система дистанційного навчання» для Порталу кафедри ВНЗ. Математичне та імітаційне моделювання систем. МОДС '2016: тези доповідей 11 міжнар. наук.-практ. конф. Чернівці: ЧНТУ, 2016. 324 – 325.
11. Дмитренко Т.А., Дмитренко А.О., Деркач Т.М. Технологія розробки системи дистанційного навчання. Integration of scientific bases into practice: abstracts of IV th International scientific and practical conference. Stockholm, Sweden, 2020. 480–484. URL: <https://isg-konf.com/uk/integration-of-scientific-bases-into-practice-ua/>

Received (Надійшла) 04.03.2025

Accepted for publication (Прийнята до друку) 14.05.2025

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 – Ph.D., 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>.

Клочко 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>.

Моделювання та реалізація системи дистанційного навчання: етапи, вимоги, функціональність

Т. М. Деркач, А. О. Дмитренко, Л. А. Клочко

Анотація. У статті розглянуто ключові аспекти організації та реалізації системи дистанційного навчання на сучасному етапі. Проаналізовано основні етапи впровадження дистанційної освіти – від планування до експлуатації системи. Представлено специфікацію вимог до функціонування системи, зокрема технічні, функціональні та організаційні аспекти. Надано характеристику основних груп користувачів, із визначенням їхніх потреб і ролей у системі. Особливу увагу приділено вимогам до інтерфейсу користувача, що має забезпечувати інтуїтивність, доступність та адаптивність. Розглянуто вимоги до структури бази даних для зберігання інформації про користувачів, курси, результати навчання. Визначено основні модулі системи, включно з модулем реєстрації, адміністрування, навчального контенту, тестування та аналітики. Запропонований підхід сприяє створенню ефективного та зручного середовища для дистанційного навчання, що відповідає сучасним освітнім потребам.

Ключові слова: дистанционное обучение, компьютерные технологии, модульное программное обеспечение.

Kuilen Do¹, Iryna Klymova¹, Elen Naumova¹, Mykhailo Herevych², Oleksandr Yankovskyi¹

¹ Kharkiv National University of Radio Electronics, Kharkiv, Ukraine

² Uzhhorod National University, Uzhhorod, Ukraine

DATA PROCESSING AND ANALYSIS METHODS IN IOT USING MACHINE LEARNING

Abstract. Relevance. The growing integration of Internet of Things (IoT) technologies into all areas of human life – from intelligent households to smart city infrastructure – is accompanied by an exponential increase in the volume of data being collected, transmitted, and processed in real time. When combined with artificial intelligence technologies, this data becomes the foundation for making autonomous decisions, predicting user behavior, and adapting environments to the needs of specific individuals. However, it is precisely in this context that the critically important issue of personal data protection arises. Many IoT devices operate in uncontrolled environments, have limited resources for cryptographic protection, and are vulnerable to cyberattacks and unauthorized data collection. Meanwhile, artificial intelligence algorithms used to analyze this data often exhibit the “black box” problem, where it is impossible to fully explain how and why a particular decision was made based on personalized data. The lack of transparency, combined with broad access to sensitive information, threatens fundamental human rights to privacy. The relevance of this topic is driven by the need to find balanced technical solutions that enable both effective analysis of large-scale data in IoT environments and a high level of data security. In this regard, the study of modern methods for processing, analyzing, and protecting data in IoT systems, adapted to the requirements of ethical artificial intelligence and digital privacy standards, represents one of the key challenges of contemporary digital science. **The object of research:** the processes of data collection, processing, analysis, and protection in IoT systems, particularly those components related to the use of users' personal information and its processing through artificial intelligence methods. **Purpose of the article:** research of modern methods for data processing and analysis in IoT systems. The objective of the work is to identify the most effective approaches to secure data handling, characterize existing privacy threats, and assess the potential for integrating protected analytical algorithms that meet both the technical and ethical requirements of the digital environment. **Research results.** A comprehensive analysis of modern approaches to data collection, processing, analysis, and protection in IoT systems has been conducted, particularly in the context of the growing role of artificial intelligence. The technological foundations of IoT functionality were examined, key architectural components identified, and their role in creating digital ecosystems for monitoring, management, and decision-making across various sectors – from household systems to critical infrastructure – was investigated. Special attention was paid to data preprocessing methods, which help reduce information load, improve the quality of analysis, and adapt data flows to the requirements of intelligent algorithms. It was demonstrated that the use of edge processing and local-level aggregation enhances both system performance and security. The main types of databases for IoT – especially those optimized for time series – were analyzed, along with tools for handling large volumes of data in cloud and hybrid environments. **Conclusions.** Data collection methods in IoT are multilayered and closely linked to the requirements for energy efficiency, security, latency, and system scalability. The quality and reliability of the collected information form the foundation for subsequent processing, analysis, and decision-making; therefore, the selection of sensors, communication protocols, and architectural models is of strategic importance for any IoT system. Preprocessing and efficient data storage in IoT are critical stages that ensure the quality, security, and usability of information for further analysis. They determine not only the accuracy of analytics but also the stability, scalability, and compliance with regulatory standards. This creates a demand for the development of adaptive, intelligent data processing and storage systems capable of dynamically responding to changes in device operation context and user requirements. The successful implementation of secure IoT solutions requires an integrated approach that combines technical expertise, legal knowledge, and ethical responsibility.

Keywords: IoT, artificial intelligence; data processing; data analysis; privacy; personal data protection; edge computing; differential privacy; machine learning; intelligent systems.

Introduction

IoT has, over the past decade, evolved from an experimental concept into a fundamental component of modern information systems. The essence of IoT lies in creating a network of physical objects – devices equipped with sensors, communication modules, computational elements, and software – that can interact with each other, with the external environment, and with centralized computing platforms. These devices continuously generate large volumes of data about physical processes, user behavior, the condition of technical systems, or the surrounding environment. As a result, the issues of efficient data processing and analysis become extremely relevant, as they allow raw information streams to be transformed into practically valuable knowledge, recommendations, or actions.

The significance of this problem becomes particularly evident against the backdrop of widespread implementation of artificial intelligence, which greatly expands the capabilities of IoT – from intelligent control of industrial processes to autonomous management of smart city systems. However, alongside technical advantages come serious challenges. Among them is the protection of personal data, which is increasingly being processed by systems embedded in everyday human life: gadgets, vehicles, household appliances, and more. The combination of large volumes of information, automatic AI-based data processing, and distributed architecture creates risks of data leakage, unauthorized access, and misuse of personal information.

Thus, in the context of rapid digitalization and the widespread adoption of IoT, the study of data processing and analysis methods becomes critically im-

portant – not only to ensure high performance and accuracy but also to meet the requirements of confidentiality, ethical data handling, and security. IoT data processing systems must be capable of working effectively with distributed, heterogeneous, and potentially sensitive datasets while maintaining data integrity, reliability, and privacy.

The purpose of this work is a comprehensive analysis of modern methods for data processing and analysis in IoT ecosystems, considering the requirements for personal data protection, which is particularly relevant in the context of implementing machine learning in digital infrastructures. This approach allows consideration not only of the technical aspects of efficient IoT system performance but also of the legal, social, and ethical components that shape trust in technology in the digital age.

Analysis of publications. In recent years, the issue of data processing and analysis in IoT systems has become one of the central topics in the fields of applied informatics, artificial intelligence, and computer security. The analysis of scientific sources indicates active development in areas such as the optimization of computational processes at the edge, the integration of machine learning for intelligent data analysis, and the development of mechanisms for ensuring privacy in IoT environments.

One of the fundamental reviews is presented in [1], which outlines the general principles of IoT architecture and identifies key problems related to the analysis of large data streams. This publication laid the foundation for further research in the field of data processing, particularly in the context of distributed computing and cloud integration.

Research [2] focuses on the practical use of data analysis methods for managing urban infrastructure, including sensor monitoring, traffic flow analysis, and environmental control. Clustering methods, neural network approaches, and elements of predictive analytics are widely applied in this work.

A separate category includes studies devoted to the application of artificial intelligence in IoT, particularly deep learning. For example, work [3] examines classification and anomaly detection algorithms in data streams from wireless sensor networks, which are foundational for many IoT solutions.

A significant number of publications are also dedicated to privacy and information protection issues. For instance, article [4] analyzes the threats of personal data leakage and considers pseudonymization, anonymization, and differential privacy technologies. These aspects become even more important in the context of automated data processing using artificial intelligence. The authors emphasize that privacy and security issues remain key challenges for IoT. Despite the rapid adoption of IoT in critical sectors (healthcare, energy, infrastructure), the level of threats remains high due to poor device protection, limited resources, and weak security governance at various architectural levels. Several major threats are highlighted: insufficient firmware and software updates for IoT devices, the absence of effective and reliable security protocols, low user awareness

of privacy risks, and real-time monitoring of active devices, which creates potential data leakage points. These factors make IoT environments vulnerable to numerous cyberattacks, certificate spoofing, node impersonation, or interception of unencrypted traffic. The proposed architecture allows for effective distribution of computational tasks and protection of data at every stage of its transmission through the system.

Modern sources such as [5,6] demonstrate a growing interest in combining IoT with machine learning, which requires new solutions for data storage, preprocessing, cleansing, and interpretation. The authors describe the synergy of artificial intelligence and IoT as a transformational approach to data processing and decision-making. It enables deeper integration of digital technologies into human life and business processes. However, alongside the advantages, this integration also brings new challenges – particularly in the context of protecting sensitive data, ensuring transparency of decisions, and the ethical use of information. For this reason, future research should focus not only on optimization algorithms but also on models for safe, responsible, and interpretable integration of artificial intelligence into the IoT environment.

In general, the analysis of scientific publications reveals a trend toward an interdisciplinary approach that combines computer science, telecommunications, applied mathematics, and cybersecurity. This approach enables the creation of not only efficient IoT data processing systems but also systems that ensure the protection of confidential information, which is a key factor in building user trust in emerging technologies.

Main part

IoT is a concept of a global network of interconnected physical objects that can collect, exchange, and process data via the Internet with no or minimal human intervention. IoT systems consist of many devices – sensors, actuators, controllers, computing modules – that interact with each other and with centralized or distributed platforms. A distinguishing feature of such systems is their ability to operate autonomously, responding to changes in environmental parameters, storing and transmitting the collected data in real time.

Unlike classical information systems, which operate within a clearly defined architecture, IoT is a dynamic, scalable environment that encompasses objects in the physical world – from household devices to infrastructure objects in industry and urban space. These objects can be mobile, distributed, limited in resources, yet still form a single digital ecosystem.

A significant advantage of IoT is the ability to obtain high-frequency and detailed data from the surrounding environment. This information is used for: monitoring the condition of systems (for example, energy networks or production equipment), optimizing resources (water supply, transport, lighting), automated decision-making (smart homes, Industry 4.0), and personalizing user interaction (healthcare, consumer services).

The integration of IoT with modern information technologies, particularly cloud computing and artificial

intelligence, has enabled a qualitatively new level of efficiency in data usage. There has emerged the possibility to implement smart systems that not only transmit information but also interpret it using machine learning algorithms, predict user behavior or environmental changes, and make optimization decisions.

In the modern information space, IoT performs a key function a link between the physical and digital worlds. This link is the foundation for building: smart cities that respond to traffic flow dynamics or environmental conditions; healthcare systems that allow continuous monitoring of patients' vital signs; logistics systems that enable real-time tracking of goods; energy systems capable of balancing loads in the grid depending on demand.

At the same time, the rapid increase in the number of connected devices (already measured in tens of billions) gives rise to a range of technical and ethical problems. Among the main ones is the security problem, since each IoT node is a potential entry point for an attacker, as well as the privacy problem, since a significant portion of the data is personalized or sensitive. These problems require new approaches to architecture, interaction protocols, data processing methods, and trust models in the IoT environment.

In conclusion, the Internet of Things is not only a technological platform for data collection, but also a strategic element of society's digital transformation. Its significance continues to grow in the context of automation, intelligence, and decentralization of control systems at all levels – from household to national.

The data collection process is a fundamental component of the functioning of any IoT system, as it is precisely through it that devices can detect changes in the surrounding environment, interact with each other, and transmit information for further analysis. The quality, volume, and structure of the collected data directly affect the effectiveness of decision-making in smart systems, especially in the context of using artificial intelligence for their interpretation. Figure 1 presents the existing data collection methods in IoT.

At the first level of data collection are sensors that can detect physical or chemical environmental parameters: temperature, humidity, pressure, motion, illumination, presence of gases, biometric indicators, etc. Sensor modules can be active – with their own power source; passive – activated by an external probing signal

Depending on the application, different types of sensors are used: temperature, motion, biometric, vibration, optical, and acoustic sensors. Data collection is provided by sensor nodes. Ready-made boards such as Raspberry Pi, Arduino, ESP32, STM32, NXP, and TI Launchpad are often used. These devices are capable not only of collecting but also of preprocessing data to reduce traffic volume.

The data collected by sensors must be transmitted to higher processing levels – either to edge devices or to the cloud. Various network protocols are used for this purpose, considering energy consumption, latency, reliability, and bandwidth limitations: MQTT, CoAP, HTTP/HTTPS, ZigBee, Z-Wave, LoRa, NB-IoT. To reduce the load on cloud services and the network, the

concept of edge computing – processing data as close as possible to the point of collection – is gaining increasing importance.

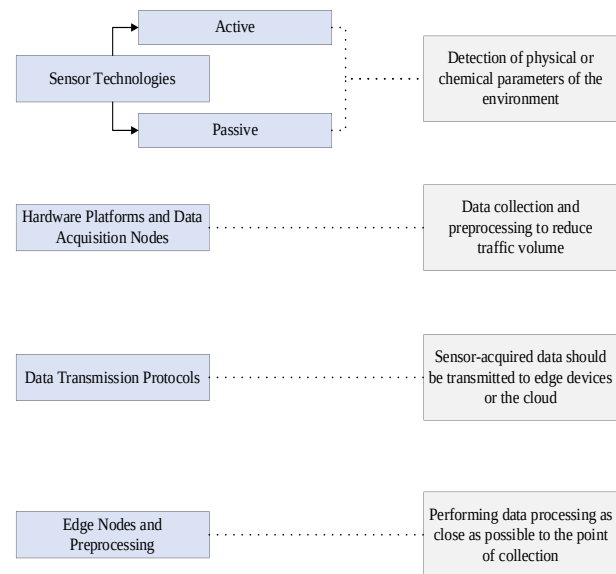


Fig. 1. Existing Data Collection Methods in IoT

Edge nodes filter out noise and data errors, perform basic analysis, and transmit only aggregated or critical data to higher levels. This helps reduce latency, decrease network load, and enhance security, as fewer incomplete or noisy data are sent to the cloud. Particular attention in modern IoT systems is paid to data protection at the collection level. Already at the sensor level, traffic encryption, device authentication, and data access control can be applied to prevent unauthorized devices from interfering with the communication channel. The data collected by IoT system sensors have no value by themselves until they are structured, cleaned, aggregated, and stored for further analysis. Since IoT data is received continuously, at high frequency, and from many sources, their preprocessing – bringing them into a consistent and analyzable format – becomes a key challenge. In this process, the data storage strategy also plays an important role: from local caching to cloud archiving. Fig. 2 presents the methods of data preprocessing and storage in IoT.

Raw data often contains noise, duplicates, incorrect or incomplete values. Therefore, at the stage of initial processing, the following methods are applied: noise filtering, missing data handling, outlier and anomaly detection, and format standardization. This stage is performed either directly on sensor nodes or on edge gateways that collect data streams from multiple devices.

To reduce the volume of information transmitted to the cloud, data aggregation is often used: calculation of averages, medians, and extreme values over a certain period; sensor data fusion; construction of synthetic indicators or indexes. Semantic data fusion from different sources is also possible – for example, combining air temperature with geolocation coordinates and video surveillance data. At the stage of preliminary intelligent processing, basic analysis algorithms can be executed: state classification, anomaly detection in the data

stream, and local storage of important events instead of the full data stream.

This reduces traffic volume, increases system autonomy, and minimizes the risks of confidential data loss.

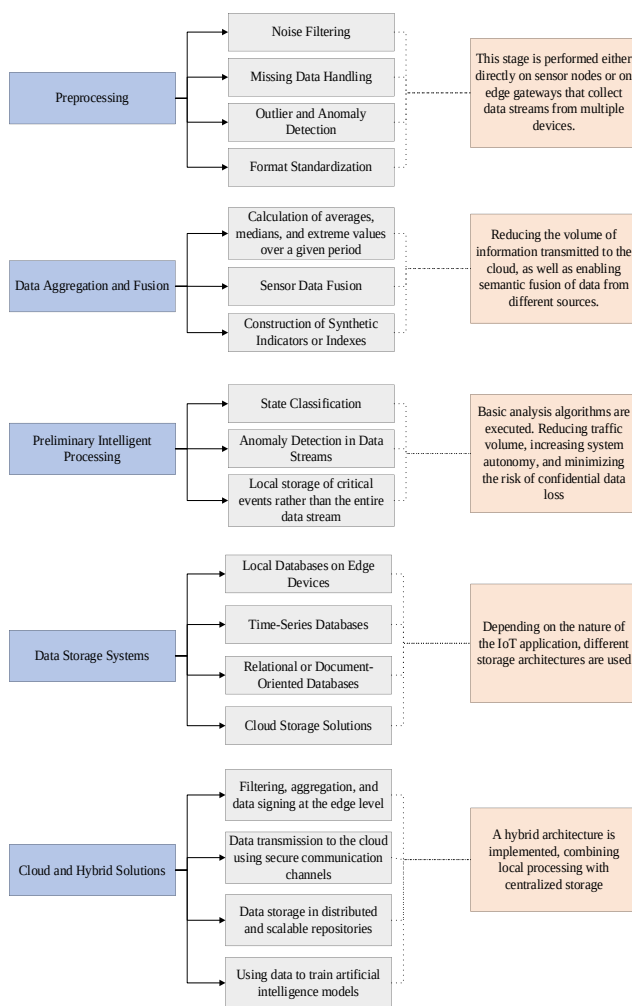


Fig. 2. Preprocessing and Data Storage Methods in IoT

Depending on the nature of the IoT application, various storage architectures are used: local databases on edge devices for caching or autonomous operation; time-series databases optimized for streaming data, with built-in time indexing mechanisms; relational or document-oriented databases for storing structured and semi-structured information; and cloud storage solutions for scalable long-term storage with backup and access to machine learning services.

Most modern systems implement hybrid architecture that combines local processing with centralized storage. In this model, data at the edge level are filtered, aggregated, and signed; transmitted to the cloud via secure channels; stored in distributed, scalable repositories; and used to train artificial intelligence models. A key requirement for data storage is protection – implementation of encryption mechanisms, access control, access logging, and compliance with GDPR standards or other local regulations.

After the stages of data collection, cleansing, and storage, the next key phase in IoT data processing is

analysis – a process that enables the extraction of useful information, recognition of patterns, detection of anomalies, and formulation of well-founded decisions. Due to the large volume, speed, and diversity of data generated by IoT devices, traditional analysis methods are often insufficient. That is why statistical analysis methods, machine learning, and deep learning are at the forefront.

At the first level of analysis, descriptive statistics are used: calculation of means, variances, medians, quartiles; correlation analysis between variables; generation of histograms, time series, and trend lines. These methods help identify general trends, seasonal fluctuations, and provide initial interpretation of parameter dynamics in the sensor environment.

When there is a need to identify categories or similar groups within a dataset, classification or clustering methods are applied. Since most IoT data is time-based, time-series analysis methods play an important role: ARIMA, SARIMA, Prophet; construction of seasonal models; detection of trends and cycles; signal smoothing and change point detection.

IoT systems integrated with artificial intelligence increasingly rely on automatic learning from data. This enables training behavioral models for automatic state response; anomaly detection; real-time control parameter optimization; and interpretation of complex relationships between input parameters and outcomes. Tools for this include TensorFlow, Scikit-learn, PyTorch, Edge AI SDK, and others.

In cases involving large and complex datasets, deep learning is applied. These models work effectively in real time, if edge computing or GPU acceleration is supported.

Data visualization is an integral part of analytics: using dashboards for interactive monitoring; trend charts, heatmaps, event maps; generating reports for real-time decision-making. Also important is the implementation of interpretable models – especially in critical domains where it is necessary to explain why the system made a particular decision.

During data collection, transmission, and analysis in IoT, several privacy-related threats arise: interception of unencrypted traffic, especially in wireless networks; unauthorized access to devices or processing servers; misuse of data; data leaks due to software vulnerabilities in IoT devices; man-in-the-middle attacks that allow data modification or duplication during transmission. These threats are particularly dangerous in domains such as healthcare, smart homes, and personal vehicles, where consequences can be both social and legal.

Many countries have legislative acts regulating the processing of personal data. In the context of IoT, this means that all devices and systems must have built-in mechanisms for access control, encryption, and user control over their own information.

Protecting personal data in the IoT environment is a complex task that combines legal, technical, and architectural approaches.

Given the growing autonomy of systems and their deep integration into users' personal lives, embedded data protection must be an essential part of any IoT solution.

The future of such systems lies in the combination of transparency, responsibility, and artificial intelligence capable of making ethical and well-founded decisions without threatening privacy.

In the process of modeling complex networks, the choice of an appropriate mathematical model is of critical importance. Each model has its own advantages, limitations, structural properties, and domain of effective application.

The analysis has shown that there is no universal model suitable for all types of networks. The selection of a model should be based on the core characteristics of the system under study – such as the nature of the degree distribution, the presence of a clustered structure, the dynamic nature of connections, and spatial or resource constraints. In complex applied problems, the best results are often achieved by hybrid or adaptive models that combine elements of several approaches.

Conclusions

As a result of the conducted research, a comprehensive analysis was carried out on modern approaches to data collection, processing, analysis, and protection in IoT systems, especially in the context of the growing role of artificial intelligence.

The technological foundations of IoT functionality were examined, key architectural components were identified, and their role was explored in the creation of digital ecosystems for monitoring, control, and decision-making across various domains – from household systems to critical infrastructure. Special attention was paid to data preprocessing methods that help reduce information load, improve the quality of analysis, and adapt data flows to the needs of intelligent algorithms. It was shown that the application of edge processing and local-level aggregation improves both the performance and security of systems. The main types of databases for IoT

were analyzed, including those optimized for time series, as well as tools for processing large volumes of data in cloud and hybrid environments.

Data collection methods in IoT are multilayered and closely linked to the requirements for energy efficiency, security, latency, and system scalability. The quality and reliability of the collected information form the foundation for further processing, analysis, and decision-making, making the selection of sensors, communication protocols, and architectural models of strategic importance for any IoT system.

Preprocessing and efficient data storage in IoT are critical stages that ensure the quality, security, and usability of information for further analysis. They determine not only the accuracy of analytics but also the stability, scalability, and compliance with regulatory requirements.

Therefore, there is a growing need to develop adaptive, intelligent data processing and storage systems capable of dynamically responding to changing device contexts and user requirements.

Data protection mechanisms were also studied, including encryption, differential privacy, anonymization, federated learning, and access-controlled architectures. The key emphasis was placed on the need to combine effective data processing with built-in mechanisms for protecting personal information, especially in the context of system scaling and AI implementation. It was determined that the successful implementation of secure IoT solutions requires an integrated approach that combines technical expertise, legal awareness, and ethical responsibility.

Thus, the results of this work can be used to develop practical recommendations for building secure, scalable, and intelligent IoT systems focused both on data processing efficiency and on adherence to digital privacy principles.

REFERENCES

1. J.Gubbi, R. Buyya, S. Marusic, M. Palaniswami. Internet of Things (IoT): A vision, architectural elements, and future directions. *Future Generation Computer Systems*. Vol.29, iss.7, Elsevier, 2013. P. 1645-1660. <https://doi.org/10.1016/j.future.2013.01.010>
2. Zanella, A., Bui, N., Castellani, A., Vangelista, L. and Zorzi, M. (2014) Internet of Things for Smart Cities. *IEEE Internet of Things Journal*, 1, 2014. P. 22-32. <https://doi.org/10.1109/JIOT.2014.2306328>
3. M. Alsheikh; S. Lin; D. Niyato; H. Tan. Machine Learning in Wireless Sensor Networks: Algorithms, Strategies, and Applications. *IEEE Communications Surveys & Tutorials*, Vol. 16, 2014. P. 1996 – 2018. <https://doi.org/10.1109/COMST.2014.2320099>
4. L. Tawalbeh, F. Muheidat, M. Tawalbeh, M. Quwaider. IoT Privacy and Security: Challenges and Solutions. *Applied Sciences*, Vol. 10(12), 2020. 17 p. <https://doi.org/10.3390/app10124102>
5. J.P. Singhal. A Survey on AI enabled IoT Applications. *International Journal of New Media Studies*, vol. 9, 2022. P. 42-46.
6. O. Aouedi, T. Vu, A. Sacco, D. Nguyen, K. Piamrat, G. Marchetto, Q. Pham. A Survey on Intelligent Internet of Things: Applications, Security, Privacy, and Future Directions. *IEEE Communications Surveys & Tutorials*, 2024. 56 p. <https://doi.org/10.1109/COMST.2024.3430368>

Received (Надійшла) 20.02.2025

Accepted for publication (Прийнята до друку) 23.04.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Нгок До Куєн – студент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Kuien Do – student, Department of Electronic Computers, Kharkiv National University of Radio Electronics Kharkiv, Ukraine;

e-mail: nhok.do1@nure.ua; ORCID Author ID: <https://orcid.org/0009-0009-7824-3233>.

Климова Ірина Миколаївна – асистентка кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Iryna Klymova – assistant of Department of Electronic Computers, Kharkiv National University of Radio Electronics Kharkiv, Ukraine;

e-mail: iryana.klymova@nure.ua; ORCID Author ID: <https://orcid.org/0000-0003-0455-6180>.

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57216950622>

Наумова Олена Валеріївна – студент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Elen Naumova – student, Department of Electronic Computers, Kharkiv National University of Radio Electronics Kharkiv, Ukraine;

e-mail: olena.naumova@nure.ua; ORCID Author ID: <https://orcid.org/0009-0009-3561-3687>.

Геревич Михайло Олександрович – доктор філософії, доцент кафедри теорії та історії держави і права, Ужгородський національний університет, Ужгород, Україна;

Mykhailo Herevych – PhD, Associate Professor of Department of Theory and History of State and Law, Uzhhorod, Ukraine;

e-mail: mykhailo.herevych@uzhnu.edu.ua; ORCID Author ID: <https://orcid.org/0000-0002-0842-2828>.

Янковський Олександр Аркадійович – кандидат технічних наук, доцент, доцент кафедри ЕОМ, Харківський національний університет радіоелектроніки, Харків, Україна;

Oleksandr Yankovskyi – PhD, Assistant Professor Assistant Professor of the Department of electronic computers, Kharkiv National University of radio electronics, Kharkiv, Ukraine;^[1]

e-mail: oleksandr.yankovskyi@nure.ua; ORCID Author ID: <https://orcid.org/0000-0002-1268-0029>.

Методи обробки та аналізу даних в IoT з використанням машинного навчання

Нгок До Куєн, І. М. Климова, О. В. Наумова, М. О. Геревич, О. А. Янковський

Анотація. Актуальність. Зростаюча інтеграція технологій Інтернету речей (IoT) у всі сфери людського життя - від інтелектуального побуту до інфраструктури «розумного міста» – супроводжується експоненціальним зростанням обсягу даних, що збираються, передаються та обробляються в реальному часі. У поєднанні з технологіями штучного інтелекту ці дані перетворюються на основу для прийняття автономних рішень, прогнозування поведінки користувачів, а також адаптації середовища до потреб конкретної особи. Однак саме в цьому контексті постає надзвичайно критичне питання захисту персональних даних. Багато IoT-пристроїв працюють у неконтрольованому середовищі, мають обмежені ресурси для криптографічного захисту, а також схильні до кібератак і несанкціонованого збору інформації. Натомість алгоритми штучного інтелекту, які використовуються для аналізу цих даних, нерідко демонструють проблему «чорної скриньки», коли неможливо повною мірою пояснити, як і чому було прийнято певне рішення на основі персоналізованих даних. Відсутність прозорості у поєднанні з широким доступом до чутливої інформації ставить під загрозу базові права людини на приватність. Актуальність теми зумовлена необхідністю пошуку балансованих технічних рішень, що дозволяють одночасно ефективно аналізувати великі масиви даних в IoT-середовищі й забезпечувати високий рівень їхньої безпеки. У зв'язку з цим, вивчення сучасних методів обробки, аналізу та захисту даних в системах IoT, адаптованих до вимог етичного штучного інтелекту та цифрових стандартів конфіденційності, є одним із ключових викликів сучасної цифрової науки. **Об'єкт дослідження:** процеси збору, обробки, аналізу та захисту даних у системах IoT, зокрема ті їхні компоненти, які пов'язані з використанням персональної інформації користувачів та її обробкою за допомогою методів штучного інтелекту. **Мета статті:** дослідження сучасних методів обробки та аналізу даних в системах IoT. Робота має на меті виявити найбільш ефективні підходи до безпечної обробки даних, охарактеризувати існуючі загрози конфіденційності та оцінити потенціал інтеграції захищених алгоритмів аналізу, які відповідають як технічним, так і етичним вимогам цифрового середовища. **Результати дослідження.** Здійснено комплексний аналіз сучасних підходів до збору, обробки, аналізу та захисту даних в системах IoT, особливо в умовах зростаючої ролі штучного інтелекту. Розглянуто технологічні основи функціонування IoT, визначено ключові архітектурні складові та досліджено їхню роль у створенні цифрових екосистем для моніторингу, управління та прийняття рішень у різних галузях – від побутових систем до критичної інфраструктури. Особливу увагу було приділено методам попередньої обробки даних, що дозволяють знизити інформаційне навантаження, підвищити якість аналізу та адаптувати потоки інформації до потреб інтелектуальних алгоритмів. Було показано, що застосування периферійної обробки та агрегації на локальних рівнях підвищує як продуктивність систем, так і їхню безпеку. Проаналізовано основні типи баз даних для IoT, зокрема оптимізовані для часових рядів, а також інструменти для обробки великих обсягів даних у хмарних та гібридних середовищах. **Висновки.** Методи збору даних в IoT є багатошаровими й тісно пов'язаними з вимогами до енергоефективності, безпеки, затримки та масштабованості систем. Якість і надійність зібраної інформації закладає основу для подальшої обробки, аналізу й прийняття рішень, тому вибір сенсорів, протоколів та архітектурної моделі має стратегічне значення для будь-якої IoT-системи. Попередня обробка та ефективне зберігання даних в IoT – критичні етапи, що забезпечують якість, безпеку та придатність інформації для подальшого аналізу. Вони визначають не лише точність аналітики, а й стабільність, масштабованість і відповідність нормативним вимогам. У зв'язку з цим постає потреба в розробці адаптивних, інтелектуальних систем обробки та зберігання, здатних динамічно реагувати на зміну контексту роботи пристроїв і користувацькі вимоги. Успішна реалізація безпечних IoT-рішень потребує інтегрованого підходу, який поєднує технічну компетенцію, правові знання та етичну відповідальність.

Ключові слова: IoT, штучний інтелект; обробка даних; аналіз даних; конфіденційність; захист персональних даних; edge computing; диференційна конфіденційність; машинне навчання; інтелектуальні системи.

О. Ю. Заковоротний, А. В. Хулап

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

АНАЛІЗ МОДЕЛЕЙ ВИЯВЛЕННЯ ВТОРГНЕНЬ НА ОСНОВІ НЕЙРОМЕРЕЖ У СИСТЕМАХ ІНТЕРНЕТУ РЕЧЕЙ

Анотація. Стрімкий ріст популярності систем інтернету речей робить проблему підвищення безпеки все більш значущою. Але через те, що більшість сенсорів та виконавчих пристроїв базуються на енергоефективних мікроконтролерах, використання популярних моделей виявлення вторгнень на основі нейромереж практично неможливе для таких пристроїв. **Мета** даної роботи полягає в аналізі існуючих систем виявлення вторгнень на основі мережевої інформації з точки зору використання в енергоефективних мікроконтролерах з обмеженими обчислювальними можливостями. **Отримані наступні результати:** переважна більшість досліджень пропонують масштабні моделі нейромереж, які неможливо використати у малих мікроконтролерах. Проводяться тільки дослідження ефективності, але не швидкодії або споживання ресурсів. Але у декількох роботах запропоновано зниження кількості параметрів мережевих протоколів, на основі яких проводиться розпізнавання, та проведено аналіз впливу цього зниження на ефективність розпізнавання. Також визначені найпопулярніші загальнодоступні набори даних, які можна використовувати для порівняння ефективності виявлення вторгнень. **Висновки:** моделі нейромереж, які використовують скорочений набір мережевих параметрів для розпізнавання, в поєднанні з оптимізацією обчислень нейромереж є перспективним напрямом подальших досліджень. При використанні загальнодоступних наборів даних можна проводити порівняння з рішеннями, які вже існують.

Ключові слова: нейромережа, розпізнавання втручань, мережевий трафік, набори даних, аналіз.

Вступ

Інтернет речей (IoT) швидко розвивається, охоплюючи різноманітні аспекти нашого життя. Завдяки технологіям 4G, 5G і покращенню стандартів зв'язку, IoT стає дедалі доступнішим. У 2020 році кількість IoT-пристроїв складала 15,1 мільярда, а до 2030 року прогнозується зростання до 29 мільярдів [1]. Основними напрямками розвитку є промислове застосування, медичні пристрої та розумні будинки. У 2023 році понад 57% організацій у Європі та Північній Америці вже впровадили IoT [2]. Очікується, що до 2025 року в середньому в домогосподарствах США буде 20 підключених пристроїв [3].

Розумні будинки включають пристрої, такі як системи безпеки, термостати та освітлення, якими можна керувати дистанційно або автоматично. У промисловості IoT використовується для моніторингу обладнання, покращення ефективності і оптимізації виробничих процесів. У медичній сфері IoT допомагає в моніторингу стану пацієнтів, управлінні хронічними захворюваннями та наданні своєчасної медичної допомоги.

Таким чином, інтернет речей (IoT) відкриває нові можливості для автоматизації та підвищення ефективності, однак також створює нові загрози для безпеки. У дослідженнях [4..9] наведені основні типи атак на IoT через мережу. Ці атаки можуть дозволити зловмисникам проникнути до внутрішньої мережі та отримати доступ до конфіденційної інформації або засобів управління пристроями. Також хакери можуть використовувати спеціально створені експлойти для несанкціонованого доступу, запуску шкідливих команд або повного порушення роботи системи.

Системи виявлення вторгнень (Intrusion Detection System, IDS) можна загалом поділити на два основні типи: виявлення вторгнень на основі хосту та мережеве виявлення вторгнень. Виявлення на основі хосту відстежує специфічні дії на хості, такі

як доступ до системних файлів або запуск програм. Мережеве виявлення вторгнень зосереджується на моніторингу потоку даних між комп'ютерами та виявленні аномального мережевого трафіку між різними комп'ютерами в мережі.

У мережі існують дві основні категорії поведінки: нормальна та аномальна. Нормальна поведінка відповідає певним стандартам щодо обсягу трафіку, використовуваних програм та типів даних, якими обмінюється мережа. Аномалії мережі можна поділити на дві основні групи: мережеві збої, такі як перевантаження мережі або поламака файлових серверів, а також атаки на безпеку мережі, зокрема DDoS-атаки та інші шкідливі дії, що здійснюються зловмисниками.

Машинне навчання (Machine Learning, ML) має важливу роль у сучасних системах виявлення мережевих вторгнень (IDS), оскільки дозволяє ефективно аналізувати великі обсяги мережевих даних і виявляти аномалії, які можуть бути ознаками кібератак. До традиційних належать дерева рішень, наївний баєсівський класифікатор та метод опорних векторів (SVM). Ці алгоритми дозволяють класифікувати трафік на нормальний і підозрілий, ґрунтуючись на аналізі характеристик даних. Наприклад, дерева рішень забезпечують високу точність і добре масштабуються на великі обсяги інформації. Глибоке навчання, зокрема конволюційні (CNN) та рекурентні нейронні мережі (RNN), демонструє високу ефективність у виявленні атак завдяки здатності виявляти складні просторові та часові залежності в даних. Це дозволяє моделювати поведінкові шаблони, характерні для різних типів вторгнень. Генеративні змагальні мережі (GAN) використовуються для вирішення проблеми дисбалансу даних, що є актуальним у випадках рідкісних атак. Штучно згенеровані зразки підвищують ефективність навчання IDS-моделей. Гібридні підходи поєднують переваги кількох методів, наприклад, CNN та RNN, що дозволяє враховувати як просторову, так і часову структуру трафіку. Крім того, застосування методів обробки природної мови (NLP) дає

зможу аналізувати текстові компоненти мережеских даних, такі як лог-файли чи заголовки повідомлень. Таким чином, використання машинного навчання у виявленні вторгнень дозволяє значно підвищити точність та адаптивність систем кіберзахисту в умовах постійно змінюваного середовища загроз.

Попри значні досягнення, використання машинного навчання в IDS стикається з певними викликами. Одним з основних є необхідність у великих обсягах даних для навчання моделей. Крім того, моделі машинного навчання можуть бути вразливими до атак, таких як отруєння даних, коли зловмисники вводять шкідливі дані для навчання моделі. Важливо також забезпечити інтерпретацію та візуалізацію результатів, щоб користувачі могли розуміти, чому модель приймає певні рішення.

Однак найбільша проблема полягає в необхідності великих ресурсів, таких як обчислювальна потужність і великий обсяг пам'яті, що є недоступними для більшості пристроїв Інтернету речей. У цій статті проводиться аналіз рішень IDS, які вже існують, їх відповідність обчислювальним можливостям IoT та можливості подальшої оптимізації цих IDS.

Постановка проблеми

Сучасні кіберзагрози стають дедалі складнішими, що ставить нові вимоги до систем виявлення мережеских вторгнень (NIDS). Для забезпечення ефективності цих систем необхідні актуальні й реалістичні набори даних мережевого трафіку, здатні відображати реальні умови. Основною проблемою є те, що існуючі стандартні набори даних не завжди відповідають вимогам реального світу; вони можуть не містити необхідних характеристик або містити зайві й непотрібні дані.

У статтях [10..13] розглядаються важливі аспекти збору даних для побудови систем виявлення мережеских вторгнень, зокрема для застосування методів машинного навчання. Автори підкреслюють необхідність створення стандартного набору важливих характеристик для мережевого трафіку, що дозволяє ефективно застосовувати машинне навчання у реальних умовах. Дослідження вказують на важливість мінімізації обсягу даних, що збираються, для досягнення високої ефективності в реальному часі. Встановлення оптимального набору характеристик і мінімального обсягу даних дозволяє значно знизити витрати на збір та обробку даних, забезпечуючи при цьому точність і ефективність моделей виявлення вторгнень.

Окрім основних проблем і підходів, пов'язаних із використанням машинного навчання для виявлення вторгнень, у дослідженнях цієї галузі акцентується увага на низці важливих аспектів, що визначають ефективність практичних рішень. Одним із таких аспектів є використання даних на основі потоків (flow-based), таких як NetFlow, IPFIX або sFlow. Поточкові дані дозволяють зменшити обсяг інформації для обробки та забезпечують ефективний аналіз у високошвидкісних мережах, що є критично важливим для виявлення вторгнень у реальному часі.

Значною проблемою є постійно змінюваний ландшафт кіберзагроз. Нові типи атак швидко з'являються, роблячи застарілими як існуючі набори даних,

так і моделі виявлення, що вимагає розробки адаптивних систем, здатних працювати з новими, ще невідомими загрозами. Також підкреслюється необхідність оптимізації процесу збору даних, адже надмірний обсяг навчальної інформації призводить до зростання витрат ресурсів. Завдання полягає у мінімізації обсягу вхідних даних без втрати ефективності побудованої моделі.

Одним із критичних етапів моделювання є вибір характеристик (feature selection). Коректне визначення інформативних ознак дозволяє зменшити складність моделі та підвищити її точність. При цьому надмірна кількість непотрібних або корельованих ознак може призвести до погіршення якості класифікації та зниження швидкодії. Окремо розглядається проблема недостатньої доступності маркованих даних. У більшості реальних мереж важко отримати чітко класифікований трафік, що значно ускладнює навчання алгоритмів у контрольованому режимі. У таких умовах зростає роль методів напівконтрольованого та безконтрольного навчання.

Важливим викликом є перенесення теоретичних рішень у реальні умови. Моделі, розроблені у лабораторних середовищах, не завжди ефективно працюють у продуктивних мережах, де додатковими обмеженнями є питання конфіденційності, точності збору та затримки в обробці даних. Крім того, визначення мінімально необхідного обсягу навчальних даних для побудови ефективної моделі дозволяє оптимізувати процеси збору та знизити ресурсоемність впровадження IDS у практичні мережі. Нарешті, важливою задачею є стандартизація наборів ознак для виявлення атак. Визначення загального підходу до формування ефективних характеристик дозволяє спростити інтеграцію рішень та забезпечити можливість використання технік перенесення навчання, що зменшує потребу у створенні окремих моделей для кожного нового випадку.

Ці теми підкреслюють важливі аспекти, пов'язані з практичним застосуванням машинного навчання у кібербезпеці, а також необхідність оптимізації процесів збору та обробки даних для досягнення ефективних результатів.

Основна частина роботи

У галузі виявлення мережеских вторгнень (NIDS) вирішальну роль відіграють відкриті набори даних, що використовуються для навчання та тестування різноманітних моделей та алгоритмів. Наявність якісних та різноманітних даних дозволяє ефективно перевіряти нові системи в реальних умовах. У цьому контексті порівняння найпоширеніших публічних наборів даних для NIDS є важливим етапом у виборі найбільш підходящих для подальших досліджень. У аналізі [14.. 16] розглядаються найбільш вживані набори даних, їхні особливості та недоліки, а також висновки щодо їхньої ефективності для порівняння нових систем виявлення вторгнень.

Серед усіх наборів слід відзначити наступні:

- KDD Cup 1999: Один з найвідоміших наборів даних для IDS, створений на основі DARPA 1998 року. Він містить 41 ознаку та понад 4.8 млн рядків даних, що описують атаки в чотирьох категоріях:

DoS, R2L, U2R та probing. Однак він має деякі недоліки, такі як дублювання між навчальним та тестовим набором, відсутність важливих ознак (наприклад, IP-адрес), застарілість даних, що було зібрано понад 20 років тому, і синтетичність даних.

- NSL-KDD 2009: Поліпшений набір даних для виправлення недоліків KDD Cup 1999, що зменшив кількість дублюючих записів, зберігаючи більшу частину важливих даних. Незважаючи на це, він все одно містить дані, зібрані ще в 1998 році, що обмежує його актуальність та застосування до сучасних атак.

- UNSW NB15 IDS: Зібраний за допомогою генератора трафіку IXIA, цей набір містить дані у форматах pcap, BRO, Argus і CSV. Він є оновленим набором, що включає нові типи атак. Однак дані згенеровані синтетично, що знижує їхню реалістичність порівняно з реальним трафіком.

- UGR'16: Зібраний з кількох колекторів netflow v9 в мережі іспанського інтернет-провайдера, цей набір має поділ на калібрувальні та навчальні набори. Більшість трафіку позначено як "фоновий", що може бути як аномальним, так і безпечним. Суміш синтетичних атак і реального трафіку також впливає на якість даних.

- CIDDs-001: Цей набір даних зібраний для оцінки аномалій в IDS та містить обмежену кількість атак (DoS, Brute Force, Port Scanning). Недостатня різноманітність атак обмежує можливості для оцінки ефективності систем у широкому спектрі атак.

- CICIDS'17: Набір даних включає різні типи атак, що відбувалися протягом п'яти днів, із зазначенням конкретних атак на кожен день. Трафік на основі реального використання, хоча й містить певні застарілі дані, але цей набір більш актуальний порівняно з попередніми.

- CSE-CIC-IDS2018: Набір включає 50 машин у мережі, де присутні атаки різних типів з використанням В-профілю для моделювання поведінки. Однак використання штучно створених профілів може знизити точність даних.

- LITNET-2020: Зібраний в Литві, цей набір містить реальні мережеві атаки в національній мережі. Оскільки дані зібрані в специфічному регіоні, вони можуть бути менш застосовними для досліджень в інших країнах або на глобальніших мережах.

- MAWILab: Набір даних містить мережевий трафік між Японією та США, охоплюючи кілька років і включаючи різноманітні аномалії. Однак він може бути менш актуальним для сучасних мереж, оскільки використовуються старі мережеві лінії.

Аналізуючи наведені набори даних, можна зробити висновок, що найбільш придатними для подальших досліджень є наступні набори. CICIDS'17 та CSE-CIC-IDS2018: Обидва набори містять більш актуальні дані про атаки, що відбуваються в сучасних мережах. Вони охоплюють широкий спектр атак, що дозволяє проводити різноманітні дослідження ефективності систем в реальних умовах. NSL-KDD 2009: Попри застарілість, цей набір все ж може бути корисним для порівняння нових методів з давно відомими базами, оскільки він надає структуровані й очищені дані. Для порівняння ефективності нових систем, що

використовують машинне навчання, необхідно враховувати не лише актуальність даних, а й специфічні характеристики трафіку в мережах. Зокрема, використання реальних даних (наприклад, у CICIDS'17 та LITNET-2020) буде значно кориснішим для тестування нових IDS на сучасних реальних атаках.

З огляду на наведене, для подальших досліджень найдоцільніше зосередитися на CICIDS'17, CSE-CIC-IDS2018 та NSL-KDD 2009, оскільки вони дозволяють порівняти різні методи на актуальних даних та проводити ефективне тестування нових підходів до виявлення вторгнень.

Попри високий потенціал нейронних мереж, особливо моделей глибинного навчання, їхнє застосування у системах виявлення вторгнень супроводжується низкою обмежень, які необхідно враховувати при розробці та впровадженні таких рішень.

Однією з головних проблем є значні обчислювальні вимоги. Глибокі моделі потребують високопродуктивного апаратного забезпечення, зокрема графічних (GPU) або тензорних процесорів (TPU), що робить їх навчання ресурсомістким та дорогим процесом [17]. Паралельно з цим, великі обсяги моделі призводять до підвищеного споживання пам'яті [18], що особливо критично в умовах обмежених ресурсів, як-от на периферійних пристроях або в реальному середовищі мережевого моніторингу.

Ще одним важливим обмеженням є тривалий час навчання. Моделі глибинного навчання вимагають значного часу для обробки великих обсягів даних, що ускладнює оперативне оновлення моделей при появі нових типів атак. Цей фактор також знижує здатність систем до швидкої адаптації. Важливою вимогою до мережевих IDS-систем є здатність працювати в режимі реального часу. Проте висока складність моделей може спричиняти затримки в обробці трафіку, що обмежує їхню ефективність у виявленні загроз у момент їх виникнення.

Окрему увагу слід приділити проблемі перенавчання. У випадку навчання на незбалансованих наборах даних, де частка атак є незначною, нейронні мережі можуть демонструвати хороші результати на навчальному наборі, але погано узагальнювати знання на нові або раніше невідомі загрози. Зі збільшенням масштабів мереж також ускладнюється завдання масштабованості та адаптації. У таких умовах підтримка високої точності потребує регулярного оновлення моделей або їхнього часткового перенавчання, що додатково збільшує навантаження на систему та її ресурси [19].

Системи виявлення мережевих вторгнень (NIDS), що базуються на нейронних мережах, надають розширені можливості для виявлення складних та еволюціонуючих загроз. Однак їх використання часто супроводжується численними проблемами, пов'язаними з використанням ресурсів, продуктивністю та ефективністю. Зокрема, високі обчислювальні вимоги, значне споживання пам'яті, тривалі часи навчання моделей та затримки при обробці даних у реальному часі ускладнюють їхнє впровадження. Крім того, проблеми перенавчання, складність у масштабуванні та потреба в регулярному перенавчанні моделей знижують загальну

ефективність систем. Ці фактори підкреслюють важливість оптимізації процесів збору та обробки даних для досягнення ефективних результатів у реальному часі, що залишається суттєвим викликом у використанні NIDS, заснованих на нейронних мережах. Виявлення вторгнень у комп'ютерних мережах є важливою складовою забезпечення кібербезпеки. Останні десятиліття для вирішення цієї проблеми активно використовуються алгоритми машинного навчання. Зокрема, методи, такі як опорні вектори (SVM), дерева рішень (DT), глибинне навчання демонструють високу ефективність у задачах класифікації та виявлення атак. Однак ефективність цих методів може значно знижуватися в умовах ворожого середовища, що створює нові виклики для розвитку систем виявлення вторгнень. Стаття [15] розглядає основні підходи до виявлення вторгнень, зокрема алгоритми машинного навчання та їх застосування в умовах реальних атак.

Для оцінки ефективності системи використовуються п'ять ключових метрик: точність, прецизійність, відгук, помилковий тривога та F-оцінка. Ці метрики дозволяють оцінити здатність системи виявляти атаки та розрізняти нормальний і шкідливий мережевий трафік. Застосування таких методів у контексті мережевої безпеки підтверджено у кількох роботах, таких як [15, 20, 21].

Точність (accuracy) вимірює загальну частку правильних класифікацій. Висока точність є бажаною, але вона може не відображати якості класифікації в умовах нерівномірних даних. Прецизійність (precision) показує частку правильно виявлених атак серед усіх виявлених атак. Ця метрика є критично важливою для мінімізації помилкових тривог. Відгук (recall) відображає частку правильно виявлених атак серед усіх реальних атак. Високий відгук свідчить про те, що система виявляє значну частину атак, але не враховує помилкові позитиви. Рівень помилкової тривоги (false alarm) вимірює кількість нормальних мережевих подій, які були помилково класифіковані як атаки. Низький рівень помилкових тривог є важливим для забезпечення точності системи. F-оцінка (f-score) комбінує прецизійність та відгук, дозволяючи знайти оптимальний баланс між ними. Вона є корисною при роботі з нерівномірними наборами даних.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}; \quad (1)$$

$$Precision = \frac{TP}{TP + FP}; \quad (2)$$

$$Recall = \frac{TP}{TP + FN}; \quad (3)$$

$$Falsealarm = \frac{FP}{FP + TN}; \quad (4)$$

$$F - score = \frac{2 * Precision * Recall}{Precision + Recall}, \quad (5)$$

де TP – кількість істинно позитивних результатів (true positives), TN – істинно негативних результатів (true negatives), FP – помилкових позитивів (false positives), FN – помилкових негативів (false negatives).

Використання цих п'яти метрик — точності, прецизійності, відгуку, помилкових тривог та F-оцінки — дає змогу детально оцінити ефективність системи виявлення мережевих вторгнень на основі глибоких згорткових нейронних мереж. Кожна метрика підкреслює різні аспекти роботи системи, забезпечуючи всебічну оцінку ефективності методу.

Далі буде розглянуто декілька наукових праць, які пропонують різні варіанти моделей нейромереж для виконання задач розпізнавання мережевих атак. Всі вони мають різні підходи та використовують загальнодоступні набори даних, які були розглянуті вище.

У дослідженні [22] було розроблено глибоку нейронну мережу яка проводила розпізнавання на основі 41 параметра, які були або напряду взяті з мережевих даних, або сформовані на основі цих даних. Точність розпізнавання склала більше 98%, окрім класу атак U2R (User to Root attack). Але це може бути спричинено відносно невеликою кількістю записів типу U2R у використаних наборах даних. Для експерименту використовувався ПК з процесором Intel Core i-7-5820k, 16 GB RAM та відеокартою Geforce GTX1060. Результати представлені у табл. 1 та 2.

Таблиця 1 – Ефективність на наборі KDD99

Тип вторгнення	Accuracy	Recall	F-score
PROBE	0.9818	0.9896	0.9909
DOS	0.999	0.9997	0.9998
U2R	0.8182	0.9091	0.8333
R2L	0.9706	0.9804	0.9756
NORMAL	0.9997	0.9995	0.9997

Таблиця 2 – Ефективність на наборі NSL-KDD

Тип вторгнення	Accuracy	Recall	F-score
PROBE	0.9773	0.9935	0.9927
DOS	0.9867	0.994	0.9959
U2R	0.8182	0.9091	0.6452
R2L	0.9694	0.9796	0.9412
NORMAL	0.9984	0.9935	0.9959

У статті [23] проведено три експерименти використовуючи набори даних CICIDS2017 та CTU. У першому експерименті були використані тільки просторові параметри наборів даних. У другому — тільки часові параметри. А у третьому — і просторові, і часові параметри. Результати експерименту приведені у табл. 3 та 4. Точність розпізнавання — більше 99% для усіх типів атак. Для експерименту був використаний ПК з Intel(R) Xeon(R) CPU E5-26206 32 GB RAM та відеоадаптером Geforce GTX1080ti 11GB.

Таблиця 3 – Ефективність на наборі CICIDS2017

Тип н. мережі	Accuracy	Precision	Recall	F-score
CNN2	0.9985	0.9976	0.9985	0.9980
LSTM2	0.9942	0.9944	0.9903	0.9923
CNN+ LSTM2	0.9991	0.9985	0.9993	0.9989
CNN11	0.9978	0.9994	0.9995	0.9995
LSTM11	0.9927	0.9972	0.9971	0.9972
CNN+ LSTM11	0.9981	0.9985	0.9998	0.9992

Таблиця 4 – Ефективність на наборі NSL-KDD

Тип н. мережі	Accuracy	Precision	Recall	F-score
CNN2	0.9982	0.9985	0.9989	0.9987
LSTM2	0.9942	0.9984	0.9932	0.9957
CNN+ LSTM2	0.9982	0.9991	0.9984	0.9988
CNN11	0.9701	0.9984	0.9921	0.9952
LSTM11	0.9639	0.9975	0.9916	0.9945
CNN+ LSTM11	0.9877	0.9984	0.9981	0.9983

У експерименті, описаному у [24], використовується набір даних NSL-KDD CUP. Він містить 2 піднабори: KDDTest+ та KDDTest-21. Результати експерименту приведені у табл. 5. Точність визначення — від 79 до 95% в залежності від моделі. Найнижчі результати були отримані для рекурсивної нейронної мережі (RNN). Для експерименту використовувалася ПК з наступною конфігурацією: Intel Core i7-7700, 16 GB RAM, Geforce GTX1050.

Таблиця 5 – Ефективність на наборі KDDTest

Набір даних	Модель	Accuracy	Recall	False alarm
KDD Test+	IBWNIDM	0.9536	0.9555	0.0076
	LeNet-5	0.8454	0.9131	0.0203
	DBN	0.9245	0.9568	0.0085
	RNN	0.9308	0.9439	0.0092
KDD Test ⁻²¹	IBWNIDM	0.9091	0.9242	0.0396
	LeNet-5	0.8236	0.8793	0.0324
	DBN	0.8869	0.8948	0.0446
	RNN	0.7947	0.8058	0.0175

Таблиця 8 – Ефективність на наборі CIC-IDS2017

Модель	К-ть шарів CNN	Accuracy	Precision	Recall	F-score	False alarm
Basic 1D-CNN	1	0.9799	0.9854	0.9975	0.9437	0.0237
	2	0.9808	0.9248	0.9651	0.9445	0.016
	3	0.9766	0.8838	0.992	0.9348	0.0265
Basic 2D-CNN	1	0.981	0.913	0.9809	0.9458	0.019
	2	0.9761	0.9192	0.9414	0.9301	0.0168
	3	0.9774	0.8864	0.9937	0.937	0.0259
Hybrid LTSM with 1D-CNN	2	0.9802	0.9038	0.9881	0.9441	0.0214
Hybrid LTSM with 2D-CNN	2	0.9755	0.8765	0.9954	0.9321	0.0285
Hybrid Autoencoder with CNN	2	0.9524	0.9076	0.7997	0.8502	0.0166

Таблиця 9 – Ефективність на наборі NSL-KDD

Модель	К-ть шарів CNN	Accuracy	Precision	Recall	F-score	False alarm
Basic 1D-CNN	1	0.8267	0.9262	0.7558	0.8323	0.0796
	2	0.849	0.9334	0.7913	0.8565	0.0747
	3	0.8887	0.9493	0.8498	0.8968	0.0599
Basic 2D-CNN	1	0.8522	0.9335	0.7971	0.8599	0.0751
	2	0.8559	0.9718	0.7692	0.8587	0.0295
	3	0.8443	0.9348	0.7809	0.8509	0.072
Hybrid LTSM with 1D-CNN	2	0.8993	0.9547	0.8642	0.9072	0.0542
Hybrid LTSM with 2D-CNN	2	0.8205	0.9302	0.7403	0.8244	0.0734
Hybrid Autoencoder with CNN	2	0.8439	0.9335	0.7813	0.8507	0.0735

І тому взагалі не підходять до інтернету речей, де потрібно розпізнавання у реальному часі при дуже обмежених ресурсах. Наприклад, обчислення нейромережі, запропонованої у [25], займало від 3.59 с до 4.17 с. Це дуже великий час для аналізу одного пакета. Через це було проведено дослідження [27] по

дослідження, яке описане у [25], пропонує декілька моделей нейромереж та використовує набори даних NSL-KDD та UNSW-NB15. На наборі даних NSL-KDD запропоновані моделі показали точність вище 99.5%. Але на наборі даних UNSW-NB15, точність розпізнавання була дещо нижчою — 93%, але, все одно, це дуже гарний результат. Результати цього дослідження наведені у таблицях 6 та 7.

Таблиця 6 – Ефективність на наборі NSL-KDD

Модель	Accuracy	Precision	Recall	F-score
MDNN	0.995	0.9953	0.995	0.9951
MCNN	0.995	0.995	0.995	0.995
MCNN-DFS	0.997	0.995	0.998	0.996
NDAE	0.8542	1.00	0.8542	0.8737

Таблиця 7 – Ефективність на наборі UNSW-NB15

Модель	Accuracy	Precision	Recall	F-score
BDNN	0.9321	0.94	0.93	0.93
BCNN	0.9442	0.95	0.94	0.94
BCNN-DFS	0.9571	0.96	0.96	0.96

Дослідники, які проводили огляд моделей розпізнавання, які базуються на згорткових нейронних мережах [26], також провели і свій експеримент, запропонувавши декілька своїх моделей розпізнавання. На наборах даних CIC-IDS2017 та NSL-KDD була досягнута точність розпізнавання у 97% та 82% відсотки відповідно. Результати експерименту наведені у табл. 8 та 9.

Всі моделі нейромереж, наведені вище, мають дуже високу точність розпізнавання, але, на жаль, вимагають значних обчислювальних ресурсів.

визначенню мінімально достатнього набору параметрів мережевих протоколів для забезпечення якісного розпізнавання вторгнень. Експерименти показали, що на всіх наборах даних є спільні найбільш інформативні ознаки. Кількість обраних ознак, які не призводять до зниження здатності моделі машинного

навчання до розпізнавання, коливається навколо 10 із 33 доступних ознак у наборах. Остаточний список ознак був складений на основі результатів відбору ознак і порівняний з результатами коефіцієнта кореляції між ознаками, після чого ознаки, які були сильно корельовані між собою, були видалені. Цей список приведений у табл. 10.

Таблиця 10 – Зменшений перелік мережевих параметрів

Назва параметру	Значення параметру
FLOW_DURATION_MILLISECONDS	Тривалість потоку даних в мілісекундах
TCP_WIN_MAX_IN	Максимальне TCP вікно (src-dst)
DURATION_OUT	Тривалість потоку від клієнта до сервера (мсек)
MAX_TTL	Максимальний TTL (Time To Live) потоку
L7_PROTO	Протокол 7-го рівня (числ. код)
SRC_TO_DST_AVG_THROUGHPUT	Байт в секунду від джерела до призначення
SHORTEST_FLOW_PKT	Найкоротший пакет потоку (байт)
MIN_IP_PKT_LEN	Довжина min IP-пакету потоку
TCP_WIN_MAX_OUT	Максимальне TCP вікно (dst-src)
OUT_BYTES	Кількість вихідних байтів

Висновки та перспективи подальших досліджень

Моделі нейронних мереж, що використовують повний набір параметрів мережевих протоколів, стикаються з суттєвими обмеженнями при впровадженні в інтернеті речей (IoT) через незначні обчислювальні ресурси та обсяги пам'яті кінцевих пристроїв. Такі

моделі вимагають значних обчислювальних потужностей для обробки даних у реальному часі, що ускладнює їх використання в умовах IoT, де час виконання є критичним. Крім того, зберігання великої кількості параметрів та проміжних результатів потребує великих обсягів оперативної пам'яті, що часто є неприпустимим для малих і середніх пристроїв з обмеженими ресурсами.

Натомість моделі нейронних мереж, що використовують скорочений набір ключових мережевих параметрів (10 параметрів), здатні значно зменшити вимоги до ресурсів при збереженні високої ефективності у завданнях класифікації та виявлення аномалій. Поеднання таких моделей із методами оптимізації обчислювальних процесів [28] може сприяти значному покращенню їх продуктивності та масштабованості в реальному часі. Подібні підходи відкривають перспективи для більш ефективного застосування нейронних мереж в умовах IoT, де обмеження з боку ресурсів є ключовим фактором.

Використання загальнодоступних наборів даних, таких як KDD Cup або NSL-KDD 2009, дозволяє проводити порівняння з існуючими методами та алгоритмами, що вже довели свою ефективність в області виявлення вторгнень та аномалій. Це дає змогу систематично оцінювати нові підходи, аналізувати їх точність та швидкодію, а також вдосконалювати методи машинного навчання для забезпечення більш стабільного та надійного захисту мереж в умовах обмежених ресурсів. Подальші дослідження у цьому напрямку є важливими для оптимізації існуючих моделей та їх адаптації до специфічних вимог IoT-систем.

СПИСОК ЛІТЕРАТУРИ

- Khanna, A., Kaur, S. (2020), "Internet of Things (IoT), Applications and Challenges: A Comprehensive Review", *Wireless Pers Commun* 114, 1687–1762, doi: <https://doi.org/10.1007/s11277-020-07446-4>
- Kumar, S., Tiwari, P. & Zymbler, M. (2019), "Internet of Things is a revolutionary approach for future technology enhancement: a review", *J Big Data* 6, 111, doi: <https://doi.org/10.1186/s40537-019-0268-2>
- Choudhary, A. (2024), "Internet of Things: a comprehensive overview, architectures, applications, simulation tools, challenges and future directions", *Discov Internet Things* 4, 31, doi: <https://doi.org/10.1007/s43926-024-00084-3>
- (2023), "IoT Security: 15 Types of Attacks with Real-World Examples", *Inovasense*, url: <https://www.inovasense.com/15-types-of-attacks-with-real-world-examples/>
- Chakraborty, S., Pandey, S.K., Maity, S. (2024), "Detection and Classification of Novel Attacks and Anomaly in IoT Network using Rule based Deep Learning Model", *SN COMPUT. SCI.* 5, 1056, doi: <https://doi.org/10.1007/s42979-024-03429-5>
- Bensaoud, A., Kalita, J. (2025), "Optimized detection of cyber-attacks on IoT networks via hybrid deep learning models", *Ad Hoc Networks* 170: 103770, doi: <https://doi.org/10.48550/arXiv.2502.11470>
- Almohaimeed, M.; Albalwy, F. (2024), "Enhancing IoT Network Security Using Feature Selection for Intrusion Detection Systems", *Appl. Sci.*, 14, 11966, doi: <https://doi.org/10.3390/app142411966>
- Newman, R.C. (2009), "Computer Security: Protecting Digital Resources", *Jones & Bartlett Learning*, url: <https://www.jblearning.com/catalog/productdetails/9780763759940?srsltid=AfmBOop8BLj-VJGp7IKaVkbLmnhpqhMjpdof-8hPzKwLAPy-fAaoVcuUQ>
- Thottan M., Chuanyi Ji. (2003), "Anomaly detection in IP networks", *IEEE Transactions on Signal Processing*, vol. 51, no. 8, pp. 2191-2204, doi: <https://doi.org/10.1109/TSP.2003.814797>
- Dong, H., Kotenko, I. (2025), "Cybersecurity in the AI era: analyzing the impact of machine learning on intrusion detection", *Knowl Inf Syst*, doi: <https://doi.org/10.1007/s10115-025-02366-w>
- Ajagbe, S.A., Awotunde, J.B. & Florez, H. (2024), "Intrusion Detection: A Comparison Study of Machine Learning Models Using Unbalanced Dataset", *SN COMPUT. SCI.* 5, 1028, doi: <https://doi.org/10.1007/s42979-024-03369-0>
- Hussein Ali, A., Charfeddine, M., Ammar, B. et al. (2024), "Unveiling machine learning strategies and considerations in intrusion detection systems: a comprehensive survey", *Frontiers in Computer Science*, doi: <https://doi.org/10.3389/fcomp.2024.1387354>
- Ring, M., Wunderlich, S., Grödl, D. et al. (2017), "Flow-based benchmark data sets for intrusion detection", *ECCWS 2017 16th European Conference on Cyber Warfare and Security. Academic Conferences and publishing limited*, url: https://books.google.com.ua/books?hl=uk&lr=&id=uFA8DwAAQBAJ&oi=fnd&pg=PA361&dq=intrusion+detection+based+on+network+data&ots=YUq_mFVv5m5C&sig=J12NbOfJq4fMxm7JXE-JWvFkmgM&redir_esc=y#v=onepage&q=intrusion%20detection%20based%20on%20network%20data&f=false

14. Mohammadpour, L., Ling, T. C., Liew, C. S., & Aryanfar, A. (2022), "A Survey of CNN-Based Network Intrusion Detection", *Applied Sciences*, 12(16), 8162, doi: <https://doi.org/10.3390/app12168162>
15. Zhen Yang, Xiaodong Liu, Tong Li, Di Wu, Jinjiang Wang, Yunwei Zhao, Han Han. (2022), "A systematic literature review of methods and datasets for anomaly-based network intrusion detection", *Computers & Security*, Volume 116, doi: <https://doi.org/10.1016/j.cose.2022.102675>
16. Chou D., Jiang, M. (2021), "A Survey on Data-driven Network Intrusion Detection", *ACM Computing Surveys (CSUR)*, Volume 54, Issue 9, doi: <https://doi.org/10.1145/3472753>
17. Abdi, M., Razzak, M. I., & Zaib, A. (2020), "Neural Networks for Network Intrusion Detection Systems: A Survey", *Computers, Materials & Continua*, 63(3), 1401-1420, doi: <https://doi.org/10.32604/cmc.2020.011216>
18. Zhang, Y., Wang, S., & Jin, X. (2019), "Deep Learning for Network Intrusion Detection Systems: A Survey." *Future Generation Computer Systems*, 101, 755-763, doi: <https://doi.org/10.1016/j.future.2019.07.046>
19. Siddiqui, M. K., Khan, M. A., & Zhang, L. (2020), "Challenges in the Real-Time Application of Neural Network-Based Intrusion Detection Systems.", *International Journal of Computer Science and Network Security*, 20(9), 56-63, doi: <https://doi.org/10.22937/IJCSNS.2020.20.9.56>
20. L. Nie, Z. Ning, X. Wang, X. Hu, J. Cheng and Y. Li (2020), "Data-Driven Intrusion Detection for Intelligent Internet of Vehicles: A Deep Convolutional Neural Network-Based Method", *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 4, pp. 2219-2230, 1 Oct.-Dec. 2020, doi: <https://doi.org/10.1109/TNSE.2020.2990984>
21. Muneer, S., Farooq, U., Athar, A., Ahsan Raza, M., Ghazal, T.M., & Sakib, S. (2024), "A Critical Review of Artificial Intelligence Based Approaches in Intrusion Detection: A Comprehensive Analysis", *Journal of Engineering*, doi: <https://doi.org/10.1155/2024/3909173>
22. Jia, Y., Wang, M. and Wang, Y. (2019), "Network intrusion detection algorithm based on deep neural network", *IET Inf. Secur.*, 2019, Vol. 13 Iss. 1, pp. 48-53, doi: <https://doi.org/10.1049/iet-ifs.2018.5258>
23. Zhang Y., Chen X., Jin L., Wang X., Guo D. (2019), "Network Intrusion Detection: Based on Deep Hierarchical Network and Original Flow Data", *IEEE Access*, vol. 7, pp. 37004-37016, doi: <https://doi.org/10.1109/ACCESS.2019.2905041>
24. Yang H., Wang F. (2019), "Wireless Network Intrusion Detection Based on Improved Convolutional Neural Network", *IEEE Access*, vol. 7, pp. 64366-64374, 2019, doi: <https://doi.org/10.1109/ACCESS.2019.2917299>
25. Al-Turaiki, I., Altwaijry, N. (2021), "A Convolutional Neural Network for Improved Anomaly-Based Network Intrusion Detection", *Big Data*, 2021, Vol. 9, No. 3, pp. 233-252. doi: <https://doi.org/10.1089/big.2020.0263>
26. Mohammadpour, L., Ling, T.C., Liew, C.S., Aryanfar, A. A (2022), "Survey of CNN-Based Network Intrusion Detection", *Appl. Sci.* 2022, 12, 8162, doi: <https://doi.org/10.3390/app12168162>
27. Komisarek, M., Pawlicki, M., Kozik, R., Hołubowicz, W., & Choraś, M. (2021), "How to Effectively Collect and Process Network Data for Intrusion Detection?", *Entropy*, 23(11), 1532, doi: <https://doi.org/10.3390/e23111532>
28. Заковоротний, О.Ю., Хулап, А.В. (2024), "Оптимізація обчислення нейромереж за допомогою використання цілочисельної арифметики", *Системи управління, навігації та зв'язку*, Том 2 № 76, doi: <https://doi.org/10.26906/SUNZ.2024.2.090>

Received (Надійшла) 20.02.2025

Accepted for publication (Прийнята до друку) 23.04.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Заковоротний Олександр Юрійович - доктор технічних наук, професор, завідувач кафедри комп'ютерної інженерії та програмування, Національний технічний університет "Харківський політехнічний інститут", Харків, Україна;
Oleksandr Zakovorotnyi - Doctor of Technical Sciences, Professor, Head of Department of Computer Engineering and Programming, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;
e-mail: oleksandr.zakovorotnyi@khpi.edu.ua; ORCID Author ID: <http://orcid.org/0000-0003-4415-838X>;
Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57201613700>.

Хулап Андрій Валерійович - аспірант кафедри комп'ютерної інженерії та програмування, Національний технічний університет "Харківський політехнічний інститут", Харків, Україна;

Andrii Khulap - PhD student, Department of Computer Engineering and Programming, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;
e-mail: Andrii.Khulap@cs.khpi.edu.ua; ORCID Author ID: <https://orcid.org/0000-0002-4103-7972>.

Optimization of neural network computation using integer arithmetic

Oleksandr Zakovorotnyi, Andrii Khulap

Abstract. The rapid growth in the popularity of Internet of Things (IoT) systems makes the issue of improving security increasingly important. However, since most sensors and actuators are based on energy-efficient microcontrollers, the use of popular neural network-based intrusion detection models is practically infeasible for such devices. **The aim of this work** is to analyze existing intrusion detection systems based on network information in terms of their applicability to energy-efficient microcontrollers with limited computational capabilities. **The following results were obtained:** the vast majority of studies propose large-scale neural network models that cannot be used on small microcontrollers. Most research focuses solely on detection accuracy, without considering performance or resource consumption. However, several studies propose reducing the number of network protocol parameters used for detection, and analyze the impact of this reduction on detection effectiveness. The most popular publicly available datasets suitable for evaluating and comparing intrusion detection efficiency were also identified. **Conclusions:** neural network models that use a reduced set of network parameters for recognition, combined with neural network computation optimization, represent a promising direction for future research. Publicly available datasets enable comparison with existing solutions.

Keywords: neural network, intrusion detection, network traffic, datasets, analysis.

І. В. Ільїна, Р. В. Артюх, М. М. Зимогляд

Харківський національний університет радіоелектроніки, Харків, Україна

АНАЛІЗ ДАНИХ ТА МАШИННЕ НАВЧАННЯ У ХМАРНИХ ТА ТУМАННИХ ПЛАТФОРМАХ ДЛЯ ЕФЕКТИВНОЇ ПЕРЕДАЧІ ДАНИХ

Анотація. У даній статті досліджується ефективність використання традиційних методів оптимізації передачі даних та запропонованої моделі на основі алгоритмів машинного навчання. В умовах стрімкого і постійного зростання обсягу інформації, розширення мережевої інфраструктури та зростаючих вимог до продуктивності систем з'являється необхідність адаптивного підходу до управління трафіком і зниженню енерговитрат. Особлива увага приділена інтеграції технологій машинного навчання з метою підвищення ефективності передачі даних у хмарних та туманних обчисленнях. Запропонована модель поєднує у собі методи прогнозування мережевого навантаження (LSTM, XGBoost), технології стиснення даних (VAE, K-Means) та програмно-визначені мережі (SDN) для балансування трафіку. У межах дослідження було проведено тестування комбінацій алгоритмів у розподілених середовищах із різними рівнями навантаження, що дало можливість оцінити ефективність кожного за критеріями продуктивності, надійності та енергоспоживання. Даний підхід забезпечує адаптивність до динамічних змін у трафіку і сприяє зменшенню затримок у передачі даних. Отримані результати досліджень підтверджують доцільність машинного навчання у застосуванні до оптимізації мережевої взаємодії і відкривають перспективи для подальших досліджень у сфері інтелектуального управління туманними обчисленнями. Запропоноване рішення може використовуватися у різних потребах, включаючи автономні транспортні системи, медичні інформаційні платформи, промисловий Інтернет речей та критично важливі інфраструктури.

Ключові слова: машинне навчання, туманні обчислення, хмарні обчислення, LSTM, XGBoost, VAE, K-Means, SDN, оптимізація навантаження, стиснення даних, затримка передачі, енергоефективність.

Вступ

Постановка проблеми. Сучасний етап розвитку інформаційних технологій характеризується стрімким зростанням обсягів даних, що генеруються різноманітними пристроями, системами та користувачами. Інтернет речей (IoT), розподілені обчислювальні платформи, автономні транспортні системи, медичні IoT-рішення та інші інноваційні технології вимагають миттєвої обробки та передачі інформації в умовах обмежених ресурсів. Це ставить нові виклики перед інфраструктурою хмарних і туманних обчислень, які покликані забезпечити ефективне функціонування розподілених систем. Однак традиційні підходи до управління мережевим трафіком, балансування навантаження та оптимізації енергоспоживання все частіше виявляються недостатніми для задоволення вимог сучасних додатків, де навіть мінімальні затримки або втрати даних можуть призвести до критичних наслідків.

Особливу актуальність набуває інтеграція методів машинного навчання (МН) та аналізу даних у хмарні й туманні середовища. Туманні обчислення, як проміжна ланка між периферійними пристроями та хмарою, дозволяють зменшити затримки та підвищити локалізовану обробку даних. Проте їхній потенціал залишається недостатньо розкритим у контексті динамічної адаптації до змін навантаження, оптимізації трафіку та енергоефективності. Існуючі рішення часто ґрунтуються на статичних алгоритмах, які не враховують нелінійні залежності та часові характеристики мережевих процесів. Це обмежує їхню здатність до масштабування в умовах експоненційного зростання кількості підключених пристроїв і різноманітності типів трафіку.

Метою даного дослідження є розробка адаптивної моделі оптимізації передачі даних у

хмарних і туманних платформах, яка поєднує передові методи машинного навчання, стиснення даних та програмно-визначених мереж (SDN).

Основна увага приділяється створенню гібридного підходу, що інтегрує прогнозування навантаження за допомогою LSTM (Long Short-Term Memory) та XGBoost (Extreme Gradient Boosting), стиснення даних на основі варіаційних автокодувальників (VAE) і кластеризації K-Means, а також динамічного балансування трафіку через SDN. Таке поєднання дозволяє не лише зменшити затримки та втрати пакетів, але й оптимізувати використання пропускну здатності мережі та енергоресурсів.

Важливим аспектом роботи є також аналіз взаємодії між різними компонентами моделі. Наприклад, комбінація LSTM і XGBoost дозволяє ефективно працювати як з часовими рядами, так і зі структурованими даними, тоді як інтеграція VAE та K-Means забезпечує гнучке стиснення без втрати критичної інформації. Використання SDN додає шаг динамічного управління, що дозволяє системі адаптуватися до змін у реальному часі. Актуальність роботи підкріплюється результатами експериментів, проведених у середовищах Google Cloud Platform (GCP), AWS Lambda та OpenStack.

Мета статті – представлення результатів дослідження, спрямованого на розробку та експериментального тестування адаптивної моделі оптимізації передачі даних у туманних обчисленнях, яка базується на методах машинного навчання, зокрема XGBoost, LSTM та Autoencoder.

Аналіз останніх досліджень та публікацій. Сучасні дослідження в галузі хмарних і туманних обчислень демонструють критичну роль інтеграції машинного навчання для оптимізації передачі даних.

Так, Bonomi та співавтори у статті [1] ввели концепцію туманних обчислень як проміжного шару між

хмарою та пристроями IoT. Це дозволяє суттєво зменшити затримки в розподілених системах. Ця робота стала фундаментом для подальших інновацій у сфері обробки даних у реальному часі.

Дослідження Shi та ін. у роботі [2] розширили цю концепцію, довівши ефективність туманних обчислень у поєднанні з edge computing для IoT-додатків.

У контексті машинного навчання ключову роль відіграють алгоритми прогнозування мережевого навантаження.

Наприклад, у статті [3] акцентує увагу на ефективності методів на кшталт XGBoost і LSTM, тоді як Chen та Guestrin у статті [4] емпірично довели переваги XGBoost у задачах класифікації та регресії завдяки унікальній архітектурі дерев рішень.

У роботі Hochreiter та Schmidhuber [5], які запропонували модель LSTM для аналізу часових рядів у розподілених системах, продемонстровано високу точність у прогнозуванні динамічних навантажень.

Важливим аспектом оптимізації трафіку є стиснення даних. Kingma та Welling у статті [6] запропонували теоретичні основи варіаційних автокодувальників (VAE), що відкрило можливості їх застосування в реальних мережах.

Для підвищення ефективності стиснення використовувався метод K-Means, ефективність підтверджується дослідженнями Jain [7] у контексті кластеризації даних.

Програмно-визначені мережі (SDN) були і залишаються ключовим інструментом для динамічного балансування навантаження. За дослідженням Feamster та ін. [8], архітектура SDN дозволяє нам адаптивно керувати трафіком шляхом централізованого контролю, що критично для систем із високими вимогами до часу відгуку.

Енергоефективність при впровадженні методів машинного навчання підтверджує робота [9], де інтелектуальний розподіл навантаження сприяє зменшенню енергоспоживання у хмарних середовищах.

Виклад основного матеріалу

Туманні та хмарні обчислення забезпечують ефективну передачу даних у розподілених середовищах, але вони створюють проблеми, що пов'язані із затримкою, балансуванням навантаження та оптимізацією пропускної здатності мережі. Наприклад, затримка передачі даних у медичних системах IoT може мати критичні наслідки для пацієнтів, а в автономному транспортуванні навіть невеликі затримки можуть спричинити надзвичайні ситуації.

Оскільки дедалі більше пристроїв генерують великі обсяги даних, потрібні передові методи керування мережею. Використовуючи машинне навчання, можливо передбачити навантаження на мережу та запобігати перевантаженням. Затримку передачі можливо зменшити за допомогою інтелектуального балансування трафіку, а використання пропускної здатності можливо оптимізувати за допомогою стиснення даних і кластеризації.

В роботі запропонована модель на основі машинного навчання, що включає в себе три основні складові (рис. 1).

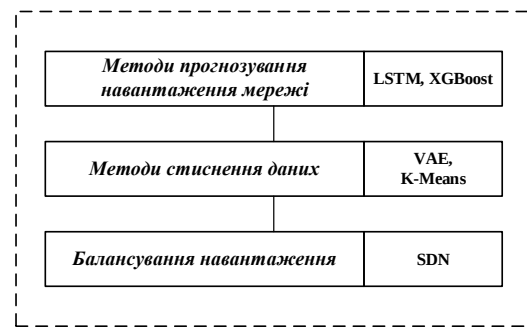


Рис. 1. Запропонована модель

Використовуючи синтетичні та реальні набори даних, експериментальні дослідження проводилися в середовищах Google Cloud Platform (GCP), AWS Lambda та OpenStack.

Прогнозування. Прогнозування навантаження виконувалось за допомогою LSTM і XGBoost для аналізу мережевого трафіку (рис. 2, 3).

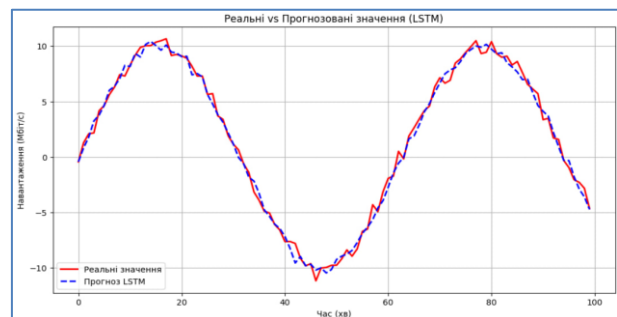


Рис. 2. Графік реальних та прогнозованих значень для LSTM

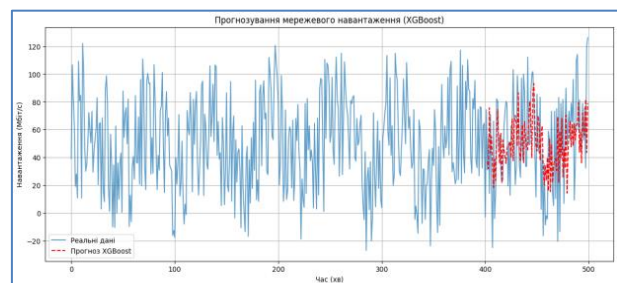


Рис. 3. Графік реальних та прогнозованих значень для XGBoost

ARIMA була обрана як базова модель для порівняння через її широке використання в аналізі часових рядів.

Модель ARIMA (AutoRegressive Integrated Moving Average) є класичним статистичним методом для аналізу часових рядів і широко використовується в задачах прогнозування навантаження. Її було включено до порівняння, що дозволило оцінити переваги сучасних методів машинного навчання (LSTM, XGBoost) над традиційними підходами. Наприклад, ARIMA не враховує нелінійні залежності та потребує ручного налаштування параметрів, що обмежує її ефективність у динамічних середовищах.

LSTM була обрана через здатність аналізувати довгострокові залежності в часових рядах, що є критичним для прогнозування мережевого трафіку.

XGBoost було використано через його швидкість, стійкість до шуму та можливість роботи зі структурованими даними.

Обидва методи дозволяють автоматично адаптуватися до змін у навантаженні, на відміну від ARIMA. Використання LSTM, XGBoost пояснюється їх високою ефективністю в задачах аналізу часових рядів і прогнозування динамічних змін мережі. Під час реалізації моделі використовувалися оптимізовані гіперпараметри:

- для LSTM – двошаровий підхід із розміром комірки 50 та використанням функції активації ReLU;

- для XGBoost – кількість дерев рішень 100, глибина 5, швидкість навчання 0,1.

Основні результати підтвердили ефективність інтеграції машинного навчання точності прогнозування навантаження мережі для підвищення продуктивності туманних обчислень (рис. 4).

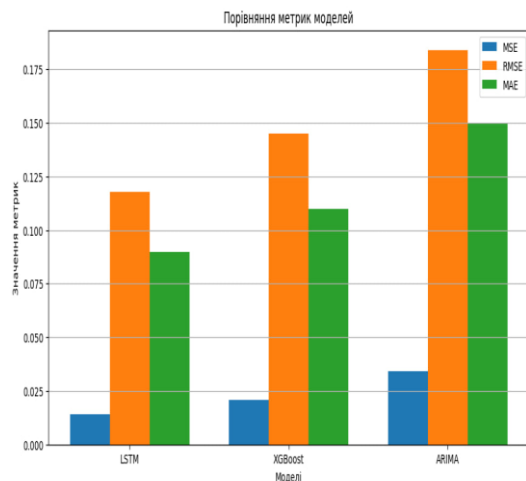


Рис. 4. Порівняння метрик (MSE, RMSE, MAE) між методами при прогнозуванні навантаження

Метрики, які були використані для оцінки методів прогнозування:

1. Середньоквадратична похибка (MSE) [10] для прогнозування навантаження розрахована:

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2, \quad (1)$$

2. Корінь середньоквадратичної похибки (RMSE) розраховувався за формулою:

$$RMSE = \sqrt{MSE}, \quad (2)$$

3. Середня абсолютна похибка (MAE) розраховувалась таким чином:

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|, \quad (3)$$

де y_i – реальне значення, $\hat{y}_i$ – прогнозоване, N – кількість спостережень.

Дослідження показали, що в порівнянні зі ARIMA методи LSTM та XGBoost мають менші значення метрик, що доводить їх здатність точніше прогнозувати навантаження мережі. Наприклад, MSE

для LSTM (0.014) значно нижчий, ніж у ARIMA (0.034), що підтверджує вищу точність сучасних методів. А LSTM показала найвищу точність завдяки здатності аналізувати довгострокові залежності (табл. 1).

Таблиця 1 – Результати порівняльного аналізу методів за метриками MSE, RMSE, MAE

Метод	MSE	RMSE	MAE
LSTM	0.014	0.118	0.090
XGBoost	0.021	0.145	0.110
ARIMA	0.034	0.184	0.150

Стиснення. Стиснення даних виконувалося за допомогою варіаційного автокодувальника (VAE), що зменшує обсяг переданих даних. Архітектура VAE включала прихований рівень із 64 елементами та простором затримки розміром 32.

Застосування методу K-Means для стиснення дозволяє скоротити надмірний обсяг даних шляхом залишення кількох найбільш характерних представників від кожного кластера.

Поеднанням стиснення на основі VAE та K-Means, вдалося зменшити переданий трафік значно знизивши навантаження на мережеву інфраструктуру, що свідчить про синергію між кластеризацією та глибоким навчанням (табл. 2).

Оптимізація пропускну здатності за рахунок зменшення розміру даних розрахована:

$$R = \left(\frac{\text{Original Size} - \text{Compressed Size}}{\text{Original Size}} \right) \times 100\%, \quad (4)$$

де *Original Size* – обсяг даних до стиснення; *Compressed Size* – обсяг даних після стиснення.

Таблиця 2 – Результати порівняння обсягу даних, переданих у різних сценаріях

Метод стиснення	Обсяг переданих даних (%)	Скорочення трафіку (%)
Без стиснення	100	0
K-Means	72	28
VAE	55	45
K-Means+VAE	43	57

Балансування. Балансування навантаження виконувалося за допомогою програмно-визначеної мережі (SDN), яка дозволяє динамічно перерозподіляти потоки даних залежно від поточного навантаження на вузли мережі. Балансування маршрутизації в SDN базується на алгоритмі Shortest Path First з додатковим зважуванням для затримок у вузлі.

Було виявлено, що балансування навантаження через SDN зможе оптимізувати пересилання запитів

у реальному часі, і це вирішальне значення для роботи інтелектуальних систем IoT.

Результати тестування підтвердили, що застосування запропонованих підходів дозволяє ефективно керувати навантаженням у туманних і хмарних обчисленнях, зменшує споживання ресурсів і покращує загальну продуктивність системи.

У порівнянні з іншими підходами, такими як традиційні методи балансування за фіксованими правилами, дослідження показало, що запропонований підхід може покращувати розподіл трафіку та зменшувати середню затримку.

Середня затримка в мережі розраховувалась за параметрами [11]:

$$Latency\ avg = \frac{1}{N} \sum_{i=1}^N (T_{Response} - T_{Request}), \quad (5)$$

де $T_{Response}$ - час отримання відповіді; $T_{Request}$ - час відправлення запиту; N - кількість спостережень.

Для комплексної оцінки ефективності запропонованої моделі було проведено порівняльний аналіз зі звичайними методами оптимізації. Традиційні методи включають ARIMA для прогнозування, статичне балансування та базове стиснення даних. Енергоспоживання розраховано на основі експериментів у середовищі OpenStack з використанням стандартних інструментів моніторингу.

Результати порівняння представлені в табл. 3.

Таблиця 3 – Результати порівняльного аналізу звичайних методів оптимізації трафіку за допомогою машинного навчання

Метрика	Традиційні методи	Запропонована модель
Точність прогнозування навантаження (%)	85,2	95,6
Скорочення обсягу переданих даних (%)	22,1	57
Зменшення середньої затримки (%)	18,4	48,6
Енергоспоживання (%)	100	60

Точність прогнозування навантаження досягла 95,6% за допомогою методу LSTM.

Запропонована модель включає в себе кластеризацію даних за допомогою K-Means і використання варіаційного автокодувальника (VAE) для стиснення трафіку. Ефективність цих підходів оцінювалась шляхом порівняння обсягу даних, переданих у різних сценаріях. За результатами проведених експериментів, комбінований підхід K-Means і VAE зменшив обсяг переданих даних на 57%, значно підвищивши ефективність використання пропускну здатності мережі.

Виявлено, що середня затримка була зменшена на 48,6% за допомогою адаптивного балансування.

Отримані результати підтверджують, і дають зрозуміти, що інтеграція методів машинного навчання може значно підвищити продуктивність туманних і

хмарних платформ. Зокрема, основними перевагами запропонованого підходу є скорочення затримки на 48,6% і зниження споживання енергії на 40%.

Це відкриває нові можливості для впровадження в системах реального часу, таких як медичні моніторингові платформи, автономний транспорт та промислові IoT-мережі, де ефективність передачі даних безпосередньо впливає на безпеку та продуктивність.

Висновки

Сучасні технології, такі як туманні та хмарні обчислення, відкривають нові можливості для ефективної передачі даних, але водночас ставлять складні виклики, пов'язані з затримками, енергоефективністю та обробкою великих обсягів інформації.

Це дослідження показало, що поєднання машинного навчання з інноваційними методами оптимізації може суттєво покращити роботу таких систем. Наприклад, модель LSTM, яка аналізує часові ряди, дозволила передбачати навантаження з рекордною точністю, а XGBoost прискорила обробку даних, навіть коли частина інформації була відсутня.

Окремо варто відзначити ефективність стиснення даних: завдяки поєднанню кластеризації K-Means та глибокого навчання (VAE) вдалося значно скоротити обсяг переданої інформації. Це не лише зменшило навантаження на мережу, але й заощадило енергію, що критично важливо для пристроїв IoT з обмеженими ресурсами. Крім того, програмно-визначені мережі (SDN) допомогли динамічно розподіляти трафік, знизивши середню затримку майже на половину. Такі результати особливо актуальні для медичних систем, де кожна мілісекунда може мати вирішальне значення, або для автономних автомобілів, де стабільність передачі даних безпеко-критична.

Важливим аспектом роботи стала також енергоефективність. Запропоновані рішення дозволили знизити споживання енергії на 40%, що робить їх привабливими не лише з технічної, а й з екологічної точки зору.

Дослідження показало, що даний підхід не лише теоретично обґрунтований, але може бути впроваджений у реальні системи передачі даних, де швидкість, точність та енергоощадність стають запорукою технологічного прориву.

У статті підкреслюється важливість поєднання нових рішень із існуючими, щоб зменшити навантаження на канали зв'язку та адаптувати їх до зростаючих вимог.

Запропонована модель адаптивної оптимізації для передачі даних покращує точність прогнозування навантаження, зменшує затримку передачі даних, оптимізує пропускну здатність та зменшує енергоспоживання у поєднанні традиційних методів з використанням інтелектуального розподілу навантаження між обчислювальними ресурсами. Також, можна зробити висновки, що цей підхід може бути застосований у розподілених системах у режимі реального часу, що дозволить покращити ефективність їх використання.

Подальші дослідження можуть включати впровадження квантових алгоритмів для прискорення обчислень, гібридних архітектур нейронних мереж та

розширення методів стиснення для різних типів даних. Умови, створені стрімким розвитком технологій, вимагають постійного вдосконалення таких систем, і проведені дослідження це підтверджують.

СПИСОК ЛІТЕРАТУРИ

1. Bonomi, F. Fog computing and its role in the internet of things [Text] / F. Bonomi, R. Milito, J. Zhu, S. Addepalli // Proceedings of the MCC Workshop on Mobile Cloud Computing. – 2012. – P. 13–16. DOI:10.1145/2342509.2342513.
2. Shi, W. Edge computing: Vision and challenges [Text] / W. Shi, J. Cao, Q. Zhang, Y. Li, L. Xu // IEEE Internet of Things Journal. – 2016. – Vol. 3, No. 5. – P. 637–646. DOI:10.1109/JIOT.2016.2579198.
3. Stock-Price Forecasting Based on XGBoost and LSTM / P. Hoang Vuong et al. Computer Systems Science and Engineering. 2022. Vol. 40, no. 1. P. 237–246. DOI.org/10.32604/csse.2022.017685
4. Chen, T. XGBoost: A scalable tree boosting system [Text] / T. Chen, C. Guestrin // Proceedings of the 22nd ACM SIGKDD International Conference. – 2016. – P. 785–794. DOI:10.1145/2939672.2939785.
5. Hochreiter, S. Long short-term memory [Text] / S. Hochreiter, J. Schmidhuber // Neural Computation. – 1997. – Vol. 9, No. 8. – P. 1735–1780. DOI:10.1162/neco.1997.9.8.1735.
6. Kingma, D. P. Auto-encoding variational Bayes [Text] / D. P. Kingma, M. Welling // arXiv preprint. – 2013. – arXiv:1312.6114.
7. Jain, A. K. Data clustering: 50 years beyond K-means [Text] / A. K. Jain // Pattern Recognition Letters. – 2010. – Vol. 31, No. 8. – P. 651–666. DOI:10.1016/j.patrec.2009.09.011.
8. Feamster, N. The Road to SDN: An Intellectual History of Programmable Networks [Text] / N. Feamster, J. Rexford, E. Zegura // ACM SIGCOMM Computer Communication Review. – 2014. – Vol. 44, No. 2. – P. 87–98. DOI:10.1145/2602204.2602219.
9. Volk M.O. Models optimization of resources in cloud computing: a hybrid approach to automation of operations and energy saving / M. O. Volk et al. Scientific notes of Taurida National V.I. Vernadsky University. Series: Technical Sciences. 2024. Vol. 1, no. 5. P. 91–96. URL: <https://doi.org/10.32782/2663-5941/2024.5.1/15>.
10. Belas A. O., Bidyuk P. I. CHOOSING A QUALITY CRITERION FOR EVALUATING THE FORECAST OF NONLINEAR NON-STATIONARY PROCESSES. KPI Science News. 2021. No. 2. URL: <https://doi.org/10.20535/kpissn.2021.2.236936>
11. Kreutz, D. Software-Defined Networking: A Comprehensive Survey [Text] / D. Kreutz, F. M. V. Ramos, P. E. Verissimo // Proceedings of the IEEE. – 2015. – Vol. 103, No. 1. – P. 14–76. DOI:10.1109/JPROC.2014.2371999.

Received (Надійшла) 25.03.2025

Accepted for publication (Прийнята до друку) 14.05.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Ільїна Ірина Віталіївна – кандидат технічних наук, доцент, доцент кафедри Електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Iryna Ilina – candidate of technical sciences, Associate Professor, Associate Professor of the Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: iryna.ilina@nure.ua; ORCID Author ID: <https://orcid.org/0000-0003-3132-7949>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57202223262>.

Артюх Роман Володимирович – кандидат технічних наук, доцент кафедри комп'ютерно-інтегрованих технологій, автоматизації та робототехніки, Харківський національний університет радіоелектроніки, Харків, Україна;

Roman Artiukh – candidate of technical sciences, Associate Professor of the Department of Computer-Integrated Technologies, Automation and Robotics, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: roman.artiukh77@gmail.com; ORCID Author ID: <http://orcid.org/0000-0002-5129-2221>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57208344824>.

Зимогляд Микола Миколайович – студент кафедри Електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Mykola Zymohliad – student, Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: nickzymoglyad@gmail.com; ORCID Author ID: <https://orcid.org/0009-0001-7136-1968>;

Data analysis and machine learning in cloud and fog platforms for efficient data transmission

Iryna Ilina, Roman Artiukh, Mykola Zymohliad

Abstract. This article investigates the effectiveness of traditional methods for optimizing data transmission and the proposed model based on machine learning algorithms. In the context of rapid and constant growth in the volume of information, expansion of network infrastructure and increasing requirements for system performance, there is a need for an adaptive approach to traffic management and reducing energy consumption. Special attention is paid to the integration of machine learning technologies in order to increase the efficiency of data transmission in cloud and fog computing. The proposed model combines network load prediction methods (LSTM, XGBoost), data compression technologies (VAE, K-Means) and software-defined networks (SDN) for traffic balancing. As part of the study, combinations of algorithms were tested in distributed environments with different load levels, which made it possible to evaluate the effectiveness of each according to the criteria of performance, reliability and energy consumption. This approach provides adaptability to dynamic changes in traffic and helps reduce data transmission delays. The obtained research results confirm the feasibility of machine learning in the application to the optimization of network interaction and open up prospects for further research in the field of intelligent control of fog computing. The proposed solution can be used for various needs, including autonomous transportation systems, medical information platforms, industrial Internet of Things and critical infrastructures.

Keywords: machine learning, fog computing, cloud computing, LSTM, Extreme Gradient Boosting) Variational Autoencoder, K-Means, SDN, load optimization, data compression, transmission latency, energy efficiency.

А. А. Коваленко, В. О. Чхеїдзе, О. М. Севостьянова, О. О. Фомічов

Харківський національний університет радіоелектроніки, Харків, Україна

ПІДВИЩЕННЯ ТОЧНОСТІ АНАЛІЗУ ТА ОБРОБКИ СКЛАДНОСТРУКТУРНИХ ЗОБРАЖЕНЬ

Анотація. Актуальність. Складноструктурні зображення (ССЗ) тяжіють до консолідації графіків різного типу та гетерархічності семантичних планів, що призводить до того, що необхідно використовувати гібридні та ітеративні алгоритми, специфічні для такого типу зображень. **Об'єктом** дослідження є процес аналізу та обробки растрових ССЗ. **Предметом** дослідження є моделі та підвищення точності та продуктивності алгоритмів аналізу та обробки ССЗ. **Мета** дослідження полягає у підвищенні точності та продуктивності алгоритмів аналізу та обробки ССЗ. **Результати дослідження.** У ході роботи було проведено аналіз сучасного стану аналізу та обробки растрових складноструктурних зображень. Проведені дослідження підходів до локалізації та визначення типів сегментованих об'єктів. Проведені дослідження підходів до розпізнавання образів та їх угруповання. Удосконалений та досліджений метод підвищення точності аналізу та обробки ССЗ. **Висновок.** Запропонований підхід дозволив підвищити точність аналізу та обробки складноструктурних зображень за рахунок сегментації, локалізації, розпізнавання та групування точкових, лінійних і площинних об'єктів, що засновані на комплексному використанні відомих методів.

Ключові слова: складноструктурне зображення, точковий об'єкт, площинний об'єкт, сегментація.

Вступ

В даний час відбувається прискорений розвиток засобів обробки та аналізу інформації внаслідок накопичення великої кількості даних про різні об'єкти і процеси. Одними з важливих складових таких засобів є програмні комплекси аналізу та обробки растрових складноструктурних зображень, що поєднують у собі числові та графічні дані. Метою таких програмних комплексів є отримання інформації про просторово-розподілені об'єкти для подальшої обробки. Основними підходами є векторизація та розуміння семантичних зв'язків між об'єктами.

Прикладами (ССЗ) є діаграми, номограми, креслення, схеми, архітектурні проекти, карти місцевостей. У цілому нині дані зображення можна охарактеризувати наявністю багатьох семантичних верств, накладених друг на друга у візуальному плані. Наприклад, топографічні карти, які можна назвати типовим представником ССЗ, містять найбільш повну інформацію про місцевість, тому з них можна отримати інформацію для геоінформаційної системи (ГІС), які можна використовувати як підтримку прийняття рішення на етапі проектування для вирішення різних завдань, наприклад, оптимального розміщення станцій мобільного зв'язку, логістичних центрів, транспортних розв'язок і транспортних розв'язок. Також ГІС є корисним інструментом аналізу для здійснення, наприклад, кадастрових робіт, передбачення потенційних зон пожеж та затоплень. Згідно зі звітом «Geographic Information System (GIS) Market», опублікованому Allied Market Research, загальносвітовий ринок ГІС зростає до 25,5 млрд. дол.

При цьому ручна обробка великої кількості ССЗ стає економічно неефективною. Проблема розпізнавання та аналізу образів набула визначного значення в умовах інформаційних переваг, коли людина не справляється з розумінням вступників до нього все зростаючих обсягів даних. В результаті його мозок переключається на режим одночасного сприйняття і мислення, якому таке розпізнавання

властиво, але негативно впливає на якість роботи в умовах небажаності помилок. Тому сучасний період розвитку інформаційних технологій характеризується широким впровадженням інтелектуальних методів у процес обробки графічної інформації, а автоматична обробка та аналіз зображень, що володіють складною структурою, є одним з найбільш перспективних напрямів розвитку.

Метою дослідження є розробка підходів до підвищення точності та продуктивності алгоритмів аналізу та обробки складноструктурних зображень..

1. Аналіз та обробка складноструктурних зображень

Інформацію у вигляді цифрового зображення можна подати у вигляді такого відображення

$$f: E \rightarrow P, \quad (1)$$

де E – деякий опис зовнішнього світу; P – цифрове зображення, задане у вигляді матриці.

Таким чином, основною метою аналізу та обробки ССЗ є розробка зворотного відображення

$$f^{-1}: P \rightarrow E_1, \quad E_1 \subseteq E, \quad (2)$$

де E_1 – необхідна частина опису зовнішнього світу, наприклад, властивості та розташування певних об'єктів на зображенні.

Вхідне відображення f не є оборотним у загальному випадку, оскільки, по-перше, є функцією згортки, що перетворює вхідні дані довільного розміру на матрицю встановленого. Тим не менш, для подальшої роботи (наприклад, прийняття рішень) отримання повної та точної інформації про зовнішній світ часто не потрібно: достатньо знати її з певною похибкою. По-друге, правила відображення f можуть мати як явний, так і неявний характер. Під явними правилами мається на увазі однозначне відображення об'єктів зовнішнього світу та їх властивостей візуального плану, наприклад, зображення будинку на карті у вигляді умовної піктограми або спрощеного кадастрового плану. Під

неявними правилами маємо на увазі як багатозначні відносини, коли елементи зображення не мають конкретної прив'язки в просторі растру, або виражені у вигляді неявних функцій, що зв'язують частини

$$P : R(p_1, \dots, p_k) = 0, \quad (3)$$

де p_i – елемент зображення (ЕЗ), або його частини;

$$E : R(e_1, \dots, e_k) = 0, \quad (4)$$

де e_i – об'єкт або система об'єктів зовнішнього світу, або складного відношення $R(p_1, \dots, p_k, e_1, \dots, e_k) = 0$.

Чим більше відношення явних правил у відображенні f над неявними, тим простіше виконати зворотне відображення. Неявні правила можуть залежати від багатьох речей. Наприклад, ці правила можуть бути залежними від переваг оператора, що відповідає за створення зображення P або від складності зовнішнього світу, що відображається (як приклад, вирішення завдань ідентифікації по фото паспорту та ідентифікації по камерах відеоспостереження в метрополітені).

Для мінімізації неявних правил і спрощення їх вирішення необхідно нівелювати вплив ЕЗ одне на одного, що досягається зменшенням їх розмірів, які розглядаються, за допомогою декомпозиції. Для цього використовуватимемо висхідний підхід до створення загальної методики обробки ССЗ: аналіз та обробка відбуватиметься від окремих пікселів до зображення в цілому. Узагальнену методику показано нарис. 1.

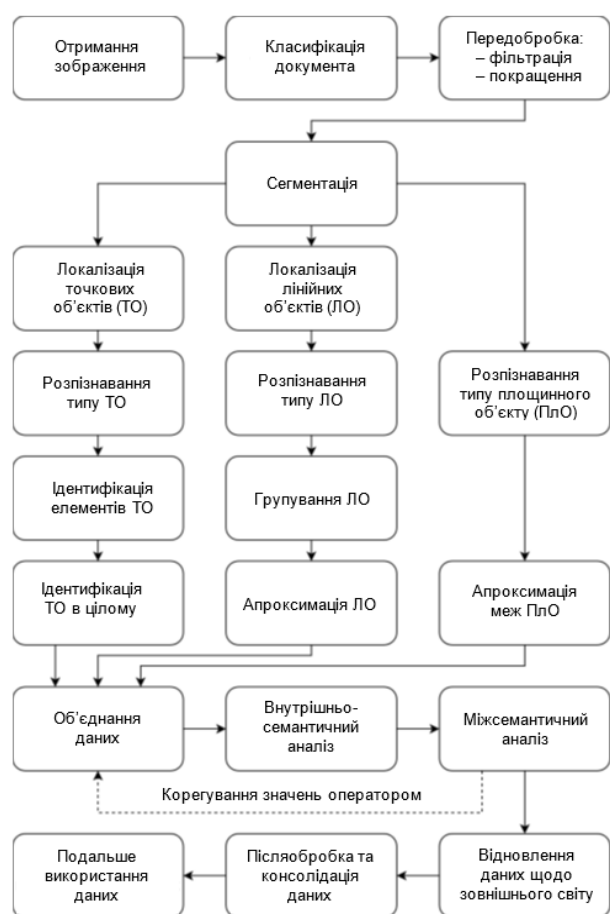


Рис. 1. Узагальнена методика програмного аналізу та обробки складноструктурних зображень

2. Метод багатометкової сегментації складноструктурних зображень

При обробці складноструктурних зображень потрібно перейти до отримання та аналізу об'єктів. Історично склалися два підходи до реалізації такої «об'єктною» сегментації:

- екземплярна, що заснована на поділі об'єктів,
- семантична, що заснована на поділі класів об'єктів.

Ні той, ні інший підхід не може надати достатньої якості при аналізі ССЗ. Перший підхід спирається на ідею обробки об'єктів - речей (things), які можуть в значному ступені накладатися друг на друга (немає строгого відокремлення). Другий підхід спирається на ідею обробки об'єктів - регіонів, які на ССЗ можуть виявлятися з великою варіативністю.

Відтепер активно формується паноптичний підхід, який не поділяє два вищезазначених поняття і тому є більш гнучким. Основна його ідея полягає в використанні найпростіших обчислювачів для формування непересічних суперпікселів. При цьому все одно вважається, що окремий піксель може належати тільки одному класу, що не відповідає властивості гетерархічності ССЗ, тому пропонується метод багатометкової сегментації, при якому піксель відповідає класам об'єктів з деякою мірою.

Для опису окремого пікселя на зображенні у випадку, коли кожен об'єкт визначається одним кольором, використовується така формула:

$$p = \sum_{i=1}^n w_i \cdot C_i + E, \quad (5)$$

де p – вектор кольору пікселя у просторі RGB, координати якого приймають дискретні значення, наприклад, $[0, 255]$; n – кількість класів об'єктів зображення; C_i – вектор кольору i -го класу об'єкта; E – вектор зашумленості (може мати від'ємні координати); w_i – вагові коефіцієнти при векторах кольорів.

Хоча даний випадок і є найпростішим, але на ССЗ з великим числом перегинів його можна використовувати в якості основи, так як навіть складні об'єкти часто представляються у вигляді набору одноколірних об'єктів. Завданням багатометкової сегментації є визначення вагових коефіцієнтів для кожного типу об'єкта. Працюючи з кольоровими зображеннями це завдання можна розв'язати загалом за допомогою формальних математичних методів не більше, ніж для трьох класів об'єктів.

Якщо об'єкт не можна представити у вигляді сукупності об'єктів одного кольору, то можна висунути більш загальну формулу, в якій враховується різниця координат між пікселем, що розглядається, і центром об'єкта та інші властивості об'єкта, наприклад, кут нахилу і масштаб:

$$p = g \left(\sum_{i=1}^n w_i \cdot C_i (\Delta x_i, \Delta y_i, a_i, s_i) + E \right), \quad (6)$$

де g – загальна функція.

З аналізу формули (6) можна зробити висновок щодо неможливості проведення сегментації без

знання інформації про розташування об'єктів. Хоча ця інформація згідно розробленої методики (рис. 1) виходить при наступних етапах обробки, але вона може бути отримана і без проведення етапу сегментації. Це підтверджується використанням алгоритмів R-CNN і YOLO при обробці фото- і відеоматеріалів. Проблема даних алгоритмів полягає в їх локальності: вони можуть бути застосовані тільки для знаходження ТО, але не для ЛО або ПЛО.

Враховуючи попередні зауваження щодо якості ССЗ, наявності на них об'єктів певних типів та дослідження методів сегментації, було зроблено висновок про необхідність використання інтелектуальних методів для сегментації, що є класифікацією окремих пікселів. Так як для визначення, чи є даний піксель частиною деякого об'єкта, необхідно знати оточення даного пікселя, то для аналізу можна використовувати метод ковзного вікна. Вікно розміром $a \times a$ ($a \in \mathbb{N}$, $a > 1$) пікселів призведе до створення $3a^2$ значень вхідних даних, які необхідно проаналізувати з урахуванням різних напрямків. Квадратне вікно вибрано, виходячи з можливості зображення об'єктів під будь-яким кутом.

Оскільки необхідно визначити, чи належить цей піксель деякому i -му типу об'єктів, то вихідний шар моделі має бути функцією активації *softmax*:

$$\sigma_i = e^{z_i} / \sum_{k=1}^K e^{z_k}, \quad (7)$$

де K – кількість класів, z – дектор розмірності K .

Значення цієї функції (що може лежати в інтервалі $[0, 1]$) для кожного i визначає міру того, що піксель належатиме класу об'єктів i . Сума всіх значень σ дорівнює одиниці. Таким чином, зображення після проведення даної багатометкової сегментації перетворюється на напівтонових зображень, що пред'являє підвищені вимоги до пам'яті системи за наявності великої кількості класів. Зберігання даних растрів можна скоротити, якщо використовувати розряджені матриці, до яких можна перейти, якщо переніціалізувати низькі значення нуля.

На рис. 2 приклад роботи алгоритму проведення багатометкової сегментації, який використовує раніше навчену модель.

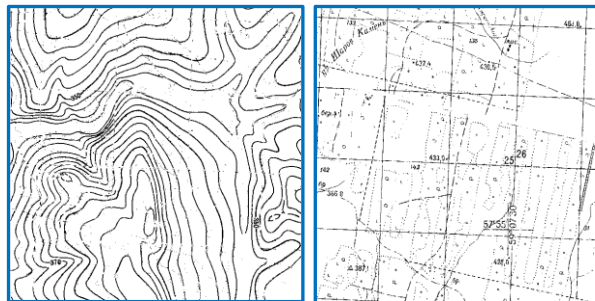


Рис. 2. Приклад багатометкової сегментації різних типів об'єктів під час обробки топографічних карт

3 Алгоритм для класифікації площинних об'єктів

Однією із ознак ССЗ є нерегулярність, тому розпізнавання великих площинних об'єктів (ПЛО) стає

проблемою, так як вони можуть бути представлені великою різноманітністю форм та текстур.

Існує підхід до розпізнавання площинних об'єктів як точкових при достатньо великих по розміром навчальних зображення, використовуючи згорткові нейронні мережі (CNN)

Головним недоліком використання CNN є необхідність тривалого навчання і швидкість роботи вже навченої моделі. Видно, що результати розпізнавання ПЛО далекі від аналогічних результатів при розпізнаванні ТО. Одним з підходів є використання гістограм: при ньому по ковзаючому вікну обчислюється гістограма (наприклад, яскравості для напівтонового зображення) і далі робляться висновки щодо знаходження об'єкту. Недоліком даного підходу є те, що такі гістограми по суті є інтегральними характеристиками, і в них можуть проявитися власні особливості, якщо об'єкти розташовані під різними кутами.

Хоча при відображенні ПЛО на ССЗ можуть з'явитися значущі шуми, але в цілому слід відзначити статистичну сталість в варіативності відхилення кольору і її частоті навіть в невеликих областях, що дозволяє значно обмежити розмір ковзаючих вікон на відмінність від, наприклад, CNN, для яких розмір вікон для точкової класифікації часто перевищує сотні пікселів. Для вирішення задачі визначення ПЛО на ССЗ потрібно очистити її від об'єктів, знайдених на етапі обробки ТО і ЛО, які зображуються поверх ПЛО, що є для них тілом. Для таких пікселів неможливо в принципі визначити, що знаходилося під образами переднього візуального плану, але можна припустити деяку континуальність і протяжність об'єктів заднього фону. Дана властивість можна назвати принципом локальності, який можна виразити в існуванні такої функції, що визначає клас пікселя за класами пікселів його сусідів і задовольняє наступній умові для майже кожного пікселя (під майже кожним пікселем тут розуміється, що ця умова може виконуватися для $O(1)$ пікселів кожного площинного об'єкта):

$$class(p) = f_{local}(\{class(p') | d(p,p') \leq d_{max}\}), \quad (8)$$

де $class(p)$ – клас фонового зображення поточного пікселя p , що визначається через свої координати растру, p' – пікселі, для яких виконується умова сусідства, d – метрика, за якою розраховується відстань між двома пікселями, d_{max} – максимальне значення даної метрики. Приклад реалізації розробленого алгоритму класифікації площинних об'єктів представлений на рис. 3.

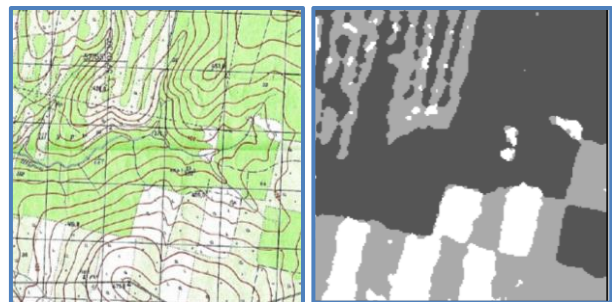


Рис. 3. Приклад класифікації площинних об'єктів

Висновки

У ході роботи було проведено аналіз сучасного стану аналізу та обробки растрових ССЗ. Проведені дослідження підходів до локалізації та визначення типів сегментованих об'єктів. Проведені дослідження підходів до розпізнавання образів та їх угрупування.

Удосконалений та досліджений метод підвищення точності аналізу та обробки ССЗ. Запропонований підхід дозволив підвищити точність аналізу та обробки складноструктурних зображень за рахунок сегментації, локалізації, розпізнавання та групування точкових, лінійних і площинних об'єктів, що засновані на комплексному використанні відомих методів.

СПИСОК ЛІТЕРАТУРИ

1. Zhang, Zh. MKA Net: An Efficient Network with Sobel Boundary Loss for Land-Cover Classification of Satellite Remote Sensing Imagery. *Remote Sens.* 2022. No. 18:4514. doi: <https://doi.org/10.3390/rs14184514>
2. Li, H. Intelligent Map Reader: A Framework for Topographic Map Understanding With Deep Learning and Gazetteer. *IEEE Access.* 2018. V. 6. P. 25363-25376. doi: <https://doi.org/10.1109/ACCESS.2018.2823501>
3. Petrovska, I., Kuchuk, H. (2023). Adaptive resource allocation method for data processing and security in cloud environment. *Advanced Information Systems*, 7(3), 67–73. doi: <https://doi.org/10.20998/2522-9052.2023.3.10>
4. Zakharchenko, A., Stepanets, O. (2023). Digital twin value in intelligent building development. *Advanced Information Systems*, 7(2), 75–86. doi: <https://doi.org/10.20998/2522-9052.2023.2.11>
5. Li, G., Liu, Y., Wu, J., Lin, D., Zhao, Sh. (2019). Methods of Resource Scheduling Based on Optimized Fuzzy Clustering in Fog Computing. *Sensors*, MDPI, 19(9). doi: <https://doi.org/10.3390/s19092122>
6. Кучук Н. Г., Мерлак В. Ю., Скороделов В. В. Метод зменшення часу доступу до слабкоструктурованих даних. *Сучасні інформаційні системи*. 2020. Т. 4, № 1. С. 97-102. doi: <https://doi.org/10.20998/2522-9052.2020.1.14>
7. Zuev, A., Karaman, D., Olshevskiy, A. (2023). Wireless sensor synchronization method for monitoring short-term events. *Advanced Information Systems*, 7(4), 33–40. doi: <https://doi.org/10.20998/2522-9052.2023.4.04>
8. Moreno-Garcia C. F., Elyan E., Jayne C. New trends on digitisation of complex engineering drawings. *Neural Computing and Applications*. 2019. V. 31(6). P. 1695-1712. doi: <https://doi.org/10.1007/s00521-018-3583-1>

Received (Надійшла) 27.02.2025

Accepted for publication (Прийнята до друку) 17.04.2025

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>.

Чхеїдзе Важа Олександрович – магістрант кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Vazha Chkheidze – master's student of the Department of electronic computers, Kharkiv National University of radio electronics, Kharkiv, Ukraine;

e-mail: Vazha.Chkheidze@khp.edu.ua; ORCID Author ID: <http://orcid.org/0009-0004-3424-481X>.

Севостьянова Олена Миколаївна – старший викладач кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Olena Sevostianova – Senior lecturer of the Department of electronic computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: olena.sevostianova@nure.ua; ORCID Author ID: <http://orcid.org/0009-0008-2595-5133>.

Фомічов Олександр Олександрович – кандидат технічних наук, старший викладач кафедри електронних обчислювальних машин Харківського національного університету радіоелектроніки, Харків, Україна;

Oleksandr Fomichov – Candidate of Technical Sciences, lead lecturer of the Department of electronic computers, Kharkiv National University of radio electronics, Kharkiv, Ukraine;

e-mail: oleksandr.fomichov@nure.ua; ORCID author ID: <https://orcid.org/0000-0001-9273-9862>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57219773778>.

Improving the accuracy of analysis and processing of complex structured images

Andriy Kovalenko, Vazha Chkheidze, Olena Sevostianova, Oleksandr Fomichov

Abstract. Topicality. Complex-structured images (CSI) tend to consolidate graphs of various types and hetero-hierarchy of semantic plans, which leads to the need to use hybrid and iterative algorithms specific to this type of images. **The object** of the study is the process of analyzing and processing raster CSI. **The subject** of the study is models and increasing the accuracy and productivity of CSI analysis and processing algorithms. **The purpose of the study** is to increase the accuracy and productivity of CSI analysis and processing algorithms. **Research results.** In the course of the work, an analysis of the current state of analysis and processing of raster complex-structured images was conducted. Research was conducted on approaches to localization and determination of types of segmented objects. Research was conducted on approaches to pattern recognition and their grouping. An improved and researched method for increasing the accuracy of analysis and processing of CSI. **Conclusion.** The proposed approach allowed to increase the accuracy of analysis and processing of complex structural images due to segmentation, localization, recognition and grouping of point, linear and planar objects, which are based on the integrated use of known methods.

Keywords: complex structural image, point object, planar object, segmentation.

Г. К. Кожевніков¹, А. О. Кравченко¹, А. О. Стадник², Д. В. Гриньов¹

¹ Національний технічний університет «Харківський політехнічний інститут», Харків, Україна

² Національний аерокосмічний університет «Харківський авіаційний інститут», Харків, Україна

ФОРМУВАННЯ СТІЙКОГО МОБІЛЬНОГО КЛАСТЕРУ ТУМАННОГО ШАРУ СИСТЕМИ ПІДТРИМКИ ІНТЕРНЕТУ РЕЧЕЙ

Анотація. Актуальність. Туманні обчислення забезпечують виконання оперативних транзакцій, мінімізуючи затримки передачі даних. Також туманні вузли можна реалізовувати, задіяючи мобільні пристрої IoT. Але при наявності в кластері мобільних вузлів його структура може порушитися внаслідок виходу якогось мобільного вузла за межі досяжності центрального вузла. Отже, для мереж IoT речей з високою щільністю при формуванні кластерів з мобільними вузлами треба враховувати стійкість структури. **Мета статті** – підвищення стійкості системи Інтернету речей з високою щільністю шляхом формування стійких мобільних кластерів туманного шару. **Результати дослідження.** Запропонована архітектура туманного шару з мобільними компонентами. Проведена оцінка трафіку туманних обчислень в мережі підтримки Інтернету речей. Розроблена модель мережі мобільних туманних обчислень. Сформований стійкий кластер мобільного туманного шару. **Висновок.** Запропонований підхід до забезпечення стійкості кластера мобільного туманного шару дозволяє використовувати у якості туманних вузлів мобільні пристрої. Мобільний кластер формується із найбільш стійких мобільних вузлів IoT, що суттєво зменшує можливість руйнування архітектури кластера.

Ключові слова: Інтернет речей, туманні обчислення, хмарна система, стійкий кластер, мобільна компонента.

Вступ

Останнім часом стрімко розвивається концепція Інтернету речей (IoT) [1–3]. IoT базується на хмарній платформі, доступ до центрів обробки даних (ЦОД) якої пристрої Інтернету речей отримують за допомогою мережі Інтернет [4]. Але при прямому зв'язку пристроїв та шлюзів IoT із ЦОД хмарної платформи при потребі виконання оперативних транзакцій виникають неприємними часові затримки при передачі даних, що провокує недотримання вимог QoS [5]. Тому інфраструктура системи підтримки IoT була розширена додатковими периферійними шарами: туманним та граничним [6].

Туманні обчислення забезпечують виконання оперативних транзакцій, мінімізуючи затримки передачі даних [7]. Також туманні вузли можна реалізовувати, задіяючи мобільні пристрої IoT. Доволі часто траєкторія руху мобільних пристроїв є випадковою величиною, отже такі мобільні туманні вузли можуть вийти за границі досяжності у мережі [8]. У мережах IoT з високою щільністю або з ультрамалими затримками, зазвичай, туманні вузли об'єднуються у кластери, доволі часто критерієм кластеризації є територіальна близькість туманних вузлів кластера [9, 10]. Але при наявності в кластері мобільних вузлів його структура може порушитися внаслідок виходу якогось мобільного вузла за межі досяжності центрального вузла кластера [11]. Отже, для мереж Інтернету речей з високою щільністю при формуванні кластерів з мобільними вузлами треба враховувати стійкість структури кластера.

Метою дослідження є підвищення стійкості системи Інтернету речей з високою щільністю шляхом формування стійких мобільних кластерів туманного шару.

1. Архітектура туманних обчислень

Принципи побудови хмарних обчислень збереглися і при формуванні концепції туманних обчислень. Але побудова архітектури туману є більш складною задачею, тому що цей шар децентралізований,

має суттєво обмежені обчислювальні та мережеві ресурси, і крім того, допускається рухливість вузлів.

Розгортання вузлів у туманному шарі зазвичай відбувається за такими моделями [12]:

- 1) приватний вузол туманних обчислень, що є вузлом індивідуального використання організацією;
- 2) вузол туманних обчислень спільноти ексклюзивно використовується певною спільнотою споживачів, котрі мають якісь спільні інтереси;
- 3) громадський вузол туманних обчислень зазвичай призначається для використання провайдером туманних обчислень;
- 4) гібридний вузол туманних обчислень є композицією декількох туманних вузлів, пов'язаних між собою технологією, котра повинна забезпечувати переносимість застосунків і даних.

Узагальнена схема архітектури туманних обчислень представлений на рис. 1.

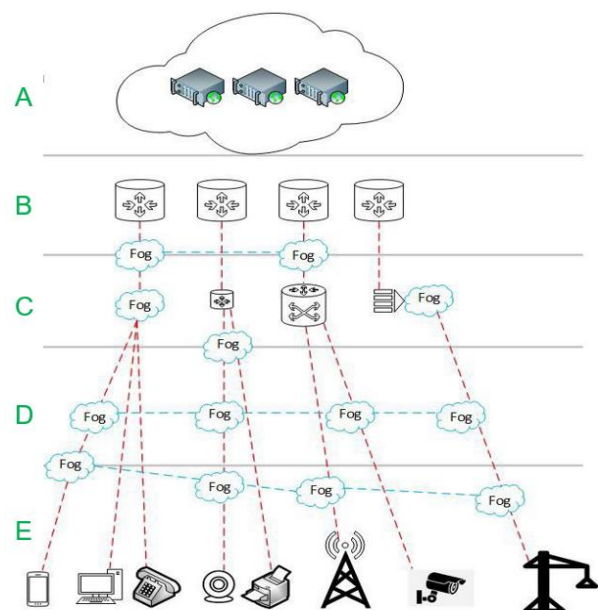


Рис. 1. Архітектура туманних обчислень

На рис. 1 позначені такі рівні архітектури: А – хмара, В – децентралізоване ядро туману, С – пристрої доступу, D – шлюзи, Е – пристрої IoT. У туманному шарі всі пристрої об'єднуються в єдину мережу доступу з вузлами, котрі мають вільний обчислювальний потенціал.

2. Оцінка трафіку туманних обчислень

Модель мережі підтримки IoT, котра використовує шар туманних обчислень схематично представлена на рис. 2.

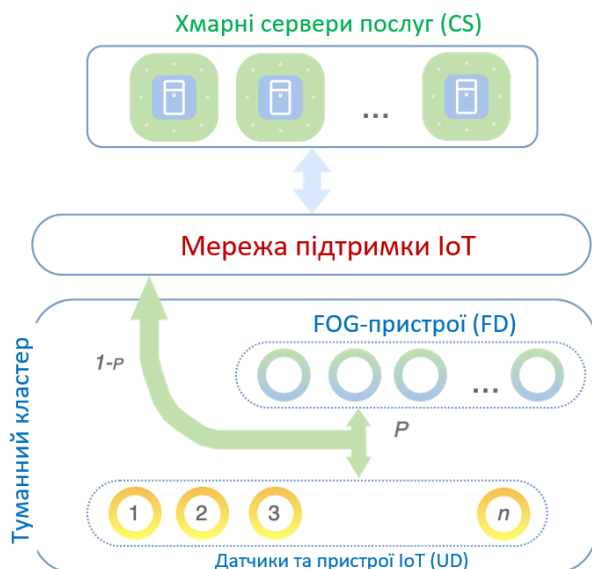


Рис. 2. Модель мережі підтримки IoT

Послугу надає один із центральних серверів CS, використовуючи канал передачі даних між пристроєм користувача UD і сервером, в результаті взаємодії в обидві сторони передається деякий обсяг даних [13].

Нехай пристрої туманних обчислень обслуговують n_i користувачів туманного кластера, крім того, частка трафіку, що оброблена в межах туманного кластера, буде пропорційна частці послуг, які мігрують в пристроях туманних обчислень. Нехай a_{i0} – трафік IoT для i -ої послуги. Розрахуємо загальний трафік, вироблений n датчиками та пристроями IoT туманного кластера, який надає m послуг:

$$a_{\Sigma} = n \cdot \sum_{i=1}^m a_{i0}. \quad (1)$$

Визначимо змінну p_i як ймовірність такої події, при котрій i -та послуга реалізується пристроями туманного кластера. Тоді інтенсивність трафіку, що функціонує в межах розглядаємого кластера буде такою:

$$a_{int} = n \cdot \sum_{i=1}^m p_i \cdot a_{i0}. \quad (2)$$

Отже інтенсивність зовнішнього трафіку є такою:

$$a_{ext} = n \cdot \sum_{i=1}^m (1 - p_i) \cdot a_{i0}. \quad (3)$$

Частка трафіку γ , яка оброблена у туманному кластері, є випадковою величиною. Вона залежить від

низки факторів та специфіки транзакцій IoT, можливостей пристроїв туманних обчислень кластера та інших. Її значення наближається до такої величини:

$$\gamma = \frac{a_{int}}{a_{int} + a_{ext}}. \quad (4)$$

Використання туманних обчислень дозволяє розвантажити хмарне ядро мережі підтримки IoT та скоротити цифровий розрив.

3. Модель мережі мобільних туманних обчислень

Пропонуємо модель описує функціонування мережі підтримки IoT при русі туманних пристроїв. В більшості випадків при моделювання оперують групами пристроїв – роями, в яких мобільні пристрої прагнуть зберегти загальні параметри руху. Формуючи динамічні туманні кластери систем Інтернету речей великої щільності розглядають групи мобільних щільно розміщених пристроїв. Такі пристрої намагаються зберегти своє положення щодо інших членів групи.

Для опису взаємного положення елементів мережі динамічних туманних обчислень підтримки Інтернету речей пропонується використовувати модель точкового процесу (рис. 3).

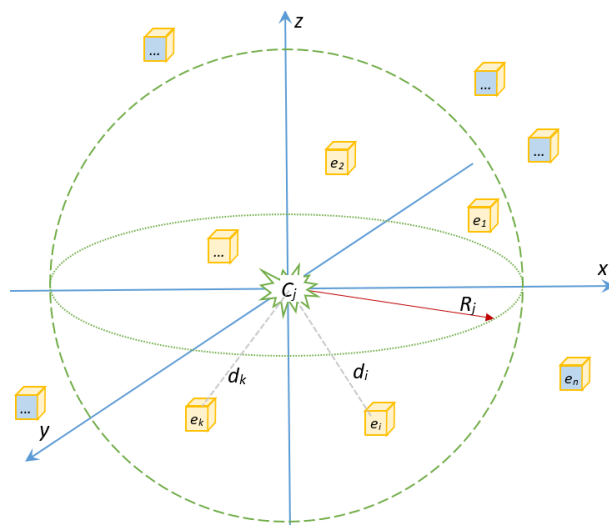


Рис. 3. Розташування пристроїв IoT у тривимірному просторі

На рис. 3 окремі динамічні пристрої IoT $e_1, \dots, e_k$, та інші входять до кластера K_j та розташовані в поточних позиціях, котрі формуються правилами організації для пристроїв IoT та в конкретний момент часу фіксовані щодо центру кластера C – головного вузла даного кластера. Всі елементи кластера, як і головний вузол, переміщуються з приблизно рівними швидкостями не міняючи напрямку, тобто елементи кластера є приблизно нерухомими відносно його центру.

Отже, відстань між мобільним пристроєм IoT і центром кластера K_j в тривимірному просторі можна визначати за таким виразом:

$$d_{ji0} = \sqrt{(x_{ji} - x_{jc})^2 + (y_{ji} - y_{jc})^2 + (z_{ji} - z_{jc})^2}, \quad (5)$$

де $x_{ji}, x_{jc}, y_{ji}, y_{jc}, z_{ji}, z_{jc}$ – поточні координати мобільних пристроїв туманного шару та відповідно координата центру кластера K_j – точки C_j .

В процесі реального руху як швидкості, так і напрямки руху деяких динамічних елементів можуть відрізнятися від заданих, тому вони відхиляються від плануємих позицій в кластері, тобто координати $x_{ji}, x_{jc}, y_{ji}, y_{jc}, z_{ji}, z_{jc}$ – це зазвичай випадкові величини, хоч і мобільні пристрої туманного шару все ж таки у більшості випадків зберігають положення щодо інших членів кластера. Якщо випадкові координати мобільних пристроїв туманного шару мають нормальний розподіл, у системі координат кластера K_j початкове положення i -го пристрою задається точкою A_{ji0} ($x_{ji0}, y_{ji0}, z_{ji0}$), а поточне положення i -го пристрою A_{ji} (x_{ji}, y_{ji}, z_{ji}), то відхилення пристрою від початкового положення розраховується таким чином:

$$\xi_{ji} = \sqrt{(x_{ji} - x_{ji0})^2 + (y_{ji} - y_{ji0})^2 + (z_{ji} - z_{ji0})^2}, \quad (6)$$

де ξ_{ji} не залежить від системи координат, отже $\xi_{ji} = \xi_i$ $\forall i$ та розподілена за нормальним законом із функцією щільності ймовірності

$$f(\xi_i) = \frac{1}{\sigma_i \sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2}(\xi_i / \sigma_i)^2\right), \quad (7)$$

де σ_i – середньоквадратичне відхилення ξ_i від початкового положення i -го мобільного пристрою.

Нехай пристрій IoT e_i , належить кластеру K_j з радіусом зони досяжності R_j , а d_{ji0} – відстань між початковим положенням e_i та центром кластера K_j . Рухаючись, пристрій e_i може відхилитися від початкового положення не більше, ніж на $3\sigma_i$, тобто всі поточні положення пристрою не вийдуть за межі сфери радіусом R_j з центром у точці K_j , тобто ймовірність знаходження пристрою IoT e_i буде у зоні досяжності кластера K_j є такою:

$$P_{ji}^+ \rightarrow P_{ji}^{(theor+)} = \begin{cases} 1, & \text{else } 3\sigma_i + d_{ji0} \leq R_j; \\ 0.5 + \Phi(3\sigma_i + d_{ji0} - R_j), & \\ \text{else } 3\sigma_i + d_{ji0} > R_j, \end{cases} \quad (8)$$

$$\text{де } \Phi(\xi_i) = \frac{1}{\sqrt{2\pi}} \int_0^{\xi_i} \exp(-t^2/2) dt.$$

Тоді ймовірність руйнування структури кластера K_j з k_j мобільними елементами буде такою:

$$P_j^- = 1 - \sum_{i=1}^{k_j} P_{ji}^+. \quad (9)$$

В якості критерію формування кластера застосовуємо його максимальну стійкість, тобто ймовірність збереження його структури.

Тоді цільовою функцією при формуванні кластера буде такий вираз:

$$\max_j \left(\min_{pos(C_j)} \left(1 - \sum_{i=1}^{k_j} P_{ji}^+ \right) \right), \quad pos(C_j) \in D, \quad (10)$$

де D – зона обслуговування; $pos(C_j)$ – позиція центру кластера в тривимірному просторі; n – розмір групи.

4. Забезпечення стійкості кластера мобільного туманного шару

При формуванні кластера як критерій використовуємо мінімізацію величини сумарної відстані від центра до елементів кластера. Мінімізація проводиться за рахунок вибору центрів кластерів, котрі визначаються як координати центру мас із врахуванням вагових коефіцієнтів w_i , котрі пов'язані з умовною масою відповідного елемента. Якщо ваговий коефіцієнт якогось елемента набагато більший ніж вагові коефіцієнти інших елементів, то центр кластера зміщується до цього елемента (рис. 4).

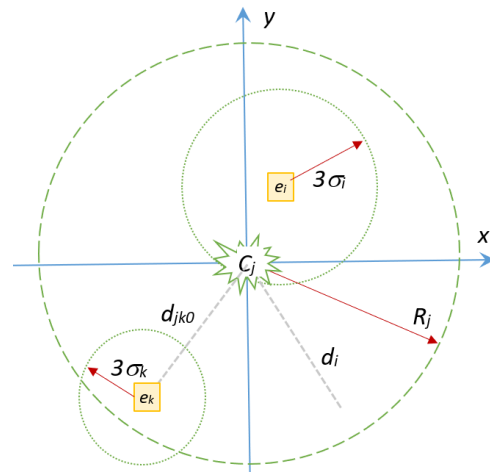


Рис. 4. Проекція кластера туманних обчислень

Вибір позиції координатора базується на центроїдному алгоритмі кластеризації FOREL. Модифікація алгоритму полягає в тому, що до поточного кластеру відбираються тільки ті об'єкти IoT, для яких ймовірність виходу об'єкту із кластера мінімальна.

Висновки

У статті запропонований підхід до формування стійкого мобільного кластера туманного шару підтримки IoT. Наведена архітектура туманного шару з мобільними компонентами. Проведена оцінка трафіку туманних обчислень в мережі підтримки Інтернету речей. Розроблена модель мережі мобільних туманних обчислень. Сформований стійкий кластер мобільного туманного шару. Запропонований підхід дозволяє використовувати у якості туманних вузлів мобільні пристрої. Мобільний кластер формується із найбільш стійких мобільних вузлів IoT, це зменшує можливість руйнування архітектури кластера.

СПИСОК ЛІТЕРАТУРИ

- Chalapathi, G.S.S., Chamola, V., Vaish, A., Buyya, R. (2022). [Industrial internet of things applications of edge and fog computing: A review and future directions](https://doi.org/10.1007/978-3-030-57328-7_12). *Advances in Information Security*, 83, 293–325. doi: https://doi.org/10.1007/978-3-030-57328-7_12
- Petrovska, I., Kuchuk, H. (2023). Adaptive resource allocation method for data processing and security in cloud environment. *Advanced Information Systems*, 7(3), 67–73. doi: <https://doi.org/10.20998/2522-9052.2023.3.10>

3. Schulz, A.S. (2023). User Interactions with Internet of Things (IoT) Devices in Shared Domestic Spaces. *ACM International Conference Proceeding Series*, 577–579. doi: <https://doi.org/10.1145/3626705.3632615>
4. 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. *IMM*, 104, 258–275. doi: <https://doi.org/10.1016/j.indmarman.2022.04.012>
5. Zakharchenko, A., Stepanets, O. (2023). Digital twin value in intelligent building development. *Advanced Information Systems*, 7(2), 75–86. doi: <https://doi.org/10.20998/2522-9052.2023.2.11>
6. 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. doi: <https://doi.org/10.1186/s13677-022-00345-y>
7. Li, G., Liu, Y., Wu, J., Lin, D., Zhao, Sh. (2019). Methods of Resource Scheduling Based on Optimized Fuzzy Clustering in Fog Computing. *Sensors*, MDPI, 19(9). doi: <https://doi.org/10.3390/s19092122>
8. Кожевніков Г. К., Черниш Д. С., Матяш О. Ю. Онтологічний підхід до перерозподілу навантаження Інтернету речей. Системи управління, навігації та зв'язку. Полтава : Національний університет «Полтавська політехніка імені Юрія Кондратюка», 2024. Вип. 2(76). С. 91–94. doi: <https://doi.org/10.26906/SUNZ.2024.4.92>
9. Заковоротний О., Орлова Т. Порівняльний аналіз хмарних та туманних середовищ Інтернету речей. Системи управління, навігації та зв'язку. Збірник наукових праць. Полтава: ПНТУ, 2023. Вип. 3(73). С. 152–154. doi: <https://doi.org/10.26906/SUNZ.2023.3.152>
10. Кучук Н. Г., Мерлак В. Ю., Скороделов В. В. Метод зменшення часу доступу до слабкоструктурованих даних. *Сучасні інформаційні системи*. 2020. Т. 4, № 1. С. 97–102. doi: <https://doi.org/10.20998/2522-9052.2020.1.14>
11. Zuev, A., Karaman, D., Olshevskiy, A. (2023). Wireless sensor synchronization method for monitoring short-term events. *Advanced Information Systems*, 7(4), 33–40. doi: <https://doi.org/10.20998/2522-9052.2023.4.04>
12. Кучук Г.А. Управління трафіком мультисервісної розподіленої телекомунікаційної мережі / Г.А. Кучук // Системи управління, навігації та зв'язку. – К.: ЦНДІ НІУ, 2007. – Вип. 2. – С. 18–27.
13. Gomathi, B., Saravana Balaji, B., Krishna Kumar, V., Abouhawwash, M., Aljahdali, S., Masud, M. and Kuchuk, N. (2022), “Multi-Objective Optimization of Energy Aware Virtual Machine Placement in Cloud Data Center”, *Intelligent Automation and Soft Computing*, Vol. 33(3), pp. 1771–1785, doi: <http://dx.doi.org/10.32604/iasc.2022.024052>

Received (Надійшла) 20.02.2025

Accepted for publication (Прийнята до друку) 24.04.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Кожевніков Георгій Костянтинович – кандидат технічних наук, доцент, професор кафедри мультимедійних та Інтернет технологій і систем, Національний технічний університет “Харківський політехнічний інститут”, Харків, Україна;

Georgii Kozhevnikov – Candidate of Technical Sciences, Associate Professor, Professor of Multimedia and Internet Technologies and Systems Department, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;

e-mail: heorhii.kozhevnikov@khpі.edu.ua; ORCID ID: <http://orcid.org/0000-0002-6586-6767>;

Scopus ID: <https://www.scopus.com/authid/detail.uri?authorId=6701803684>.

Кравченко Антон Олександрович – магістрант кафедри мультимедійних та Інтернет технологій і систем, Національний технічний університет “Харківський політехнічний інститут”, Харків, Україна;

Anton Kravchenko – master's student of Multimedia and Internet Technologies and Systems Department, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;

e-mail: anton.kravchenko@khpі.edu.ua; ORCID ID: <http://orcid.org/0000-0003-0797-3588>.

Стадник Анастасія Олександрівна – кандидат технічних наук, доцент кафедри комп'ютерних систем, мереж і кібербезпеки, Національний аерокосмічний університет “Харківський авіаційний інститут”, Харків, Україна;

Anastasia Stadnik – Candidate of Technical Sciences, Associate Professor, Department of Computer Systems, Networks and Cybersecurity, National Aerospace University “Kharkiv Aviation Institute”, Kharkiv, Ukraine;

e-mail: a.stadnyk@csn.khai.edu; ORCID ID: <https://orcid.org/0000-0002-7717-4346>;

Scopus ID: <https://www.scopus.com/authid/detail.uri?origin=resultslist&authorId=56332781400>.

Гриньов Деніс Валеріович – кандидат технічних наук, доцент кафедри комп'ютерної інженерії та програмування, Національний технічний університет “Харківський політехнічний інститут”, Харків, Україна;

Denys Grynov – Candidate of Technical Sciences, Associate Professor of computer engineering and programming Department, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;

e-mail: dgrynov@gmail.com; ORCID ID: <http://orcid.org/0009-0007-3092-9397>;

Scopus ID: <https://www.scopus.com/authid/detail.uri?authorId=55822619300>.

Formation of a sustainable mobile fog layer cluster of the Internet of Things support system

Georgii Kozhevnikov, Anton Kravchenko, Anastasia Stadnik, Denys Grynov

Abstract. Topicality. Fog computing ensures the execution of operational transactions, minimizing data transmission delays. Fog nodes can also be implemented using mobile IoT devices. But if there are mobile nodes in the cluster, its structure may be disrupted due to the fact that some mobile node goes beyond the reach of the central node. Therefore, for high-density IoT networks, when forming clusters with mobile nodes, it is necessary to take into account the stability of the structure. **The purpose of the article** is to increase the stability of the high-density Internet of Things system by forming stable mobile fog layer clusters. **Research results.** A fog layer architecture with mobile components is proposed. An assessment of fog computing traffic in the Internet of Things support network is carried out. A model of a mobile fog computing network is developed. A stable mobile fog layer cluster is formed. **Conclusion.** The proposed approach to ensuring the stability of a mobile fog layer cluster allows using mobile devices as fog nodes. The mobile cluster is formed from the most stable mobile IoT nodes, which significantly reduces the possibility of destruction of the cluster architecture.

Keywords: Internet of Things, fog computing, cloud system, stable cluster, mobile component.

О. С. Кравченко, О. А. Галуза, А. О. Савченко, С. В. Погорелов, А. І. Роговий

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

МОДЕЛЮВАННЯ ТОПОГРАФІЇ МІКРОПОДРЯПИН НА ДЗЕРКАЛЬНИХ ПОВЕРХНЯХ

Анотація. Предметом дослідження є моделювання топографії дефекту типу мікроподряпини на плоскій дзеркальній поверхні із застосуванням сучасних алгоритмічних підходів, які забезпечують відтворення гладких профілів дефектів із високою відповідністю реальним даним. **Мета роботи** — розробити алгоритм синтезу топографічних карт для моделювання мікроподряпин із гладкими та реалістичними профілями, що враховує природні флуктуації глибини й області впливу дефекту. **Вирішуються завдання:** 1) розробка математичної моделі, що враховує ключові характеристики мікроподряпин; 2) забезпечення плавного переходу між дискретними точками профілю мікроподряпини; 3) моделювання флуктуації глибини та області впливу дефекту. **Використовуються методи:** алгоритмічне моделювання цифрових карт висот, яке базується на синтезі поперечного перерізу подряпини з заданими параметрами; кубічна сплайнова інтерполяція для забезпечення гладкості профілів і точного відтворення контурів; синусоїдальна модуляція глибини з урахуванням амплітуди і частоти коливань та нелінійні функції затухання (експоненціальне й косинусне) для детального опису розподілу впливу дефекту в просторі. Також застосовано пакетний режим генерації синтетичних даних, що дозволяє варіювати ключові параметри моделі (максимальне заглиблення, ширину області впливу, коефіцієнт амплітуди модуляції, частоту коливань та параметри затухання), що забезпечує отримання широкого спектру варіантів топографії мікроподряпин. **Результати:** синтетичні карти висот показали високу відповідність експериментальним даним, підтверджуючи ефективність моделі в реальних умовах і її переваги для систем автоматичного контролю оптичних поверхонь. **Висновки:** розроблена модель має значний потенціал для інтеграції в системи аналізу дефектів, сприяє підвищенню надійності оптичних пристроїв та є перспективною для подальших досліджень, зокрема, для інтеграції з методами глибокого навчання та подальшої експериментальної валідації.

Ключові слова: моделювання, топографія, мікроподряпини, алгоритм, інтерполяція.

Вступ

У сучасних умовах розвитку прецизійної оптики, де застосовуються дзеркальні поверхні найвищої якості, навіть мікроскопічні дефекти, зокрема мікроподряпини, можуть суттєво впливати на робочі характеристики оптичних систем. Такі пошкодження розмірами в нанометровому чи субмікронному діапазоні викликають розсіювання світла, інтерференційні ефекти, оптичні аберації, погіршення формування зображень, локальний нерівномірний розподіл теплових потоків, що може спричинити перегрів, а також скорочення терміну служби приладів [1, 2]. Це робить детальне вивчення топографії мікроподряпин важливим завданням як із науково-теоретичної, так і з практичної точки зору, особливо для підвищення надійності та довговічності оптичних систем, що використовуються в астрономії, лазерних системах, медичній діагностиці та інших галузях.

Метою цього дослідження є розроблення математичної моделі топографії мікроподряпини, яка дає змогу з високою точністю відтворювати реалістичні розподіли заглиблень і забезпечувати плавні переходи вздовж усієї довжини дефекту. Така модель має враховувати не лише локальні особливості формування мікроподряпини, а й передбачати варіювання параметрів дефекту для відображення різноманітних типів пошкоджень, що виникають на етапі виробництва (наприклад, під час полірування) чи внаслідок експлуатації оптичних приладів [3].

Для досягнення поставленої мети у цій роботі розв'язуються такі завдання:

- Визначити ключові параметри, що впливають на формування топографії мікроподряпини, зокрема максимальну глибину дефекту, ширину зони

впливу та розподіл заглиблень уздовж траєкторії пошкодження.

- Розробити математичний алгоритм, який дозволяє моделювати профіль мікроподряпини з урахуванням реалістичних розподілів заглиблень і плавних переходів між різними ділянками дефекту.

- Обґрунтувати можливість застосування розробленої моделі для створення синтетичних карт висот, придатних для навчання нейромережових систем виявлення дефектів.

Вибір теми дослідження зумовлений необхідністю створення високоякісних синтетичних карт висот мікроподряпин, які адекватно відображають їхні реальні фізичні особливості для подальшої побудови інтерференційних картин таких поверхонь. Ці картини слугуватимуть навчальною вибіркою для нейромережових систем розпізнавання та класифікації дефектів. Традиційні методи моделювання геометричних об'єктів часто обмежуються дискретними підходами, що призводить до різких переходів і недостатньої гладкості профілю дефекту [4]. Тому розроблення математичної моделі, яка враховує реалістичний розподіл заглиблень і забезпечує плавні переходи, є важливим кроком для підвищення точності прогнозування та аналізу пошкоджень дзеркальних поверхонь.

Отже, поставлене завдання спрямоване на розв'язання актуальної проблеми сучасної оптики — моделювання топографії мікроподряпин, що має пряме практичне значення для покращення характеристик оптичних систем і подовження терміну їхньої служби. Надалі розроблена модель буде інтегрована в систему генерації синтетичних даних для тренування моделей глибокого навчання, що дасть змогу автоматизувати процес виявлення та класифікації дефектів у прецизійних оптичних приладах.

1. Аналіз проблеми й наявних методів

Сьогодні цифрове моделювання топографії дзеркальних поверхонь є важливим інструментом для вивчення мікроскопічних дефектів, таких як мікроподряпини, які суттєво впливають на якість оптичних систем. Ця проблема набуває особливої актуальності через зростання попиту на синтетичні дані, необхідні для навчання нейромережових систем, що автоматично виявляють і класифікують дефекти. Розуміння існуючих підходів до такого моделювання допомагає визначити їхні сильні сторони та обмеження, а також обґрунтувати потребу в нових математичних рішеннях, які пропонує ця робота [5].

За останні роки оптична наука та інженерія досягли значного прогресу в розробці методів для створення цифрових карт висот поверхонь. Такі карти традиційно отримують експериментально за допомогою методів профілометрії чи атомної силової мікроскопії (AFM), які дозволяють детально відтворювати рельєф із урахуванням найменших нерівностей. Однак ці підходи мають суттєві недоліки, коли йдеться про підготовку великих навчальних вибірок для навчання нейромереж. Профілометрія обмежена повільністю вимірювань і складністю масштабування, тоді як AFM, хоча й високоточний, є вкрай трудомістким і дорогим процесом, що робить його непрактичним для генерації тисяч або мільйонів зразків, необхідних для ефективного навчання нейромереж [6]. Водночас дедалі частіше застосовуються алгоритмічні методи, що синтезують поверхні на основі математичних моделей. Проте традиційні підходи, зокрема дискретні алгоритми, часто стикаються з проблемою різких переходів у профілях дефектів, що не відображає їхньої природної форми, сформованої механічними процесами, такими як полірування чи експлуатація [7].

У класичному підході до моделювання поверхонь базова геометрична модель зазвичай представлена у вигляді дискретного набору точок, полігональної сітки або примітивних елементів, де сусідні точки з'єднуються відрізками [8]. У такій моделі дефект описується як послідовність окремих кроків: для кожної точки траєкторії визначається локальне заглиблення, а спад впливу моделюється за допомогою простої лінійної функції затухання з максимумом у центрі подряпини. Такий підхід є достатньо простим у реалізації, проте недостатньо гнучким, оскільки призводить до різких перепадів між сусідніми точками, що суперечить фізичній природі дефектів, які формуються шляхом неперервних змін під впливом механічних або термічних факторів.

Крім того, ця модель не враховує випадкові або періодичні коливання, характерні для реальних мікроподряпин, що можуть виникати через неоднорідність матеріалів чи інструментів [9].

Ці недоліки стають особливо помітними, коли синтетичні карти висот використовуються для навчання нейромережових систем. Реалістичність таких даних критично впливає на ефективність алгоритмів розпізнавання дефектів, адже ідеалізовані профілі, отримані класичними методами, погано корелюють

із реальними умовами. Дослідження показують, що низька якість синтетичних даних може знизити точність класифікації на 20–30% у реальних оптичних системах [10]. Саме тому виникає потреба в удосконаленні моделі, яка б забезпечувала природніші розподіли заглиблень і плавні переходи.

Для подолання цих обмежень пропонується внести кілька математичних коригувань. По-перше, сплайн-інтерполяція дозволяє побудувати неперервну криву на основі дискретних точок, забезпечуючи гладкість профілю, що відображає реальні фізичні процеси. По-друге, синусоїдальна модуляція глибини додає періодичні коливання, імітуючи природні нерівності, зумовлені механічними впливами. По-третє, нелінійне затухання, наприклад, експоненціальне чи косинусне, моделює плавний спад впливу дефекту від центру до країв, на відміну від жорсткого лінійного спаду.

Отже, аналіз літератури показує, що хоча сучасні методи моделювання мають значний потенціал, їхні обмеження в точності відтворення дефектів потребують нових рішень. Запропоновані коригування відкривають шлях до створення якісніших синтетичних карт висот, які стануть основою для ефективного навчання алгоритмів виявлення мікроподряпин. У наступному розділі буде детально розкрито методологію розроблення цієї моделі, орієнтованої на імітацію реальних пошкоджень оптичних поверхонь.

2. Методологія

Розробка реалістичного алгоритму моделювання мікроподряпин є ключовим етапом у створенні синтетичних карт висот, які можуть ефективно використовуватися для навчання нейромережових систем виявлення дефектів на дзеркальних поверхнях. У цьому розділі детально описано покроковий процес розробки такого алгоритму, включаючи математичні формули та процедури, що дозволяють усунути недоліки традиційних методів і забезпечити природність профілю дефектів [11]. Початок роботи з моделювання передбачає ініціалізацію процесу, де обирається стартова точка на поверхні, наприклад, (0, 0), або генерується випадкова позиція, а також визначається кількість кроків (N) у межах від мінімального N_{min} до максимального N_{max} значення, що задає довжину дефекту.

Далі алгоритм переходить до обчислення профілю дефекту.

Для кожного кроку $i=0,1,\dots,N-1$ розраховується параметр прогресу:

$$progress_i = \frac{i}{N-1}, \quad (1)$$

який змінюється від 0 (початок) до 1 (кінець траєкторії).

На основі цього параметра визначається базовий рівень заглиблення:

$$depth_i = scratch_depth \times fade_factor(progress_i), \quad (2)$$

де $scratch_depth$ – максимальна глибина, а функція затухання $fade_factor(\cdot)$ задає розподіл глибини вздовж траєкторії.

Щоб врахувати локальну зону впливу, яка визначається як область навколо центра подряпини, в межах якої глибина дефекту зменшується від максимального значення до нуля відповідно до заданої функції затухання, причому ширина цієї області визначається параметром *scratch_width*, для кожного кроку визначається коло з центром у поточній позиції та радіусом *scratch_width*.

У межах цього кола обчислюється відстань *d* від центра подряпини, а коефіцієнт впливу *infl(d)*, який є безрозмірною величиною, що визначає ступінь зменшення глибини дефекту в точці, розташованій на відстані *d* від центра подряпини, і змінюється від 1 (у центрі подряпини) до 0 (на межі зони впливу) відповідно до обраної функції затухання, розраховується за допомогою двох альтернативних нелінійних функцій затухання:

– косинусне затухання:

$$\text{infl}(d) = 0,5 \left[1 + \cos \left(\pi \frac{d}{\text{scratch_width}} \right) \right], \quad (3)$$

$$d \leq \text{scratch_width};$$

– експоненційне затухання:

$$\text{infl}(d) = \exp \left[-\alpha \left(\frac{d}{\text{scratch_width}} \right)^2 \right], \quad (4)$$

$$d \leq \text{scratch_width};$$

де α – параметр, що визначає швидкість затухання.

Коригування висоти поверхні в поточній точці розраховується як:

$$\Delta z = \text{depth}_i \times \text{infl}(d), \quad (5)$$

що дозволяє плавно розподілити ефект дефекту. Після цього поточна позиція оновлюється з урахуванням базового кроку та випадкових зміщень, які генеруються за допомогою Perlín-шумів. Perlín-шуми — це алгоритмічно згенерована псевдовипадкова функція, що забезпечує плавні та природні варіації значень у просторових координатах і широко використовується для моделювання органічних текстур і траєкторій у комп'ютерній графіці та симуляціях [12]. Їхнє застосування додає природної варіативності траєкторії, хоча основна увага в алгоритмі зосереджена на вдосконаленні глибини та зони впливу.

Щоб усунути різкі переходи, характерні для класичного підходу, до моделі інтегрується кубічна сплайн-інтерполяція. На основі отриманих точок $\{(x_i, z_i)\}_{i=0}^{N-1}$, де z_i відображає обчислене значення глибини в точці x_i , будується суцільна крива. Кубічний сплайн будується у вигляді набору поліномів $S_i(x)$ на кожному інтервалі $[x_i, x_{i+1}]$, які задаються так:

$$S_i(x) = a_i + b_i(x - x_i) + c_i(x - x_i)^2 + d_i(x - x_i)^3, \quad (6)$$

при цьому коефіцієнти a_i, b_i, c_i та d_i визначаються таким чином, щоб забезпечити неперервність самої функції $S(x)$ та її перших двох похідних $S'(x)$ і $S''(x)$ на

всій області визначення [13]. Такий підхід дозволяє отримати гладку криву дефекту, що відповідає фізичним умовам формування мікроподряпин і значно підвищує якість моделювання порівняно з класичним дискретним методом.

Реальні мікроподряпини характеризуються природними флуктуаціями глибини, які можуть бути спричинені неоднорідністю матеріалу або випадковими механічними впливами. Для імітації цих ефектів у модель вводиться синусоїдальна модуляція глибини вздовж траєкторії дефекту.

Глибина для кожного кроку тепер визначається за формулою:

$$\text{depth}_i =$$

$$= \text{scratch_depth} \times \text{fade_factor}(\text{progress}_i) \times$$

$$\times [1 + \epsilon \sin(2\pi f \cdot \text{progress}_i + \phi)],$$

де ϵ – коефіцієнт амплітуди модуляції (зазвичай обирається в діапазоні 0.1–0.2), f – частота модуляції, що визначає кількість періодів коливань уздовж дефекту, ϕ – фазовий зсув, який визначає зміщення початкової фази синусоїдальної функції в аргументі, що дозволяє регулювати положення максимального заглиблення вздовж траєкторії.

Введення синусоїдальної модуляції надає профілю дефекту більш природного вигляду, імітуючи дрібні варіації, характерні для реальних мікроподряпин [14].

У класичних моделях стандартно використовується лінійне затухання коефіцієнта впливу в міру віддалення від центра дефекту, що призводить до різких змін на краях зони впливу. Для отримання більш реалістичного розподілу застосовується нелінійне затухання, яке може бути описане функціями експоненційного (4) чи косинусного (3) затухання.

Обидві функції забезпечують максимальне значення коефіцієнта в центрі зони $d \approx 0$ і плавний спад до нуля в міру наближення до межі зони впливу. Застосування нелінійного затухання дозволяє точніше моделювати розподіл заглиблення, наближаючи синтетичні карти до експериментальних даних [15].

Розроблений алгоритм реалізовано у вигляді програмного комплексу, який дає змогу моделювати як окремі мікроподряпини, так і генерувати великі набори даних синтетичних карт висот (пакетний режим) для подальшого навчання нейромережових систем.

Програмна реалізація включає такі ключові компоненти:

- Ініціалізація: вибір стартової точки та генерація випадкового числа кроків N .
- Обчислення профілю: на кожному кроці розраховується базовий рівень заглиблення з урахуванням функції затухання.
- Застосування модуляції: синусоїдальна модуляція вносить варіації в розподіл глибини вздовж траєкторії.
- Розрахунок затухання зони: нелінійне затухання (експоненційне або косинусне) визначає розподіл впливу дефекту від центра до країв.

- Оновлення координат: на основі базового кроку та випадкових зміщень (з використанням Перлін-шумів) відбувається оновлення поточної позиції.

- Перевірка меж: контроль виходу за межі модельованої області та завершення генерації дефекту.

Пакетний режим реалізації дозволяє варіювати ключові параметри моделі, такі як: максимальна глибина *scratch_depth*, радіус зони впливу *scratch_width*, параметри синусоїдальної модуляції ϵ , f , ϕ та швидкість експоненційного затухання α . Такий підхід забезпечує генерацію різноманітних варіантів топографії мікроподряпин, що є суттєвою перевагою під час створення великих навчальних вибірок для систем машинного навчання [16].

Для наочного представлення алгоритму подано блок-схему (рис. 1), яка демонструє послідовність основних етапів.



Рис. 1. Блок-схема алгоритму моделювання мікроподряпини

Ця блок-схема ілюструє модульність алгоритму та послідовне виконання всіх етапів, що дає змогу легко адаптувати модель до різних умов моделювання та оптимізувати процес генерації синтетичних даних.

У сукупності описані методи забезпечують отримання високореалістичних синтетичних карт топографії мікроподряпин, придатних для подальшого аналізу та застосування в системах виявлення дефектів оптичних поверхонь.

3. Експериментальна частина

Для оцінки ефективності розробленої моделі моделювання мікроподряпин було проведено серію експериментів, спрямованих на аналіз реалістичності отриманих топографічних карт і впливу запропонованих коригувань на кінцевий профіль дефекту. Основну увагу зосереджено на оцінці трьох ключових удосконалень: кубічної сплайн-інтерполяції, синусоїдальної модуляції глибини та нелінійного затухання зони впливу. Результати порівнювалися з базовою дискретною моделлю, щоб кількісно та якісно продемонструвати переваги нового підходу.

Для проведення експериментів було змодельовано двовимірну область, що являє собою ділянку дзеркальної поверхні.

Розмір симульованої області обрано таким чином, щоб охопити масштаби дефектів, які виникають у реальних умовах. Зокрема, для моделювання поверхні використано область розміром 1024×1024 пікселі з роздільною здатністю 100 нм на піксель, що відповідає фізичному розміру 102.4×102.4 мкм.

У процесі моделювання варіювалися ключові параметри алгоритму, щоб дослідити їхній вплив на профіль мікроподряпини.

Кількість кроків N змінювалася від 50 до 200, визначаючи довжину дефекту.

Максимальна глибина *scratch_depth* обиралася в діапазоні від 100 до 600 нм, що відповідає типовим експериментальним спостереженням.

Ширина зони впливу *scratch_width* варіювалася від 1 до 7 мкм, охоплюючи як вузькі, так і ширші дефекти.

Параметри синусоїдальної модуляції: коефіцієнт амплітуди ϵ – обирався в діапазоні від 0.1 до 0.2, що дає змогу додати помірні флуктуації в профіль; частота f – підбиралася так, щоб на довжині дефекту було від 1 до 3 періодів коливань; фазовий зсув ϕ – генерувався випадковим чином для кожного експерименту. Параметр експоненційного затухання α : значення підбиралися в діапазоні від 2 до 5 для забезпечення плавного спаду впливу від центра дефекту до його країв.

Моделювання виконувалося за чотирма варіантами реалізації алгоритму, щоб порівняти вплив кожного коригування. Перший варіант — базова дискретна модель без удосконалень, яка характеризується різкими переходами між кроками та однорідним розподілом заглиблень. Рис. 2, а ілюструє типовий профіль базової моделі, де чітко видно різкі зміни висоти, що не відповідають реальним топографіям. Другий варіант включав лише сплайн-інтерполяцію, що забезпечує гладкість профілю завдяки неперервності перших двох похідних. Рис. 2, б демонструє приклад такого профілю, де плавність досягається завдяки сплайн-інтерполяції. Третій варіант додавав синусоїдальну модуляцію глибини, дозволяючи відтворити природні флуктуації, спричинені неоднорідністю механічного впливу. Четвертий варіант інтегрував нелінійне затухання зони впливу (експоненційне або косинусне), що забезпечує плавний спад заглиблення від центра до країв.

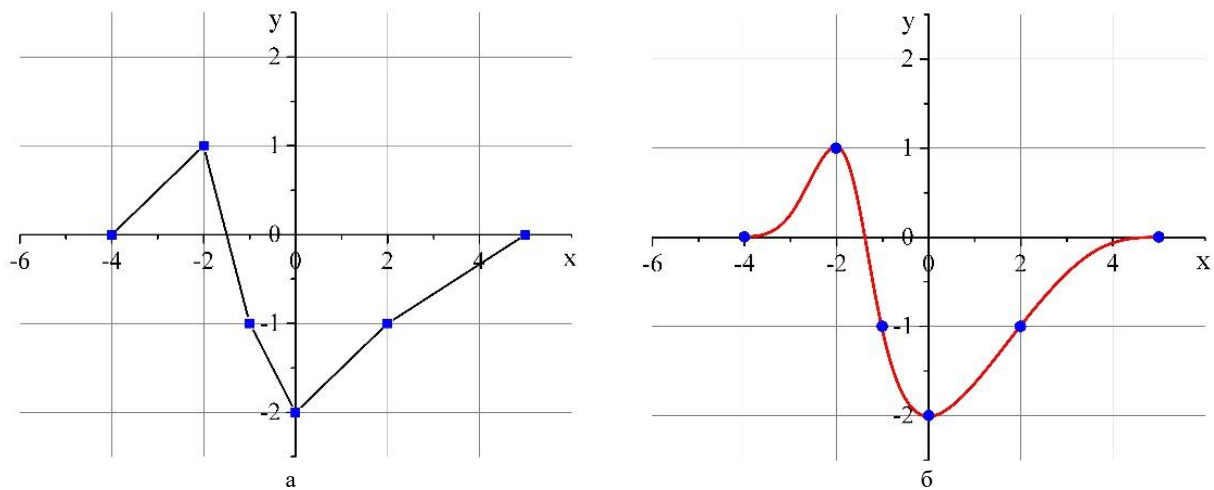


Рис. 2. Профіль дефекту:
для базової дискретної моделі, що демонструє різкі переходи висоти (а),
для моделі з застосуванням сплайн-інтерполяції, що ілюструє плавний перехід висоти (б)

Реалістичність отриманих топографічних карт оцінювалася за кількома критеріями. По-перше, проводилося порівняння з експериментальними даними, отриманими методами атомної силової мікроскопії, таких параметрів як максимальна глибина, ширина дефекту та форма розподілу заглиблень. По-друге, застосовувався статистичний аналіз, зокрема розрахунок показника гладкості профілю як середньоквадратичного відхилення різниць між сусідніми значеннями висоти, а також аналіз розподілу глибин уздовж траєкторії. По-третє, проводився візуальний аналіз 2D- і 3D-візуалізацій для оцінки наявності

різких переходів і неоднорідностей.

Результати візуального аналізу наочно демонструє рис. 3, де представлено 3D-візуалізацію профілю мікроподряпини з двох ракурсів: зверху (а) та знизу (б). На зображенні видно, що застосування сплайн-інтерполяції та синусоїдальної модуляції створює плавний профіль із природними коливаннями глибини, без різких стрибків, характерних для базової моделі. Колірна шкала відображає глибину від 0 (червоний) до -100 нм (синій), підкреслюючи плавний перехід заглиблення та варіативність профілю, що відповідає реальним фізичним умовам.

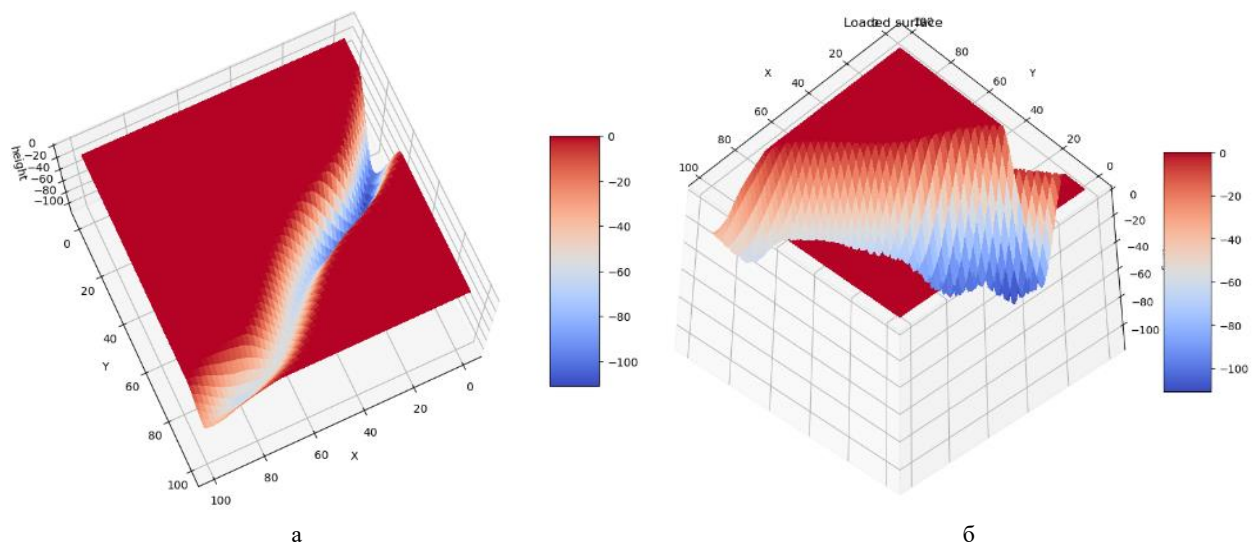


Рис. 3. 3D-візуалізація профілю мікроподряпини:
вид згори (а) та вид знизу (б)

Візуальний аналіз отриманих моделей показав, що базова дискретна модель генерує дефекти з вираженими різкими переходами між послідовними кроками, що не відповідає спостережуваній топографії мікроподряпин у реальних оптичних системах.

Застосування сплайн-інтерполяції суттєво покращує гладкість профілю, усуваючи різкі зміни висоти.

Додатково, синусоїдальна модуляція глибини вносить природні коливання, характерні для мікроподряпин, які утворюються під впливом неоднорідного механічного впливу. Нелінійне згасання зони впливу, у свою чергу, дозволяє створити плавний перехід заглиблення від центра дефекту до його країв, що сприяє отриманню більш реалістичних топографічних карт.

Для об'єктивної оцінки реалістичності моделі було проведено кількісну оцінку. Вона підтвердила візуальні спостереження.

Показник гладкості профілю, визначений як середньоквадратичне відхилення різниць між сусідніми значеннями висоти, знизився на 30% у моделі зі сплайн-інтерполяцією порівняно з дискретною моделлю.

Максимальна глибина в усіх варіантах відповідала заданому параметру *scratch_depth*, але моделі з синусоїдальною модуляцією показали варіації в межах $\pm 10\text{--}15\%$, що відображає природні флуктуації.

Статистичний аналіз розподілу глибин також виявив, що нелінійне затухання зменшує розкид значень у периферійних зонах, забезпечуючи рівномірніший спад заглиблення.

Дослідження впливу параметрів моделі показало, що збільшення амплітуди модуляції ϵ призводить до виражених коливань у профілі, але значення вище 0.2 можуть спричинити нереалістичні стрибки, тому оптимальний діапазон становить 0.1–0.2.

Частота f , яка визначає кількість періодів коливань, при значеннях вище 3 створює надмірну "шумність", тому рекомендується обмежуватися 1–3 періодами.

Параметр α впливає на швидкість експоненційного спаду: діапазон 2–5 забезпечує плавний, але виразний перехід від центра до країв, що узгоджується з експериментальними даними.

Таким чином, експерименти підтвердили, що запропонована модель із застосуванням сплайн-інтерполяції, синусоїдальної модуляції та нелінійного затухання значно підвищує реалістичність синтезованих топографічних карт мікроподряпин.

Отримані профілі не лише відповідають фізичним характеристикам реальних дефектів, а й забезпечують достатню варіативність для створення великих навчальних вибірок, що є необхідною умовою для ефективного навчання нейромережових систем виявлення дефектів у прецизійних оптичних приладах.

Далі буде розглянуто практичне застосування розробленої моделі для аналізу дефектів і її інтеграцію в автоматизовані системи.

4. Застосування моделі в задачах аналізу дефектів

Розроблений алгоритм дозволяє генерувати великі обсяги синтетичних карт висот, на яких детально відтворено топографію мікроподряпин. Завдяки пакетному режиму можливо автоматично змінювати такі ключові параметри моделі, як:

- максимальна глибина *scratch_depth*,
- ширина зони впливу *scratch_width*,
- амплітуда модуляції ϵ ,
- частота f ,
- параметр експоненційного затухання α .

Наприклад, під час генерації набору даних із 10 000 карт висот параметр *scratch_depth* варіювався від 100 до 600 нм, а ϵ набувало значень у межах 0.1–0.2, що дало змогу охопити різні стадії формування дефекту – від початкових мікроподряпин до більш

виражених пошкоджень, які виникають при тривалій експлуатації оптичних приладів. Такий підхід забезпечує отримання синтетичних даних, здатних адекватно відображати різноманітність реальних дефектів.

Різнорозмірність синтетично згенерованих карт висот має важливе значення для підвищення якості навчання нейромереж. Використання широкого спектра варіантів дефектів дозволяє навчальній моделі адаптуватися до різних умов експлуатації оптичних систем.

Попередні дослідження показали, що нейромережа, навчена на синтетичному наборі даних за розробленою моделлю, досягла точності розпізнавання дефектів на рівні 92.3%, що перевищує результати навчання на наборі даних за дискретною моделлю (77.6%).

Покращення якості класифікації зумовлене більш повним охопленням параметричного простору дефектів, що дозволяє моделі успішно обробляти як типові, так і менш поширені варіанти пошкоджень.

Було проведено серію експериментів із навчання згорткової нейромережі (CNN) для автоматичного виявлення мікроподряпин.

Для навчання використовувався синтетично згенерований набір даних, що складався з 10 000 зображень, отриманих за описаним алгоритмом. Короткий підсумок результатів досліджень:

- Точність класифікації: досягнуто точність 92.3% на тестовій вибірці.
- Час навчання: середній час навчання становив близько 3 годин на стандартному GPU.
- Порівняння з експериментальними даними: візуальний і статистичний аналіз показав високу схожість синтетичних карт із вимірними топографіями мікроподряпин, що підтверджує реалістичність згенерованих даних.

Ці результати свідчать про високий потенціал використання синтетично згенерованих даних для навчання нейромережових систем виявлення дефектів, що відкриває нові перспективи для автоматизованого контролю якості оптичних поверхонь.

Рис. 4 ілюструє загальний процес.

На першому етапі синтетичні карти висот генеруються з використанням пакетного режиму, що дозволяє варіювати параметри моделі.

Далі дані проходять стадію передобробки, яка включає нормалізацію, фільтрацію та аугментацію, після чого використовуються для навчання нейромережової моделі, що призводить до підвищення точності та надійності системи виявлення мікроподряпин.

Таким чином, запропонована модель і пакетний режим генерації синтетичних даних дозволяють створювати різноманітні та високореалістичні набори даних, які значно покращують навчання нейромережових систем.

Застосування цього підходу відкриває нові можливості для автоматизованого контролю якості оптичних поверхонь і може стати основою для розроблення складніших систем діагностики дефектів у прецизійній оптиці.



Рис. 4. Схема використання синтетичних даних для навчання моделі глибокого навчання

5. Результати досліджень та їхнє обговорення

У цьому розділі проводиться порівняльний аналіз розробленої моделі з традиційними підходами до моделювання мікроподряпин, обговорюються її переваги, обмеження та перспективи подальших удосконалень.

Переваги запропонованої моделі полягають у значно покращеній гладкості профілю, що досягається завдяки застосуванню кубічної сплайн-інтер-

поляції. Такий підхід забезпечує неперервність першої та другої похідних, усуваючи різкі переходи, притаманні дискретним моделям, і дозволяє отримати фізично обґрунтований плавний профіль дефекту.

Реалістичний розподіл глибини досягається за рахунок упровадження синусоїдальної модуляції вздовж траєкторії, що імітує природні коливання, спричинені неоднорідністю матеріалу та варіативністю механічних впливів.

Застосування нелінійних функцій згущання, таких як експоненційне або косинусне, для опису зони впливу дефекту забезпечує плавний спад заглиблення від центра до периферії, що є важливим для точного відтворення реальної топографії пошкоджень.

Проте розроблена модель не позбавлена певних спрощень. Зокрема, для спрощення розрахунків не враховуються складні фізико-хімічні процеси, такі як динаміка корозії, поверхневі хімічні реакції або мікроструктурні зміни, що виникають у матеріалі внаслідок тривалого впливу навантажень. Також відсутня інтеграція впливу зовнішніх факторів, наприклад, температурних і вологісних умов, що може обмежувати застосування моделі в деяких специфічних сценаріях.

Ці обмеження, однак, сприймаються як відправна точка для подальших досліджень, спрямованих на розширення параметричного простору моделі та її адаптацію до реальних умов експлуатації оптичних систем.

Наглядне порівняння основних характеристик дискретної моделі та запропонованої вдосконаленої моделі представлено в табл. 1 нижче.

Таблиця 1 – Порівняння характеристик дискретної моделі та розробленої моделі

Характеристика	Дискретна модель	Вдосконалена модель
Гладкість профілю	Низька, наявні різкі переходи	Висока, неперервні зміни завдяки сплайн-інтерполяції
Реалістичний розподіл глибини	Однорідний розподіл із лінійним згущанням	Нелінійний розподіл із синусоїдальною модуляцією та нелінійним згущанням
Урахування випадкових флуктуацій	Мінімальне врахування шуму	Ефективне моделювання випадкових коливань за рахунок синусоїдальної модуляції
Застосовність для навчання нейромережі	Обмежена, не охоплює всі реальні сценарії	Висока, генеруються різноманітні синтетичні дані, наближені до експериментальних

Аналіз представлених даних дозволяє зробити висновок, що запропонована модель має значний потенціал для інтеграції в системи автоматизованого контролю якості оптичних поверхонь.

Отримані синтетичні карти висот характеризуються високим ступенем реалістичності, що суттєво підвищує якість навчальних вибірок для нейромережних систем виявлення та класифікації мікроподряпин.

Перспективи подальших досліджень пов'язані з розширенням параметричного простору моделі за рахунок урахування динаміки зношування та впливу зовнішніх умов, а також із інтеграцією розробленого

підходу з методами глибокого навчання для автоматизації процесу виявлення дефектів. Таким чином, незважаючи на деякі припущення та обмеження, запропонована модель демонструє значний прогрес порівняно з традиційними методами та становить перспективну основу для подальших розробок у галузі аналізу та контролю якості оптичних систем.

Висновки

У результаті проведеного дослідження було розроблено математичну модель топографії мікроподряпини, призначену для синтезу реалістичних цифрових карт висот дзеркальних поверхонь з дефектами.

Основний внесок роботи полягає у значному покращенні гладкості профілю дефекту порівняно зі класичними підходами, що досягається завдяки застосуванню кубічної сплайн-інтерполяції. Такий підхід забезпечує неперервність першої та другої похідних, що дозволяє усунути різкі переходи між дискретними точками та отримати фізично обґрунтований плавний профіль мікроподряпини.

Додаткове підвищення реалістичності моделі досягається за рахунок введення синусоїдальної модуляції глибини вздовж траєкторії дефекту. Це рішення імітує природні флуктуації, зумовлені неоднорідністю матеріалу та варіативністю механічних впливів, що суттєво підвищує схожість синтетично згенерованих профілів із реальними дефектами.

Застосування нелінійних функцій затухання, таких як експоненційне та косинусне, для опису розподілу впливу дефекту дає змогу змодельовати плавний спад заглиблення від центра дефекту до його країв, що є важливим для адекватного відтворення реальної топографії пошкоджень.

Практична значущість розробленої моделі полягає в її здатності генерувати високоякісні синтетичні дані, придатні для навчання нейромережових систем, призначених для автоматичного виявлення та класифікації мікроподряпин. Можливість пакетного ре-

жиму генерації, що дозволяє варіювати ключові параметри моделі, забезпечує отримання різноманітних варіантів дефектів, які охоплюють широкий спектр реальних сценаріїв пошкоджень.

Це, у свою чергу, сприяє підвищенню точності та надійності оптичних систем, оскільки навчені на таких моделях нейромережі зможуть ефективніше виявляти навіть незначні дефекти, що впливають на якість роботи приладів.

Перспективи подальших досліджень пов'язані з інтеграцією розробленої моделі з методами глибокого навчання, що дозволить розробляти та навчати нейромережові системи на основі синтетичних даних із подальшою експериментальною валідацією.

Крім того, планується розширення параметричного простору моделі за рахунок уключення додаткових факторів, таких як динаміка зношування та вплив зовнішніх умов, що може суттєво підвищити точність відтворення складних фізичних процесів, які відбуваються на оптичних поверхнях.

Таким чином, запропонована модель демонструє високий потенціал для підвищення надійності оптичних систем завдяки точнішому виявленню дефектів і становить перспективну основу для подальших досліджень у галузі автоматизованого контролю якості в прецизійній оптиці.

СПИСОК ЛІТЕРАТУРИ

1. Stover J. C. (2012), "Optical Scattering: Measurement and Analysis", SPIE Press, Bellingham, WA. <https://doi.org/10.1117/3.975276>
2. [Zhongkai Liu, Jincheng Wang, Rongkuan Leng, Xiaokun Wang, Min Zhang, Jing Wang, Mengxue Cai, Wenhan Li, Bin Liu, Lingzhong Li, Qiang Cheng, Longxiang Li, Xiao Luo, Xuejun Zhang. Impact of mirror local defects on system scattering in telescopes // Results in Physics. – 2024. – Vol. 56. – 107265. <https://doi.org/10.1016/j.rinp.2023.107265>
3. Cao H., Peng X., Shi F., Tian Y., Kong L., Chen M., Hao Q. Advances in Subsurface Defect Detection Techniques for Fused Silica Optical Components: A Literature Review // Journal of Materials Research and Technology. – 2025. – Vol. 35. – P. 809-835. <https://doi.org/10.1016/j.jmrt.2025.01.045>
4. Galuza A., Shkoda M., Tevyasheva O., Belyaeva A., Savchenko A., Kolenov I. Modeling and Synthesis of Monochrome Interference Patterns of Flat Optical Surfaces with Typical Defects for Automatic Surface Quality Control // 2020 10th International Conference on Advanced Computer Information Technologies (ACIT): Conference Proceedings, Deggendorf, Germany, 16-18 Sept. – 2020. – P. 344-347. <https://doi.org/10.1109/ACIT49673.2020.9208908>
5. Monnet J., Petrovic O., Herfs W. Investigating the generation of synthetic data for surface defect detection: A comparative analysis // Procedia CIRP. – 2024. – Vol. 130. – P. 767–773. <https://doi.org/10.1016/j.procir.2024.10.162>
6. [Mei L., Guan G. Profilometry and atomic force microscopy for surface characterization // Nano TransMed. – 2023. – Vol. 2, Issue 1. – e9130017. <https://doi.org/10.26599/NTM.2023.9130017>
7. Gutierrez P., Luschkova M., Cordier A., Shukor M., Schappert M., Dahmen T. Synthetic training data generation for deep learning based quality inspection // Proc. of SPIE. – 2021. – Vol. 11794. – 1179403. <https://doi.org/10.1117/12.2586824>
8. Discrete Geometry for Computer Imagery / Ed.: Couprie M., Cousty J., Kenmochi Y., Mustafa N. – Springer, Cham, 2019. – 496 p. <https://doi.org/10.1007/978-3-030-14085-4>
9. Bosch C., Pueyo X., Mérillou S., Ghazanfarpour D. A Physically-Based Model for Rendering Realistic Scratches // Computer Graphics Forum. – 2004. – Vol. 23, Issue 3. – P. 361-370. <https://doi.org/10.1111/j.1467-8659.2004.00767.x>
10. Shorten C., Khoshgoftaar T. M. A Survey on Image Data Augmentation for Deep Learning // Journal of Big Data. – 2019. – Vol. 6, No. 1. – P. 60. <https://doi.org/10.1186/s40537-019-0197-0>
11. H. Müser M. H., Nicola L. Modeling the surface topography dependence of friction, adhesion, and contact compliance // MRS Bulletin. – 2022. – Vol. 47. – P. 1221-1228. <https://doi.org/10.1557/s43577-022-00468-2>
12. Lagae A., Lefebvre S., Cook R., DeRose T., Drettakis G., Ebert D.S., Lewis J.P., Perlin K., Zwicker M. A Survey of Procedural Noise Functions // Computer Graphics Forum. – 2010. – Vol. 29, Issue 8. – P. 2579-2600. <https://doi.org/10.1111/j.1467-8659.2010.01827.x>
13. Lyche T., Manni C., Speleers H. Foundations of Spline Theory: B-Splines, Spline Approximation, and Hierarchical Refinement. In: Lyche T., Manni C., Speleers H. (eds) Splines and PDEs: From Approximation Theory to Numerical Linear Algebra. Lecture Notes in Mathematics, vol 2219. Springer, Cham. https://doi.org/10.1007/978-3-319-94911-6_1
14. Pawlus P., Reizer R. Profilometric measurements of wear scars: A review // Wear. – 2023. – Vol. 534–535. – 205150. <https://doi.org/10.1016/j.wear.2023.205150>

15. Kwon T.-Y., Ramachandran M., Park J.-G. Scratch formation and its mechanism in chemical mechanical planarization (CMP) // Friction. – 2013. – Vol. 1, Issue 4. – P. 279–305. <https://doi.org/10.1007/s40544-013-0026-y>
16. Zhang J. Classification and Comparison of Data Augmentation Techniques // Transactions on Computer Science and Intelligent Systems Research. – 2024. – Vol. 6. – P. 180–187. <https://doi.org/10.62051/7e91md96>

Received (Надійшла) 11.02.2025

Accepted for publication (Прийнята до друку) 21.05.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Кравченко Олександр Сергійович – аспірант кафедри комп'ютерної математики і аналізу даних, Національний технічний університет “Харківський політехнічний інститут”, Харків, Україна;

Oleksandr Kravchenko – PhD Student of the Department of Computer Mathematics and Data Analysis, National Technical University “Kharkiv Polytechnic Institute”, Kharkiv, Ukraine;

e-mail: Oleksandr.Kravchenko@cs.khpi.edu.ua; ORCID Author ID: <https://orcid.org/0000-0002-6169-1250>.

Галуза Олексій Анатолійович – доктор фізико-математичних наук, професор, професор кафедри комп'ютерної математики і аналізу даних, Національний технічний університет “Харківський політехнічний інститут”, Харків, Україна;

Oleksii Haluza – Doctor of Physical and Mathematical Sciences, Professor, Professor at the Department of Computer Mathematics and Data Analysis, National Technical University “Kharkiv Polytechnic Institute”, Kharkiv, Ukraine;

e-mail: Oleksii.Haluza@khpi.edu.ua; ORCID Author ID: <https://orcid.org/0000-0003-3809-149X>;

Scopus ID: <https://www.scopus.com/authid/detail.uri?authorId=6603590390>.

Савченко Алла Олександрівна – кандидат фізико-математичних наук, доцент, доцент кафедри фізики, Національний технічний університет “Харківський політехнічний інститут”, Харків, Україна;

Alla Savchenko – Candidate of Physical and Mathematical Sciences, Associate Professor, Associate Professor at the Department of Physics, National Technical University “Kharkiv Polytechnic Institute”, Kharkiv, Ukraine;

e-mail: Alla.Savchenko@khpi.edu.ua; ORCID Author ID: <https://orcid.org/0000-0001-9309-1797>;

Scopus ID: <https://www.scopus.com/authid/detail.uri?authorId=56637062400>.

Погорелов Станіслав Вікторович – доктор фізико-математичних наук, професор, професор кафедри комп'ютерної математики і аналізу даних, Національний технічний університет “Харківський політехнічний інститут”, Харків, Україна;

Stanislav Pohorielov – Doctor of Physical and Mathematical Sciences, Professor, Professor at the Department of Computer Mathematics and Data Analysis, National Technical University “Kharkiv Polytechnic Institute”, Kharkiv, Ukraine;

e-mail: Stanislav.Pohorielov@khpi.edu.ua; ORCID Author ID: <https://orcid.org/0000-0002-0189-8655>;

Scopus ID: <https://www.scopus.com/authid/detail.uri?authorId=9270293800>.

Роговий Антон Іванович – кандидат технічних наук, доцент, доцент кафедри управління проєктами в інформаційних технологіях, Національний технічний університет “Харківський політехнічний інститут”, Харків, Україна;

Anton Rogovyi – Candidate of Technical Sciences, Associate Professor, Associate Professor at the Project Management in Information Technology Department, National Technical University “Kharkiv Polytechnic Institute”, Kharkiv, Ukraine;

e-mail: Anton.Rogovyi@khpi.edu.ua; ORCID Author ID: <https://orcid.org/0000-0002-8178-4585>;

Scopus ID: <https://www.scopus.com/authid/detail.uri?authorId=57196221255>.

Modeling the topography of microscratches on mirror surfaces

Oleksandr Kravchenko, Oleksii Haluza, Alla Savchenko, Stanislav Pohorielov, Anton Rogovyi

Abstract. The subject of the study is the modeling of the topography of a microscratch-type defect on a flat mirror surface using modern algorithmic approaches that ensure the reproduction of smooth defect profiles with high fidelity to real data. **The aim of the work** is to develop an algorithm for synthesizing topographic maps to model microscratches with smooth and realistic profiles, taking into account natural depth fluctuations and the defect's area of influence. **The tasks addressed** include: 1) developing a mathematical model that accounts for the key characteristics of microscratches; 2) ensuring a smooth transition between discrete points of the microscratch profile; 3) modeling depth fluctuations and the defect's area of influence. **The methods used** include: algorithmic modeling of digital height maps based on the synthesis of a scratch cross-section with specified parameters; cubic spline interpolation to ensure profile smoothness and accurate contour reproduction; sinusoidal depth modulation considering amplitude and oscillation frequency, as well as nonlinear attenuation functions (exponential and cosine) for a detailed description of the defect's influence distribution in space. A batch mode for generating synthetic data was also employed, allowing variation of key model parameters (maximum depth, width of the influence area, modulation amplitude coefficient, oscillation frequency, and attenuation parameters), which enables the generation of a wide range of microscratch topography variants. **Results:** The synthetic height maps demonstrated high correspondence with experimental data, confirming the model's effectiveness in real-world conditions and its advantages for automated optical surface inspection systems. **Conclusions:** The developed model holds significant potential for integration into defect analysis systems, contributes to improving the reliability of optical devices, and is promising for further research, particularly for integration with deep learning methods and subsequent experimental validation.

Keywords: modeling, topography, microscratches, algorithm, interpolation.

М. В. Ліпчанський, О. В. Ліпчанська, Г. І. Молчанов, К. А. Христофоров

Національний технічний університет «Харківський політехнічний інститут», Україна, Харків

РОЗРОБКА API ДЛЯ АДАПТИВНОГО ТЕСТУВАННЯ В СИСТЕМАХ УПРАВЛІННЯ НАВЧАЛЬНИМИ КУРСАМИ

Анотація. Стаття присвячена розробці API для системи управління навчальними курсами з інтегрованою підтримкою адаптивного тестування. Актуальність теми зумовлена зростаючою потребою в індивідуалізації освітнього процесу та підвищенні ефективності оцінювання знань у сучасних умовах цифрової трансформації освіти. Метою статті є дослідження принципів та обґрунтування підходів до створення ефективного, захищеного та гнучкого API для управління навчальними курсами. Особливістю API є підтримка адаптивного тестування з урахуванням рівня знань користувача. У рамках дослідження проведено аналітичний огляд систем діагностики знань, адаптивних моделей та стратегій тестування. На основі отриманих результатів спроектовано архітектуру системи, обрано сучасні технології, зокрема Spring Boot та PostgreSQL, реалізовано бізнес-логіку. Розроблено програмний код, що демонструє роботу алгоритмів адаптивного тестування, включаючи механізм вибору тестових завдань на основі попередніх відповідей користувача. API забезпечує можливість інтеграції адаптивного тестування в різноманітні освітні платформи, що підвищує ефективність та індивідуалізацію навчального процесу. Впровадження API відкриває нові можливості для персоналізації навчання та вдосконалення якості освітніх послуг.

Ключові слова: API, REST, IRT, база даних, адаптивне тестування, навчальний курс.

Вступ

У сучасних умовах цифрової трансформації освіти зростає значення технологій, які сприяють індивідуалізації навчального процесу та підвищенню ефективності оцінювання знань. Однією з таких технологій є адаптивне тестування, яке дає змогу створювати персоналізовані тести, що враховують рівень підготовки користувача.

Таким чином, у сучасному цифровому часі сфера освіти зазнає значних змін під впливом технологічного прогресу. Зростає потреба в ефективних інструментах, які б дозволяли не лише надавати знання, а й об'єктивно оцінювати рівень їх засвоєння кожним окремим студентом.

Постановка проблеми. Традиційні методи оцінювання, такі як іспити та контрольні роботи, часто не здатні повною мірою врахувати індивідуальні особливості навчального процесу та можуть бути суб'єктивними.

Адаптивне тестування є однією з інноваційних технологій, що дозволяє вирішити зазначені проблеми. Воно передбачає індивідуалізацію процесу тестування, коли складність завдань підлаштовується під рівень знань кожного конкретного користувача, що призводить до більш точних оцінок знань, зменшує час тестування та підвищує мотивацію студентів. Проте, для ефективного впровадження адаптивного тестування в освітній процес необхідні відповідні програмні рішення. Одним з ключових елементів таких рішень є API (Application Programming Interface) – інтерфейс, що дозволяє різним системам та платформам взаємодіяти між собою. Розробка якісного API для системи управління навчальними курсами з підтримкою адаптивного тестування є складною задачею, що вимагає врахування багатьох факторів.

Таким чином, існує нагальна потреба у розробці ефективного, масштабованого, безпечного та зру-

чного в інтеграції API для систем управління навчальними курсами з підтримкою адаптивного тестування. Вирішення цієї проблеми сприятиме підвищенню якості освіти та впровадженню сучасних технологій в навчальний процес.

Відповідно, **метою статті** є вивчення принципів та обґрунтування методології побудови API, який буде ефективним, безпечним і гнучким. Це дозволить керувати навчальними курсами з підтримкою адаптивного тестування для врахування індивідуального рівня знань студента.

Аналіз останніх досліджень і публікацій. Сучасні реформи у вищій освіті спрямовані на створення сприятливих умов для розвитку та реалізації інтелектуального потенціалу суспільства. В [1] підкреслюється важливість контролю якості освіти та оцінювання навчальних досягнень студентів вищих навчальних закладів для забезпечення якісного навчального процесу та як ключового елемента розвитку інтелектуального потенціалу суспільства. Також підкреслюється значущість розробки та впровадження ефективних систем оцінювання у навчальний процес та необхідність постійного вдосконалення методів контролю.

У [2] особлива увага приділяється історичному розвитку контролю знань у навчальному процесі. Автор проводить аналіз еволюції форм і методів оцінювання, підкреслюючи, що головна мета оцінювання – визначення рівня засвоєння навчального матеріалу – залишається незмінною протягом історії. З цього можна зробити висновок про те, що розуміння історичного контексту є важливим для глибшого осмислення сучасних викликів у сфері оцінювання.

У [3] детально аналізується поняття "тест" у педагогічному контексті, приділяючи увагу його практичному застосуванню як методу та інструменту вимірювання. Автори підкреслюють необхідність розробки та використання тестів з дотриманням на-

укових стандартів. Вони пропонують усвідомлений підхід до тестування, що базується на знанні його теоретичних основ. Ці пропозиції є корисними для розробників та користувачів тестів, які прагнуть забезпечити якість та ефективність процесу оцінювання, зокрема в контексті адаптивного тестування, де розуміння сутності та методології тестування є критично важливим для створення валідних та надійних інструментів оцінювання.

У роботі [4] розглядаються перші наукові спроби вимірювання індивідуальних відмінностей, здійснені Френсісом Гальтоном. Автор аналізує методи, які Гальтон використовував для дослідження чутливості та інших характеристик. При цьому виділяються три ключові принципи тестування, сформульовані Гальтоном, які зберігають свою актуальність і в сучасній практиці. Підкреслюється фундаментальний внесок Гальтона у становлення теорії тестів та важливість статистичного підходу в цій галузі.

У контексті історичного розвитку педагогічного тестування, у [5] приділяється увага значній ролі Альфреда Біне в сфері діагностики інтелектуальних здібностей. Праці Біне є важливою відправною точкою для подальшого прогресу в галузі вимірювань у психології та освіті. У зв'язку з цим, пропонується застосовувати пропозиції Біне при розробці сучасних інструментів оцінювання.

Таким чином, у сучасній освіті контроль знань студентів є ключовим для управління навчальним процесом та досягнення позитивних результатів. Проведений аналіз літератури показав, що контроль знань є важливою складовою освітнього процесу, а адаптивне тестування має значний потенціал для його вдосконалення. Однак, недостатньо дослідженою залишається проблема розробки ефективних API для інтеграції адаптивного тестування в існуючі системи управління навчанням. Це дослідження спрямоване на розв'язання цієї проблеми шляхом розробки API, що забезпечує гнучкість, масштабованість та безпеку при впровадженні адаптивного тестування. Серед різноманітних технологій контролю знань особливу увагу привертає адаптивне комп'ютерне тестування, яке належить до найсучасніших підходів у цій сфері.

Виклад основного матеріалу

Розробка API для адаптивного тестування в системі управління навчальними курсами відбувається в декілька етапів.

Аналіз вимог та проєктування системи. На першому етапі визначаються функціональні і нефункціональні вимоги до системи, відбувається проєктування архітектури системи, розробляються моделі даних та проєктування бази даних, аналіз існуючих систем діагностики знань та сучасних підходів до адаптивного тестування.

Під час визначення функціональних і нефункціональних вимог до системи проведено аналіз потреб користувачів та визначені основні функції, які повинна виконувати система. Функціональні вимоги включають:

- реєстрація та автентифікація користувачів;
 - керування навчальними курсами (створення, редагування, видалення);
 - проведення адаптивних тестування;
 - реалізація алгоритму адаптивного вибору тестових завдань;
 - обробка результатів тестування.
- Нефункціональні вимоги визначають:
- продуктивність (час відгуку API на запити);
 - безпека (захист даних користувачів та системи);
 - масштабованість (здатність системи обробляти велику кількість користувачів та даних);
 - надійність (стабільність роботи системи та мінімізація збоїв);
 - зручність використання (простота інтеграції API з іншими системами).

Під час проєктування архітектури системи було визначено ключові компоненти системи, принципи їхньої взаємодії, а також технології, які були використані для реалізації розробки. У рамках даного дослідження було прийнято рішення використовувати наступні технології та підходи:

- Spring Boot для розробки серверної частини;
- RESTful API для забезпечення взаємодії між клієнтом та сервером;
- JWT-автентифікація для захисту API.

Spring Boot обраний як основний фреймворк для розробки серверної частини застосунку. Spring Boot спрощує процес створення веб-застосунків на Java, надаючи готовий до використання набір інструментів та конфігурацій. Це дозволяє зосередитися на реалізації бізнес-логіки, а не на налаштуванні інфраструктури, що відповідає вимогам швидкої розробки та масштабованості.

Для організації взаємодії між клієнтською та серверною частинами системи використано RESTful API. RESTful API є архітектурним стилем, який забезпечує стандартизований обмін даними між системами через HTTP. Цей підхід забезпечив гнучкість, масштабованість та простоту інтеграції з різними клієнтськими застосунками.

Для захисту API та забезпечення безпеки обміну даними використано JWT (JSON Web Token) автентифікацію. JWT є стандартом для безпечної передачі інформації між сторонами у вигляді JSON-об'єктів. Цей механізм дозволив перевіряти справжність запитів та надавати доступ до ресурсів лише авторизованим користувачам.

Такий вибір технологій та архітектурних рішень, як Spring Boot, RESTful API та JWT, обумовлений необхідністю створення ефективної, масштабованої та безпечної системи, яка відповідає сучасним вимогам веб-розробки та забезпечує простоту інтеграції з іншими системами.

Наступним кроком у розробці системи є визначення моделі даних.

Розробка моделі даних та проєктування бази даних є важливим етапом у створенні будь-якої системи, що працює з інформацією. На цьому етапі визначені сутності, які будуть зберігатися в базі даних, зв'язки між ними та їх атрибути.

При розробці API для системи управління навчальними курсами з підтримкою адаптивного тестування визначені наступні сутності:

- користувачі (users): інформація про користувачів системи, включаючи їхні облікові дані та ролі (викладач, студент);
- навчальні курси (courses): дані про курси, які пропонуються в системі;
- теми (topics): розділи або теми, на які поділяються навчальні курси;
- тестові завдання (questions): запитання для тестування, включаючи їхній текст, рівень складності та інші параметри, необхідні для адаптивного тестування;
- варіанти відповідей (answer_options): можливі відповіді на тестові завдання з позначенням правильної відповіді;
- результати тестування (test_sessions): інформація про проходження тестів користувачами, включаючи їхні відповіді, час проходження та результати.

Для кожної сутності визначені атрибути, що характеризують її (наприклад, для сутності users це: id, email, password, role), а також атрибути мають тип даних (наприклад, текст, число, дата), обмеження (наприклад, унікальність, обов'язковість) та інші властивості (рис. 1).

Визначені зв'язки між сутностями. Наприклад, "один-до-багатьох" між courses і topics (один курс може містити багато тем), "багато-до-багатьох" між users і courses (студенти можуть вивчати багато курсів, і курс можуть вивчати багато студентів) та інші.

```
CREATE TYPE USER_ROLES AS ENUM ('USER',
'TEACHER');

CREATE TABLE users
(
    id          BIGSERIAL PRIMARY KEY,
    email       TEXT UNIQUE NOT NULL,
    password    TEXT        NOT NULL,
    role        USER_ROLES NOT NULL
);
```

Рис. 1. Сутність users в базі даних

Для реалізації бази даних обрано систему управління базами даних (СУБД) PostgreSQL, як потужну, надійну та розширювану реляційну СУБД з відкритим вихідним кодом, яка підтримує широкий спектр типів даних, включаючи стандартні типи SQL, а також розширені типи, такі як JSON, XML та інші. Вибір PostgreSQL обумовлений її надійністю, продуктивністю, підтримкою розширених функцій та активною спільнотою розробників. Застосування PostgreSQL забезпечило високий рівень цілісності даних, підтримання транзакцій, ACID-властивостей та багато інших функцій, важливих для сучасних вебзастосунків.

Традиційні методи оцінювання, такі як іспити та контрольні роботи, мають низку недоліків, зокрема суб'єктивність оцінювання з боку викладача та ризик недостовірності результатів. На відміну від них, адаптивне тестування являє собою діалогову тестову систему, яка динамічно змінює порядок по-

дання завдань залежно від успішності виконання попередніх. Існують різні стратегії адаптивного тестування, зокрема двокрокові та багатокрокові підходи, кожен із яких характеризується власними особливостями. До основних методів адаптивного тестування належать пірамідальне тестування, тестування з будь-якого рівня та тестування на основі банку тестових завдань, реалізація яких описана на відповідних етапах дослідження.

Реалізація базової функціональності користувачів є наступним етапом у реалізації завдання для забезпечення безпеки процесу реєстрації та авторизації, щоб захистити дані користувачів від несанкціонованого доступу. На цьому етапі виконується розробка API для реєстрації та авторизації користувачів, реалізація механізмів валідації вхідних даних, хешування паролів та управління сесіями, створення сутності User та відповідного функціоналу для взаємодії з базою даних.

Розробка API для реєстрації та авторизації користувачів включає в себе:

- розробка ендпоінтів API (створення URL-адрес, за якими клієнтські застосунки можуть надсилати запити на реєстрацію та авторизацію; визначення HTTP-методів для реєстрації або авторизації, які будуть використовуватись для цих операцій);
- обробка запитів на реєстрацію (отримання даних користувача, надісланих клієнтським застосунком; перевірка валідності цих даних; перевірка, чи не існує вже користувача з таким самим ім'ям або електронною поштою);
- створення облікових записів користувачів (збереження даних користувача в базі даних; забезпечення безпечного зберігання паролів);
- перевірка облікових даних при вході в систему (отримання облікових даних, надісланих клієнтським застосунком; пошук облікового запису користувача в базі даних; перевірка правильності введеного пароля);
- повернення токенів аутентифікації (у разі успішної авторизації – генерація унікального токена, який буде використовуватись для подальших запитів до API; відправлення цього токена клієнтському застосунку; визначення терміну дії токена) (рис. 2).

Реалізація механізмів валідації вхідних даних, хешування паролів та управління сесіями є критично важливим етапом для забезпечення безпеки та надійності будь-якої вебсистеми, особливо тієї, яка обробляє конфіденційні дані користувачів.

Валідація вхідних даних запобігає введенню некоректних, шкідливих або небезпечних даних, які можуть призвести до помилок у роботі системи, витоку інформації або навіть атак. В розробці реалізовані наступні елементи валідації:

- перевірка формату електронної пошти;
- перевірка довжини пароля;
- перевірка наявності обов'язкових полів;
- перевірка типу даних.

Реалізація валідації здійснюється як на клієнтській стороні за допомогою JavaScript, так і на серверній стороні за допомогою Java (рис. 3).

```

@Override
public String generateToken(User user) {
    ApplicationProperties.JwtInfo jwtInfo =
properties.getJwt();
    String secret = jwtInfo.secret();
    Date now = new
Date(System.currentTimeMillis());
    Date expiryDate = new Date(now.getTime() +
jwtInfo.expirationTime());
    Claims claims = Jwts.claims()
        .subject(user.getEmail())
        .add("userId", user.getId() + "")
        .add("role", user.getRole())
        .build();

    return Jwts.builder()
        .claims(claims)
        .issuedAt(now)
        .expiration(expiryDate)
        .signWith(getSignInKey(secret))
        .compact();
}

```

Рис. 2. Метод для генерації токена

```

@Override
public boolean validateToken(String token) {
    String secret =
properties.getJwt().secret();
    try {
        Jwts.parser()

.verifyWith(getSignInKey(secret))
        .build()
        .parseSignedClaims(token)
        .getPayload();
        return true;
    } catch (JwtException ex) {
        log.error(ex.getMessage());
        return false;
    }
}

```

Рис. 3. Метод для перевірки токена

Під час розробки API реалізовано механізм хешування паролів користувачів. При реєстрації система обчислює хеш пароля, який зберігається в базі даних. Під час авторизації обчислюється хеш введеного користувачем пароля, який порівнюється із хешем, збереженим у базі даних. Використання хешування замість зберігання паролів у відкритому вигляді значно підвищує безпеку системи.

Також задіяний механізм управління сесіями, який дозволяє зберігати інформацію про користувача між його запитами до сервера. Це важливо для реалізації таких функцій, як авторизація користувача, збереження даних поточного тесту, персоналізація контенту. Сервер генерує JWT-токен (JSON Web Token), який містить інформацію про користувача, і передає його клієнту. Клієнт передає цей токен серверу при кожному запиті у заголовок Authorization. Також використовуються refresh tokens для отримання нових JWT-токенів без необхідності повторної авторизації користувача. Це підвищує зручність використання системи та безпеку, оскільки JWT-токени мають обмежений термін дії.

Розробка функціональності управління навчальними курсами та тестовими завданнями включає в себе наступні складові: створення сутностей Topic та Question для представлення тем та тестових завдань; реалізація CRUD-операцій для курсів

(створення, отримання списку, оновлення, видалення курсів) забезпечення зв'язку між курсами та тестовими завданнями (рис. 4).

Створення сутностей Topic та Question для зберігання інформації про теми тестування та самі тестові запитання відбувається на базі таблиць topics (поля id, name) та questions (поля id, text, difficulty, guessing, topic_id).

Застосування зовнішнього ключа topic_id в таблиці questions дозволяє пов'язати кожне тестове завдання з конкретною темою, що є складовою частиною навчального курсу.



Рис. 4. CRUD-операції та методи REST API.

Реалізація CRUD-операцій для курсів (Create, Read, Update, Delete) забезпечує стандартний набір функцій для роботи з даними у базі даних та підтримує зв'язок між курсами і тестовими завданнями. Автоматичне створення CRUD-методів відбувається в результаті застосування концепції репозиторіїв Spring Data JPA (рис. 5).

```

public interface TopicRepository extends
JpaRepository<Topic, Long> {
    Optional<Topic> findByName(String name);
}

```

Рис. 5. Репозиторій для сутності Topic

Реалізація алгоритму адаптивного тестування передбачає застосування обраної моделі адаптивного тестування. Були розглянуті наступні моделі:

- багатоступінчасте тестування на основі моделей лінійного програмування, що використовує критерій максимуму інформації та дозволяє враховувати різноманітні обмеження для різних тестових структур;

- пірамідалне тестування як один з видів багатоступінчастого підходу, де траєкторія відповідей на запитання візуалізується у вигляді піраміди;

- постійна адаптація, у якій зміна порядку пред'явлення тестового завдання відбувається на кожному кроці тестування, тобто після кожної відповіді;

- блокова адаптація, яка передбачає, що рішення про зміну порядку проходження завдань приймається на основі аналізу результатів відповідей випробуваного на спеціальному блоці завдань. При цьому завдання згруповані в блоки, а порядок їх подання всередині кожного блоку залишається фіксованим.

Зазвичай, для адаптивного тестування використовуються одновимірні моделі IRT (Item Response Theory), де рівень знань студента може бути описаний одним параметром theta. Це параметр, що відображає рівень здібностей, знань або навичок особи, є одновимірним континуумом і зазвичай має стан-

дартизовану шкалу, чим вище значення θ , тим вищий рівень здібностей.

Перед початком тестування, θ студента ініціалізується початковим середнім значенням здібностей студента (в даній реалізації запропоновано 0.5).

Процес тестування починається зі створення сутності `TestSession` для збереження інформації про тестові сесії, розробку та реалізацію алгоритму адаптивного вибору тестових завдань на основі відповідей користувача, реалізацію механізму обчислення результатів тестування з урахуванням складності завдань, збереження історії відповідей користувача.

Реалізована структура даних `TestSession`, призначена для зберігання всієї необхідної інформації, пов'язаної з кожною окремою сесією тестування. Ця сутність включає атрибути: ідентифікатор сесії, ідентифікатор користувача, час початку та закінчення тесту, поточний стан тесту, поточний номер питання, список питань у тесті, відповіді користувача на кожне питання, оцінка за тест. Використання цієї сутності дозволяє системі коректно відстежувати прогрес кожного користувача під час тестування, зберігати результати та забезпечувати можливість продовження перерваних сесій (рис. 6).

```
@Entity
@Table(name = "test_sessions")
@Data
@NoArgsConstructor
public class TestSession {
    @Id
    @GeneratedValue(strategy =
GenerationType.UUID)
    private UUID id;
    @ManyToOne
    @JoinColumn(name = "user_id")
    private User user;
    @ManyToOne
    @JoinColumn(name = "topic_id")
    private Topic topic;
    @Column(nullable = false, precision = 4,
scale = 2)
    private BigDecimal score = BigDecimal.ZERO;
    @Column(nullable = false)
    private BigDecimal currentTheta =
BigDecimal.valueOf(0.5);
    @OneToMany(mappedBy = "session", cascade =
CascadeType.ALL)
    private List<UserAnswer> answers;

    public TestSession(User user, Topic topic)
    {
        this.user = user;
        this.topic = topic;
    }
}
```

Рис. 6. Сутність `TestSession` в Java

Адаптивне тестування базується на динамічному виборі питань залежно від відповідей користувача. Основна мета цієї логіки - поступово оцінювати рівень знань користувача, використовуючи параметри складності запитань та алгоритм оновлення рівня знань. Запропоновано метод `selectNextQuestion()`, який обирає наступне питання для тесту, керуючись рівнем знань користувача. Питання підбираються таким чином, щоб їх складність мінімально відрізнялася від поточного рівня користувача (рис. 7).

```
@Override
public QuestionResponseDto
selectNextQuestion(Long userId, UUID sessionId)
{
    TestSession session =
testSessionService.getSessionByIdAndUserId(sess
ionId, userId);
    if (session.getAnswers().size() == 10) {
        throw new
MaxQuestionsAnsweredException(String.format("User
%d answered all questions", userId));
    }

    BigDecimal theta =
session.getCurrentTheta();
    return
QuestionMapper.toResponseDto(questionService.ge
tWithMinDifficultyDifference(theta));
}
```

Рис. 7. Метод для отримання нового питання

У процесі адаптивного тестування створюється модель суб'єкта навчання, яка використовується для вибору наступних завдань, що відповідають його рівню підготовки. Таке тестування має задовольняти вимоги щодо регулювання пропорцій легких, середніх і важких завдань, тематичних розділів та рівнів складності, а також передбачати адаптивний перехід на вищі рівні. Процес тестування стає формалізованим і гнучким, що дозволяє проводити адекватну оцінку знань, а можливість регулювання складності завдань базується на використанні статистичних методів для забезпечення валідної апроксимації успішності.

Проведення тестування розробленого API виконується для перевірки його працездатності та функціонування системи управління навчальними курсами в цілому. На цьому етапі здійснюється комплексна перевірка, спрямована на забезпечення коректної роботи API та його відповідності заданим вимогам. Це включає в себе перевірку функціональності окремих компонентів, таких як реєстрація, авторизація, CRUD-операції з курсами, адаптивне тестування, а також інтеграційне тестування для перевірки взаємодії між ними. Особлива увага приділяється тестуванню алгоритму адаптивного вибору тестових завдань та обробці помилок.

На рис. 8 показано, як змінюється поточна оцінка θ в залежності від відповідей студента.

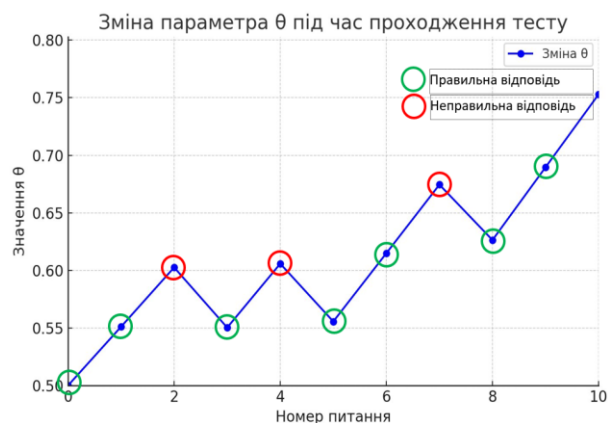


Рис. 8. Зміна оцінки θ в процесі тестування

На основі оцінки θ система обирає наступне завдання, складність якого максимально наближена до поточної оцінки θ студента. Студент відповідає на обране завдання (правильно або неправильно), що враховується у оновленому значенні оцінки θ . Якщо студент відповідає правильно на завдання, складність якого була близькою до його поточної θ , то оцінка θ зростає. Якщо студент відповідає неправильно на завдання, складність якого була близькою до його поточної θ , то оцінка θ зменшується. Масштаб зміни θ залежить від параметрів завдання. Критерієм зупинки тестування є досягнення максимальної кількості завдань.

Таким чином, θ постійно адаптується, прагнучи якомога точніше оцінити реальний рівень знань студента, на невеликій кількості заданих пи-

тань. Це дозволяє системі швидко сходиться до достовірної оцінки, не переобтяжуючи студента надто легкими або надто складними завданнями.

Також перевірено процес взаємодії користувача з системою, починаючи з реєстрації. Розглянуті сценарії успішної та неуспішної реєстрації, а також перевірка наявності облікового запису в базі даних, обробка помилок при введенні невірних даних. Перевірені генерація токенів доступу та їх оновлення.

Головна увага приділена процесу адаптивного тестування.

Детально розглянуто кроки взаємодії, отримання питань, надання відповідей, зміну складності питань залежно від відповідей, обмеження на кількість питань, обчислення та збереження результатів (рис. 9).

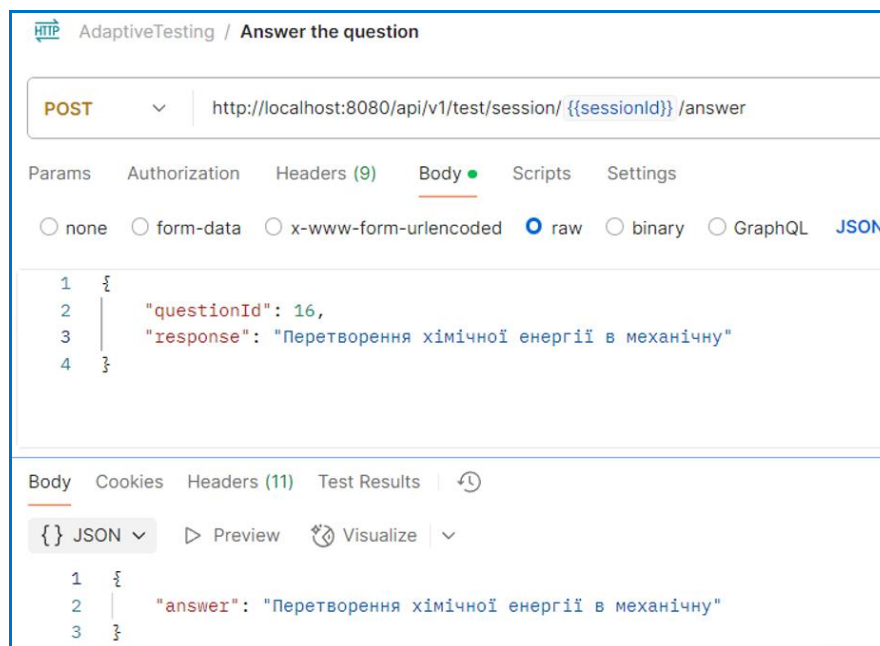


Рис. 9. Процес тестування

Висновки

У результаті проведеного дослідження обґрунтовано і реалізовано API системи управління навчальними курсами з використанням адаптивного тестування, а саме:

- здійснено аналіз сучасних підходів до комп'ютерного тестування, зокрема адаптивних методів, які забезпечують індивідуалізований підхід до оцінювання знань користувачів;
- спроектовано архітектуру системи та обрано сучасні технології (Spring Boot, REST API, PostgreSQL);
- реалізовано бізнес-логіку відповідно до функціональних вимог;
- розроблено програмний код, що демонструє роботу алгоритмів адаптивного тестування, включаючи вибір завдань на основі попередніх відповідей користувача.

Розроблене програмне забезпечення надає можливість інтеграції адаптивного тестування в освітні системи та платформи дистанційного навчання, що

сприяє індивідуалізації навчального процесу та підвищенню ефективності оцінювання знань.

Представлене дослідження має певні обмеження.

По-перше, розроблений API був зосереджений на реалізації базової функціональності адаптивного тестування, і деякі розширені можливості, такі як підтримка різних форматів тестових завдань (наприклад, аудіо- чи відеозапитання), не були включені, API був протестований на обмеженій вибірці даних, і подальша перевірка на більших обсягах може виявити додаткові аспекти, які потребують оптимізації.

По-друге, хоча в роботі розглянуто декілька моделей адаптивного тестування, реалізовано лише одну з них, що обмежує порівняльний аналіз їх ефективності.

Подальші дослідження можуть бути спрямовані на розширення функціональності API, включення підтримки інших моделей адаптивного тестування та проведення масштабніших експериментів для оцінки продуктивності та масштабованості системи.

СПИСОК ЛІТЕРАТУРИ

1. Ревуцька Н.М. "Оцінювання якості знань студентів педагогічних спеціальностей засобами тестового контролю" URL: http://www.narodnaosvita.kiev.ua/?page_id=3381 (дата звернення 30.04.2025)
2. Галузьяк В. М. Педагогічна діагностика : курс лекцій / В. М. Галузьяк, І. Л. Холковська. – Вінниця : Нілан ЛТД, 2015. – 155 с.
3. Baker, Frank B.; Kim, Seock-Ho. "Item Response Theory: Parameter Estimation Techniques". - Marcel Dekker. - 2004. – 495 p. URL: https://openlibrary.org/books/OL3383823M/Item_response_theory (дата звернення 30.04.2025)
4. Michael Bulmer. "Francis Galton—Pioneer of Heredity and Biometry". Baltimore and London: The Johns Hopkins University Press, 2003. 357 p.
5. Біографічний довідник "Видатні психологи". Київ – 2005. 120 с.
6. Ковальчук Ю.О. "Теорія освітніх вимірювань". – Ніжин 2012. – 200 с.
7. Король Т. Г. "Адаптивний тестовий контроль англomовної компетентності в читанні фахової літератури студентами немовних спеціальностей". Педагогічні науки 126' 2015 с. 86 – 94.
8. Радкевич О. "Адаптивне тестування в контексті використання електронних засобів навчання: суть, розроблення та оцінювання". Професійна педагогіка 1(26). 2023, С. 58-73.
9. О. Ф. Шевчук, А. А. Яровий, Ю. М. Паночин, С. І. Петришин, О. А. Козловський. "Моделювання адаптивного тестування знань: поріг ефективності, рівень складності та час виконання завдань" Вісник ВПІ, вип. 1, 2025. С. 104–112.
10. Patni, S., Pro RESTful APIs: Design, Build and Integrate with REST, JSON, XML and JAX-RS, 1st ed., Apress, Santa Clara, California, USA, 2017. 126 p.
11. "What is API: Definition, Types, Specifications, Documentation". 2019. URL: <https://www.altexsoft.com/blog/engineering/what-is-api-definition-types-specifications-documentation/>

Received (Надійшла) 17.04.2025

Accepted for publication (Прийнята до друку) 11.06.2025

ВІДОМОСТІ ПРО АВТОРІВ/ ABOUT THE AUTHORS

Ліпчанський Максим Валентинович – кандидат технічних наук, доцент, доцент кафедри комп'ютерної інженерії та програмування, Національний технічний університет "Харківський політехнічний інститут", Харків, Україна;

Maksym Lipchanskyi – Candidate of Technical Sciences, Associate Professor, Associate Professor of the Department of Computer Engineering and Programming, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine; e-mail: Maksym.Lipchanskyi@khpi.edu.ua; ORCID Author ID: <https://orcid.org/0000-0003-2837-0444>; Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57218514897>.

Ліпчанська Оксана Валентинівна – кандидат технічних наук, доцент кафедри комп'ютерної інженерії та програмування, Національний технічний університет "Харківський політехнічний інститут", Харків, Україна;

Oksana Lipchanska – Candidate of Technical Sciences, Associate Professor of the Department of Computer Engineering and Programming, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine; e-mail: Oksana.Lipchanska@khpi.edu.ua; ORCID Author ID: <https://orcid.org/0000-0003-4173-699X>; Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57218516877>.

Молчанов Георгій Ігоревич - старший викладач кафедри комп'ютерної інженерії та програмування, Національний технічний університет "Харківський політехнічний інститут", Харків, Україна;

Heorhii Molchanov – Senior Lecturer of the Department of Computer Engineering and Programming, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine; e-mail: Heorhii.Molchanov@epam.com; ORCID Author ID: <https://orcid.org/0000-0001-9299-461X>; Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=59921154600>.

Христофоров Кирило Андрійович – магістр кафедри комп'ютерної інженерії та програмування, Національний технічний університет "Харківський політехнічний інститут", Харків, Україна;

Kyrylo Khristoforov – Master of the Department of Computer Engineering and Programming, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine; e-mail: Kyrylo.Khristoforov@cit.khpi.edu.ua; ORCID Author ID: <https://orcid.org/0009-0008-3198-8329>.

Developing an API for adaptive testing in a learning management system

Maksym Lipchanskyi, Oksana Lipchanska, Heorhii Molchanov, Kyrylo Khristoforov

Abstract. The article is devoted to the development of an API for a learning course management system with integrated adaptive testing functionality. The relevance of the topic is due to the increasing need for individualization of the educational process and improving the efficiency of knowledge assessment in the current context of the digital transformation of education. The aim of the article is to study the principles and substantiate approaches to creating an effective, secure, and flexible API for managing learning courses. A feature of the API is its support for adaptive testing, taking into account the user's knowledge level. Within the study, an analytical review of knowledge diagnostic systems, adaptive models, and testing strategies was conducted. Based on the results obtained, the system architecture was designed, modern technologies, including Spring Boot and PostgreSQL, were chosen, and the business logic was implemented. Program code demonstrating the operation of adaptive testing algorithms, including the mechanism for selecting test tasks based on the user's previous answers, was developed. The API enables the integration of adaptive testing into various educational platforms, which increases the efficiency and individualization of the learning process. The implementation of the API opens up new opportunities for personalizing learning and improving the quality of educational services.

Keywords: API, REST, IRT, database, adaptive testing, learning course.

С. В. Носко, С. С. Бульба, О. П. Чорних, В. В. Скороделов, В. І. Панченко

Національний технічний університет «Харківський політехнічний інститут», Харків, Україна

СПОСІБ ВИСОКОЕФЕКТИВНОЇ РЕАЛІЗАЦІЇ САЙДКАР КОМПОНЕНТА З МІНІМАЛЬНИМИ ВИТРАТАМИ СИСТЕМНИХ РЕСУРСІВ

Анотація. У статті представлено спосіб щодо реалізації високоефективного сайдкар з мінімальними витратами системних ресурсів у контексті мікросервісної архітектури. Проведено аналіз існуючих рішень для реалізації сайдкару, а також виконано порівняння різних технологій з точки зору використання оперативної пам'яті та часу на запуск, що є критично важливими аспектами для сайдкар-компонентів. На основі отриманих результатів аналізу обрано Quarkus у зв'язку з тим, що його можливості компіляції ahead-of-time (AOT) на базі GraalVM забезпечують високу швидкість запуску та низьке споживання об'ємів пам'яті та центрального процесору (CPU). Для вирішення завдань із витоками пам'яті (memory leaks) запропоновано проведення оптимізації управління потоками даних, зокрема, шляхом усунення блокуючих викликів та використання асинхронних підходів. Заміна високорівневого WebClient на низькорівневий HttpClient для перенаправлення HTTP-запитів дозволило істотно зменшити використання оперативної пам'яті та уникнути помилок OutOfMemory. Вибір оптимального Garbage Collector у JVM середовищі забезпечує високу стійкість до навантажень. Розроблений та запропонований спосіб дозволяє мінімізувати додаткові витрати на інфраструктуру, зберегти переваги сайдкар-архітектури – незалежність основних мікросервісів, простоту масштабування та високу ефективність. Розроблений сайдкар стійкий до навантажень та забезпечує мінімальне використання ресурсів. Він може ефективно інтегруватися у сучасні мікросервісні системи.

Ключові слова: сайдкар, мікросервісна архітектура, оптимізація оперативної пам'яті, витік пам'яті, асинхронне програмування, мінімальне використання ресурсів, потокова передача даних, масштабованість, збірник сміття.

Вступ

Постановка проблеми. У мікросервісних системах, де використовується сайдкар-компонент, він працює поруч із основним мікросервісом та бере на себе виконання додаткових завдань, таких як: моніторинг, управління конфігурацією, забезпечення безпеки, відновлення після збоїв, комунікація з іншими компонентами системи тощо. Однак, для його розгортання та функціонування потрібні додаткові ресурси, включаючи процесорний час та оперативну пам'ять (ОП, ОЗП). У випадку мікросервісної системи, яка складається зі 100 мікросервісів, необхідно додатково розгорнути 100 сайдкарів. Це створює виклик для системи, оскільки додаткові компоненти не лише збільшують навантаження на інфраструктуру, але й можуть негативно вплинути на економічну доцільність рішення.

Проблема полягає у тому, що без ефективної реалізації сайдкар може стати значним споживачем системних ресурсів, що ускладнює масштабованість та стабільність мікросервісної архітектури. Високий рівень споживання ОЗП, зокрема через потенційні витоки пам'яті та неконтрольоване створення короткоживучих потоків, може призвести до зниження продуктивності системи та виникнення помилок OutOfMemory (OOM) при обробці великих даних (високих навантажень).

Ще одним викликом є необхідність у інтеграції додаткового компонента, який не повинен збільшувати складність або вартості інфраструктури, але, водночас, забезпечувати максимальну функціональність. Наприклад, передача HTTP-запитів через сайдкар потребує низького використання пам'яті, швидкого виконання та мінімізації витрат на серіалізацію та десеріалізацію даних.

Таким чином, актуальним науковим завданням є розробка сайдкару, який буде відповідати наступ-

ним характеристикам: мати оптимальні параметри для роботи в умовах високого навантаження; мати мінімальне споживання ОЗП та центрального процесору (CPU); бути здатним до потокової передачі даних для усунення витрат на створення проміжних буферів; бути безпечним з точки зору управління потоками та повністю позбавленим від витоків пам'яті.

Аналіз останніх досліджень і публікацій. Сайдкар патерн широко використовується у мікросервісній архітектурі.

У [1] детально описані області застосування сайдкару, його переваги і особливості, а також процес розгортання. Особлива увага приділяється високорівневим аспектам та опису мікросервісних систем на вищому рівні абстракції.

У [2] описана доцільність реалізації авторизації у сайдкарі, розглянуто популярний сервер аутентифікації і авторизації Keycloak, а також описані архітектурні діаграми взаємодії з API Gateway Kong. Розкрито проблему частого виклику Keycloak, що може стати вузьким місцем в архітектурі. Наведено ефективні і високопродуктивні рішення розв'язання проблеми з використанням Caffeine кешу.

У [3] акцентовано увагу на продуктивність сайдкарів у середовищах з високим навантаженням. Підкреслюється вплив на витік пам'яті та тривалі GC (Garbage Collection) паузи на стабільність системи. Вказується необхідність у використанні сучасних збірок сміття таких, як: ZGC та G1 GC, які забезпечують короткі паузи та ефективне очищення пам'яті.

У [4] проведено аналіз підходів до запуску сайдкарів та їх інтеграція у контейнеризованих середовищах, таких як Kubernetes.

Документація Quarkus [5] надає інструкції до розгортання швидкісних застосунків на основі цієї платформи.

Крім фреймворку Quarkus розглядалися й інші альтернативи, наприклад, Spring Boot [6]. Фреймворк демонструє значно нижчі показники ефективності у продуктивності та споживанню системних ресурсів, що є визначальним фактором при виборі технології для ефективної реалізації зазначеного функціоналу. Тому, для подальшого аналізу розглядаються модифікації стандартного Spring Boot, зокрема Spring Boot Native Image, який оптимізує продуктивність та ефективність використання ресурсів через компіляцію у нативний код.

У [7] досліджують Envoy, як високопродуктивний проху сервер, який використовується у багатьох мікросервісних рішеннях та часто виступає як стандарт де-факто для сайдкар компонентів. Основним недоліком є його складність у налаштуванні та потенційно високе споживання ресурсів.

У [8] розглядається Micronaut, який подібно до Quarkus забезпечує швидку відповідь через зниження часу на старт. Micronaut використовує Ahead-of-Time (AOT) компіляцію для вирішення залежностей та конфігурацій, що дозволяє уникнути використання рефлексії та проксі, які є основними концепціями у Spring. Основними недоліками є необхідність вивчення нового API та архітектурних підходів, зокрема, після роботи зі Spring, а також менша спільнота, що може вплинути на підтримку у майбутньому.

Мета статті – розробка способу для оптимізації ефективного сайдкар-компонента до мікросервісної архітектури, який забезпечить мінімальне споживання системних ресурсів, стійкість до високих навантажень та економічну доцільність.

Виклад основного матеріалу

За головну мету реалізації сайдкара у рамках роботи є забезпечення його високої продуктивності та економічної ефективності. Для досягнення мети використано фреймворк Quarkus у режимі native з функцією GraalVM для випереджувальної компіляції AOT, яка безпосередньо компілює код у нативний. Це дозволяє забезпечити надзвичайно швидкий час на запуск – лише кілька мілісекунд.

На відміну від традиційних Java-додатків, які компілюються у байткод Java і потім інтерпретуються або компілюються під час виконання віртуальною машиною Java (JVM), AOT-компіляція переводить додаток у нативний машинний код ще на етапі збірки. Це означає, що сайдкар не залежить від JVM для запуску та не страждає від типової затримки запуску JVM, пов'язаної з оптимізацією коду JIT-компілятором під час виконання. Натомість він стартує вже оптимізованим, забезпечуючи значно швидший час на запуск.

У режимі native додатки Quarkus компілюються у нативний виконуваний файл. Такий файл може взаємодіяти з операційною системою напряму, без посередництва JVM. У результаті, додаток починає виконання одразу після запуску, що забезпечує практично миттєвий старт.

На офіційному сайті GraalVM [9] надана діаграма, яка порівнює час на запуск додатків, написаних різними мовами програмування (рис. 1).

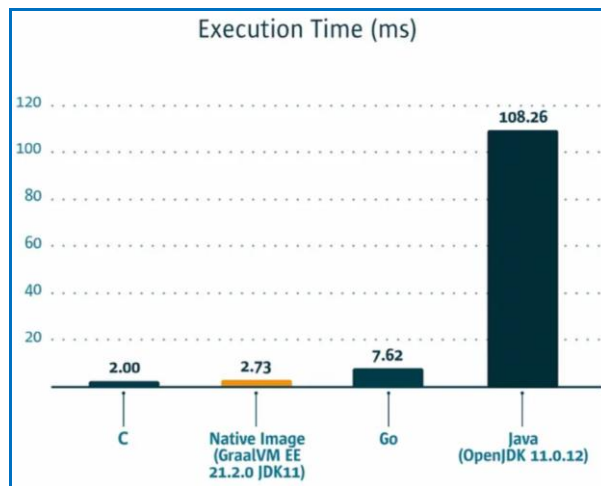


Рис. 1. Використання ОЗП різними платформами

За результатами порівняння, використання C та GraalVM показують ідентичні показники продуктивності. Однак, використання GraalVM у поєднанні з Quarkus дозволяє використовувати усі переваги високорівневої мови Java, зокрема, широкий спектр абстракцій у порівнянні з кодом на C, забезпечуючи аналогічну продуктивність до коду на C.

Для об'єктивної оцінки зазначених результатів проведено експериментальні запуски локального REST-застосунку, реалізованого за допомогою Quarkus у двох режимах роботи. Під час запуску застосунку у режимі JVM за допомогою Quarkus час на старт склав 1,07 с (рис. 2). У нативному режимі з використанням GraalVM час на запуск значно знизився до 13 мс (рис. 3), що демонструє вражаючу ефективність.

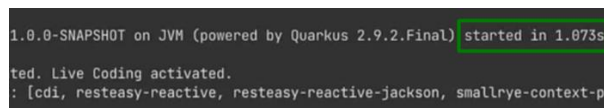


Рис. 2. Запуск Quarkus локально у JVM режимі

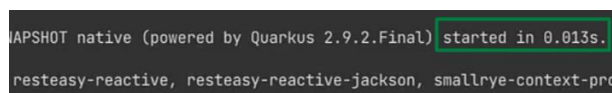


Рис. 3. Запуск Quarkus у нативному режимі

Таким чином, комбінація Quarkus та GraalVM з написанням коду на Java є оптимальним вибором, який використаний як основа для розробки високопродуктивного сайдкара.

Quarkus пропонує різноманітні внутрішні оптимізації для зменшення споживання ОЗП, що робить його дуже економічно вигідним рішенням. Серед таких оптимізацій можливо виділити такі:

- дотримання мінімальної моделі часу на виконання. Тобто, у ОЗП завантажуються компоненти, необхідні для виконання програми, що дозволяє уникнути накладних витрат на завантаження та підтримку непотрібних ресурсів у ОЗП, зменшуючи загальне споживання;
- використання стратегії вибіркового завантаження класів, де у ОЗП завантажуються лише кла-

си, які фактично використовуються програмою, що дозволяє уникнути завантаження невикористаних бібліотек та класів, що призводить до зменшення об'єму ОЗП;

– Quarkus розроблено таким чином, щоб займати мінімальний об'єм ОЗП. Він використовує такі методи оптимізації, як: стиснення метаданих JVM та видалення непотрібної інформації про налагодження, зменшуючи споживання ОЗП, необхідної для запуску програми.

На рис. 4 представлено діаграму споживання пам'яті для Quarkus, яка зображена на офіційному веб-сайті [5]. Діаграма ілюструє порівняння споживання ОЗП між різними технологічними рішеннями для розробки REST застосунків. Значна різниця у споживанні ОЗП спостерігається при використанні Quarkus у native режимі з GraalVM порівняно з традиційними підходами, що вказує на високу ефективність Quarkus для оптимізації ресурсів ОЗП.

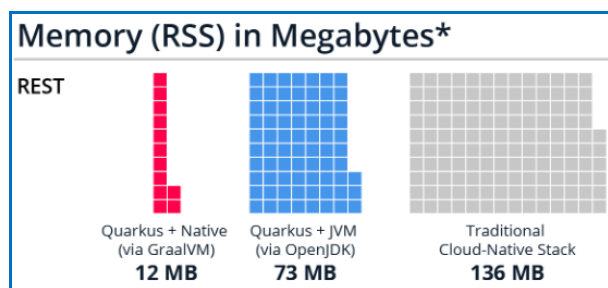


Рис. 4. Порівняння використання ОЗП у різних рішеннях

Реактивний підхід, який впроваджений у Quarkus з використанням підсистеми Vert.x, надає значні переваги для розробки легкого та високо оптимізованого сайдкара у мікросервісній архітектурі.

Сутність реактивного програмування полягає у асинхронному та неблокуючому стилі виконання, що забезпечує більш ефективне використання ресурсів порівняно з традиційним блокуючим підходом. Традиційний блокуючий підхід зображено на рис. 5.

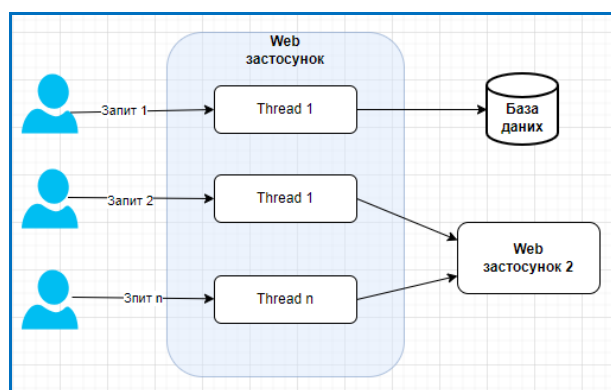


Рис. 5. Традиційний блокуючий підхід

Коли надходить запит, веб-сервер призначає створює новий потік, який буде зайнятий цим запитом, доки він не завершить процес. Процес може викликати базу даних або сторонній API, що потребує часу. За умови, якщо надходить інший запит, то він буде призначений іншому доступному потоку,

який також буде зайнятий, доки процес не завершиться. Слід зазначити, що для певного веб-сервера існує максимальна кількість потоків, які він може обробити одночасно.

На приклад, при використанні Spring Boot за замовчуванням Tomcat контейнер використовує пул на 200 потоків, розмір пулу обмежує паралельність програми. Недоліками такого підходу є те, що за умови, якщо запиту призначено свій потік, то він буде недоступним, поки запит не завершить обробку. За умови, якщо усі наявні потоки зайняті, то наступні запити, які надходять на сервер, будуть чекати поки не звільниться хоча б один потік. Крім того, кожен потік має свою вартість, використовуючи ОЗП та CPU. Великі пули потоків призводять до не легковісних застосунків, що не підходить для реалізації високоефективного та швидкого сайдкара.

Для вирішення висвітленого завдання використовується неблокуючий асинхронний підхід, де компоненти системи реагують на події замість очікування результатів операцій, що знижує потребу у блокуванні ресурсів. На рис. 6 продемонстровано один і той самий потік (Thread), який переключається на виконання Запиту 1, Запиту 2 та Запиту n. Отже, потік завжди у роботі, а ресурси використовуються найбільш ефективно без простоїв.

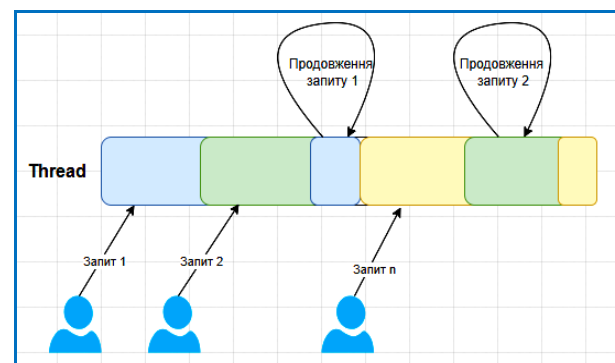


Рис. 6. Робота окремого потоку у асинхронному підході

Додатково, реалізація сайдкара з використанням фреймворку Quarkus застосовує підхід реактивного програмування, який дозволяє ефективно керувати та використовувати ресурси, на відміну від традиційного підходу із блокуючими потоками.

Реактивне програмування – це парадигма, яка спрямована на створення неблокуючих асинхронних додатків, що здатні ефективно обробляти велику кількість паралельних потоків даних. Такий реактивний підхід сприяє кращій ідентифікації та використанню усіх виділених ресурсів, забезпечуючи максимальну ефективність.

Основа Quarkus-Vert.x – це реактивний інструментарій, який дозволяє створювати легкі і швидкодіючі системи та може обробляти великі об'єми трафіку з меншою кількістю потоків ОС, що забезпечує вищий паралелізм, а навантаження на CPU і ОЗП знижується. Сайдкар, реалізований на базі Quarkus та Vert.x з транспортом Netty, не буде вимагати інтенсивного використання ресурсів для виконання своїх задач, дозволяючи головному мікросервісу

ефективно обробляти бізнес-логіку, що є особливо важливим для реалізації ефективного сайдкара.

Використання неблокуючого вводу-виводу має значні переваги, але воно не є безкоштовним. Неблокуючий ввід-вивід вводить альтернативну модель виконання, яка радикально відрізняється від традиційної блокувальної моделі у класичних фреймворків. Однак, реалізація у Quarkus зменшує таку складність, оскільки фреймворк абстрагує базовий рівень управління неблокуючими операціями за допомогою вже розробленого набору інструментів та API. Отже, розробникам не потрібно докладати додаткових зусиль для впровадження цієї моделі виконання, що сприяє більш швидкій адаптації та імplementації реактивних систем на основі Quarkus.

Проведено моніторинг роботи сайдкара під навантаженням. Для вимірювання споживання ресурсів та інших відповідних показників за різних умов/конфігурацій використано інструмент JDK Mission Control (JMC) [10]. Метою дослідження є ефективний розподіл часу та зусиль щодо оптимізації витрат та продуктивності. З метою об'єктивного порівняння роботи з GC [4], виявлення потенційних витоків пам'яті (memory leak) та оптимізації використання ОЗП, – сайдкар розгорнуто у звичайному JVM режимі замість нативного з використанням GraalVM за такими основними причинами:

- у режимі JVM надається доступ до розширених інструментів Java для профілювання, аналізу ОЗП та моніторингу таких, як: JMC, VisualVM для аналізу споживання ресурсів та ОЗП, зокрема. Така стратегія дозволяє знаходити потенційні витoki ОЗП. Нативні образи на основі GraalVM не підтримують традиційні JVM профайлерні інструменти. Оскільки GraalVM у native режимі компілює код до машинного, робота типових JVM-засобів діагностики є обмеженою;
- у JVM режимі можливо переключатися між різними збірниками сміття (Garbage Collectors) – на

приклад, G1 GC, Parallel GC, ZGC або Shenandoah GC – використовуючи відповідні опції JVM (наприклад, `-XX:+UseG1GC` або `-XX:+UseZGC`). У GraalVM native mode така можливість значно обмежена, оскільки GraalVM native образи включають лише базові GC механізми, такі як SubstrateVM garbage collector, які є більш простими та менш налаштованими.

Тестування у нативному GraalVM режимі актуальне для оцінки продуктивності та часу на запуск фінальної версії високооптимізованого сайдкара, який буде розгорнуто у production середовищі.

У дослідженні протестовано три різні типи збірників сміття для сайдкарів модулів системи: Serial GC, G1 GC та ZGC.

Метою експерименту є визначення ефективності роботи збірників сміття в умовах різних навантажень, а також – оцінки їх впливу на продуктивність та використання системних ресурсів, де вимоги до швидкодії та масштабованості постійно зростають.

Опис параметрів для оцінки наведено нижче, а детальне порівняння GC наведено у табл. 1:

- найдовша пауза – вказує на максимальну затримку, викликану збором сміття за усі спостереження. Менше – краще, оскільки це знижує ризик "фризів" додатка;
- більшість пауз – вказує типові (середні або найчастіші) затримки. Менше значення означає більш стабільну продуктивність;
- сума пауз за певний період – узагальнює загальний час, який витрачений на збір сміття за визначений проміжок часу (5 хв або 10 хв). Менші значення дозволяють приділяти більше часу на корисне виконання;
- середнє використання CPU – відображає ресурси процесора, які GC використовує, що впливає на загальну продуктивність системи.

Таблиця 1 – Тест №1 для визначення оптимального GC

Garbage Collector	Найдовша пауза	Більшість пауз	Сума пауз за певний період	Середнє використання CPU	Особливості
Serial GC	59 – 89 мс	10 – 40 мс	300 мс–3 с (на 5 хв)	3 – 5%	1. Не підходить для високих навантажень. 2. Тривалі паузи, особливо у модулях з великим memory heap. 3. Високі затримки.
G1 GC	7,5 – 45 мс	4 – 10 мс	0,1 с – 4 с (на 10 хв)	7 – 15%	1. Баланс між продуктивністю і ресурсами. 2. Помірні паузи навіть за високого навантаження. 3. Добре працює з обмеженим memory heap (64 – 160 МБ).
ZGC	78 мкс–0,5 мс (мікро-с.!)	40 – 100 мкс	2 – 100 мс (на 5 – 10 хв)	20 – 30%	1. Найшвидший збірник завдяки мінімальним паузам. 2. Ідеальний для низьколатентних систем. 3. Високе використання CPU може впливати на масштабованість.

За результатами аналізу табл. 1 можливо стверджувати, що Serial GC підходить лише для систем з малим хіпом та низькими вимогами до продуктивності. G1 GC – найкращий варіант для більшості систем, завдяки збалансованому споживанню ресурсів та невеликим затримкам. ZGC підходить для додатків із критично важливими вимогами до низької затримки, але варто враховувати підвищене споживання CPU.

Наступним етапом дослідження виконувались бізнес запити, які емулюють процес здійснення покупок у великонавантаженої системі протягом 6-х годин із 100 запитами за секунду, а також проведено аналіз поведінки сайдкара під таким навантаженням. Сайдкари продемонстрували постійно зростаючий об'єм ОЗП heap, що використовується Java-процесом під час усіх запусків, що вказує на потенційний витік пам'яті рис. 7.

Зафіксовано створення великої кількості короткоживучих потоків для екземплярів сайдкарів, що показано на рис. 8.

На рис. 9 спостерігається те, що майже уся ОЗП використовувалася об'єктами ThreadLocalMap.

За результатами проведеного аналізу виявлено, що сайдкар використовує надмірну кількість викликів блокуючого коду для доступу до AWS SSM, використовуючи метод `vertx.executeBlocking`.

Даний метод створює додаткові worker потоки, що належать до пулу потоків Vert.x. Такі виклики створюються за короткий проміжок часу.

Дана обставина пояснює створення нових потоків, які мають короткий термін існування так часто, що не встигають звільнитися так швидко.

Як рішення – усунуто надмірний виклик методу `vertx.executeBlocking` завдяки використанню кешування.

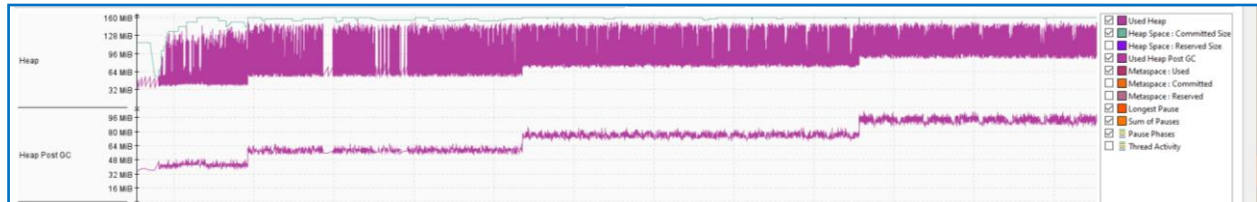


Рис. 7. Графік збільшення кількості одночасних запитів під час тесту

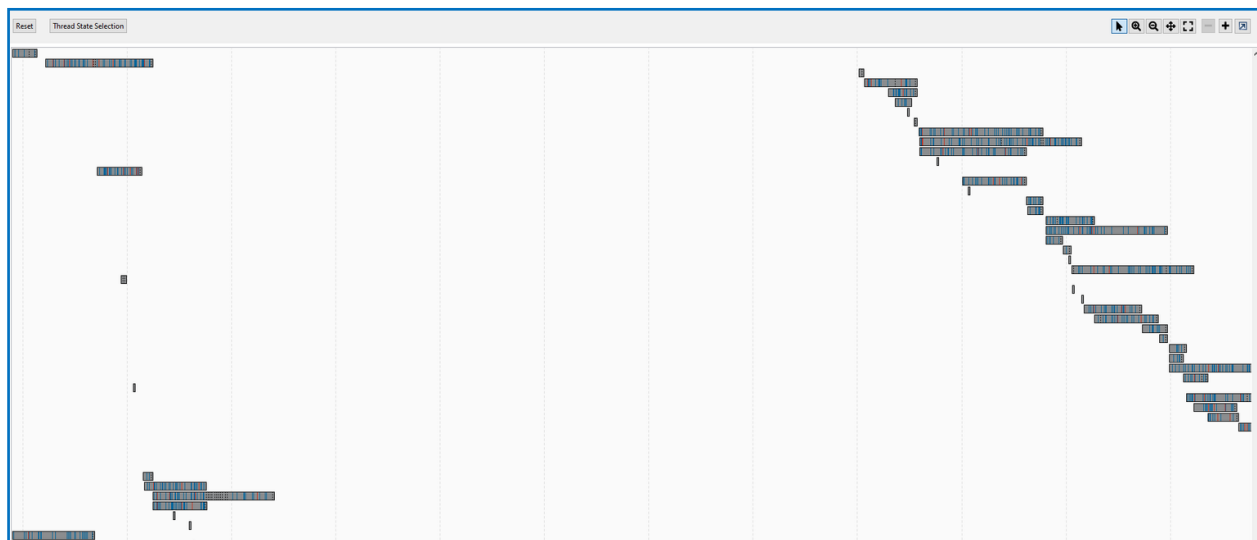


Рис. 8. Велика кількість короткоживучих потоків

i Overview default_report org.eclipse.mat.api:suspects list_objects 0xf6824698 0xf6824218 0xf6970638 0			
Class Name	Shallow Heap	Retained Heap	
<Regex>	<Numeric>	<Numeric>	
io.vertx.core.impl.VertxThread @ 0xf6d1f0f0 vert.x-eventloop-thread-1	136	8,409,392	
<Java Local> io.vertx.core.impl.VertxThread @ 0xf6d1f0f0 vert.x-eve	136	8,409,392	
threadLocalMap io.netty.util.internal.InternalThreadLocalMap @ 0xf6d1f0f0	128	8,401,888	
indexedVariables java.lang.Object[2097152] @ 0xf7e00000	8,388,624	8,400,200	
[4].... io.netty.buffer.PoolThreadCache @ 0xf6fa4ff8	56	141,992	
[9].... io.netty.buffer.PoolThreadCache @ 0xf78e2d18	56	140,496	
[2].... io.netty.channel.nio.NioEventLoop @ 0xf6d1f570	152	9,936	
[6].... java.nio.ByteBuffer[1024] @ 0xf62f96d0	4,112	4,112	
[13].... io.netty.handler.codec.CodecOutputList\$CodecOutputL	32	1,904	
[10].... io.netty.util.Recycler\$LocalPool @ 0xf6fa48e8	40	1,784	
[12].... io.netty.util.Recycler\$LocalPool @ 0xf78e2858	40	1,392	

Рис. 9. Велика кількість короткоживучих потоків

Наступною перевіркою вископродуктивної реалізації сайдкара є інтенсивна кількість операцій читання та запису даних.

Використано csv файл на 500 000 записів. У бізнес процесі імпорту такого великого файлу мікро-

сервіси під час обробки частин файлу використовують велику кількість операцій серіалізації та десеріалізації json об'єктів, які передаються через мережу.

Початкова реалізація сайдкара використовувала високорівневу реалізацію Vert.x WebClient для

усіх вхідних та вихідних викликів. До деяких переваг використання даної реалізації для комунікації між різними компонентами системи можливо віднести наступні:

1. Неблокуючий вхід-вихід. Оскільки WebClient використовує асинхронні неблокуючі I/O операції, він дозволяє сайдкару обробляти велику кількість запитів одночасно без створення вузьких місць або потреби у достатній кількості потоків для кожного запиту. Це підвищує продуктивність та зменшує загальне навантаження на систему.

2. Швидкість та ефективність реакції. Асинхронна природа WebClient, також, дозволяє сайдкару швидко реагувати на вхідні запити без затримок, призначених для обробки інших операцій.

3. Толерантність до збоїв. Використання асинхронних операцій дозволяє краще управляти помилками та збоями. WebClient може надавати механізми для повторних спроб, таймаутів та другорядних обробників помилок, що допомагає забезпечити стійкість сайдкара до збоїв у зовнішніх сервісах або мережі.

При обробці великих даних, таких як імпорт великих файлів, виникали помилки Out Of Memory (OOM). Це стало поштовхом для перегляду підходу до реалізації вихідних HTTP-запитів з використанням Vert.x WebClient та внесення оптимізацій. Vert.x WebClient забезпечує можливість сереалізації та десереалізації об'єктів, що є зручним для більшості REST-інтеграцій.

Однак, у випадку сайдкара такі абстракції зайві, оскільки основна функція сайдкара полягає у перенаправленні HTTP-запитів без будь-якої бізнес-логіки або додаткової обробки даних.

Проблема виникла через те, що у сценаріях із великими файлами Vert.x WebClient мав більш високі вимоги до ОЗП через створення буферів даних та проміжне зберігання. Це призводило до витрат ОЗП, що були неприпустимими у високонавантаженому середовищі.

З метою зменшення споживання ОЗП та забезпечення потокової передачі даних, Vert.x WebClient частково замінено на Vert.x HttpClient. HttpClient є

низькорівневим інструментом, який дозволяє напряму управляти потоками даних між серверними вхідними (server requests) та клієнтськими вихідними запитами (client requests). Зміна такого підходу базується на наступних критеріях:

1. Потреба у потоковій передачі даних (streaming): HttpClient дозволяє передавати дані напряму між запитами, оминаючи створення проміжних буферів та збільшує використання ОЗП.

2. Мінімізація витрат ресурсів: Vert.xHttpClient краще підходить для сценаріїв, де не потрібно використовувати функціонал сереалізації/десереалізації.

3. Жорстке управління ОЗП: HttpClient дозволяє уникнути витрат пам'яті на проміжну обробку, яка необхідна для WebClient.

Vert.x WebClient використовується для деяких REST-запитів таких, як: аутентифікація через Keycloak, отримання різних даних про конфігурації тощо, оскільки у даних випадках сайдкар не виступає як Proxy.

Після повторного тестування витоку пам'яті не виявлено, а збирання сміття (garbage collection) рівномірно очищає ОЗП, що представлено на рис. 10.

На рис. 11 наведено навантаження на CPU сайдкара у відсотках при 100 паралельних викликах.

Обрані технології, застосовані підходи, а також проведений аналіз продуктивності та оптимізацій дозволили створити високоефективну реалізацію сайдкара з мінімальними витратами системних ресурсів.

Висновки

Таким чином, розроблено спосіб високоефективної реалізації сайдкара компонента з мінімальними витратами системних ресурсів. Оптимізовано сайдкар, який спрямований на забезпечення мінімального використання системних ресурсів.

Проведено детальний аналіз, усунення потенційних проблем з ресурсозатратністю, а також впровадження оптимізацій, які знижують навантаження на інфраструктуру, залишаючи усі переваги сайдкар-архітектури.

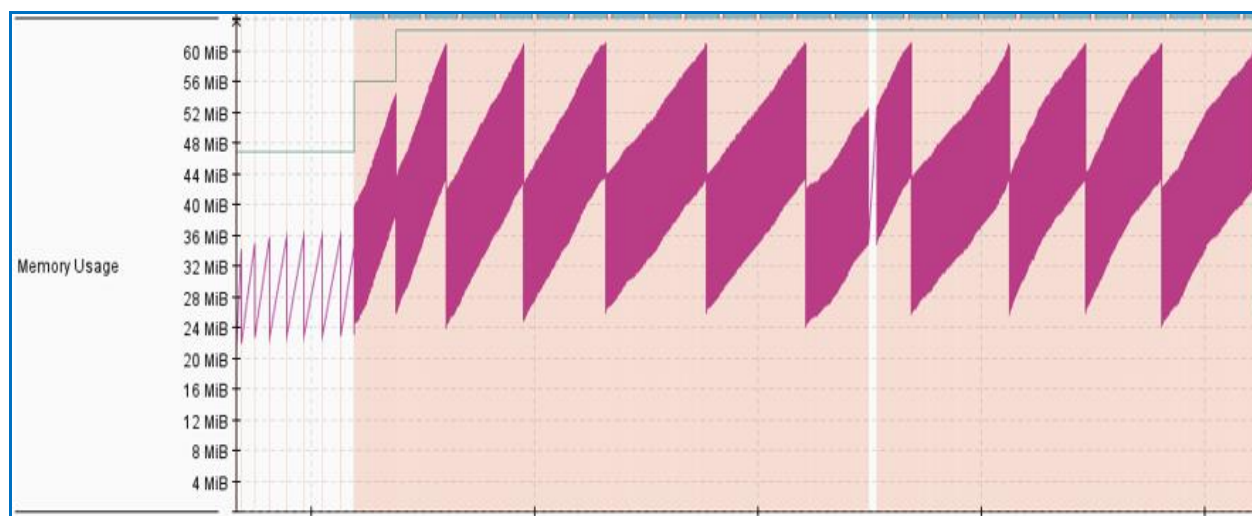


Рис. 10. Витоки пам'яті (memory leaks) не виявлено

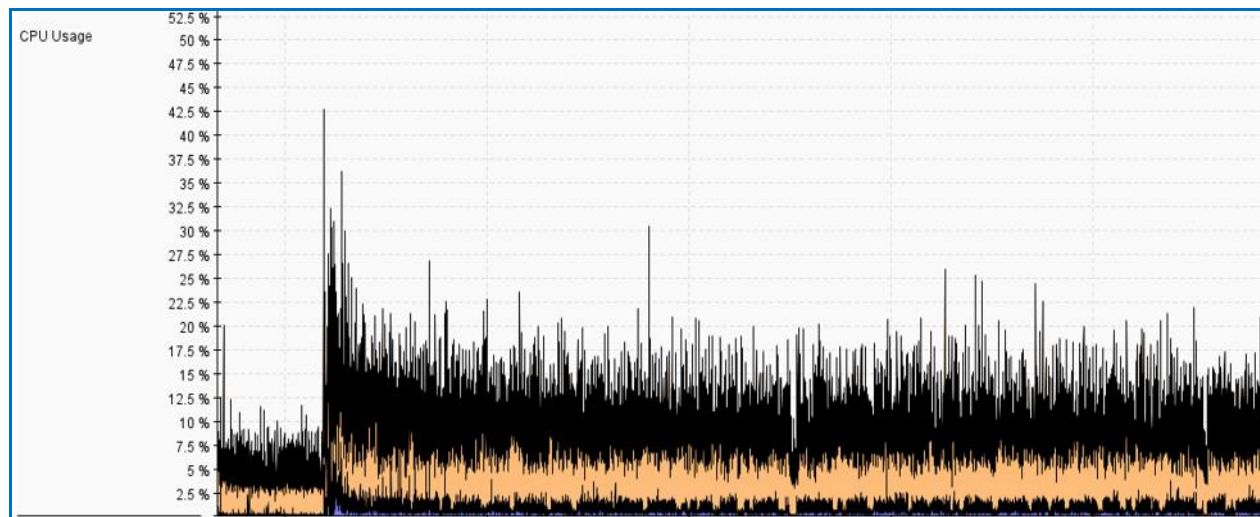


Рис. 11. CPU навантаження при 100 паралельних викликах до сайдкару

Результати показали, що завдяки обраним підходам додавання сайдкару як окремого компонента майже не вплинуло на вартість інфраструктури, але забезпечило значний функціональний та технологічний вигоди для системи.

На етапах розробки способу виявлені проблеми, які пов'язані із витоком пам'яті (memory leak) у сайдкарі. Проведено низку експериментів із використанням інструментів профілювання JVM, що підтвердило накопичення об'єктів у ОЗП, зокрема, у ThreadLocalMap та звернуло увагу на необхідність оптимізації управління потоками даних та використання ОЗП. Використання блокуючих операцій, таких як `vertx.executeBlocking`, у критичних місцях коду замінено на більш ефективні асинхронні виклики, що зменшило кількість короткоживучих потоків та оптимізувало використання ресурсів.

У процесі профілювання виявлено помилки OOM під час обробки файлів за великим об'ємом. Подальший аналіз показав, що використання високорівневого WebClient, який потребує проміжних буферів та підтримує серіалізацію/десеріалізацію даних, суттєво впливало на ресурсозатратність сайдкару. Для усунення виявленої проблеми замінено WebClient на більш низькорівневий HttpClient для усіх вихідних запитів, які не потребують додаткової обробки даних. Такий підхід реалізує повноцінний стримінг даних між сервером та клієнтом без утворення зайвих копій у ОЗП, що суттєво зменшує споживання пам'яті й уникнення помилок OOM.

Попередньо обрані технології, зокрема, використання фреймворку Quarkus у нативному режимі із застосуванням AOT compilation на основі GraalVM, забезпечили швидкий старт сайдкару та мінімізували використання ресурсів CPU та ОЗП.

За результатами проведеного тестування різних збірок сміття таких, як: Serial GC, G1 GC та ZGC, обрані оптимальні параметри роботи JVM для різних сценаріїв навантаження. Після усунення memory leaks збірка сміття стабільно очищав ОЗП, не створюючи тривалих пауз навіть за умов роботи із великим навантаженням.

Запропонований спосіб, як сукупність проведених оптимізацій продемонстрував, що сучасні підходи до реалізації сайдкару дозволяють значно зменшити потребу у ресурсах інфраструктури. Обрані технології, інструменти та методики забезпечують створення високоефективного, стійкого до навантажень сайдкару, який демонструє мінімальні витрати ресурсів та позитивно впливає на роботу системи.

Таким чином, новий спосіб дозволяє здійснити оптимізацію управління ОЗП та потоками даних, усуває блокуючі виклики, використовує більш низькорівневі інструменти для обробки HTTP-запитів, а також впроваджує потоковий механізм передачі даних.

Проведене дослідження способу демонструє, що за умови правильної реалізації сайдкар додатково покращує функціональність мікросервісної системи, майже не додаючи вартості її інфраструктури.

СПИСОК ЛІТЕРАТУРИ

1. С. С. Бульба, О. В. Коломійцев, О. І. Соловйова, С. В. Носко. Засоби побудови додаткового рівня системи комунікацій у мікро-сервісній архітектурі. Грааль науки : міжнар. наук. журнал. – Вінниця : ГО «Європейська наукова платформа», 2024. – № 46. – С. 651-659. – URL: DOI 10.36074/grail-of-science.29.11.2024.084.
2. Носко С. В., Бульба С. С., Коломійцев О. В., Лисиця Д. О., Молчанов Г. І. Пропозиції щодо авторизації в сайдкар компоненті мікросервісної архітектури. *Системи управління, навігації та зв'язку*. Полтава: НУ «ПП», 2025. № 1(7). С. 116–123. – URL: <https://journals.nupp.edu.ua/sunz/issue/view/127/68>.
3. Meadows, C., Hounsinnou, S., Wood, T., & Bloom, G. (2023). Sidecar-based Path-aware Security for Microservices. *Proceedings of the 28th ACM Symposium on Access Control Models and Technologies (SACMAT '23)*, P. 157–162. <https://doi.org/10.1145/3589608.3594742>.
4. Araldo, A., Di Stefano, A., & Di Stefano, A. (2020). Resource allocation for edge computing with multiple tenant configurations. *Proceedings of the 35th Annual ACM Symposium on Applied Computing (SAC '20)*, P. 1190–1199. <https://doi.org/10.1145/3341105.3374026>

5. Quarkus: офіційний веб-сайт. URL: <https://quarkus.io/>.
6. Parola, F., Qi, S., Narappa, A. B., Ramakrishnan, K. K., & Risso, F. (2024). SURE: Secure Unikernels Make Serverless Computing Rapid and Efficient. *Proceedings of the 2024 ACM Symposium on Cloud Computing (SoCC '24)*, 668–688. <https://doi.org/10.1145/3698038.3698558>
7. Poudel, A., Niroula, P., MacDonald, C., Gloudemans, L., & Herwig, S. (2025). Mazu: A Zero Trust Architecture for Service Mesh Control Planes. *Proceedings of the 18th European Workshop on Systems Security (EuroSec '25)*, 49–55. <https://doi.org/10.1145/3722041.3723100>
8. Basso, M., Prokopec, A., Rosà, A., & Binder, W. (2025). Improving Native-Image Startup Performance. *Proceedings of the 23rd ACM/IEEE International Symposium on Code Generation and Optimization (CGO '25)*, 689–703. <https://doi.org/10.1145/3696443.3708927>
9. Graalvm: офіційний веб-сайт. URL: <https://www.graalvm.org/native-image/>
10. JDK Mission Control: веб-сайт. URL: <https://www.oracle.com/java/technologies/jdk-mission-control.html>

Надійшла до редколегії 01.04.2025

Схвалена до друку 11.06.2025

ВІДОМОСТІ ПРО АВТОРІВ/ ABOUT THE AUTHORS

Носко Сергій Вікторович – аспірант кафедри комп'ютерної інженерії та програмування, Національний технічний університет "Харківський політехнічний інститут", Харків, Україна;

Serhii Nosko – Postgraduate of the Computer Engineering and Programming Department, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;
e-mail: Serhii.Nosko@khpi.edu.ua; ORCID Author ID: <https://orcid.org/0009-0009-5398-8650>.

Булба Сергій Сергійович – кандидат технічних наук, доцент, доцент кафедри комп'ютерної інженерії та програмування, Національний технічний університет "Харківський політехнічний інститут", Харків, Україна;

Serhii Bulba – Candidate of Technical Sciences, Associate Professor, Associate Professor of the Computer Engineering and Programming Department, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;
e-mail: bssserrega@gmail.com; ORCID Author ID: <https://orcid.org/0000-0003-0358-7516>;
Scopus ID: <https://www.scopus.com/authid/detail.uri?authorId=57216350034>.

Черних Олена Петрівна – кандидат фізико-математичних наук, доцент, професор кафедри комп'ютерної інженерії та програмування, Національний технічний університет "Харківський політехнічний інститут", Харків, Україна;

Olena Chernykh – Candidate of Physical and Mathematical Sciences, Associate Professor, Professor of the Computer Engineering and Programming Department, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;
e-mail: chernikhlena@ukr.net, Olena.Chernykh@khpi.edu.ua; ORCID Author ID: <https://orcid.org/0000-0002-4389-2446>;
Scopus ID: <https://www.scopus.com/authid/detail.uri?authorId=42461252100>.

Скородєлов Володимир Васильович – кандидат технічних наук, доцент, професор кафедри комп'ютерної інженерії та програмування, Національний технічний університет "Харківський політехнічний інститут", Харків, Україна;

Volodymyr Skorodielov – Candidate of Technical Sciences, Associate Professor, Professor of the Computer Engineering and Programming Department, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;
e-mail: Volodymyr.Skorodielov@khpi.edu.ua; ORCID Author ID: <https://orcid.org/0009-0008-1728-9089>.

Панченко Володимир Іванович – старший викладач кафедри комп'ютерної інженерії та програмування, Національний технічний університет "Харківський політехнічний інститут", Харків, Україна;

Volodymyr Skorodielov – Senior Lecturer of the Computer Engineering and Programming Department, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;
e-mail: Volodymyr.Panchenko@khpi.edu.ua; ORCID Author ID: <https://orcid.org/0000-0003-3364-3398>.

Method of the high-efficiency implementation of a sidecar component with minimal system resource consumption

Serhii Nosko, Serhii Bulba, Oleksii Kolomiitsev, Olena Chernykh, Volodymyr Panchenko

Abstract. An analysis of existing solutions for implementing a sidekick is carried out, and a comparison of different technologies is made in terms of RAM usage and startup time, which are critical aspects for sidekick components. Based on the results of the analysis, Quarkus was chosen because its GraalVM-based advance-of-time (AOT) compilation capabilities provide high startup speed and low consumption of RAM and central processing unit (CPU). To solve problems with memory leaks, it is proposed to optimize data flow management, in particular, by eliminating blocking calls and using asynchronous approaches. Replacing the high-level WebClient with the low-level HttpClient for redirecting HTTP requests significantly reduced the use of RAM and avoided OutOfMemory errors. Choosing the optimal Garbage Collector in the JVM environment provides high load resistance. Elimination of memory leaks allows the garbage collector to steadily clear RAM without creating long pauses even under heavy load. The developed and proposed method, as a set of optimizations, demonstrated that modern approaches to the implementation of sidekick can significantly reduce the need for infrastructure resources. The selected technologies, tools, and techniques ensure the creation of a highly efficient, load-resistant sidekick that demonstrates minimal resource consumption and has a positive impact on system performance. The new method allows optimizing the management of RAM and data flows, eliminates blocking calls, uses lower-level tools for processing HTTP requests, and implements a streaming data transfer mechanism. The conducted research of the method demonstrates that, if implemented correctly, the sidekick allows minimizing additional infrastructure costs, preserving the advantages of the sidekick architecture - independence of the main microservices, ease of scaling, and high efficiency. The developed sidekick is resistant to loads and ensures minimal resource utilization. It can be effectively integrated into modern microservice systems.

Keywords: sidekick, microservice architecture, RAM optimization, memory leakage, asynchronous programming, minimal resource usage, streaming data, scalability, garbage collector.

В. Д. Первєєв, Т. В. Філімончук, А. М. Бугрій, С. О. Партика

Харківський національний університет радіоелектроніки, Харків, Україна

МОДЕЛЬ МОБІЛЬНОГО ЗАСТОСУНКУ ДЛЯ ОС ANDROID

Анотація. Актуальність цього дослідження зумовлена зростаючими вимогами до продуктивності, гнучкості та можливості розширення мобільних застосунків у сучасних умовах за рахунок ефективного використання архітектурних рішень. Вибір фреймворку Flutter як середовища розробки обумовлений його можливістю забезпечувати швидкий цикл розробки, кросплатформну сумісність та інтеграцію платформозалежних сервісів. **Об'єктом дослідження** є архітектурна модель мобільного застосунку Android, яка може бути сформована обиранням окремих складових відповідно до потреб. Особливу увагу приділено можливостям розширення функціональності через впровадження додаткових модулів, таких як системи аутентифікації, аналітики та хмарних сервісів. **Предметом** дослідження виступає взаємодія компонентів мобільного застосунку в середовищі Flutter, а також вплив архітектурних рішень на продуктивність, масштабованість та ефективність розробки. У ході дослідження проаналізовано сучасні підходи до модульної архітектури мобільних застосунків, що забезпечують стабільність, швидкодію та гнучкість. **Результати** роботи демонструють, що запропонована модель здатна ефективно інтегрувати базові сервіси Android, підтримувати їх розширення та адаптацію до змінних потреб користувачів. **Висновки.** Дослідження сприяє визначенню оптимальних шляхів реалізації функціоналу мобільних застосунків, підвищенню їх продуктивності, масштабованості та конкурентоспроможності.

Ключові слова: модель мобільного застосунку, середовище розробки, Flutter, сервіси Android, модульність, продуктивність, безпека даних, масштабованість, тестування.

Вступ

Постановка проблеми. Розвиток інформаційних технологій набирає дедалі більших обертів, змінюючи різні сфери життя людини. Однією з найбільш прогресивних галузей на даний час стала розробка мобільних застосунків, які суттєво впливають на взаємодію людей із цифровими сервісами. Мобільні застосунки стали невід'ємною частиною повсякденного життя, надаючи користувачам доступ до різноманітних послуг у реальному часі. Водночас, високий рівень конкуренції на ринку програмного забезпечення вимагає від розробників забезпечення не лише інноваційного функціоналу, але й високої продуктивності, зручності та масштабованості своїх продуктів. Одним із провідних інструментів для кросплатформної розробки мобільних застосунків є фреймворк Flutter, який дозволяє створювати універсальні рішення для ОС Android та iOS. Використання Flutter надає розробникам можливість швидкої інтеграції платформозалежних сервісів, таких як геолокація, сповіщення, камера, а також реалізації унікального дизайну. Завдяки своєму інструментарію та архітектурі Flutter відкриває широкі перспективи для підвищення ефективності процесів розробки. Дана робота присвячена дослідженню складових моделі мобільного застосунку Android, яка орієнтована на використання фреймворку Flutter.

При розробці проекту мобільного застосунку головну увагу слід приділити визначенню архітектурних компонентів моделі, які реалізують логіку застосунку та відповідають за зберігання даних, а також інтерфейсу користувача та інтерфейсу програмування, який інтегрується із платформозалежними сервісами.

Аналіз останніх досліджень і публікацій. У контексті дослідження можливостей розширення моделі мобільного застосунку варто звернути увагу на кілька важливих публікацій, де розглядаються ключові аспекти архітектури мобільних застосунків,

зокрема їх модульний підхід, що дозволяє інтегрувати нові функції та адаптувати програму до сучасних вимог ринку. Особливу увагу приділено гнучкості архітектурних рішень, які забезпечують стабільність та масштабованість продукту.

У роботі [1] розглядаються основні принципи побудови архітектури мобільних застосунків з акцентом на їх ефективність та можливість масштабувати при необхідності. Автори підкреслюють значення модульного підходу, який дозволяє легко додавати нові функціональні можливості та адаптувати продукт до зростаючих потреб користувачів. Також зазначається важливість гнучкої архітектури для інтеграції нових сервісів, таких як хмарні рішення та інструменти аналітики, які забезпечують конкурентоспроможність та стійкість застосунку в умовах ринку, що швидко змінюється. У роботі [2] висвітлено процес побудови архітектури мобільних застосунків, який орієнтовано на використання модульного підходу, що забезпечує адаптивність та гнучкість системи. Автор наголошує на ролі окремих модулів, таких як інтерфейс користувача, який відповідає за візуальну складову та взаємодію з користувачем, логіку застосунку, що реалізує основний функціонал програми та базу даних, яка орієнтована на зберігання інформації. Особливе місце в роботі відведено інтерфейсу програмування, який виступає посередником між клієнтською та серверною частинами, забезпечуючи обмін даними. Також у роботі розглядається інтеграція модулів безпеки для організації конфіденційності та захисту інформації, а також пропонується можливість додавання необхідних стандартних розширень, які дозволяють легко впроваджувати новий функціонал. Завдяки такому напряду модульна архітектура спрощує масштабування застосунку, інтеграцію хмарних сервісів та підтримку змінних вимог ринку, що забезпечує конкурентоспроможність продукту. У роботі [3] розглядаються основні принципи створення мобільних застосунків із використанням архітектурних патернів для ОС Android та iOS.

Особливу увагу приділено аналізу шаблонів, таких як MVC (Model-View-Controller) та MVVM (Model-View-ViewModel), які дозволяють розробникам ефективно розподіляти функціональність між компонентами. У роботі описано переваги та недоліки цих архітектурних підходів, їх вплив на гнучкість та масштабованість кінцевих продуктів. У роботі також акцентовано на важливості модульного підходу, який забезпечує легку інтеграцію нових функціональних модулів та компонентів. Окремо розглянуто кросплатформні рішення, які дозволяють мінімізувати витрати та прискорити розробку, а також надаються рекомендації щодо вибору архітектури залежно від типу мобільного застосунку, його складності та потреб кінцевих користувачів.

У роботі [4] автори аналізують засоби та підходи до створення мобільного програмного забезпечення за допомогою мови програмування Kotlin. Особливий акцент зроблено на архітектурних рішеннях, таких як патерни MVC та MVVM, які забезпечують адаптивність, а також можливість масштабування застосунків за потреби. У роботі розглядається інтеграція базових сервісів, включаючи роботу з базами даних, інтерфейс програмного застосунку та клієнт-серверна архітектура. Автори наполягають на використанні при розробці мобільних застосунків модульного підходу, який спрощує додавання нових або модифікованих функцій та необхідного розширення можливостей додатку. Також підкреслюється важливість організації безпеки даних шляхом впровадження механізмів аутентифікації та авторизації. Завдяки використанню мови Kotlin та сучасних інструментів розробки, автори демонструють можливості створення ефективних та сучасних мобільних застосунків, що відповідають високим стандартам.

Робота [5] присвячена дослідженню архітектури мобільних застосунків з акцентом на безпеку. Автор аналізує основні вразливості архітектури, такі як небезпека втрати даних під час передачі, ненадійна аутентифікація користувачів та низька стійкість до атак. У роботі особливу увагу приділено модулю безпеки, який інтегрується в загальну архітектуру мобільного застосунку. Автор пропонує використання систем з нульовим знанням для забезпечення конфіденційності та цілісності даних, що передаються між клієнтом та сервером, що дозволяє значно підвищити рівень захисту без порушення загальної функціональності застосунку. Інші модулі архітектури, такі як інтерфейс користувача, бізнес-логіка та база даних, інтегруються із модулем безпеки для забезпечення комплексного підходу до захисту застосунку.

У роботі [6] автори аналізують сучасні підходи до побудови архітектури мобільних застосунків для платформи Android. Робота зосереджується на проектуванні модульної архітектури з використанням середовища розробки Android Studio та мови програмування Java. Особливу увагу приділено інтеграції компонентів, таких як інтерфейс користувача, база даних та бізнес-логіка, які забезпечують функціональність застосунку. Наведене дослідження також акцентує на важливості модульного підходу при розробці застосунку, який дозволяє адаптувати його до потреб користувачів. Крім того, автори розглядають особливості

проектування інтерфейсу та організації структури даних, підкреслюючи їх значення для забезпечення гнучкості та стабільності мобільного застосунку.

Проаналізувавши роботи авторів, які було наведено вище, модель мобільного застосунку можна визначити як структурований план, що включає технічні, функціональні та архітектурні елементи [7,8]. У середовищі Flutter така модель [9] може виглядати як набір компонентів (1):

$$M = \{IK, МЛЗ, МОЗД, ІПЗ\}, \quad (1)$$

де *IK* – інтерфейс користувача, *МЛЗ* – модуль логіки застосунку, *МОЗД* – модуль, що відповідає за організацію зберігання даних, *ІПЗ* – інтерфейс програмування застосунку.

Ця модель забезпечує основну взаємодію між користувачем та мобільним застосунком, але для більш складних задач вона може бути недостатньою. Наприклад, простота базової моделі може не враховувати взаємодії між компонентами, які впливають на стабільність та можуть обмежувати можливості розширення у подальшому [10]. Слід розуміти, що розробка мобільних застосунків є багатоступеневим процесом, який починається з ідеї та переходить від стадії аналізу до подальшого проектування, розробки та тестування. На кожному етапі важливу роль відіграє архітектура, яка визначає порядок реалізації встановлених задач, строки виконання проекту та його кінцеву вартість. Архітектура сприяє підвищенню продуктивності та масштабованості, що є особливо актуальним у кросплатформній розробці [11].

Метою цієї роботи є вдосконалення існуючої моделі для розробки мобільних застосунків через інтеграцію базових сервісів Android та реалізацію розширень у середовищі Flutter. Особливу увагу приділено адаптивності архітектури до змінних потреб користувачів та технологічних тенденцій, які забезпечать ефективність, гнучкість та масштабованість застосунку через впровадження модульного підходу та інтеграцію платформозалежних сервісів.

Основна частина

На основі проведеного аналізу робіт було виявлено, що більшість розглянутих моделей мобільних застосунків, орієнтовані на розробку для конкретної платформи, мають низку обмежень, зокрема складність у підтримці кросплатформного функціоналу, зростаючі витрати на розробку та недостатню гнучкість у розширенні. У зв'язку з цим у роботі запропоновано модель, яка базується на використанні фреймворку Flutter як основного середовища розробки, і складається з наступних складових (2):

$$M = \{IK, МЛЗ, МОЗД, ІПЗ, МБД, МРФМЗ, МТ, ІТ\}, \quad (2)$$

де *IK* – інтерфейс користувача, *МЛЗ* – модуль логіки застосунку, *МЗД* – модуль зберігання даних, *ІПЗ* – інтерфейс програмування застосунку, *МБД* – модуль безпеки даних, *МРФМЗ* – розширення функціональних можливостей застосунку, *МТ* – модуль тестування, *ІТ* – інструменти тестування.

У порівнянні з базовою моделлю (1) модель, що запропонована (2), значно розширена та включає нові

складові, які забезпечують повноцінну архітектуру мобільного застосунку. Інтерфейс користувача (ІК) є центральним елементом будь-якого мобільного застосунку, оскільки забезпечує взаємодію користувача із системою і будується на множині (3):

$$IK = \{DKI, OM, KI, CP, MP\}, \quad (3)$$

де *DKI* – дизайн-концепція інтерфейсу, *OM* – XML-файли для опису макетів, *KI* – компоненти інтерфейсу: кнопки, списки, анімації, які забезпечують інтуїтивність та зручність використання, *CP* – середовище розробки, *MP* – мова програмування.

У фреймворку Flutter інтерфейс користувача будується за допомогою набору віджетів, які гарантують гнучкість та адаптивність. Віджети дозволяють створювати багатофункціональні інтерфейси, що автоматично налаштовуються під різні роздільні здатності екранів, що значно покращує зручність використання. Дизайн-концепція інтерфейсу (ДКІ) охоплює створення візуального та функціонального стилю інтерфейсу. При розробці мобільних застосунків можливо використовувати UIKit або SwiftUI, які надають розробникам глибокий контроль над елементами інтерфейсу та їх поведінкою. У Flutter даний підхід реалізується за допомогою інструментів, які дозволяють створювати гармонійний дизайн із інтерактивними елементами, такими як кнопки, списки чи навігаційні панелі, забезпечуючи послідовну логіку досвіду користувача. Дуже часто при розробці мобільних застосунків використовуються XML-файли, які дозволяють здійснювати опис макетів та структур інтерфейсу, зокрема в ОС Android. За допомогою цих файлів (ОМ) у розробників є можливість задавати компоненти інтерфейсу, як-от текстові поля, кнопки, зображення, та визначати їхні взаємозв'язки, розміри та розташування, що спрощує налаштування й масштабування інтерфейсу у подальшому.

Компоненти інтерфейсу (КІ), як правило, включають базові елементи, такі як кнопки, текстові поля, графічні об'єкти, а також анімації та транзіції. У фреймворку Flutter ці компоненти представлені у вигляді віджетів, які забезпечують модульний підхід до розробки. Наприклад, віджети *AnimationController* та *Hero* дозволяють створювати плавні переходи між екранами та анімувати елементи, покращуючи досвід користувача та роблячи інтерфейс динамічним і привабливим. *Android Studio* та *Visual Studio Code* – два основні середовища розробки (СР) у Flutter. *Android Studio* забезпечує повну інтеграцію з Flutter та мовою *Dart*, має вбудований емулятор, потужні інструменти для налагодження, але споживає більше ресурсів та працює повільніше. *Visual Studio Code* легший, швидший, має гнучку систему розширень, але не містить вбудованого емулятора та менш функціональний у плані аналізу продуктивності. Якщо потрібна повноцінна інтеграція з *Android*, краще використовувати середовище розробки *Android Studio*, а для легкої та швидкої роботи з Flutter – *VS Code*. Слід також зазначити, що багато розробників комбінують обидва середовища залежно від задач, на які вони орієнтовані.

Мовою програмування (МП) у Flutter є *Dart*, яка забезпечує швидку компіляцію коду та зручну інте-

грацію з платформозалежними компонентами. Використання даної мови дозволяє реалізовувати складну логіку взаємодії між елементами інтерфейсу та забезпечує їхню інтерактивність. Іноді у розробників виникає потреба писати код за допомогою мов *Java/Kotlin* (для ОС *Android*) або *Swift/Objective-C* (для *iOS*), коли необхідно взаємодіяти з нативними API через платформні канали. Використання цього підходу зумовлено тим, що на даний час не існує Flutter-плагіну, який реалізує потрібну функціональність. Також в якості мов програмування можливо використовувати *C/C++* для роботи з низькорівневими бібліотеками (наприклад, *OpenCV*, *FFmpeg* та ін). Модуль логіки застосунку (МЛЗ) визначає бізнес-логіку, яка реалізує основний функціонал кінцевого продукту (4), включаючи обробку даних, управління бізнес-процесами та їх передачу між клієнтською (КЧ) та серверною частинами (СЧ). Також слід зазначити, що не останню роль в цьому модулі грає сервер БД (СБД):

$$MLZ = \{KCh, SCh, SBD\}. \quad (4)$$

Модуль бізнес-логіки у мобільному застосунку визначає перелік основних функцій, які забезпечують його роботу. У Flutter логіка організовується через патерни управління станом, такі як *BLoC* та *Provider*, які ізолюють бізнес-логіку від інтерфейсу користувача, що забезпечує можливість повторного використання коду, спрощує тестування та обслуговування. Наприклад, у застосунках для електронної комерції бізнес-логіка відповідає за управління коштом, обробку платежів та оновлення даних про замовлення в реальному часі. Клієнтська частина (КЧ) застосунку є фронтальною частиною, яка реалізує взаємодію користувача з системою (тут виконуються базові операції, такі як введення даних, відображення інформації та взаємодія з API). Завдяки використанню фреймворку Flutter клієнтська частина забезпечує швидке завантаження, адаптивний дизайн та інтерактивність, що робить досвід користувача зручним та привабливим. Серверна частина (СЧ) забезпечує обробку запитів, виконання бізнес-логіки на стороні серверу та управління передачею даних між клієнтською частиною та базою даних. Наприклад, сервер може обробляти запити на авторизацію, здійснювати обчислення, зберігати дані про транзакції та повертати результати в клієнтський застосунок. Висока продуктивність та надійність серверної частини є ключовою особливістю для забезпечення стабільної роботи застосунку. Сервер бази даних (СБД) відповідає за зберігання, управління та забезпечення доступу до даних. У сучасних мобільних застосунках бази даних можуть бути локальними, такими як *SQLite* або хмарними, такими як *Firebase Realtime Database*. Сервер бази даних забезпечує швидкий доступ до інформації, підтримує масштабованість та можливість обробки великих обсягів даних, що є критичним для інтерактивних та масштабованих систем.

Модуль, орієнтований на організацію зберігання даних (МОЗД) може включати наступні компоненти для управління даними (5): таблиці (Т), бази даних (БД), системи управління базами даних (СУБД), а також хмарні сховища (ХС). Використання

наведених складових забезпечує надійне зберігання, безперебійний доступ до даних та їх синхронізацію:

$$МОЗД = \{T, БД, СУБД, ХС\}. \quad (5)$$

Таблиці (Т) виступають базовими структурами, де організовуються дані у вигляді рядків та стовпців. Вони забезпечують просту та ефективну організацію даних, дозволяючи виконувати базові операції, такі як пошук, фільтрація та сортування. Бази даних (БД) є більш складними структурами, що включають множину таблиць та підтримують зв'язки між ними. Наприклад, для мобільного застосунку бази даних можуть використовуватися для зберігання профілів користувачів, історії транзакцій або інформації про товари. Сучасні мобільні застосунки зазвичай застосовують як реляційні бази даних (SQLite), так і нереляційні (Firebase Firestore) залежно від потреб проєкту.

Системи управління базами даних (СУБД) забезпечують інструменти для управління базами даних та доступу до них. Вони дозволяють створювати, змінювати, видаляти структури даних, виконувати складні запити та управляти транзакціями. Наприклад, для локальних рішень часто використовується SQLite, а для хмарних – Firebase Realtime Database або Amazon DynamoDB, які дозволяють забезпечити синхронізацію в реальному часі.

Хмарні сховища (ХС) надають можливість зберігати дані в хмарі, забезпечуючи доступ до них з будь-якого пристрою та підтримуючи синхронізацію між клієнтськими застосунками. Наприклад, Google Cloud Storage або Amazon S3 дозволяють інтегрувати зберігання файлів, мультимедіа або великих обсягів даних у мобільний застосунок, що забезпечує масштабованість та стабільність навіть за умов високих навантажень, що робить хмарні сховища невід'ємною частиною сучасних мобільних систем. Інтерфейс програмування застосунків (ІПЗ) відповідає за взаємодію між клієнтською та серверною частинами застосунку (6), забезпечуючи передачу даних та виконання відповідних операцій завдяки використанню протоколів передачі даних (ІПД), форматів даних (ФД), методів запитів (МЗ) та проміжного програмного забезпечення (ІПЗ) для аутентифікації, валідації та логування:

$$ІПЗ = \{ІПД, ФД, МЗ, ІПЗ\}. \quad (6)$$

Протоколи передачі даних (ІПД), такі як HTTP/HTTPS, WebSockets або gRPC визначають правила обміну інформацією між клієнтом та сервером. HTTPS, зокрема, гарантує безпечність передачі даних, використовуючи шифрування для захисту від стороннього доступу. Формати даних (ФД) задають структуру інформації, яка передається між компонентами системи. Найпоширенішими форматами на даний час є JSON, XML та Protocol Buffers (Protobuf). JSON забезпечує простоту та читабельність коду, що робить його популярним для роботи з API, тоді як Protobuf надає високу ефективність завдяки компактному представленню даних. Методи запитів (МЗ) дозволяють клієнту виконувати різні операції на сервері. Наприклад, метод GET використовується для отримання інформації, POST – для створення нових ресурсів, PUT або PATCH – для оновлення даних, а DELETE – для видалення. Ці методи є основою

взаємодії між клієнтом та сервером, забезпечуючи гнучкість у реалізації функціональності.

Проміжне програмне забезпечення (ІПЗ) додає додаткові рівні обробки запитів, зокрема аутентифікацію, логування та валідацію даних. Наприклад, аутентифікація гарантує, що тільки авторизовані користувачі отримають доступ до ресурсів, логування дозволяє відстежувати активність запитів та відповідей, а валідація забезпечує перевірку коректності переданих даних, запобігаючи можливим помилкам і зловживанням. Ці елементи роблять ІПЗ критично важливим компонентом для стабільної та безпечної роботи мобільного застосунку. Модуль безпеки даних (МБД) може забезпечувати захист даних у застосунку (7) через ряд інструментів: шифрування (Ш), безпечну передачу даних (БПД), захист від шкідливих атак (ЗША), аутентифікацію та авторизацію (АА), а також механізми відновлення даних (МВД):

$$МБД = \{Ш, БПД, ЗША, АА, МБД, К\}. \quad (7)$$

Модулі безпеки в архітектурі мобільних застосунків відіграють важливу роль у забезпеченні захисту даних та безпечної роботи системи. Процес шифрування (Ш) є одним із ключових елементів, який гарантує конфіденційність інформації, що передається між клієнтом та сервером, який забезпечує захист даних від несанкціонованого доступу за допомогою таких методів, як AES або RSA, що широко використовуються в сучасних мобільних застосунках.

Безпечна передача даних (БПД) забезпечується за рахунок використання протоколів, таких як HTTPS, що гарантують шифрування трафіку. Додатково, перевірка сертифікатів та мінімізація обсягу переданих даних сприяють підвищенню рівня захисту.

Захист від шкідливих атак (ЗША) включає механізми обфускації коду, мінімізацію видимості критичних компонентів та регулярний аналіз безпеки для виявлення потенційних загроз. Аутентифікація та авторизація (АА) є основою управління доступом до системи. Аутентифікація визначає особу користувача за допомогою токенів, API-ключів або багатофакторної перевірки, тоді як авторизація обмежує доступ до функцій відповідно до ролей користувачів, що запобігає несанкціонованому використанню ресурсів застосунку. Механізм відновлення даних (МВД) передбачає створення резервних копій інформації та розробку планів відновлення у разі збоїв системи, що забезпечує цілісність та доступність даних у разі критичних помилок. Кешування (К) використовується для тимчасового зберігання даних, що дозволяє зменшити навантаження на сервер та прискорити доступ до інформації, яка часто використовується.

Модуль розширення функціональних можливостей застосунку (МРФМЗ) може бути реалізовано за рахунок інтеграції нових функцій (8), таких як додаткові модулі аналітики (ДМА), інтеграція зі сторонніми сервісами (СС) або нові інтерфейсні елементи (ІЕ), що в свою чергу забезпечує гнучкість системи через інтеграцію нових модулів та інструментів:

$$МРФМЗ = \{ДМА, СС, ІЕ\}. \quad (8)$$

Під час розробки мобільного застосунку, як правило, виникають помилки в його функціонуванні.

Для вчасного їх виявлення та подальшого усунення, необхідно проводити тестування.

Комплексне тестування застосунку за рахунок додавання модулю тестування (МТ) гарантує стабільну та швидку роботу в майбутньому, сприяє ефективності та продуктивності розробки (9) за рахунок використання різноманітних типів тестів: функціональне тестування (ФТ), тестування юзабіліті (ТЮ), тестування продуктивності (ТП), тестування безпеки (ТБ):

$$MT = \{ \Phi T, TЮ, TП, TБ \}. \quad (9)$$

Функціональне тестування (ФТ) спрямоване на перевірку відповідності застосунку функціональним вимогам. Воно охоплює перевірку коректної роботи всіх модулів та функцій, таких як авторизація, обробка даних та відображення інформації в інтерфейсі користувача.

Тестування юзабіліті (ТЮ) аналізує зручність використання інтерфейсу та логіку взаємодії з користувачем. Цей етап дозволяє виявити проблеми, які можуть перешкоджати інтуїтивному розумінню та використанню застосунку. Юзабіліті-тестування забезпечує створення продукту, який відповідає потребам користувачів.

Тестування продуктивності (ТП) оцінює швидкість роботи застосунку, ефективність використання ресурсів пристрою та здатність обробляти високі навантаження, що дозволяє виявити вузькі місця, такі як повільні запити до серверів або проблеми з продуктивністю бази даних.

Тестування безпеки (ТБ) спрямоване на виявлення вразливостей у системі захисту застосунку. Воно включає перевірку надійності аутентифікації, авторизації, шифрування даних та стійкості до атак, таких як SQL-ін'єкції або XSS (міжсайтові сценарії).

Процес тестування неможливий без використання інструментів тестування (ІТ). Слід зазначити, що на даний час існує два метода тестування мобільних застосунків: ручне та автоматизоване. Для отримання найкращого кінцевого результату слід комбінувати ці підходи (9):

$$IT = \{ IAT, IPT, ITP, ITB \}, \quad (10)$$

де *IAT* – інструменти автоматизованого тестування, *IPT* – інструменти ручного тестування, *ITP* – інструменти тестування продуктивності, *ITB* – інструменти тестування безпеки.

Інструменти автоматизованого тестування (IAT) використовують для проведення трудомістких тестів, які дозволяють отримати швидкі, ефективні та точні результати. Слід зазначити, що одночасно можливо проводити декілька таких тестів на різних пристроях, що, в свою чергу, прискорює процес перевірки працездатності мобільного застосунку.

Інструменти ручного тестування (IPT) дозволяють отримати досвід користувача конкретної людини.

Також слід розуміти, що використання автоматизованих сценаріїв тестування для невеликих проєктів може виявитися надто затратним, як за людським ресурсом, так і в грошовому еквіваленті.

Інструменти тестування продуктивності (ITP)

забезпечують автоматизацію перевірок ефективності роботи застосунку. Наприклад, Firebase Performance Monitoring допомагає аналізувати швидкодію, тоді як Apache JMeter дозволяє моделювати навантаження на сервер для виявлення критичних точок.

Інструменти тестування безпеки (ITB) допомагають аналізувати вразливості застосунку. OWASP ZAP або Burp Suite можуть бути використані для перевірки безпеки мережних запитів, API та загальної стійкості системи до атак. Цей тип інструментів дозволяє своєчасно виявити та усунути потенційні загрози.

Завдяки комплексному підходу до тестування за допомогою складових модулів (9) забезпечується стабільність, продуктивність та безпека мобільного застосунку, що є критично важливим для успішного функціонування на ринку.

Висновки

В результаті проведених досліджень встановлено, що запропонована архітектурна модель мобільного застосунку, реалізована у середовищі Flutter, забезпечує високий рівень гнучкості, продуктивності та масштабованості.

Інтеграція базових сервісів Android, дозволяє створювати функціональні рішення, адаптовані до сучасних потреб користувачів.

Головною метою при розробці конкретного мобільного застосунку є визначення архітектурних компонентів моделі: для реалізації логіки застосунку, збереження даних, розробки інтерфейсу користувача та використання інтерфейсу програмування мобільних застосунків, який інтегрується із платформозалежними сервісами. Важливим аспектом моделі, що запропонована, є можливість розширення функціоналу через інтеграцію сторонніх модулів, що дозволяє швидко реагувати на зміни ринкових умов та потреби користувачів.

Використання модулю безпеки, який включає аутентифікацію, авторизацію та шифрування даних, гарантує конфіденційність та цілісність інформації, що є критичним для застосунків, які працюють із чутливими даними.

Використання інструментів моніторингу та аналітики дозволить підвищити продуктивність застосунку, виявляти можливі проблеми та своєчасно їх вирішувати.

Хмарні сервіси також надають додаткові переваги у зберіганні даних, синхронізації між пристроями та оптимізації витрат на інфраструктуру.

Загалом, запропонована модель мобільного застосунку демонструє свою ефективність у забезпеченні якісного кінцевого продукту, який відповідає вимогам користувачів та сучасним стандартам. Її гнучкість, адаптивність та надійність роблять її перспективною для широкого використання у розробці мобільних додатків на платформі Android.

Загальний висновок полягає в тому, що в залежності від цілей та направленості мобільного застосунку, який розробляється, може бути обрано різні напрями його проєктування, реалізації, інструментарію та подальшого тестування.

СПИСОК ЛІТЕРАТУРИ

1. Бідник Д.С., Цибульник С.О. (2021), "Архітектура мобільного додатку", Матеріали XIV Всеукраїнської НПК студентів, аспірантів та молодих вчених "Погляд у майбутнє приладобудування". К: КПІ ім. Ігоря Сікорського. С. 19-22.
2. Сідельнікова Д.С. (2022), "Етапи створення дизайну мобільного застосунку", Мультимедійні технології в освіті та інших сферах діяльності: науково-практична конференція з міжнародною участю. К.: НАУ. С.105-109.
3. Ставицький П.В., Войтко В.В. (2017), "Організація архітектури додатків на базі мобільних платформ", Матеріали XLVI науково-технічної конференції підрозділів ВНТУ, Вінниця. Електрон. текст. дані. URL: <https://conferences.vntu.edu.ua/index.php/all-fitki/all-fitki-2017/paper/view/2794>.
4. Козуб Г.О., Козуб Ю.Г., Могильний Г.А., Жуков А.В. (2021), "Розробка мобільного Android-додатку з застосуванням принципів Clean Architecture", *Вісник Східноукраїнського національного університету імені Володимира Даля*. №5 (269). С. 5-10. doi: <https://doi.org/10.33216/1998-7927-2021-269-5-5-10>.
5. Санак О.Є. (2018), *Захист мобільних застосунків на основі систем з нульовим знанням: магістерська дис.: 125 Кібербезпека*. Київ. 121 с.
6. Дворецький М.Л., Нездолій Ю.О., Дворецька С.В., Кандиба І.О. (2021), "Розробка мобільних застосунків для OS Android": навч. посіб. Миколаїв: Вид-во ЧНУ ім. Петра Могили. 140 с.
7. Яшина К.В., Ялова К.М., Сугаль Є.О. (2017), *Огляд гнучких методологій розробки програмного забезпечення*, *Збірник наукових праць Дніпровського державного технічного університету*. Том 1, №30. С. 153-156.
8. Вавіленкова А.І. (2021), "Аналіз гнучких методологій розробки програмного забезпечення для реалізації у командних проєктах", *Вісник Національного технічного університету "ХПІ"*. Харків: НТУ "ХПІ", 2021. №1 (7). С. 39-46.
9. Сеспедес Гарсія Н.В., Сеспедес Гарсія П.Д. (2023), "Моделі життєвого циклу розробки програмного забезпечення", *Молодий вчений*, №2 (114). С. 17-20. doi: <https://doi.org/10.32839/2304-5809/2023-2-114-4>.
10. Sommerville I. (2015), "Software Engineering". Publisher: Pearson, 10th edition. 816 p.
11. Pressman R.S. (2019), "Software Engineering: A Practitioner's Approach". Publisher: McGraw-Hill Higher Education; 9th edition. 692 p.

Received (Надійшла) 03.04.2025

Accepted for publication (Прийнята до друку) 28.05.2025

ВІДОМОСТІ ПРО АВТОРІВ/ ABOUT THE AUTHORS

Перевсєв Володимир Дмитрович – магістрант кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Volodymyr Pervieiev – master's student of the Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: volodymyr.pervieiev@nure.ua; ORCID Author ID: <https://orcid.org/0009-0004-9776-1283>.

Філімончук Тетяна Володимирівна – кандидат технічних наук, доцентка, доцент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Tetiana Filimonchuk – PhD, Associate Professor of the Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: tetiana.filimonchuk@nure.ua; ORCID Author ID: <http://orcid.org/0000-0002-4380-504X>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57190949991>.

Бугрій Андрій Миколайович – кандидат технічних наук, старший викладач кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Buhrii Andrii – Associate Professor of the Department of Electronic Computers, Senior lecturer at the Department of Electronic Computers, Kharkiv National University of Radio Electronics; Kharkiv, Ukraine;

e-mail: andrii.buhrii@nure.ua; ORCID Author ID: <http://orcid.org/0009-0002-9059-3200>.

Партика Станіслав Олександрович – старший викладач кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Stanislav Partyka – Senior lecturer of the Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: stanislav.partyka@nure.ua; ORCID Author ID: <http://orcid.org/0000-0002-7376-8980>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57204560890>.

Model of an android mobile application with integration of basic services and extensions

Volodymyr Pervieiev, Tetiana Filimonchuk, Andriy Buhrii, Stanislav Partyka

Abstract. The relevance of this research is driven by the growing demands for the performance, flexibility, and scalability of mobile applications in modern conditions. The choice of Flutter as the development environment is justified by its ability to provide a fast development cycle and cross-platform compatibility. **The object of the study** is the architectural model of an Android mobile application that includes the integration of basic services such as geolocation, notifications, and the camera, offering approaches for extending functionality through the addition of new modules. The model consists of interconnected components, including the user interface, business logic, access to APIs, and platform-specific services. Particular attention is paid to the modularity of the architecture, ensuring the model's adaptability to changing user needs and technological trends. **The subject of the study** is the interaction of components within the mobile application model in the Flutter environment. The study examines the impact of architectural decisions on performance, scalability, and development efficiency. **Results.** The proposed model demonstrates the ability to integrate basic Android services and support their extensions, ensuring stability and high product quality. **Conclusions.** Research into the mobile application model in the Flutter environment helps identify optimal ways to implement functionality, improve performance, and enhance the competitiveness of mobile applications.

Keywords: mobile application, application model, Flutter, architectural model, Android services, modularity.

Д. А. Помелуйко¹, Н. С. Єрьоміна¹, С. В. Петров², Н. В. Іщенко², Волошук О. Б.¹

¹ Харківський національний університет радіоелектроніки, Харків, Україна

² Харківський національний університет імені В.Н. Каразіна, Харків, Україна

МЕТОДИ МАСШТАБУВАННЯ KUBERNETES КЛАСТЕРУ В ХМАРНОМУ СЕРЕДОВИЩІ

Анотація. У статті розглянуто питання масштабування кластерів Kubernetes у хмарному середовищі, що є одним із ключових чинників забезпечення високої доступності мікросервісних додатків. Проаналізовано традиційні засоби та сучасні підходи, включно з механізмами вертикального та горизонтального масштабування. Розглянуто також такі методи, як передбачувальне масштабування, масштабування на основі подій та застосування спеціальних метрик. Проведено аналіз переваг і недоліків розглянутих підходів. Показано, як налаштування й оптимізація параметрів масштабування впливають на витрати на хмарні ресурси й продуктивність додатків. **Мета статті:** аналіз сучасних інструментів та підходів до масштабування Kubernetes-кластерів у хмарному середовищі AWS, виявлення переваг та недоліків кожного із наявних рішень. **Висновки.** Масштабування на основі подій є особливо корисним для навантажень з нерегулярною або імпульсною природою. Перспективними напрямками являється розробка динамічних механізмів зміни ресурсів контейнера під час виконання, одночасне використання горизонтального і вертикального масштабування, застосування гібридних підходів для мікросервісів, а також розширення досліджень на рівні інфраструктури. Вибір стратегії масштабування має базуватись на характеристиках робочих навантажень, бюджетних обмеженнях і загальних операційних цілях організації. Тонке налаштування конфігурацій масштабування та постійний моніторинг поведінки застосунків є критично важливими для досягнення ефективного використання ресурсів, підвищення продуктивності та зниження витрат у cloud-native середовищі.

Ключові слова: Kubernetes, масштабування кластеру, хмарне середовище, HPA, Horizontal Pod Autoscaler, VPA, Cluster Autoscaler, Karpenter, KEDA, Kubernetes Event Driven Scaling, Prometheus, EKS, AWS.

Вступ

Постановка проблеми. Зі зростанням попиту на динамічні та масштабовані застосунки традиційні методи розгортання та масштабування дедалі більше втрачають актуальність. Останніми роками мікросервісна архітектура стала провідним підходом до створення сучасних, гнучких та високодоступних застосунків. Проте керування мікросервісами, контейнерами та їхньою взаємодією є складним завданням, особливо у великих і складних системах.

Розвиток хмарних обчислень зумовив суттєві зміни у проєктуванні й експлуатації програмних систем. Поширення архітектури мікросервісів призвело до потреби автоматизації розгортання та обслуговування великої кількості взаємопов'язаних сервісів. Контейнерна оркестрація за допомогою Kubernetes стала стандартом для таких задач. Kubernetes сьогодні є одним із найпопулярніших інструментів для оркестрації контейнерів, який широко застосовується постачальниками хмарних сервісів. Він забезпечує керування життєвим циклом контейнеризованих застосунків. Для підтримання показників продуктивності застосунків і дотримання угоди про рівень обслуговування (SLA) упродовж часу, за умов змінних робочих навантажень або доступності ресурсів, ефективне масштабування відіграє ключову роль.

У Kubernetes реалізовано декілька вбудованих механізмів масштабування. Проте їхніх можливостей може не вистачати для належного пристосування до високодинамічних навантажень.

Метою роботи є аналіз сучасних інструментів та підходів до масштабування Kubernetes-кластерів у хмарному середовищі AWS, виявлення переваг та недоліків кожного із наявних рішень.

Основний матеріал

Із дедалі ширшим поширенням хмарно-орієнтованих архітектур і мікросервісного підходу проєктування, дедалі гостріше постає потреба ефективного масштабування застосунків, яка породжує унікальні виклики, що потрібно враховувати для забезпечення ефективності, надійності і помірної вартості рішень.

Перехід від монолітної до розподіленої архітектури збільшує складність керування застосунком. У розподілених системах багато сервісів та компонентів взаємодіють у межах різних вузлів і середовищ, що створює додаткові складнощі. Нижче наведено ключові проблеми, з якими зазвичай стикаються під час масштабування застосунків:

- Комунікація та координація. Різні сервіси мають безперешкодно взаємодіяти один із одним, що стає складнішим зі збільшенням кількості сервісів і вузлів.

- Керування станом. Підтримка цілісності стану системи в розподіленому середовищі є важливим аспектом, особливо коли йдеться про події, що відбуваються одночасно, та потребують узгодження оновлень між різними сервісами.

- Спостереження та налагодження. Пошук і виправлення помилок у розподіленій системі складніші, ніж у монолітній.

- Ефективний розподіл обчислювальних ресурсів. Один із важливих аспектів масштабування. Зі зростанням навантаження, збільшується й споживання ресурсів, що може призвести як до нестачі ресурсів, так і до їх перевитрати.

- Непередбачуваність трафіку. Багато застосунків стикаються із проблемою непредбачуваності трафіку, викликаного сезонними процесами та яви-

щами, рекламними кампаніями чи раптовою популярністю.

- Балансування навантаження. Зі збільшенням кількості копій застосунку Kubernetes використовує сервіси й механізми балансування навантаження для розподілу трафіку між подами. Однак, якщо трафік дуже нерівномірний, балансування може бути неідеальним, спричиняючи диспропорцію використання ресурсів.

Хоча Kubernetes здатен суттєво спростити ці аспекти завдяки своїм вбудованим засобам масштабування, серед яких головними є:

- Horizontal Pod Autoscaler (HPA) для масштабування кількості подів;
- Vertical Pod Autoscaler (VPA) для зміни розміру ресурсів подів;
- Cluster Autoscaler (CA) для зміни кількості вузлів в кластері.

Ці інструменти також мають й свої недоліки. Так, за даними одного дослідження, понад 50 % організацій використовують HPA, тоді як менш ніж 1 % – VPA, що обумовлено складністю конфігурації та можливими конфліктами при спільному використанні з HPA [1]. CA надійно масштабує вузли, але працює з попередньо визначеними групами вузлів і застосовує менш ефективний підхід [2], що може призводити до надлишку або недостатчі обчислювальних ресурсів в кластері.

Масштабування також часто супроводжується затримками, необхідними для запуску нових вузлів чи розгортання подів. Ця проблема відома як проблема холодного запуску (cold start problem). Вона полягає в наступних складностях:

- запуск вузла в AWS (EC2 instance) триває від кількох десятків секунд до декількох хвилин;
- завантаження контейнерних образів може займати чимало часу, в залежності від розміру образу та пропускної здатності мережі;
- ініціалізація застосунку всередині контейнера (з'єднання з БД, кешування, зовнішні залежності).

Для зменшення впливу цих проблем використовують різні механізми, зокрема метод завчасного прогрівання вузлів (warm pools), зменшення образу контейнерів (використання alpine базових образів), а також масштабування на основі подій (KEDA), що дозволяє швидше реагувати на зміни в системі.

Під горизонтальним масштабуванням Kubernetes кластеру розуміють збільшення (чи зменшення) кількості подів (Pods) або робочих вузлів (Nodes). Це дає змогу оперативно реагувати на коливання навантаження: чим більше копій, тим вище загальна пропускна здатність сервісу. Стандартний Horizontal Pod Autoscaler (HPA) передбачає ручне налаштування певних порогових значень, як-от завантаження CPU, мінімальна та максимальна кількість подів. Деякі дослідження пропонують поліпшити цей механізм, залучаючи додаткові метрики (наприклад, аналіз трафіку [3] або час відгуку [4]). Крім того, є дослідження, що додають у HPA машинне навчання чи евристичний аналіз [5, 6].

Вертикальне масштабування передбачає зміну обсягу виділених контейнерам ресурсів. У Kuber-

netes вертикальне масштабування реалізується за допомогою Vertical Pod Autoscaler (VPA). Воно дає змогу подам ефективніше використовувати доступні ресурси, однак потребує перезапуску подів при зміні конфігурації. Через таке обмеження вертикальне масштабування користується меншим менше уваги.

Гібридне масштабування поєднує горизонтальне та вертикальне масштабування. Наявні рішення [7] зазвичай виконуються поступово: спочатку визначають оптимальний обсяг ресурсів для контейнерів за допомогою вертикального масштабування, а потім динамічно змінюють кількість контейнерів методом горизонтального масштабування.

Засоби масштабування Kubernetes кластеру наступні:

- Horizontal Pod Autoscaler;
- Vertical Pod Autoscaler (VPA);
- Cluster Autoscaler (CA);
- Karpenter;
- Масштабування на основі подій;
- Масштабування на основі спеціальних метрик;
- Методи передбачувального масштабування.

Horizontal Pod Autoscaler дозволяє автоматично регулювати кількість реплік подів у Deployment або ReplicaSet на основі показників використання процесора, пам'яті або інших зазначених метрик. Контролер періодично проводить порівняння фактичних значень показників із цільовими значеннями і розраховує бажану кількість копій для підтримки заданого рівня навантаження [9]. Deployment або ReplicaSet в свою чергу відповідає за підтримання бажаної кількості копій подів. Такий процес забезпечує оптимальне використання ресурсів та відповідність наявних ресурсів потребам застосунків за різних типів навантаження. Налаштування HPA включає задання бажаного рівня завантаження процесора чи інших показників, а також визначення мінімальної і максимальної кількості копій подів. У табл. 1 наведено порівняння основних типів метрик HPA, їх переваги, недоліки та типові сценарії використання.

Vertical Pod Autoscaler (VPA) забезпечує автоматичне коригування запитів і лімітів на ресурси (CPU та пам'ять) для контейнерів у подах. На основі аналізу історичних даних про споживання ресурсів та поточного використання VPA рекомендує або автоматично призначає найоптимальніші значення ресурсів для кожного поду [10]. Основна мета використання VPA полягає в підвищенні ефективності використання ресурсів кластера, а також у покращенні продуктивності додатків завдяки динамічному регулюванню ресурсів. VPA особливо ефективний у сценаріях, коли вимоги до ресурсів є нестабільними або коли початкове налаштування ресурсів виконується вручну. Використання VPA дозволяє уникнути ситуацій недостатнього або надмірного виділення ресурсів, що призводить до зниження витрат на обслуговування кластерів. Цей підхід до масштабування є корисним для додатків із періодичними або поступовими змінами навантаження, які не потребують великої кількості ресурсів одночасно.

Таблиця 1 – Основні типи метрик для масштабування за допомогою HPA

Тип метрики	Опис	Сценарій використання	Переваги	Недоліки
CPU utilization	Відсоток завантаження CPU	Універсальні контейнери	Простота налаштування, вбудована метрика	Не завжди є показником завантаженості контейнеру
Memory utilization	Відсоток використовуваної оперативної пам'яті (ОП)	Контейнери з інтенсивним використанням ОП	Попередження помилок OutOf Memory, вбудована метрика	Можливе надлишкове виділення ресурсів
Спеціальні метрики	Специфічні показники контейнеру	Спеціалізовані робочі навантаження	Висока точність масштабування	Складність в налаштуванні
Зовнішні метрики	Показники зі зовнішніх джерел	Хмарні сервіси, черги	Масштабування за зовнішніми умовами	Залежність від зовнішніх систем

VPA може використовуватися одночасно з HPA, який відповідає за зміну кількості копій подів, тоді як VPA регулює ресурси всередині цих подів. Така інтеграція дозволяє більш точно контролювати масштабування контейнерів, ефективно використовуючи ресурси та оптимізуючи продуктивність застосунків залежно від їхніх потреб [8].

Cluster Autoscaler (CA) є важливим компонентом Kubernetes, що автоматично змінює розмір кластера шляхом додавання або видалення вузлів залежно від наявності ресурсів для розміщення подів. Він складається з головного циклу, який періодично перевіряє стан кластера та шукає поди, які неможливо розмістити через нестачу ресурсів. Потім CA взаємодіє з API відповідного хмарного провайдера для управління групами вузлів та динамічного створення чи видалення вузлів кластеру. Кожен хмарний провайдер має параметри, які впливають на роботу Cluster Autoscaler. Наприклад, типи та вартість віртуальних серверів, а також можливості API для управління групами вузлів відрізняються у провайдерів таких, як AWS, GCP чи Azure [11]. Тому налаштування Cluster Autoscaler може потребувати додаткових коригувань з метою досягнення максимальної ефективності та економічності кластера залежно від конкретного хмарного середовища.

Karpenter контролер автоматично оптимізує тип, розмір та кількість вузлів, необхідних для виконання робочих навантажень. На відміну від CA, що працює із заздалегідь визначеними статичними групами вузлів, Karpenter забезпечує вищий рівень гнучкості завдяки динамічному підбору найоптимальніших типів віртуальних серверів і оперативному масштабуванню у режимі реального часу. Контролер реагує на зміни у кластері протягом кількох секунд, тоді як реакція CA зазвичай є повільнішою. Це особливо важливо для робочих навантажень, які потребують оперативного масштабування у відповідь на різкі зміни навантаження. Karpenter створений з метою максимально ефективного використання доступних ресурсів завдяки вибору економічно вигідних доступних типів віртуальних серверів [12]. Він забезпечує розгортання лише тих ресурсів, які необхідні для запуску робочих навантажень. Це робить Karpenter особливо важливим для організації, які прагнуть зменшити витрати на хмарну інфраструктуру. При використанні Karpenter немає необхідності керувати окремими групами вузлів вручну, адже він автоматично налаштовує тип та розмір серверу відповідно до поточних потреб навантаження.

Це значно спрощує процес налаштування, роблячи його більш зрозумілим і доступним для розробників.

Масштабування на основі подій можливе за допомогою Kubernetes Event-driven Autoscaling (KEDA) – контролеру, що дозволяє масштабувати контейнери відповідно до подій, таких як повідомлення в чергах, запис у базах даних або власні спеціальні метрики [13]. Цей інструмент ефективний для сценаріїв, де навантаження є нерегулярним, наприклад, у разі безсерверних функцій чи задач обробки даних, які залежать від подій. Масштабування на основі подій ефективно використовується в реальних умовах, таких як мікросервісна архітектура, IoT та пакетна обробка даних. Наприклад, система обробки платежів може масштабуватись залежно від довжини черги транзакцій, а сервіс обробки зображень — відповідно до кількості завантажень зображень. Такі додатки мають значні переваги від цього типу масштабування через їхню непередбачувану природу навантажень.

Масштабування на основі спеціальних метрик можливе за допомогою Prometheus адаптера – інструменту, що дозволяє використовувати метрики Prometheus для масштабування Kubernetes. Він транслює запити Prometheus у формат, зрозумілий для Custom Metrics API, завдяки чому HPA може використовувати спеціальні метрики для масштабування [14]. Це дає змогу адаптувати політики масштабування до специфічних особливостей додатків, таких як швидкість запитів, довжина черг або інші бізнес-показники. Custom Metrics API розширює можливості масштабування Kubernetes, дозволяючи використовувати не лише показники CPU чи пам'яті, але й довільні метрики, специфічні для конкретних застосунків.

Розширені стратегії використання спеціальних метрик для масштабування передбачають застосування декількох метрик одночасно, використання швидкості зміни показників або відсоткових метрик для більш точного прогнозування необхідних ресурсів. Подібні стратегії потребують детального аналізу поведінки додатків та точного налаштування, але дозволяють значно підвищити ефективність процесу масштабування.

Методи передбачувального масштабування. Передбачувальне масштабування, яке ґрунтується на методах машинного навчання, використовує накопичені дані про споживання ресурсів для прогнозування майбутніх потреб мікросервісів. Для цього застосовуються регресійний аналіз, прогнозування

часових рядів та нейронні мережі [4]. Такі моделі дозволяють аналізувати історичні дані та зовнішні фактори, що впливають на навантаження, з метою прогнозування майбутніх ресурсних потреб і, таким

чином, зменшення часу реакції на зміну навантаження.

У табл. 2 наведено порівняльну характеристику методів передбачувального масштабування.

Таблиця 2 – Методи передбачувального масштабування

Модель	Техніка	Переваги	Недоліки	Складність
Лінійна регресія	Прогноз часових рядів	Простота, прозорість	Тільки лінійні залежності	Низька
ARIMA	Аналіз часових рядів	Враховує тренди та сезонність	Вимагає стаціонарності даних	Середня
Нейронні мережі	Глибоке навчання	Виявляє складні закономірності	Необхідність великих наборів даних, низька інтерпретованість	Висока
Prophet	Адитивна модель	Обробляє сезонність та пропуски даних	Можливе надмірне спрощення складних залежностей	Середня

Моделі аналізу часових рядів, такі як ARIMA, експоненційне згладжування та Prophet, дозволяють виявляти тренди, сезонність та циклічні зміни у споживанні ресурсів. Ці методи ефективно використовуються для виявлення залежностей у поведінці навантажень, що дає змогу передбачати короткострокові та довгострокові потреби у ресурсах для Kubernetes -кластерів.

Незважаючи на значні переваги, існують певні труднощі в реалізації передбачувального масштабування. Зокрема, складно забезпечити точність моделей при зміні характеру навантажень, а також балансувати між точністю прогнозів та часом їх обчислення.

Іншою проблемою є можливість появи аномалій, які можуть вплинути на точність прогнозів. Вирішення цих проблем потребує глибокого аналізу та постійного уточнення використовуваних моделей і алгоритмів прогнозування.

Висновки

У статті детально проаналізовано стандартні та передові методи масштабування Kubernetes, зокрема горизонтальне та вертикальне масштабування подів, вузлів кластера, використання спеціальних метрик, а також нові підходи до масштабування. Проведене дослідження показало, що хоча традиційні методи вертикального та горизонтального масштабування залишаються основою, нові стратегії, такі як передбачувальне масштабування, масштабування на основі подій та використання спеціальних метрик,

забезпечують додаткову гнучкість і чутливість до змін у навантаженні.

Кожен із розглянутих підходів має свої переваги й обмеження з точки зору складності впровадження, ефективності використання ресурсів та вимог до налаштування. Горизонтальне масштабування добре підходить для безстанових додатків, тоді як вертикальне є ефективнішим для ресурсомістких сервісів. Впровадження методів машинного навчання та аналізу часових рядів істотно підвищує точність передбачень в керуванні ресурсами, однак вимагає надійних моделей прогнозування та якісних даних. Масштабування на основі подій, наприклад за допомогою KEDA, є особливо корисним для навантажень з нерегулярною або імпульсною природою.

Перспективними напрямками є розробка динамічних механізмів зміни ресурсів контейнера під час виконання, одночасне використання горизонтального і вертикального масштабування, застосування гібридних підходів для мікросервісів, а також розширення досліджень на рівні інфраструктури.

Зрештою, вибір стратегії масштабування має базуватись на характеристиках робочих навантажень, бюджетних обмеженнях і загальних операційних цілях організації.

Тонке налаштування конфігурацій масштабування та постійний моніторинг поведінки застосунків є критично важливими для досягнення ефективного використання ресурсів, підвищення продуктивності та зниження витрат у cloud-native середовищі.

СПИСОК ЛІТЕРАТУРИ

1. Kubernetes Autoscaling and Best Practices for Implementations. StormForge. URL: <https://stormforge.io/kubernetes-autoscaling> (дата звернення: 09.04.2025).
2. Phuc L.H., Phan L.-A., Kim T. Traffic-Aware Horizontal Pod Autoscaler in Kubernetes-Based Edge Computing Infrastructure.
3. Choi B., Park J., Lee C., Han D. pHPA: A Proactive Autoscaling Framework for Microservice Chain. APNet 2021: Proceedings of the 5th Asia-Pacific Workshop on Networking. 2021.
4. Toka L., Dobreff G., Fodor B., Sonkoly B. Machine Learning-Based Scaling Management for Kubernetes Edge Clusters. IEEE Transactions on Network and Service Management. 2021.
5. Rudrabhatla C.K. A Quantitative Approach for Estimating the Scaling Thresholds and Step Policies in a Distributed Microservice Architecture. IEEE Access. 2020
6. Balla D., Simon C., Maliosz M. Adaptive Scaling of Kubernetes Pods. NOMS 2020: IEEE/IFIP Network Operations and Management Symposium. Apr. 2020. P. 1–5.
7. Khaleq A., Ra I. Intelligent Autoscaling of Microservices in the Cloud for Real-Time Applications. IEEE Access. 2021.
8. Baresi L., Hu D., Quattrocchi G., Terracciano L. KOSMOS: Vertical and Horizontal Resource Autoscaling for Kubernetes. ICSOC 2021: International Conference on Service-Oriented Computing. 2021

9. Nguyen T.T., Yeom Y.J., Kim T., Park D.H., Kim S. Horizontal Pod Autoscaling in Kubernetes for Elastic Container Orchestration. Sensors. 2020.
10. Rossi F. Auto-Scaling Policies to Adapt the Application Deployment in Kubernetes. ZEUS 2020: Proceedings of the Workshop on Services and Applications over Decentralized Systems.
11. Cluster Autoscaler on AWS. Kubernetes Autoscaler Documentation. GitHub. URL: <https://github.com/kubernetes/autoscaler/blob/master/cluster-autoscaler/cloudprovider/aws/README.md> (дата звернення: 09.04.2025).
12. Tarn E., Saha R. Harness the Power of Karpenter to Scale, Optimize & Upgrade Kubernetes. AWS re:Invent 2023. Presentation CON331. URL: https://d1.awsstatic.com/events/Summits/reinvent2023/CON331_Harness-the-power-of-Karpenter-to-scale-optimize-and-upgrade-Kubernetes.pdf (дата звернення: 09.04.2025).
13. KEDA Concepts. KEDA Documentation. URL: <https://keda.sh/docs/2.17/concepts/> (дата звернення: 09.04.2025).
14. Venkataraman V. Using Prometheus Adapter to Autoscale Applications Running on Amazon EKS. AWS Cloud Operations Blog. 28.09.2021. URL: <https://aws.amazon.com/blogs/mt/automated-scaling-of-applications-running-on-eks-using-custom-metric-collected-by-amazon-prometheus-using-prometheus-adapter/> (дата звернення: 09.04.2025).

Received (Надійшла) 23.02.2025

Accepted for publication (Прийнята до друку) 21.05.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Помелуйко Денис Андрійович – студент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Denys Pomeluiuko – student, Department of Electronic Computing, Kharkiv National University of Radio Electronics; Kharkiv, Ukraine;

e-mail: denys.pomeluiuko@nure.ua; ORCID Author ID: <https://orcid.org/0009-0005-6994-280X>.

Срьоміна Наталія Сергіївна – кандидат технічних наук, старший викладач кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Nataliia Yeromina - PhD, Senior Lecturer of the Department of Electronic Computing, Kharkiv National University of Radio Electronics; Kharkiv, Ukraine;

e-mail: nataliia.yeromina@nure.ua; ORCID Author ID: <https://orcid.org/0000-0002-0463-2342>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57194558513>.

Петров Сергій Валерійович – кандидат технічних наук, доцент, доцент кафедри електротехніки та електроенергетики, Харківський національний університет імені В.Н. Каразіна, Харків, Україна;

Serhii Petrov – PhD, Docent, Associate Professor of the Department of Electrical Engineering and Power Engineering, V. N. Karazin Kharkiv National University, Kharkiv, Ukraine;

e-mail: psv_topol@ukr.net; ORCID Author ID: <http://orcid.org/0000-0001-8933-9649> ;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57194556925> .

Іщенко Наталія Володимирівна – асистент кафедри автоматизації, метрології та енергоефективних технологій Харківський національний університет імені В.Н. Каразіна, Харків, Україна;

Nataliia Ishchenko – assistant of the Department of Automation, Metrology and Energy Efficient Technologies, V. N. Karazin Kharkiv National University, Kharkiv, Ukraine

e-mail: natalidocenko158@gmail.com; ORCID Author ID: <https://orcid.org/0009-0004-3899-3965>.

Волощук Олена Борисівна – кандидат технічних наук, доцент, доцент кафедри ЕОМ, Харківського національного університету радіоелектроніки, Харків, Україна;

Olena Voloshchuk – PhD, Ass.prof, Ass.prof of the Department of electronic computers, Kharkiv National University of radio electronics, Kharkiv, Ukraine;

e-mail: olena.voloshchuk@nure.ua; ORCID Author ID: <https://orcid.org/0000-0002-5912-4126>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57207762084>.

Kubernetes cluster scaling methods in cloud environments

Denys Pomeluiuko, Nataliia Yeromina, Serhii Petrov, Nataliia Ishchenko, Olena Voloshchuk

Abstract. The article explores Kubernetes cluster scaling in cloud environments, a critical factor in ensuring high availability for microservice applications. It analyzes traditional techniques and modern approaches, including vertical and horizontal scaling mechanisms. Methods such as predictive autoscaling, event-driven scaling, and the use of custom metrics are also considered. The advantages and disadvantages of these approaches are discussed. Furthermore, the paper demonstrates how the configuration and optimization of scaling parameters impact cloud resource costs and application performance. **Purpose of the article** is to analyse modern tools and approaches to scaling Kubernetes clusters in the AWS cloud environment, to identify the advantages and disadvantages of each of the available solutions. **Conclusions.** Event-based scaling is especially useful for loads with irregular or impulsive nature. Promising directions include the development of dynamic mechanisms for changing container resources at runtime, the simultaneous use of horizontal and vertical scaling, the use of hybrid approaches for microservices, and the expansion of research at the infrastructural level. The choice of scaling strategy should be based on the characteristics of workloads, budgetary constraints, and the organisation's overall operational goals. Fine-tuning scaling configurations and continuously monitoring application behaviour are critical to achieving efficient resource utilisation, increased productivity and reduced costs in a cloud-native environment.

Keywords: Kubernetes, cluster scaling, cloud environment, HPA, Horizontal Pod Autoscaler, VPA, Cluster Autoscaler, Karpenter, KEDA, Kubernetes Event Driven Scaling, Prometheus, EKS, AWS.

О. Є. Пономаренко, В. О. Горбачов

Харківський національний університет радіоелектроніки, Харків, Україна

МЕТОД ПОШУКУ ВУЗЬКИХ МІСЦЬ В СКЛАДНИХ МЕРЕЖАХ ДЛЯ ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ ОБЧИСЛЮВАЛЬНИХ РЕСУРСІВ

Анотація. В складних мережах існує проблема пошуку вузьких місць. Існує багато сфер застосування задачі виявлення вузьких місць, наприклад, алгоритми для виявлення спільнот, розуміння формування груп та зменшення перевантаження у транспортних мережах. **Метою роботи** є дослідження методу пошуку вузьких місць в складних мережах. Актуальність роботи полягає в тому, що пошук вузьких місць є масштабною задачею для складних мереж. В роботі були вирішені наступні задачі: дослідження методів пошуку вузьких місць в мережі; здійснення пошуку вузьких місць в початковій та зменшеній мережі; вимірювання часу вирішення задачі; перевірка збереження розташування вузьких місць після зменшення мережі. В результаті роботи були досліджені методи пошуку вузьких місць в мережі; був здійснений пошук вузьких місць в початковій та зменшеній мережі; було здійснено вимірювання часу вирішення задачі та була здійснена перевірка збереження розташування вузьких місць в зменшеній мережі. Дослідження дозволяють зробити **висновки**: був здійснений пошук вузьких місць в початковій та зменшеній мережі, на розглянутих прикладах час пошуку вузьких місць зменшився, та розташування вузьких місць зберіглося в зменшеній мережі. Таким чином, в результаті зменшення мережі здійснюється підвищення ефективності обчислювальних ресурсів.

Ключові слова: метод; модель; комп'ютерна мережа; структура; вузькі місця; потік.

Вступ

Пошук вузьких місць є однією з масштабних обчислювальних проблем у складних мережах [1]. Аналіз вузьких місць в мережах є важливим завданням для багатьох сфер при плануванні та інженерії. Проблема пошуку вузьких місць у транспортних мережах часто виникає в реальних задачах, наприклад, у плануванні виробництва. Однією з цілей проектування виробничої системи є виявлення недоліків у схемі на ранній стадії планування. Важливим фактором є виявлення та усунення вузьких місць [2].

Складні мережі характеризуються універсальністю для представлення майже будь-якої природної структури [3]. Висока розмірність мереж може викликати проблеми при їх моделюванні та аналізі. Для вирішення цих проблем існують методи зменшення мереж [4].

Таким чином, пошук вузьких місць в складних мережах є актуальною та важливою задачею.

Дана робота присвячена методу пошуку вузьких місць в складних мережах, здійсненню пошуку вузьких місць у початковій та зменшеній мережі, перевірці збереження розташування вузьких місць та порівнянню часу вирішення задачі.

Аналіз літератури

Багато систем структуровані у формі мереж, наприклад, великі комунікаційні системи, транспортна інфраструктура, біологічні системи [5].

Візуалізація є корисним інструментом для аналізу мереж, що дозволяє побачити важливі структурні особливості мережі. Існують різноманітні інструменти візуалізації, які допомагають структурувати, та візуалізувати мережі [6]. Та візуалізація може бути корисною лише для мереж із кількома сотнями чи тисячами вершин та невеликою кількістю дуг. Якщо в мережі дуже багато вершин або дуг,

то зображення мережі будуть надто складними для сприйняття, та ефективність візуалізації стає обмеженою. В таких випадках потрібно використовувати інші методи для аналізу мереж [7].

Складні мережі мають певні топологічні характеристики. Для аналізу та генерації складних мереж використовуються метрики, які здатні передати найбільш відповідні топологічні характеристики [3].

Складні мережі можуть мати дуже високу розмірність, що може призвести до проблем при моделюванні, аналізі, а також проблем, пов'язаних з обмеженими обчислювальними можливостями та ємністю зберігання. Ці проблеми стимулюють розробку методів зменшення складності мереж [4].

Автори в роботі [4] надають огляд досягнень у зменшенні розмірності складних мережесистем. Складність полягає у двох аспектах, а саме: великомасштабна топологія та підсистеми високої розмірності, що призводить до двох типів проблем зменшення моделі мережі. Перша проблема полягає в тому, щоб спростити складну мережесистему шляхом зменшення кількості вузлів. Методи зменшення, засновані на кластеризації та агрегації, є основними для зменшення топологічної складності. Друга проблема полягає в тому, щоб зменшити розмірність окремих підсистем у мережі.

В роботі [8] автори пропонують метод для зменшення розміру складної мережі зі збереженнями модульності, перевагою якого є покращення обчислювальних витрат.

У роботі [9] автори описують агрегацію мережі, результатом якої є початкова мережа та мережа, яка була отримана в результаті агрегації.

В роботі [10] автори пропонують алгоритм для зменшення розміру великих мереж масового обслуговування, метою якого є більш ефективне моделювання мережі. У багатьох мережах масового обслуговування показники продуктивності представляють інтерес лише для невеликої кількості вузлів мережі,

а інші вузли представляють інтерес лише в тому, як вони впливають на показники продуктивності основних вузлів. Підхід полягає в тому, щоб розділити мережу масового обслуговування на дві підмережі: підмережу, яка містить основні вузли та підмережу, яка містить решту вузлів.

Пошук вузьких місць є однією з великих проблем у складних мережах. Пошук та аналіз вузьких місць в мережах є важливим завданням для багатьох сфер, наприклад, алгоритми для виявлення спільнот, розуміння формування груп та зменшення перевантаження у транспортних мережах [1].

Вузькі місця – це точки перевантаження системи, які сповільнюють весь ланцюг роботи. Необхідно усунути вузькі місця, щоб покращити продуктивність системи. Таким чином, ідентифікація та усунення вузьких місць є важливою задачею [11].

У роботі [11] автори розглядають методи виявлення вузьких місць. Метод вимірювання середнього часу очікування виявляє пристрій з найдовшим часом очікування як вузьке місце. Метод вимірювання середнього робочого навантаження виявляє пристрій з найбільшим використанням як вузьке місце. Метод вимірювання середньої тривалості активності виявляє пристрій з найдовшою середньою тривалістю активності як вузьке місце. Метод виявлення вузьких місць зміщення виявляє пристрій з найдовшим періодом активності як вузьке місце та ідентифікує його як вузьке місце зміщення.

Дослідження проблеми контролю перевантажень для мереж базуються на тому, що продуктивність потоку визначається станом його дуги, що є вузьким місцем. Дуга, пропускна здатність якої використовується повністю, завжди є вузьким місцем принаймні одного потоку, що проходить через неї [12].

Автори у [13] описують мережі ланцюгів поставок як складні системи фірм та їхніх постачальників. Автори розглядають метрики на основі теорії мереж для виявлення вузьких місць у мережі. У теорії графів існують метрики для визначення центральності вузлів відповідно до їх положення у мережі, а саме: центральність за ступенем, центральність за посередництвом та центральність за близькістю. Ці метрики допомагають визначити відносну важливість вузлів у мережі [13].

Автори у [14] розглядають приклад використання центральності за посередництвом у телекомунікаційній мережі. Вузол із більш високим значенням центральності за посередництвом має більше контролю, та через нього може передаватись більше інформації. Таким чином, можна знайти ті вузли, які мають найбільший вплив на мережу. Автори у [14] пропонують алгоритм на основі кластеризації для вирішення проблеми ідентифікації впливових вузлів у складних мережах за допомогою центральності за посередництвом. Автори обчислюють міру центральності у підграфах, що скорочує час обчислення.

Метою роботи є дослідження методу пошуку вузьких місць в складних мережах. В ході дослідження вирішуються наступні завдання: дослідження методів пошуку вузьких місць в мережі, здійс-

нення пошуку вузьких місць в початковій та зменшеній мережі, вимірювання часу вирішення задачі, перевірка збереження розташування вузьких місць після зменшення мережі.

Опис методу пошуку вузьких місць в мережі

Визначення мережі. $N = \langle L, F, V, C \rangle$ є мережею, якщо виконуються наступні умови [12]:

- V – це множина, яка називається множиною вузлів $\{v_1, v_2, \dots, v_{|V|}\}$, $|V| = K$;
- L – це набір дуг у вигляді $\{l_1, l_2, \dots, l_{|L|}\}$, $L \subseteq V \times V$: якщо не вказано інше, дуги спрямовані, тобто $(i, j) \neq (j, i)$;
- F – множина потоків у вигляді $\{f_1, f_2, \dots, f_{|F|}\}$;
- C – набір значень у вигляді $\{c_l, \forall l \in L\}$, c_l – пропускна здатність дуги l [12].

Кожен потік f проходить через підмножину дуг $L_f \subset L$, та кожна дуга l проходить через підмножину потоків $F_l \subset F$. Кожен потік f передає дані зі швидкістю r_f . Обмеження пропускної здатності має виконуватися для всіх $l \in L$ [12]:

$$\sum_{f \in F} r_f \leq c_l. \quad (1)$$

U – це набір значень у вигляді $\{u_j, \forall j \in V\}$, де u_j – це поріг насичення вузла j , $u_j \in [0, 1]$. Кожен потік f проходить через підмножину вузлів $V_f \subset V$, та кожен вузол j проходить через підмножину потоків $F_j \subset F$. Якщо u_f – це використання вузла j для потоку f , то обмеження використання має виконуватися для всіх $j \in V$:

$$\sum_{f \in F} u_f \leq u_j. \quad (2)$$

Визначення дуги як вузького місця. $N = \langle L, F, V, C \rangle$ є мережею, де кожен потік $f \in F$ передає дані зі швидкістю r_f . Дуга l є вузьким місцем потоку f якщо виконуються наступні умови [12]:

- потік f перетинає дугу l ;
- $\partial r_f / \partial c_l \neq 0$, тобто швидкість передачі потоку f змінюється при невеликих зменшеннях пропускної здатності дуги l ;
- обмеження (1) не виконується.

Визначення вузла як вузького місця. $N = \langle L, F, V, C \rangle$ є мережею, де кожен потік f проходить через підмножину вузлів $V_f \subset V$, та через кожен вузол j проходить підмножина потоків $F_j \subset F$. Вузол j є вузьким місцем потоку f якщо виконуються наступні умови:

- потік f перетинає вузол j ;
- $\partial r_f / \partial u_j \neq 0$, тобто швидкість передачі потоку f змінюється при невеликих зменшеннях використання вузла j ;
- обмеження (2) не виконується.

Для розрахунку використання вузла можна використати закон використання [15]:

$$U_j = X_j S_j, \quad (3)$$

де X_j – пропускна здатність вузла j , S_j – середній час обслуговування запитів у вузлі j , U_j – використання вузла j , X_j , S_j , U_j – середні значення за період спостереження T .

Рівняння з урахуванням балансу потоків [15]:

$$U_i = \lambda_i S_i. \quad (4)$$

Визначення критичного насичення вузла. Вузол $j \in V$, де V – це набір системних вузлів, є критично насиченим протягом періоду спостереження T , якщо його використання перевищує поріг насичення u_j для цього інтервалу T .

Модель балансу потоків мережі. Проблема, яка розглядається, відноситься до оцінки параметрів продуктивності окремого вузла або всієї мережевої системи. Оцінка параметрів продуктивності полягає у зв'язуванні параметрів продуктивності з певними рівняннями. Для даного випадку це система рівнянь балансу потоків.

Змінні та обмеження моделі збалансованого потоку. Частота маршрутизації q_{ij} у квадратній матриці Q з елементом q_{ij} визначається як [15]:

$$q_{ij} = \begin{cases} \frac{C_{ij}}{C_i}, & \text{if } i = \overline{1, K} \\ \frac{A_{0j}}{A_0}, & \text{if } i = 0 \end{cases}, \quad j = \overline{0, K}, \quad (5)$$

де C_{ij} – це кількість разів, коли запит надходить на обслуговування до вузла j відразу після завершення обслуговування у вузлі i .

Для будь-якого вузла i [15]:

$$q_{i0} + q_{i1} + \dots + q_{iK} = \sum_{j=0}^K q_{ij} = 1, \quad i = \overline{0, K} \quad (6)$$

Закон балансу потоків запитів для вузла i . Для кожного вузла i X_i дорівнює загальній вхідній швидкості до вузла i .

Коли потік запитів збалансований, X_i є пропускну здатністю вузла. Принцип балансу у вигляді рівняння [15]:

$$C_j = A_j = \sum_{i=0}^K C_{ij}, \quad i = \overline{0, \dots, K}, \quad (7)$$

де A_j – кількість вхідних запитів протягом періоду спостереження T у вузлі j .

Баланс потоку запитів дозволяє замінити C_{0j} на A_{0j} , та вважаючи на (5) [15]:

$$C_j = \sum_{i=0}^K C_i q_{ij} \quad (8)$$

Зважаючи на те, що $X_i = C_i/T$, рівняння балансу потоків запитів [15]:

$$X_j = \sum_{i=0}^K X_i q_{ij}, \quad j = \overline{0, K} \quad (9)$$

Підхід, який розглядається в статті, складається з трьох кроків:

- розв'язання системи лінійних рівнянь балансу потоків (9) та отримання пропускну здатність вузла X_j ;
- обчислення коефіцієнту використання вузла U_j враховуючи (3).
- знаходження підмножини вузлів, для яких коефіцієнт використання перевищує відповідний поріг насичення:

$$U_b \subset U, \quad (10)$$

де U_b є підмножиною вузлів, для яких значення коефіцієнту використання U_j перевищує відповідний поріг насичення u_j .

Загалом вибір вузла j у підмножину U_b можна описати виразом:

$$U_b = P(U_1, \dots, U_k), \quad (11)$$

де символ P описує різні стратегії формування підмножини U_b .

Пошук вузьких місць в початковій та зменшеній мережі

У роботі [16] була представлена програмна платформа з використанням якої можна згенерувати мережу, здійснити агрегацію мережі, здійснити статистичну обробку даних та вирішити задачу про пошук максимального потоку.

У цій роботі до програмної платформи додається модуль пошуку вузьких місць в мережі. Таким чином, програмна платформа складається з п'яти модулів: модуль генератора мережі, модуль агрегації, модуль статистичної обробки даних, модуль пошуку максимального потоку та модуль пошуку вузьких місць в мережі.

Структура програмної платформи представлена на рис 1.

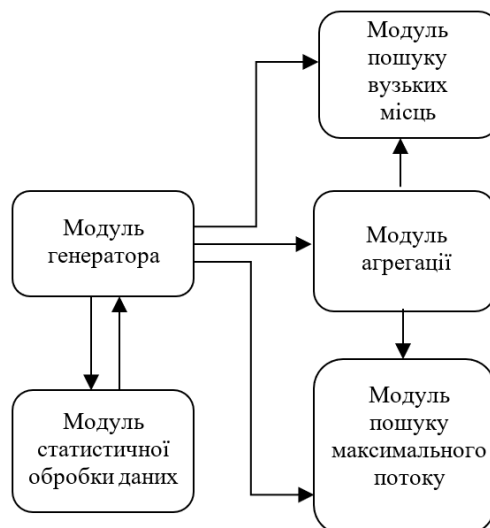


Рис. 1. Структура програмної платформи

- В модулі генератора мережі створюється структура мережі з заданими параметрами;
- в модулі агрегації здійснюється зменшення розмірності мережі;
- в модулі пошуку максимального потоку здійснюється пошук максимального потоку в початковій та зменшеній мережі;
- в модулі статистичної обробки даних здійснюється обчислення кількості дуг у графі та метрики розподілу ступенів;
- в модулі пошуку вузьких місць здійснюється пошук вузьких місць в початковій та зменшеній мережі.

Алгоритм експерименту пошуку вузьких місць в початковій та зменшеній мережі полягає у наступному:

- генерація мережі, що містить вузли $n_1, n_2, \dots, n_m$, здійснюється в модулі генератора програмної платформи;

– зменшення розмірності мережі здійснюється в модулі агрегації програмної платформи;

– пошук вузьких місць в початковій та зменшеній мережі здійснюється в модулі пошуку вузьких місць програмної платформи.

Пошук вузьких місць в мережі складається з наступних етапів:

– обчислення пропускної здатності вузлів X_i наступним чином: на основі графу складається система лінійних рівнянь, здійснюється вирішення системи лінійних рівнянь;

– обчислення коефіцієнтів використання вузлів по формулі (3).

Результат експерименту. Початкова мережа містить 30 вузлів.

Результат розрахунків коефіцієнтів використання вузла в початковій мережі представлений на рис. 2.

U11 = 0,665
U5 = 0,63
U1 = 0,595
U7 = 0,595
U3 = 0,56
U9 = 0,56
U13 = 0,525
U12 = 0,232
U27 = 0,161

Рис. 2. Результат розрахунків коефіцієнтів використання вузла в початковій мережі

У даному прикладі на початковій мережі вузли p_{11} та p_5 мають найвищі значення коефіцієнту використання вузла.

Мережа містить 15 вузлів після здійснення зменшення її розмірності.

Результат розрахунків коефіцієнтів використання вузла в зменшеній мережі представлений на рис. 3.

U6 = 0,871	component contains nodes: 11, 12
U3 = 0,72	component contains nodes: 5, 6
U5 = 0,661	component contains nodes: 9, 10
U7 = 0,632	component contains nodes: 13, 14
U1 = 0,595	component contains nodes: 1, 2
U4 = 0,595	component contains nodes: 7, 8
U2 = 0,56	component contains nodes: 3, 4

Рис. 3. Результат розрахунків коефіцієнтів використання вузла в зменшеній мережі

У даному прикладі на зменшеній мережі вузли p_6 та p_3 мають найвищі значення коефіцієнту використання вузла.

Підсистема p_6 містить вузол p_{11} початкової мережі, підсистема p_3 містить вузол p_5 початкової мережі, які є вузькими місцями в початковому графі.

Час пошуку вузьких місць на початковій мережі – 20 мс, час пошуку вузьких місць в зменшеній мережі – 0,78 мс. Час пошуку вузьких місць в початковій та зменшеній мережі для декількох прикладів продемонстровано на рис. 4.

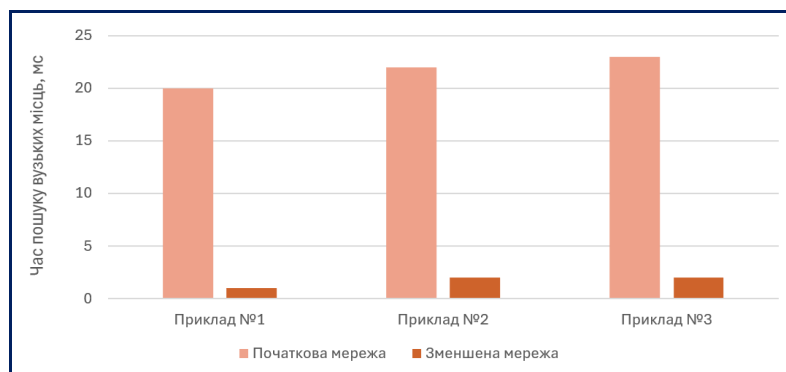


Рис. 4. Час пошуку вузьких місць в початковій та зменшеній мережі

Отже, в результаті зменшення розмірності мережі на продемонстрованих прикладах зберігається розташування вузьких місць в мережі, та час пошуку вузьких місць в мережі зменшується. Таким чином, можна зробити висновок, що в результаті зменшення розмірності мережі є підвищення ефективності обчислювальних ресурсів.

У мережах масового обслуговування метрики продуктивності представляють інтерес лише для невеликої кількості вузлів у мережі, інші вузли представляють інтерес у тій мірі, в якій вони впливають на показники продуктивності в основних вузлах [10]. У роботі автори розділяють мережу на дві підмережі: перша містить основні вузли, що представляють інтерес, друга містить інші вузли. При такому розділенні мережі на підмережі визначення розташування вузьких місць може бути корисним для аналізу мережі. Це може допомогти визначити,

чи є вузькі місця в підсистемі, яка містить вузли, що представляють інтерес.

Висновки

У роботі було здійснено дослідження методу пошуку вузьких місць у складних мережах. Також був здійснений пошук вузьких місць у початковій та зменшеній мережі, було здійснено вимірювання часу вирішення задачі та була здійснена перевірка збереження розташування вузьких місць у зменшеній мережі.

На розглянутих прикладах час пошуку вузьких місць зменшився, та розташування вузьких місць зберіглося у зменшеній мережі.

Таким чином, можна зробити висновок, що в результаті зменшення розмірності мережі є підвищення ефективності обчислювальних ресурсів. Подальша робота полягає у вирішенні наступних задач:

дослідження пошуку вузьких місць для мереж із більшою кількістю вершин, дослідження інших за- дач для аналізу складних мереж з метою підвищення ефективності обчислювальних ресурсів.

СПИСОК ЛІТЕРАТУРИ

1. Chalupa, D., Hawick, K. A. & Walker, J. A. (2018). Hybrid Bridge-Based Memetic Algorithms for Finding Bottlenecks in Complex Networks. *Big Data Research*, 14, 68–80. <https://doi.org/10.1016/j.bdr.2018.04.001>
2. Post, T., Hamann, B. & Aurich, J. (2017). Ensemble Visualization of Bottlenecks in Planar Flow Networks. *Applied Mechanics and Materials*, 869, 234–243. <https://doi.org/10.4028/www.scientific.net/AMM.869.234>
3. Costa, L. da F., Rodrigues, F. A., Travieso, G. & Villas Boas, P. R. (2007). Characterization of complex networks: A survey of measurements. *Advances in Physics*, 56(1), 167–242. <https://doi.org/10.1080/00018730601170527>
4. Cheng, X. & Scherpen, J. M. A. (2021). Model Reduction Methods for Complex Network Systems. *Annual Review of Control, Robotics, and Autonomous Systems*, 4, 425–453. <https://doi.org/10.1146/annurev-control-061820-083817>
5. Barrat, A., Barthelemy, M., Pastor-Satorras, R. & Vespignani, A. (2004). The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences*, 101(11), 3747–3752. <https://doi.org/10.1073/pnas.0400087101>
6. Hackl, J. (2017). tikz-network: a LaTeX library for vizualizing complex networks. 6th International Conference on Complex Networks & Their Applications, Lyon, France, 318–319.
7. Newman, M. (2010). *Networks: An Introduction*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199206650.001.0001>
8. Arenas, A., Duch, J., Fernandez, A. & Gomez, S. (2008). Size reduction of complex networks preserving modularity. *New Journal of Physics*, 10(5), 053039. <https://doi.org/10.1088/1367-2630/10/5/053039>
9. Hirata, H. & Ulanowicz, R. E. (1985). Information Theoretical Analysis of the Aggregation and Hierarchical Structure of Ecological Networks. *Journal of Theoretical Biology*, 116(3), 321–341. [https://doi.org/10.1016/S0022-5193\(85\)80271-X](https://doi.org/10.1016/S0022-5193(85)80271-X)
10. Shortle, J. F., Mark, B. L. & Gross, D. (2009). Reduction of Closed Queueing Networks for Efficient Simulation. *ACM Transactions on Modeling and Computer Simulation*, 19(3), 10. <https://doi.org/10.1145/1540530.1540531>
11. Wang, Y., Zhao, Q. & Zheng, D. (2005). Bottlenecks in production networks: An overview. *Journal of Systems Science and Systems Engineering*, 14(3), 347–363. <https://doi.org/10.1007/s11518-006-0198-3>
12. Ros-Giralt, J., Amsel, N., Yellamraju, S., Ezick, J. R., Lethin, R. A., Jiang, Y., Feng, A. & Tassioulas, L. (2022). A quantitative theory of bottleneck structures for data networks. *arXiv preprint arXiv:2210.03534*. <https://arxiv.org/abs/2210.03534>
13. Mizgier, K., Jüttner, M., & Wagner, S. (2012). Bottleneck identification in supply chain networks. *International Journal of Production Research*, 51(1), 1–14. <https://doi.org/10.1080/00207543.2012.695878>
14. Szabari, B. & Kiss, A. (2020). Performance Evaluation of Betweenness Centrality Using Clustering Methods. *Studia Universitatis Babeş-Bolyai Informatica*, 65(1), 51. <https://doi.org/10.24193/subbi.2020.1.05>
15. Denning, P. J. & Buzen, J. P. (1978). The operational analysis of queueing network models. *ACM Computing Surveys*, 10(3), 225–261. <https://doi.org/10.1145/356733.356735>
16. Пономаренко, О., Горбачов, В. (2023). Програмна платформа для оцінювання ефективності агрегації структурної моделі складних систем. *Сучасний стан наукових досліджень та технологій в промисловості*, (3(25)), 79–87. <https://doi.org/10.30837/ITSSI.2023.25.079>

Received (Надійшла) 18.03.2025

Accepted for publication (Прийнята до друку) 11.06.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Пономаренко Ольга Євгенівна – аспірантка кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Olha Ponomarenko – PhD student, Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: olha.ponomarenko@nure.ua; ORCID Author ID: <https://orcid.org/0009-0002-4634-6552>.

Горбачов Валерій Олександрович – кандидат технічних наук, професор, професор кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Valeriy Gorbachov – PhD of Engineering Sciences, Professor, Professor of Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: valeriy.gorbachov@nure.ua; ORCID Author ID: <https://orcid.org/0000-0003-3423-2371>.

Method for finding bottlenecks in complex networks to increase efficiency of computing resources

Olha Ponomarenko, Valeriy Gorbachev

Abstract. In complex networks, there is a problem of finding bottlenecks. There are many areas of application for the bottleneck detection problem, for example, algorithms for detecting communities, understanding group formation, and reducing congestion in transport networks. **The goal of the work** is to study the bottlenecks search method for complex networks. The relevance of the work lies in the fact that the bottlenecks searching is a wide problem for complex networks. The following tasks were solved in the work: research of methods for finding bottlenecks in the network; searching for bottlenecks in the initial and reduced network; measurement of the time to solve the problem; verification of the preservation of the bottleneck location after network reduction. As a result of the work, methods for finding bottlenecks in the network were investigated; searching for bottlenecks in the initial and reduced network was carried out; measurement of the time to solve the problem was carried out, and verification of the preservation of the location of bottlenecks on the reduced network was carried out. The studies allow us to **conclude**: searching for bottlenecks was carried out on the initial and reduced network; in the considered examples the time for bottlenecks searching was decreased, and the location of bottlenecks was preserved on the reduced network. Thus, as a result of the reduction of the network, the efficiency of computing resources is increased.

Keywords: method; model; computer network; structure; bottlenecks; flow.

Ievgen Samborskyi¹, Yevhen Peleshok²

¹ State University "Kyiv Aviation Institute", Kyiv, Ukraine

² Research Institute of Military Intelligence, Kyiv, Ukraine

SYNTHESIS OF LOGICAL-DYNAMIC INFORMATION AND EVENT MANAGEMENT SYSTEMS FOR COMPUTER SECURITY STRUCTURES

Abstract. The article considers one of the possible approaches to solving a relevant scientific problem dedicated to the synthesis of logical-dynamic information and event management systems for the security of computer structures. An analysis of modern approaches to the creation of an integrated management system has been made, the implementation of which will allow for effective management of information arrays of security events in the conditions of changed threats to computer structures. Special attention is paid to the development of models that combine logical and dynamic aspects of management, which allow ensuring high adaptability and functional stability of computer structures in the conditions of intensification of security event activity. The implementation of such an approach will allow to automatically respond to these events, optimize the processes of detecting and eliminating the consequences of threats, as well as ensure the constant stability of the operation of computer structures. The article proposes a method for synthesizing a logical-dynamic model that synergizes not only technical, but also organizational-cybernetic aspects of the security of computer structures. The developed approaches can be recommended for use in creating more effective and secure computer tools capable of withstanding various incidents.

Keywords: control system, logical-dynamic system, information security, security event, synthesis, analysis, theory of logical-dynamic systems, computer structure, incident, array of security events.

Introduction

Statement of the problem. Modern computer structures (systems, networks), especially critical infrastructure objects, are constantly under the influence of numerous threats to the security of their functioning. This situation requires the constant creation of effective methods for managing information and security events of these computer structures. Traditional approaches to monitoring, analyzing and compensating for the destructive consequences of security events (incidents), as a rule, are often not effective enough due to the growing complexity of these information infrastructures, the high level of dynamism of threats and the need for a quick response to their unauthorized effects. A quick response to threats is possible due to the optimal adaptation of computer structures to security situations, which is possible in the case of implementing control processes of these programmable means using new approaches based on the theory of logical-dynamic systems. Control systems synthesized using this theory allow combining formal analysis methods with adaptive control algorithms, which will provide automated detection, assessment, and optimal response to security incidents. However, there are a number of unsolved problems associated with the synthesis of such systems. These problems are related to the formalization of logical-dynamic models in incident analysis, the development of algorithms for adaptive response to security events, and the flexible integration of such systems into modern computer infrastructures. Thus, the scientific and applied task of synthesizing logical-dynamic information and security event management systems arises, which requires the development of new models, methods, and algorithms to increase the level of security, interference, and functional stability of computer structures.

Analysis of recent research and publications. The analysis of the results of scientific research, which are reflected in the latest scientific publications, conducted by the authors, shows that the process of information and

security event management should be organized using the theory of logical-dynamic systems. This theory allows you to model and analyze the behavior of large organizational complex systems, which are computer structures, and to take into account the features of logical and dynamic processes in these electronic computing devices. Such an approach will contribute to more effective detection and optimal response to threats in modern and promising information infrastructures.

To understand the role and place of the theory of logical-dynamic systems, let us consider a number of scientific works that outline theoretical and practical aspects of systemic approaches to organizing information security management of a computer structure, and methods for synthesizing information management systems and security events in these complex information structures.

The basic principles of systems theory and the methodology of scientific research, which are the foundations of a systems approach to information security management as a complex logical-dynamic process are set out in [1]. The scientific work considers methods of analysis and forecasting the conditions for ensuring the stability of systems, statistical data processing, mathematical models of systems, methods of nonlinear dynamics, methods of controlling the functioning of systems under conditions of unpredictability of their states, systems analysis of complex structures, as well as methods and approaches to the synthesis of control systems and the laws of functioning of these structures.

In the work [2] special emphasis is placed on the synthesis of automated and automatic information systems for managing the cyber security of complex objects. Possible approaches and principles for building these systems for organizing the security stability of information infrastructure are also formulated.

The main threats to information security that arise in modern computer structures are considered in the scientific article [3]. It pays considerable attention to the issue of optimizing information protection processes in computer

networks through the functioning of the information security management system. It is indicated that the basis of the information security management system is based on risk assessment and management of these destructive events. The process of security risk management allows you to achieve maximum management efficiency at a given level of risk. The implementation of information security risk management allows you to use effective control of their security status in distributed corporate networks by reducing the risk to the desired level.

In the work [4] the possibility of increasing the efficiency of information security management of a computer structure through the use of multi-agent technologies is investigated. The authors emphasize that information security of a computer structure is a complex problem of management organization, which requires prompt response to incidents and flexible integration of various subsystems of a computer structure to ensure information protection from possible threats. A multi-agent management system is described, which is able to detect threats, monitor and analyze information resources, and also implement mechanisms of prompt response in real time to incidents (security events).

In the works [5–7] the authors comprehensively consider the methodology for creating information security management systems. These scientific articles analyze various approaches to information security risk management and propose the implementation of this methodology in a computer model. The practical implementation of this model helps to develop and implement safe and effective information and security event management systems. It is proposed to consider the structures of information security management systems as a set of subsystems, complexes, components, and the relationships between them. This approach allowed this management system to be represented by a tree of modular units. The properties of information and the model of interaction levels of open systems during its transmission and protection in a computer network are considered by ensuring a given level of management quality. The approaches to creating a methodology for building an information security management system, initiated in works [5–7] were improved by the authors in scientific works [8] and [9]. For this purpose they used a functional analysis of information security management systems. Based on this analysis, the main approaches to the design and implementation of synthesized management systems are investigated. Special emphasis is placed on the importance of an integrated approach to risk management, as well as the role of technical, organizational and cybernetic measures in ensuring the security of information systems. A scientific publication [10] is devoted to the method of analyzing logical-dynamic processes of argument transformation in computing nodes of digital control systems. The limitations and shortcomings of formal methods used to describe data processing processes in digital control systems are identified, and a method for describing argument transformation processes using graph-analytic models is proposed. One of the possible approaches to the practical application of an information security management system is given in [11]. The authors note that the use of systemic approaches to ensuring

information security allows identifying weaknesses and shortcomings in the system and eliminating them in a timely manner which will contribute to ensuring the effectiveness, optimality and cost-effectiveness of information security measures. Fundamental aspects of the synthesis of control models for complex dynamic objects taking into account their security events are given in [12]. The authors comprehensively analyze the impact of dynamic processes on the functioning of complex objects and also develop methods for flexible integration of security events into the control model to increase the reliability and stability of the system. Particular attention is paid to approaches to timely response to threats to the security of the computer structure. Along with theoretical developments the work also offers practical recommendations for the development of effective control strategies which in the future should be implemented in information infrastructures using the theory of a logical-dynamic system.

The purpose of the article is to develop theoretical and practical (applied) principles for creating effective information and security event management systems in a computer structure. Their implementation will ensure the efficiency, accuracy and optimality of data processing in the arrays of security events of a computer structure. For this purpose, it is planned to synthesize logical and dynamic models of automation processes for detecting and preventing information security threats.

The main goal of the article is to propose new and effective approaches to the synthesis of logical-dynamic systems that provide reliable protection of the computer structure from current existing threats and in the future, from zero-day threats.

Presentation of the main material

The organization of the process of information and security event management in a computer structure is characterized by complexity and dynamism, which require the use of fundamental theoretical research, which will become the basis for creating optimal algorithmic support for these programmable structures. The results of the analysis of the well-known theoretical developments, which are represented in scientific works [12–22] regarding the features and specifics of processes related to security events in a computer structure, indicate that it is appropriate and desirable to use the theory of logical-dynamic systems to describe them. Systems of this class are characterized by the capabilities and properties of describing hyper-complex processes in information and security event management in a computer structure (and especially those that ensure the functional stability of critical infrastructure objects) [19]. To achieve the research goal we will consider the fundamental provisions of the theory of logical-dynamic systems, their axiomatics, capabilities and specific properties. The axiomatic definition, description and formalization of these systems is as follows. A logical-dynamic system is a complex, large, organizational, hypercomplex synergistic structure. In formalized mathematical models, the components of their structure that describe the logic of processes (variations of final states) and the components that characterize the dynamics of changes in the parameters of these systems (rapid change in real time of its state) are synergistically

linked. At the same time, in logical-dynamic systems, both the rules of logic and variations of these rules or changes in the dynamic states of the system, which functionally depend on a number of external and internal factors can be interconnected. We focus on the peculiarities of the logical-dynamic system, which consists in the fact that changes are described according to clearly established logical rules. But in many situations, these rules can be dynamically (quickly) adjusted (changed) a priori by some influences, for example, such as feedback in the system or flexible adaptation to external disturbances. It follows that this dynamic system belongs to the class of controlled systems (processes in which change in real time in accordance with clearly established rules or logic), which are characterized by adaptive changes in its internal structure, associated with variations in the constituent elements of this structure. It is proposed to describe models of individual possible processes and states of such systems using classical control theory, taking into account that transitions between changes in these states occur in accordance with logical (mathematical) rules. Models of states of logical-dynamic systems are mathematical models that are used to describe and analyze the behavior of complex systems that change their state in time coordinates. In this case, logical-dynamic systems combine logical expressions of discrete mathematics with equations of system dynamics. This approach allows modeling both discrete and continuous variations of changes in the system.

Note that for modeling processes in a logical-dynamic structure, it is advisable to use the theories of:

- mathematical logic (discrete mathematics). (used for mathematical interpretation of the functioning of discrete elements of the system);
- differential equations (used to describe continuous processes that are components of the general process of system functioning);
- graphs. (interprets the structure of the system, namely: an ensemble (tuple) of elements of the logical-dynamic structure and the variation of the relationships between them);
- mathematical modeling of the functioning of finite state machines (describes discrete states of the logical-dynamic structure and transitions between them).

The results of a systematic analysis of existing approaches to the application of the main provisions of these theories confirm the feasibility and necessity of mandatory consideration of the specific features of the logical-dynamic structure. We characterize these features by a number of factors, namely:

- controllability – a feature that characterizes the ability of a logical-dynamic system to respond to external influences and promptly and adequately change its state in accordance with the synthesized control algorithm;
- system functionality – this feature takes into account the hyper-complexity of the logical-dynamic system which consists in the fact that the specified system combines structurally diverse subelements each of which clearly performs only its own specific task within the framework of the overall structure of the system;
- the presence of a general law of operation of control processes (target control function) – a logical-dynamic system that operates in accordance with

established laws that determine their behavior. These laws can be either discrete (based on the implementation of logical operations), continuous (based on the application of a system of mathematical equations) or discrete-continuous (synergistically combine the properties of discreteness and continuity of processes in this system).

It should be noted that scientific works [18–22] focus special attention on the above-mentioned features of the logical-dynamic system. These features allow us to fully take into account the specifics of processes in organizing the management of information and events of computer system security. The theory of logical-dynamic systems will be used to analyze the stages of ensuring the security of a computer system when formalizing and modeling the processes of such provision. These stages include:

Intrusion detection. Logical-dynamic systems should be used to synthesize subsystems for detecting possible intrusions. These subsystems will detect abnormal conditions in the entire computer network or on its individual computers which will allow for prompt detection and optimal identification of intrusions (security incidents).

Risk analysis. At this substage, the logical-dynamic system is proposed to be used to assess the security risks of a computer structure (based on the analysis of various factors, for example, the number of vulnerabilities, the frequency of attacks and possible losses from these incidents, etc.), as well as to form strategies for minimizing them or compensating for the destructive consequences of dangerous incidents that will interfere with the normal functioning of the computer system.

Incident Management. Implementation of management processes based on algorithms synthesized using logical-dynamic systems models will allow for the automation of reliable and prompt response to computer system security incidents, for example, blocking malicious traffic or complete “isolation” of infected computers in a computer system.

Predicting possible threats. Using logical-dynamic systems models to predict based on the analysis of accumulated a posteriori data in security event arrays about specific features of incidents in a computer system will ensure optimal identification of future incidents.

System analysis of the proposed stages shows that the implementation of logical-dynamic systems models will ensure the management of information and security events of a computer system. The use of logical-dynamic systems will also allow to realize a number of significant advantages over the known traditional methods of managing information and security events of a computer system, namely: flexibility, accuracy, automation of management and forecasting of the development of processes in these tools. A methodology for organizing the management of security events of a computer system has been developed which is based on the use of the theory of logical-dynamic systems and is given in Table 1.

The dominant component of the proposed methodology is the synthesis of a system for managing information and security events of a computer system. Let us consider, as an example, one of the components of this management system in more detail, namely: a subsystem for detecting unauthorized intrusions into the software environment of a computer system.

Table 1 – Methodology for organizing information and event management of computer security structures

Stages of the methodology	Algorithms for implementing the stages of the methodology
Formalization of the description of the system and its subsystems	Description of the computer structure model (network with servers and workstations), its structure and specifics of functioning
Identification of possible (probable) security incidents and events	Formation of a list of security events that taking into account the probability of their occurrence will occur in the system: • login of an unauthorized user; • attempt to gain access to confidential information; • appearance of various viruses; • attack on the server to further create a denial of service to users; • attack on the server for unauthorized removal, distortion or destruction of information; • appearance of new events (zero-day events).
Development of the logical component of the model	The logical component of the computer structure model is synthesized which describes the structural relationships between the logical elements (states) of the system: • login to the system by an unauthorized user – logical transition: attempt to gain unauthorized access to confidential information; • detection of a virus on a computer – logical transition: attack on other computers in the system (network).
Development of a dynamic component model	The dynamic component of the model is synthesized. Its structure consists of equations that describe the temporal variations of incidents (security events), namely: • the time required for guaranteed detection of intrusions into the system; • the speed of virus propagation in the system; • the duration of the attack on the system.
Control system synthesis	A security information and event management system is synthesized, the structural components of which are: • a subsystem for detecting unauthorized intrusions into the software environment; • a subsystem for analyzing possible risks; • a subsystem for predicting situations of possible threats; • a subsystem for managing incidents (security events).
Testing and tuning	• the information and security event management system is tested on a real or virtual model (digital twin) of the computer structure; • the information and security event management system is configured in accordance with the requirements for the given security level of the computer structure.

Applying the theory of logical-dynamic systems to form an algorithm of functioning we will consider the following functional elements:

Data collection. Collects data about all current computer system states and events that occurred on the network, such as changes in IP addresses, MAC addresses, port states, use of management protocols, etc.

Data analysis and optimal processing. Analyzes data collected in security event arrays and optimally processes it to further identify anomalies in network behavior that indicate an intrusion into a computer system.

Intrusion detection and identification. Detects anomalies according to the intrusion identification template, generates notifications (if necessary) and forms control actions to compensate for the destructive consequences of these incidents.

Intrusion response. Implementation of control actions for automatic control of the process of optimal and prompt response to unauthorized intrusions (for example, forms effective control actions for emergency blocking of the attacker's IP address or for complete and reliable isolation of an infected computer).

To implement the methodology for organizing the management of information and events of computer system security using the theory of logical-dynamic class systems, we will consider specific aspects of its axiomatics. In the future, we will use the obtained theoretical results in the synthesis of control systems for ensuring the security of computer systems.

It should be noted that the modeling of information and security event management processes of a computer system using the theory of logical-dynamic systems differs significantly in essence and content from classical approaches to the organization of management using the theory of automatic control systems, for example, in the management of pulse systems, systems with modal control or multi-connected control systems, etc. Considering that the control object - a security event of a computer system is described, as a rule, by models of complex systems that combine two fundamental classes of automatic systems, namely logical and dynamic. We will use a number of new categories and concepts of a mathematical nature, and we will conduct their further interpretation and adaptation for the synthesis of a system for managing information and security events of a computer system. To do this, we will consider a possible approach to the mathematical description of a logical-dynamic system, which will be the basis for creating a methodology for the synthesis of complex hypercomplex control structures [24], [25]. Under a dynamic system, we will consider a system that is generally described axiomatically by a mathematical model represented by an ordered set (tuple):

$$SD = (T, X, U, \Omega, Y, \Gamma, \Phi, Z). \quad (1)$$

Applying a number of necessary additional definitions in model (1) and using the appropriate interpretation for describing processes in complex large

organizational systems, we will describe the models of the component subsystems of the dynamic system SD , namely:

a) finite dynamic subsystem

$$SD_{[0,1]} = (\Omega, Y, Q, \lambda, \delta), \quad (2)$$

б) диференціальної динамічної субсистеми

$$SD_{[0,\infty]} = (T, U, X, Y, \varphi, \eta),$$

в) symbiotic (synergistic) combination of mathematical expressions (1) and (2) will describe a model of a logical-dynamic system

$$\begin{cases} SD_{[0,1]} = (\Omega, Y, Q, \lambda, \delta), \\ SD_{[0,\infty]} = (T, U, X, Y, \varphi, \eta), \\ \Omega_{[0,1]} = (T, Y, U, Z). \end{cases} \quad (4)$$

Components of a tuple (1) which are included in (2), (3), (4) describe the variable logical and dynamic states of the system logical-dynamic system. We will describe these components in more detail below.

Note that the set of time instants T is considered in (4) as an ordered subset of the set of real numbers. For finite dynamical systems this set of time instants is T will be plural Z which corresponds to an ordered abelian group of integers. Multiple input influences Ω not empty. Therefore, we can specify a set of instantaneous values of input influences U and instantaneous values of the output quantities Y .

Write the system state transition function, namely:

$$\varphi: T \times T \times T \times X \times \Omega \rightarrow X. \quad (5)$$

The values of this function describe its following possible states:

$$x(t) = \varphi(t, \tau, x, \omega) \in X. \quad (6)$$

The given function (6) has the properties of directed time, consistency of the mapping (4), causality, and also semigroup properties

$$\varphi(t_3, t_1, x, \omega) = \varphi(t_3, t_2, \varphi(t_2, t_1, x, \omega), \omega), \quad (7)$$

for any values $t_1 < t_2 < t_3$ and $x \in \Omega$, $t_1 < t_2 < t_3$.

Let's write down the values $\eta: T \times X \rightarrow Y(t)$. This expression defines the output values $y(t) = \eta(t)$ by the values of the state transition function, namely: $y(t) = \eta(t, x(t))$.

Let us make some clarifications to adapt the mathematical models (5), (6), (7) to the applied terminology of real systems. Then (τ, x) where $\tau \in T$, $x \in X$ is called a "system event". In this case the phase space of events will be the set $T \times X$. System state transition function φ is interpreted as its trajectory (variations of the system) in the form of a solution to an ordinary differential equation. Control (inputs to the system) ω or transforms a state describing system events x into a state that is described as $\varphi(t, \tau, x, \omega)$. We emphasize that considering variations of the system

SD in the context of the logical-dynamic approach we mean solutions φ along a certain path of controlled discrete transitions on a finite set of states – the system outputs. These states adequately describe the current events in the system. In this case let us assume that for models of logical-dynamic systems $Q = Y$.

A dynamic differential system is defined in such a way that its models have the properties of smoothness, reversibility and completeness [1]. Thanks to these properties we can formalize the concepts of reachability, controllability and observability, on the basis of which it is intended to consider the problem of optimal control of the state or events in the specified system. When synthesizing complex control systems which include the system of information and event management of the security of a computer structure we propose to apply the theory of dynamic differential systems to solve a narrow range of control problems within the framework of individual aspects of the functioning of these systems. But this theory is a convenient mathematical apparatus, the use of which allows formalizing models of control processes in a simplified formulation. The specified axiomatics is characteristic of the stage of system synthesis, the purpose of which is, firstly, to build the structure of the system, and secondly, to create a general law of the functioning of the control system and individual laws of the functioning of its subsystems. It is also advisable to carry out the synthesis of the information and event management system of the security of a computer structure using the theory of finite automata [23]. We will perform it based on the given law of the functioning of the discrete (logical) part of this control system. For this purpose we will use the applied apparatus of Boolean algebra which contains means of analyzing finite dynamic systems that are of an iterable nature, in connection with which taking into account all the volumes of information processing is one of the main difficulties in solving the applied problems of this synthesis. We will also take into account other difficulties in the synthesis of modern systems for managing security events of a computer structure, which are associated with a number of the following manifestations in the processes of its functioning. Let us consider these manifestations in more detail. The intensification of the processes of functioning of modern computer structures generates incompatibility of restrictions that act as contradictory requirements between the provisions of the normative theory and the limiting physical constants. This is what incompatibility of constraints on elements means $u(t)$, $x(t)$ plurals U and X where $u(t) \in U$, $x(t) \in X$. The unification of automation tools for managing information processing processes in a computer structure, as a rule, establishes a finite class of functions, limiting physical constants (for example, the volumes of the computing environment of these structures and the volumes of information processed, etc.), generates on the set X type restriction system

$$x_i(t) \leq x_{0\max}(t), \quad (8)$$

Within the framework of one model, these requirements (8) are incompatible. This is due to the fact that the equation

$$x_i(t) = f_i(x, u, t), \quad i = 1, 2, \dots, n \quad (9)$$

for $u(t) \in U$ and $x_i(t) \leq x_{0\max}(t)$ are valid only on a finite subinterval $[t_0, t_1]$.

To implement the necessary process of system functioning on the entire set of values T the parameters of the object itself must also be controllable. This requirement is expressed in the discretely varying functions in equation (7). From this requirement it follows that the automatic information and event management system of the security of a computer structure is synthesized as a coordinate-parametric control system [24]. Sets of constraints on sets U and X have the volumes ("power") of finite sets. This in turn means that a discrete controlled process on a set of functions $f_{il}, i = 1, 2, \dots, n$ (by index $l = 1, 2, \dots, s$) makes sense in processes controlled by finite state machines. The functioning of this object occurs in a wide range of various heterogeneous states (different phases, system state, different nature of interaction in the information environment, etc.). Let us take into account that complex objects, as a rule, have a multi-purpose nature of functioning. This implies the need to implement a controlled change in the structure of interacting subsystems. In the process of achieving a goal or some planned sequence of sub-goals the control object undergoes structural variations in unidentified and critical situations. As a result the set of these conditions requires the analysis and synthesis of a system with coordinate-structural control. The structure of the system must be controllable $SD_{[0, \infty]} = (T, U, X, Y, \varphi, \eta)$ which implies a variation in the original mathematical model $F_k \in F$ functions $f_{(il)k} \in F_k, k = 1, 2, \dots, K$.

In all considered cases, which assume the presence of a controllable system structure, we will use the following categories of sets, namely:

- finite sets of threshold and other constraints on the phase coordinates of the system state;
- finite sets of normalized (unified) functions and quantities;
- finite set of subgoals and their relative ordering;
- finite set of homogeneous operating conditions, information environments and their relative ordering;
- finite set of system structures generated by abnormal situations (events) controlled by processes and operations.

We will define the elements of these finite sets formally. They are mathematical objects that model discrete transitions on finite sets of coordinate functions $u(t) \in U$, parameters $f_{il} \in F_k$ and structures $F_k \in F$ in the laws of functioning of complex systems with coordinate-structural control. To formalize the law of control of the functioning of a complex system, it is necessary, in addition to determining the elements of finite sets U, F_k and F , interpreted as discrete states of

the system, we will also describe the probable paths of transitions on these sets. The proposed approach models, together with the transition conditions the order of changes in the states of the system or the law of functioning of a finite dynamic system. Transition conditions on finite sets are attributed to the class of deterministic conditions – two-valued predicates. The domain of definition of the subject variables of such predicates is vector spaces with a scalar or probability measure, countable sets of time cycles and other heterogeneous sets of events. The domain of values of the specified predicates is described by the two-element set $\{0, 1\}$. Then the calculation of two-valued inhomogeneous predicates transfers the synthesis of the model of functioning of complex coordinate-structural control systems into the domain of finite analysis. This analysis is based on Boolean algebra which will allow identifying the logical part of the laws of functioning of such systems by finite state machines. Based on the considered approach, a mathematical description of systems of the logical-dynamic class as models of functioning with structural and parametric control is obtained [25]. This approach to the organization of control in systems of the specified class is successfully algorithmized, since the functions of predicates are optimally implemented in the computing environment of the computer structure. In the established methodology for the synthesis of information and event management systems for computer structure security which is based on the theory of a logical-dynamic system described by a mathematical model (3) we distinguish the following stages:

Systemic. At this stage, the architectures (structures) of the system are synthesized and the laws of its functioning are formed.

Structural-logical. This stage involves the formation of a structure for algorithmic support of computer structure management processes.

Software-technical. At this stage, the creation of separate functional blocks of the software for the information and event management system for computer structure security and their implementation in these information structures is envisaged.

It should be noted that each of the above stages is accordingly based on independent and inherent mathematical models and approaches to system analysis and optimization of management processes, as well as on corresponding algorithmic procedures.

The results of theoretical studies show that the complication of the laws of system functioning can lead to the fact that it will be problematic to describe the synthesis object as a single (indivisible) whole by the models of the system stage. As a result, it is difficult to perform the operation of quantitative comparison of alternatives or to interpret the synthesis object and its functioning processes by numerical indicators of modeling. The synthesis process itself using models of the system, structural-logical and software-technical stages is problematic to combine functionally. The problem of the complexity of a single structural procedure arises, since the results of the synthesis of control laws obtained at the system stage require

flexible adaptation to the model of the software-technical stage. We will take this problem into account in the future when ensuring the quality of management.

The criteria of quality of management, the main of which are the speed of working out of controlling influences and the minimum time spent on the implementation of management will be represented by elements of models in sets of vector spaces. But the peculiarity of these spaces is that they are almost completely “opened up” as a single system of assessments in the categories of models of the structural-logical and program-technical stages. This indicates that the criteria of quality of management are constructive within the models and are invariant with respect to the criteria of the efficiency of the functioning of the system being synthesized. Therefore, we will take into account this feature that the synthesis process as a

decisive procedure is mathematically open-ended. Closure involves the representation of the main estimates and algorithmic and programmatic components of the object by categories of models of the system stage, which will be supplemented at the next stages of the implementation of the models. The above allows us to conclude that the step-by-step solution of synthesis tasks from the point of view of mathematical correctness is consistent. At the same time the ability of the logical-dynamic system to model complex, dynamic environments and include both logical rules and dynamic behavior makes them suitable for solving current problems of information management and security events in modern and promising computer structures. The results of solving these problems are proposed to be implemented in algorithms the structure of which is presented in Table 2.

Table 2 – Structure of algorithms for managing information and security events of a computer structure

Algorithm components	Algorithm substeps	Functions that are implemented using the theory of logical-dynamic systems
Adaptive security management	Real-time threat detection	- models the dynamic nature of incidents, adapting to variations in new security events and vulnerabilities of the computer structure; - detects anomalies and potential violations in real time (thanks to logical security policy rules and dynamic models that describe the behavior of the computer structure).
	Responding to security incidents and events	- promptly generates the necessary control actions to ensure security based on predefined rules and results of analysis of the current state of the system; - initiates automated (automatic) countermeasures, such as isolating affected systems or blocking malicious traffic.
Risk assessment and minimization	Vulnerability analysis	- models the relationships between different components in a computer structure; - models approaches to identifying potential vulnerabilities in a computer structure; - models possible attack scenarios to assess the risks associated with variations in vulnerabilities; - models the hierarchy of priorities for efforts to mitigate or compensate for the consequences of incidents.
	Dynamic risk management	- constantly monitors the security status of the system; - promptly adjusts security event management in accordance with the current risk level; - automatically increases the level of security monitoring in the event of a significant increase in the flow of incidents; - implements if necessary more stringent access control to the software environment of the computer structure.
Ensuring compliance with security policies	Formal modeling of security policies	- modeling security policies for use in a computer structure; - forming logical rules that interpret access control policies; - adjusting information array protection rules in accordance with current security requirements.
	Maintaining compliance with security policies	- checking the system configuration for compliance with defined security policies; - identifying abnormal deviations in the system configuration from the reference (defined, specified) security policies; - ensuring (guaranteeing) the system's compliance with established security standards and rules.
Intrusion detection and prevention	Anomaly detection	- evaluates the behavior of the system for compliance with established norms (rules); - promptly detects deviations that indicate unauthorized intrusion; - synergizes logical rules that determine the given behavior of the system for compliance with its dynamic models of functioning; - detects non-specific (unconventional, unknown) anomalies.
	Responding to an intrusion	- automatically responds to detected intrusions; - automatically blocks malicious information transmission traffic; - blocks suspicious processes in the information environment and promptly generates danger alerts.

To create a methodology for organizing information and security event management in a computer structure (Table 2), we adapt through specification and necessary additional definitions the theory of logical-dynamic systems and the mathematical model (3) for the further implementation of process management in order to significantly mitigate or compensate for the consequences

of computer incidents. When synthesizing the processes of information and event management of computer security in the mathematical model, a number of formalization descriptions of its constituent (structural) components were used. The proposed approach provides the possibility, through decomposition (3), to present the model of a logical-dynamic system in the form of:

$$S = (X, Y, Z, f, g', L, h). \quad (10)$$

Taking into account the axiomatics of the logical-dynamic system we will further define the mathematical expressions (1)–(9) which makes it possible to formalize its components in (10).

Let us consider the components of the set structure (10) in detail.

1) $X = \{x_1, x_2, \dots, x_n\}$ – a set that describes the current states of a computer structure. These states are formalized by the following indicators:

- use of computer structure resources (processor, RAM, data storage devices, etc.);
- the number of network traffic patterns used to ensure information security at the physical level;
- the number of processes involved for secure information processing;
- indicators of the status of security software (firewalls, antiviruses, etc.);
- system vulnerability status indicator, etc.

2) $Y = \{y_1, y_2, \dots, y_m\}$ – a set that comprehensively describes security events in a computer structure, namely:

- unauthorized intrusion attempts;
- presence of malicious software;
- denial of service attacks;
- leakage of confidential information;
- system errors in information processing, etc.

3) $Z = \{z_1, z_2, \dots, z_k\}$ – a set of actions for implementing security measures (security management). This set describes the following actions:

- blocking IP addresses;
- implementation of quarantine dossiers;
- emergency (safe) completion of information processing processes;
- correction (adjustment) of the consequences of vulnerabilities;
- prompt notification of computer security personnel.

4) $f: X \times Y \rightarrow X$ – a function that formally describes changes in system states. This function characterizes changes in the system state depending on the destructive effects of security events. Taking into account the metric of the current system state $x \in X$ and security events $y \in Y$, function $f(x, y) = x' \in X$ where x' – a new state of the system that characterizes variations in system dynamics depending on incidents.

5) $g: X \times Y \rightarrow Z$ – a function of selecting security management approaches based on the current state of the system and security events that have occurred in the computer structure. It determines what actions must be taken promptly in response to a security event to ensure the functional stability of the computer structure. Current state is set $x \in X$ and security event $y \in Y$, $g(x, y) = z \in Z$. This function reflects the security policy or decision-making process for

significantly mitigating the destructive consequences of incidents or fully compensating for them.

6) $L = \{l_1, l_2, \dots, l_p\}$ – a set of logical conditions for ensuring the safe state of the system. These conditions can be used to refine approaches to choosing an action in the case of:

- CPU load over 80%;
- the firewall is disabled;
- the number of unsuccessful login attempts is more than three, etc.

7) $h: X \rightarrow \{True, False\}_p$ – a function that is described by a set of logical conditions of the states of a computer structure (which are true or false) $h(x) = \{h_1(x), h_2(x), \dots, h_p(x)\}$, where $h_i(x)$ – truth value of a logical condition l_i in possible states x .

8) $g': X \times Y \times \{True, False\}_p \rightarrow Z$ – an extended action selection function that takes into account not only the system state and event, but also the truth value of logical conditions $g'(x, y, h(x)) = z \in Z$. Its application allows for the implementation of more complex and context-dependent security responses in the computer structure. We will present the developed dynamics of incident compensation in the form of a formalized model that implements the phased steps of this compensation and display it in the form of an algorithm:

Step 1: /Initial state/ *The system is (starting) in the initial state* $x_0 \in X$.

Step 2: /Occurrence of an event – the occurrence of a security incident/ *A security event is occurring.* $y \in Y$.

Step 3: /Change (transition) of states/ *Change (transition) of states* y_i *transitions to a new state* $x'_i = f(x_i, y_i)$.

Step 4: /Condition assessment/ *Logical conditions* L *are displayed with the new status:* $h(x') = \{b_1, b_2, \dots, b_p\}$, where $b_i \in \{True, False\}$.

Step 5: /Formation of management actions (influences)/ *Based on the new state* x'_i *and security events* y'_i *an action is being formed to manage security information and events* $z_i = g(x'_i, y_i, h(x'))$.

Step 6: /Implementation of management actions/ *Selected action* z_i *is processed by the system and affects the change in its state.*

Step 7: /Repetition/ *When new (subsequent) incidents (security events) occur, the process from step 2 to step 6 is repeated cyclically.*

Present the implementation of the application of model (9) taking into account the proposed step-by-step stages. In this example, if the system is in a secure state and an attempt is made to log in to the system, the system remains in a secure state, and the security management action consists in blocking access. If a security event (incident) is detected, the system enters a vulnerable state, and the security management action involves compensating for the destructive consequences

of the security event. The formalization of the specified approach is presented in the form of a variation of the states of the computer structure:

$X = \{safe, vulnerable\}$ – system states.

$Y = \{login\ attempt, incident\ detected\}$ – security events.

$Z = \{block\ access, compensate\ for\ the\ consequences\ of\ the\ incident\}$ – safety management actions.

State transition function f is defined as follows:

$f(secure, login\ attempt) = safe$.

$f(safe, security\ event\ detected) = impressionable$.

$f(vulnerable, login\ attempt) = impressionable$.

$f(вразливий, виявлено\ вірус) = impressionable$.

Action selection function g introduce it like this:

$g(secure, login\ attempt) = block\ access$.

$g(safe, incident\ detected) = compensate\ for\ the\ consequences\ of\ the\ incident$.

$g(vulnerable, login\ attempt) = block\ access$.

$g(vulnerable, incident\ detected) = compensate\ for\ the\ consequences\ of\ the\ incident$.

Taking into account this formalization, the incident detection algorithm can be presented in a simplified form as follows:

$X = \{Normal, high, attacked\}$.

$Y = \{LoginFailure, MalwareDetected\}$.

$Z = \{BlockIP, QuarantineFile, LogAlert\}$.

$L = \{FailedLoginAttempts > 3\}$.

Transitions and actions:

$f(Normal, malware\ detected) = attacked$.

$f(Normal, LoginFailure) = HighCPU$ (if there were many failed logins).

$g'(HighCPU, LoginFailure, True) = BlockIP$ (if $FailedLoginAttempts > 3$ true).

$g'(Attacked, MalwareDetected) = QuarantineFile$ (regardless of other conditions).

Conclusions

Thus, the paper proposes an approach to the synthesis of a logical-dynamic system for managing information and security events of a computer structure. It is noted that by synergistically integrating logical and dynamic components of the processes of functioning of a computer structure into models, it is possible to synthesize systems that will not only detect and respond to security threats, but also predict and prevent potential vulnerabilities. The synthesized systems have increased adaptability and functional resistance to security threats. The implementation of logical rules provides accurate correlation and analysis of events, and dynamic modeling facilitates the prediction of system behavior and operational compensation of destructive factors. Further scientific research is proposed to be carried out with an emphasis on the practical implementation and evaluation of the results obtained in real systems for managing information and security events of a computer structure, as well as on improving dynamic models and automating the adaptation of logical rules.

REFERENCES

- Hrytsiuk P. M., Dzhoshi O. I. ta Hladka O. M. (2021) Osnovy teorii system i upravlinnia. NUVHP, Rivne. 272 s.
- Yashchuk V.I. Proektuvannia avtomatyzovanykh informatsiinykh system upravlinnia kibernetichnoi bezpekoii ob'ektiv krytychnoi infrastruktury Ukrainy, Aktualni problemy upravlinnia informatsiinoiu bezpekoii derzhavy: zb. tez nauk. dop. nauk.-prakt. konf. (Kyiv, 26 bereznia 2021 r.). [Elektronne vydannia]. – Kyiv : NA SBU, 2021. C. 233–235.
- Hrebeniuk A.M. Osnovy upravlinnia informatsiinoiu bezpekoii: navch. posib. / A.M. Hrebeniuk, L.V. Rybalchenko. – Dnipro: DDUVS, 2020. – 144 s. <http://er.dduvs.edu.ua/handle/123456789/5717>.
- Hnusov Yu. V., Mozhaiev O. O. Rozrobka multyagentnoi systemy upravlinnia informatsiinoiu bezpekoii. Pravo i bezpeka. 2022. No 4 (87). S. 171–183. <https://doi.org/10.32631/pb.2022.4.14>.
- Mokhor V., Tsurkan V. (2021). Kompiuterna model rozrobky system upravlinnia informatsiinoiu bezpekoii. Zbirnyk "Informatsiini tekhnologii ta bezpeka", 9 (1), C. 80–90. <https://doi.org/10.20535/2411-1031.2021.9.1.249814>.
- Tsurkan V. Metod syntezy struktury systemy upravlinnia informatsiinoiu bezpekoii // Ukrainskyi naukovyi zhurnal informatsiinoi bezpeky, 2020, tom 26, vypusk 2, C. 116–122. DOI: 10.18372/2225-5036.26.14926.
- Tsurkan V., Shapoval O. (2022). Analiz metodiv otsinky ryzykiv bezpeky kompiuternoi merezhi. Zbirnyk "Informatsiini tekhnologii ta bezpeka", 10 (2), C. 204–215. <https://doi.org/10.20535/2411-1031.2022.10.2.270437>.
- Volodymyr Mokhor, Vasyi Tsurkan Metodolohiia pobudovy system upravlinnia informatsiinoiu bezpekoii zakhyst informatsii, tom 23, № 4, zhovten-hruden 2021, C. 200–212. DOI: 10.18372/2410-7840.23.16766.
- Tsurkan V. (2020). Metod funktsionalnoho analizuvannia system upravlinnia informatsiinoiu bezpekoii. Elektronne fakhove naukove vydannia "Kiberbezpeka: osvita, nauka, tekhnika", 4(8), C. 192–201. <https://doi.org/10.28925/2663-4023.2020.8.192201>.
- Ushkarenko O.O. Metod analizu protsesiv peretvorennia danykh v obchysliuvalnykh vuzlakh tsyfrovyykh system upravlinnia Vcheni zapysky Tavriiskoho NU imeni V.I. Vernadskoho. Seriia: Tekhnichni nauky Tom 32 (71) № 4 2021, stor. 162...168, <https://doi.org/10.32838/2663-5941/2021.4/25>
- Skitsko Oleksii, Shyrshov Roman. Systema upravlinnia informatsiinoi bezpeky yak instrument pidvyshchennia zakhyshchenosti ta efektyvnosti ob'ektiv krytychnoi infrastruktury. International Science Journal of Engineering & Agriculture. Vol. 2, No. 6, 2023, pp. 12–22. doi: 10.46299/j.isjea.20230206.02
- I. Samborskyi, Ie. Samborskyi, V. Hol, Y. Peleshok, S. Sholokhov .,(2024) Synthesis of the model of management of complex dynamic objects taking into account the events of their security. Information Technology and Security. January-June 2024. Vol. 12. Iss. 1 (22) P. 4–16. <https://doi.org/10.20535/2411-1031.2024.12.1>
- Sholokhov S.M., Pavlenko P.M., Nikolaienko B.A., Samborsky I.I., Samborsky E.I.,(2023/2024) The method of optimizing the distribution of radio suppression means and destructive software influence on computer networks. Radio Electronics, Computer Science, Control. 2023/2024. № 4 (67). P. 16–29. URL: <https://doi.org/10.15588/1607-3274-2023-4-2>
- Ye. Samborskyi, I. Samborskyi, Osoblyvosti obrobky masyviv podii bezpeky dlia upravlinnia zakhystom elektronnykh komunikatsiinykh system. Materialy II Mizhnarodnoi naukovo-praktychnoi konferentsii "Kiberbezpeka derzhavnykh instytutsii ta podolannia kryzovykh staniv" : v 2 t. Kyiv : ISZZI KPI im. Ihoria Sikorskoho, 2023. T. 1.s. 293–294.

15. P. Pavlenko, Ye. Samborskyi, Modeli obrobky podii bezpeky dlia upravlinnia zakhystom kompiuternykh system. Materialy KhVI mizhnarodnoi naukovo-tekhnichnoi konferentsii "AVIA-2023". – K.: NAU, 2023 C. 13.8–13.11
16. Ye. I. Samborskyi, Rozrobka modeli upravlinnia podiiamy bezpeky kompiuternykh system, Materialy X Mizhnarodnoi naukovo-tekhnichnoi Internet-konferentsii "Suchasni metody, informatsiine, prohramne ta tekhnichne zabezpechennia system keruvannia orhanizatsiino-tekhnichnymy ta tekhnolohichnymy kompleksamy", 24 lystopada 2023 [Elektronnyi resurs]. – K: NUKhT, 2023. – 224 s. – Rezhym dostupu: <https://nuft.edu.ua/naukova-diyalnist/naukovi-konferencii>, s. 213.
17. Ye. Samborskyi, I. Samborskyi, Model otsinky ta upravlinnia podiiamy bezpeky kompiuternykh system, Materialy VI Naukovo-praktychna konferentsiia kursantiv (studentiv), aspirantiv, doktorantiv ta molodykh uchenykh "Aktualni pytannia zastosuvannia spetsialnykh informatsiino komunikatsiinykh system", Kyiv : ISZZI KPI im. Ihoria Sikorskoho, 2023, s. 38–39.
18. Sholokhov S.M., Peleshok Ye.V., Samborskyi Ye.I. Syntez synerhetychnoi modeli optymalnogo upravlinnia protsesamy zabezpechennia informatsiinoi bezpeky kompiuternykh system. Perspektyvy rozvytku ozbroiennia ta viiskovoi tekhniki Sukhoputnykh viis: Zbirnyk tez dopovidei Mizhnarodnoi naukovo-tekhnichnoi konferentsii (Lviv, 15-16 travnia 2024 roku), Lviv, NASV, 2024, C. 283–284.
19. Pavlenko P.M., Samborskyi Ye. I. Syntez systemy upravlinnia podiiamy bezpeky dlia zabezpechennia zakhystu informatsiinoho seredovyscha infrastruktury krytychnykh ob'ektiv, // Materialy KhI Vsesvitnogo konhresu "Aviatsiia v XXI stolitti" – "Bezpeka v aviatsiino-kosmichnykh tekhnolohiiakh", 25–27 veresnia 2024, Kyiv: NAU, 2024, stor.11.16–11.19.
20. Ye. Samborskyi, I. Samborskyi, Osoblyvosti zastosuvannia tsyfrovyykh dviinykh dlia upravlinnia podiiamy bezpeky kompiuternykh system krytychnykh infrastruktur, // Materialy III Mizhnarodnoi naukovo-praktychnoi konferentsii "Kiberbezpeka derzhavnykh instytutsii ta podolannia kryzovykh staniv", 14 veresnia 2024, Kyiv : ISZZI KPI im. Ihoria Sikorskoho, 2024. 294 c.
21. Samborskyi Ye. I., Zastosuvannia tsyfrovyykh dviinykh dlia optymizatsii upravlinnia podiiamy bezpeky kompiuternykh system, 22 lystopada 2024 roku. Materialy II Mizhnarodnoi naukovo-praktychnoi Internet-konferentsii "Innovatsii ta perspektyvni shliakhy rozvytku informatsiinykh tekhnolohii" (IPShRIT-2024), m. Cherkasy, 2024. C. 105–106.
22. Pavlenko P.M., Samborskyi Ye.I. Model upravlinnia podiiamy bezpeky kompiuternykh system krytychnoi informatsiinoi infrastruktury na osnovi tsyfrovyykh dviinykh, Materialy II mizhnarodnoi naukovo-praktychnoi konferentsii "Suchasni aspekty didzhitalizatsii ta informatyzatsii v prohramnii ta kompiuternii inzhenerii", 19–21 hrudnia 2024 roku., m. Kyiv, Derzhavnyi universytet informatsiino-komunikatsiinykh tekhnolohii , 2024, C. 181–184.
23. Krychevets O.M., Zastosuvannia elementiv teorii kintsevykh avtomativ dlia doslidzhen dynamichnykh vlastyvostei obchysliuvalnykh komponentiv vymiriuvalnykh system. Vymiriuvalna tekhnika ta metrolohiia 77 (2016): C. 121–126.
24. Rabchun D. I, Lohiko-dynamichna model protsesu upravlinnia resursamy zakhystu v umovakh informatsiinoho protystoiannia. Suchasnyi zakhyst informatsii, №3, 2016, s.62–67.
25. Zhuk K. D. Pytannia aksiomatychnoho pidkhotu do pobudovy teorii lohiko-dynamichnykh system keruvannia, Avtomatyka, № 3-6, 1971, s. 62–74.

Received (Надійшла) 21.02.2025

Accepted for publication (Прийнята до друку) 16.04.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Самборський Євген Іванович – аспірант кафедри організації авіаційних перевезень Національного університету "Київський авіаційний інститут", Київ, Україна;

Ievgen Samborskyi – Postgraduate student, Department of Air Transportation Organization, National University "Kyiv Aviation Institute", Kyiv, Ukraine.

e-mail: seinauedu@gmail.com; ORCID Author ID: <https://orcid.org/0000-0003-4441-1947>.

Пелешок Євген Володимирович – кандидат технічних наук, старший науковий співробітник Науково-дослідного інституту воєнної розвідки, Київ, Україна;

Yevhen Peleshok – Candidate of Technical Sciences (PhD), Senior Researcher, Research Institute of Military Intelligence, Kyiv, Ukraine.

e-mail: pel85@ukr.net; ORCID Author ID: <https://orcid.org/0000-0003-0033-1160>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57209822977>.

Синтез логіко-динамічних систем управління інформацією та подіями безпеки комп'ютерних структур

Є. І. Самборський, Є. В. Пелешок

Анотація. У статті розглядається один з можливих підходів до вирішення актуального наукового завдання, присвяченого синтезу логіко-динамічних систем управління інформацією та подіями безпеки комп'ютерних структур. Зроблено аналіз сучасних підходів до створення інтегрованої системи управління, реалізація якої дозволить ефективно керувати інформаційними масивами подій безпеки в умовах змінених загроз комп'ютерним структурам. Окремо увага приділяється розвитку моделей, які об'єднують логічні та динамічні аспекти управління, що дозволяють забезпечити високу адаптивність і функціональну стійкість комп'ютерних структур в умовах інтенсифікації активності подій безпеки. Реалізація такого підходу дозволить автоматично реагувати на ці події, оптимізувати процеси виявлення та усунення наслідків загроз, а також забезпечити постійну стабільність роботи комп'ютерних структур. У статті пропонується метод синтезу логіко-динамічної моделі, яка синергетизує не лише технічні, але й організаційно-кібернетичні аспекти безпеки комп'ютерних структур. Розроблені підходи можуть бути рекомендовані для використання при створенні більш ефективних та безпечних комп'ютерних засобів, здатних протистояти різноманітним інцидентам.

Ключові слова: система управління, логіко-динамічна система, інформаційна безпека, подія безпеки, синтез, аналіз, теорія логіко-динамічних систем, комп'ютерна структура, інцидент, масив подій безпеки.

Д. Тереник, В. С. Харченко

Національний аерокосмічний університет «Харківський авіаційний інститут», Харків, Україна

АВТОМАТИЗАЦІЯ СТВОРЕННЯ КЛАСИФІКАТОРІВ: АЛГОРИТМИ І ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ ДЛЯ ФОРМУВАННЯ МНОЖИНИ СТРАТЕГІЙ РОЗГОРТАННЯ ТА ЗАБЕЗПЕЧЕННЯ НАДІЙНОСТІ РОЮ БПЛА

Анотація. Актуальність. Забезпечення надійності та автономності безпілотних літальних апаратів (БПЛА), а також їх роїв (РБА) потребує ефективних методів розроблення і вибору стратегій їх розгортання, а також забезпечення функціональної надійності в складних умовах застосування. Кількість таких стратегій зростає, що вимагає відповідного інструментарію їх розроблення, класифікації та вибору. **Об'єкт дослідження:** процеси класифікації та автоматизації побудови класифікаторів стратегій розгортання і забезпечення надійності (СРЗН) мобільних систем, зокрема РБА. **Мета статті:** розроблення методу автоматизації процесу створення та модифікації (розширення і зміни структури) класифікаторів СРЗН мобільних систем у різних предметних областях, демонстрація його застосування, на прикладі РБА, задля підвищення ефективності прийняття рішень за наявності великої кількості чинників предметної області, які треба брати до уваги. **Результати дослідження.** Проведено аналіз існуючих методів, запропоновано концепцію та алгоритми автоматизації процесів класифікації та моделювання СРЗН, розроблено і валідовано програмне забезпечення, надано приклади використання, які демонструють його здатність підтримувати рішення щодо формування множини і видалення суперечливих або практично неприйнятних стратегій. **Висновки.** Основним науковим внеском даного дослідження є запропонований і реалізований метод автоматизації, що базується на фасетно-ієрархічному описі ознак і декартових моделях для формування можливих конфігурацій стратегій, що може підвищити надійність та ефективність застосування РБА у складних та мінливих умовах завдяки формуванню повних і внутрішньо несуперечливих множин для вибору СРЗН.

Ключові слова: класифікатор стратегій, розгортання рою БПЛА, забезпечення надійності, фасетна класифікація, автоматизація прийняття рішень.

Вступ

Автономні мобільні системи, зокрема, рої безпілотних літальних апаратів (РБА), стрімко набувають поширення в таких сферах, як пошуково-рятувальні операції, моніторинг критичної інфраструктури, військові застосування, комунікація у складних умовах під час техногенних катастроф. Здатність швидко адаптуватися до динамічних умов, зокрема, до впливу фізичних та кіберзагроз, робить РБА важливим інструментом вирішення різноманітних задач у надзвичайних ситуаціях.

Забезпечення надійності та автономності РБА потребує ефективних алгоритмів і методів розроблення і вибору стратегій їх розгортання, а також забезпечення функціональної надійності при використанні [1]. Кількість і різноманітність таких стратегій зростає для задач моніторингу, подолання, забезпечення комунікацій та зменшення рівня впливу кіберфізичних впливів у небезпечних просторах (великі промислові споруди з руйнуваннями та підвищеним рівнем шкідливих речовин, території суходолу та акваторії, забруднені вибухонебезпечними предметами (ВНП), зони лісових пожеж та ін. [2, 3]). Вибір стратегій має передувати обстеженню предметної області та формуванню їх множини. Множина стратегій розгортання і забезпечення надійності (СРЗН) має бути повною, несуперечливою, детальною та систематизованою. Ця задача є важливою для забезпечення ефективності та надійності, важко формалізованою і потребує відповідних процедур та засобів програмної підтримки.

Мета статті – розроблення концепції, алгоритмів і програмного забезпечення для автоматизації процесів створення та модифікації класифікаторів

СРЗН для різних систем (на прикладі РБА), використовуючи фасетно-ієрархічний підхід для структурування ознак і врахування обмежень.

Сучасний стан досліджень. Проаналізуємо різні методи побудови класифікаторів стратегій.

Ієрархічна класифікація: базується на традиційному підході, за яким множина об'єктів послідовно декомпонується на підпорядковані підмножини на основі певної ознаки, утворюючи деревоподібну структуру. Перевагами цього підходу є відносна простота побудови. Однак недоліком є слабка гнучкість структури, обумовлена фіксованою основою поділу, що не допускає включення нових об'єктів без переробки всієї класифікаційної схеми [4].

Фасетна класифікація: дозволяє організувати складні предметні області за незалежними атрибутами (фасетами). Цей метод забезпечує гнучкість та масштабованість, особливо у випадках, коли додавання нових атрибутів не порушує загальної структури. У дослідженні [5] запропоновано фасетну класифікацію особливих випадків під час польоту БПЛА або РБА. Автори виділяють такі категорії, як потрапляння повітряного судна у небезпечні метеорологічні умови, втрата орієнтування, втрата просторового орієнтування та незаконне втручання. Цей підхід дозволяє систематизувати та ефективно аналізувати різноманітні ситуації, що можуть виникнути під час експлуатації БПЛА, забезпечуючи більш структурований підхід до управління ризиками та підвищення безпеки польотів.

Декартові моделі: використовуються для генерації множини варіантів шляхом комбінаторного аналізу. З математичного погляду, декартовий добуток дозволяє отримати повну множину варіантів рішень, у кожному з яких кожна ознака приймає

одне зі своїх допустимих значень. Водночас декартові моделі потребують обов'язкової фільтрації неефективних чи конфліктних комбінацій, оскільки не всі теоретично згенеровані варіанти виявляться допустимими або корисними.

Кластерне групування: використовується для групування об'єктів у кластери на основі схожості їхніх характеристик. Зокрема, алгоритм k-means є популярним підходом, який дозволяє ефективно класифікувати дані. У дослідженні [6] запропоновано вдосконалення цього алгоритму шляхом визначення початкових центроїдів та автоматичного розрахунку кількості кластерів, що підвищує точність класифікації. k-means передбачає, що кластери мають сферичну форму та однаковий розмір. Це обмеження може бути неприйнятним для даних зі складною структурою кластерів. Також алгоритм чутливий до різних масштабів ознак, тому необхідно проводити попереднє нормалізування даних, щоб уникнути домінування ознак з більшими значеннями.

Онтологічний підхід: забезпечує семантичну основу для організації знань, дозволяє моделювати взаємозв'язки між об'єктами, атрибутами та їх залежностями, що виходить за рамки класичної класифікації. Використання мови OWL (Web Ontology Language) дозволяє автоматично формувати нові знання.

Зазначимо, що вибір типу класифікатора обумовлюється, в першу чергу, особливостями предметної області (типом систем, умовами застосування, задачами використання тощо) і тому, зазвичай, можуть реалізовуватися комбіновані схеми. Зважаючи на необхідність гнучкої та масштабованої системи класифікації, яка дозволяє враховувати різноманітні комбінації атрибутів, а також з огляду на особливості РБА у цьому дослідженні обрано поєднання фасетної класифікації та декартових моделей. Такий підхід забезпечує можливість ефективної систематизації та аналізу стратегій завдяки врахуванню всіх можливих комбінацій незалежних атрибутів без порушення загальної структури класифікації.

Постановка проблеми. При роботі систем з поетапним виконанням завдань, місія є успішною за умов успішності на будь-якому з етапів. Отже, важливо забезпечити надійність системи на всіх етапах використання РБА [4]. Основна проблема полягає в тому, що планування розгортання та забезпечення надійності систем відбувається за умов невизначеності характеристик середовища, доступних ресурсів та обмежень, які можуть швидко змінюватися. З одного боку, необхідно формалізувати множину різних ознак (наприклад, параметри комунікацій, можливість ремонту тощо). З іншого, – чітко визначити обмеження і залежності, пов'язані з реальною експлуатацією (наприклад, варіанти резервування з дробовою чи цілою кратністю, вимоги з безперервності зв'язку тощо). Тому постає завдання розробити уніфіковану систему для підтримки опису ознак і обмежень, що охоплює різні сценарії застосування.

Автоматизація створення класифікаторів стратегій розгортання і забезпечення надійності є багатоаспектною задачею, що охоплює:

- *Формалізацію вимог* до класифікатора, яка передбачає структурування ознак та визначення їх залежностей та обмежень.

- *Забезпечення гнучкості оновлення* класифікаційних структур без порушення загальної схеми.

- *Розроблення алгоритмів* для автоматизації процесу створення множини стратегій, базуючись на заданому списку ознак та налаштувань залежностей.

- *Генерацію валідних стратегій*, що відповідають заданим обмеженням, на основі структурованих даних про ознаки та їхні обмеження.

Зважаючи на ці вимоги, необхідно розробити автоматизовану систему, яка дозволяє на основі структурованих даних про ознаки та обмеження генерувати множину валідних стратегій, усуваючи неефективні або суперечливі комбінації.

Крім того, перспективними є наступні задачі, які в даному дослідженні тільки окреслені:

- *Інтеграція програмного забезпечення* для моделювання процесу розгортання в реальному часі.

- *Побудова системи критеріїв*, що дозволяють оцінювати, ранжувати та обирати найкращі варіанти стратегій для конкретних умов.

Основні завдання розроблення алгоритмів і програмного забезпечення

Розроблення автоматизованої системи класифікації та генерації конфігурацій передбачає створення уніфікованих структур для опису ознак, обмежень і логічних зв'язків між ними.

Ця система має бути універсальною, придатною для використання в різних предметних областях, включаючи, але не обмежуючись, стратегічним плануванням, управлінням ресурсами та оптимізацією процесів.

Основними завданнями на даному етапі є наступні:

1. *Формалізація структур даних:* використання фасетно-ієрархічного підходу для опису ознак і обмежень включає:

- розробка уніфікованої схеми опису ознак;
- створення уніфікованого формату для обмежень;
- формалізація вихідних результатів.

2. *Автоматизація генерації конфігурацій*, що включає:

- розробка алгоритму створення множини можливих конфігурацій шляхом застосування декартового добутку ознак;
- фільтрація невідповідних (неприйнятних за певних причин) конфігурацій за визначеними обмеженнями, що забезпечує отримання тільки валідних варіантів.

3. *Реалізація програмного забезпечення*, що включає:

- розробка JSON-структури для зберігання ознак та обмежень у стандартизованому форматі;
- імплементація алгоритмів обробки конфігурацій для побудови множини допустимих конфігурацій;
- зберігання результатів.

Методологія побудови класифікаторів і розроблення структури даних

Розробимо методологію побудови класифікаторів і уніфіковану структуру даних для ознак (features) та обмежень (constraints), а також загальні правила формування валідних конфігурацій для подальшої генерації стратегій.

1. Методологія побудови класифікаторів

Процес побудови класифікаторів стратегій у межах цього дослідження спирається на такі положення:

1. Аналіз предметної області. Виділяється й формалізується множина можливих класифікаційних ознак (Set of Classification Features - SCF):

$$SCF = \{CF_i | i = 1, 2, \dots, n\}.$$

2. Встановлення відношень між ознаками. Запроваджуються фасетні, ієрархічні, матричні або змішані відношення, що утворюють множину взаємозв'язків ознак (Set of Feature Relationships - SFR). Отже, визначається як саме ознаки поєднуються або успадковуються.

3. Упорядкування ознак. З урахуванням відношень із SFR визначається кортеж класифікаційних ознак (Classification Features Tuple - CFT) з огляду на їх пріоритетність і логіку слідування.

4. Побудова підмножин сутностей. Для кожної ознаки CF_i формується підмножина сутностей (Subset of the Set of Entities - SSE_i) SSE_i. Сукупність усіх таких підмножин можна розглядати як множину сутностей класифікації (Set of Classification Entities - SCE).

5. Декартовий добуток. Повну множину класифікованих сутностей (Full set of Classified Entities FCE) розглядають як декартовий добуток усіх SSE_i:

$$FCE = SSE_1 \times SSE_2 \times \dots \times SSE_n.$$

6. Задання обмежень. Формально обмеження можна подати як сукупність множини (Conditions - C),

де кожне обмеження $ck \in C$ описує певне правило:

- Domain-обмеження:

$$ck: CF_i \in D_i,$$

де D_i - дозволений піднабір (домен) можливих значень ознаки CF_i.

- Conditional-обмеження (умовний тип):

$$ck: \text{якщо}(CF_i = v) \Rightarrow CF_j \in A_{ij},$$

де v - конкретне значення ознаки CF_i; A_{ij} - дозволений набір для CF_j за умови CF_i= v .

7. Застосування обмежень (фільтрація). Для кожного елемента $x \in FCE$ (тобто для кожної комбінації ознак) перевіряємо виконання всіх правил із C . Якщо x порушує бодай одне з обмежень, такий елемент вилучається. Формально це можна відобразити так:

$$FCE^* = x \in FCE \forall ck \in C: ck(x) = True,$$

тут $ck(x)=True$ означає, що комбінація x не порушує обмеження ck .

8. Фінальне відсіювання. У результаті залишається кінцева множина FCE^* , елементи якої відповідають класифікованим сутностям, що не суперечать жодному з визначених обмежень і мають практичну реалізованість.

У подальших розділах наведені приклади формату опису ознак CF_i та обмежень ck . В якості формату зберігання обрано JSON, оскільки він забезпечує зручність обробки, легкість інтеграції з іншими компонентами та можливість швидкого розширення структури.

2. Уніфікація опису ознак

Для кожної ознаки (feature) було визначено: назву, короткий чи розширений опис, тип даних (enum чи boolean), перелік можливих значень (domain), а також опційно групу. У табл. 1 наведено назву, опис та доступні значення полів сутності ознаки.

Таблиця 1 - Опис полів для features

Назва поля	Особливості	Допустимі значення	Опис
name	Обов'язкове Унікальне	Строка	Унікальний ідентифікатор ознаки
description	Не обов'язкове	Строка	Додатковий опис сутності ознаки (пояснює призначення або особливості цієї ознаки).
type	Обов'язкове	Одне зі значень: • enum; • bool	Вказує, як трактувати можливі значення: enum – перелік можливих опцій (наприклад, "ПН", "ПП"); bool – логічне значення, де допустимими значеннями будуть [true, false].
domain	Обов'язкове	Для типу enum це масив унікальних строкових полів ["Значення1", "Значення2", ...]. Для типу bool - це масив [true, false].	Перелік доступних значень для ознаки.
group	Не обов'язкове	Строка	Дає змогу об'єднувати ознаки за темами (наприклад, "Комунікаційні", "Надійності", "Архітектурні" тощо).

На рис. 1 наведено приклади опису ознак у форматі JSON.

Порядок, у якому зазначено ознаки у JSON (масив "features"), є критично важливим. Ознаки, розташовані вище у списку, вважаються більш важливими або базовими, і саме їх значення використовуються для формування конфігурації наступних за списком ознак. Таким чином, ключові параметри впливають на подальший вибір додаткових ознак.

```
{
  "name": "Можливість переривання зв'язку",
  "description": "Можливість переривання зв'язку",
  "type": "enum",
  "domain": [
    "ПН",
    "ПП"
  ],
  "group": "Комунікаційні"
}
```

Рис. 1. Приклади опису ознак

Розробник повинен уважно обрати порядок розміщення ознак у JSON, оскільки цей порядок буде застосовуватися при обчисленні комбінацій.

Вихідний рядок конфігурації також формується, використовуючи визначений порядок (список назв ознак). Це дає однозначне представлення кожного варіанту.

3. Уніфікація опису обмежень

Було визначено, що обмеження (constraints) може бути двох основних типів:

- *domain*: безпосередньо встановлює звужений набір дозволених значень для певної ознаки;

- *conditional*: описує залежності між декількома ознаками за принципом "якщо – тоді".

У табл. 2 наведено повний список всіх полів з описом та доступними значеннями.

На рис. 2-3 наведено приклади опису domain (для уточнення домену конкретної ознаки) та conditional (за однією або кількома умовами застосовуються певні дії щодо іншої ознаки) обмежень у форматі JSON.

Таблиця 2 - Опис полів для обмежень

Назва поля	Особливості	Допустимі значення	Опис
id	Обов'язкове, унікальне	Строка	Унікальний ідентифікатор обмеження
description	Не обов'язкове	Строка	Додатковий опис сутності обмеження (пояснює логіку застосування)
rule_type	Обов'язкове	Одне зі значень: • conditional; • domain	Визначає, як обробляти обмеження: domain – обмеження на домен однієї ознаки; conditional – умовне обмеження (з умовами і діями).
feature	Обов'язкове для rule_type = domain	Строка	Посилання на назву ознаки, до якої застосовується обмеження.
allowed_values	Обов'язкове для rule_type = domain	Список значень	Перелік дозволених значень для ознаки.
conditions	Обов'язкове для rule_type = conditional	Список об'єктів з наступною структурою: feature : строка (назва ознаки); value : строка (значення, яке активує правило).	Перелік умов, які повинні бути виконані для активації обмеження.
actions	Обов'язкове для rule_type = conditional	Список об'єктів. Кожен об'єкт може містити такі поля: feature : строка (назва ознаки, на яку впливає дія); mode : строка (режим дії, наприклад, "null" або "block"); allowed_values : список значень (використовується для режиму "block").	Дії, що виконуються при виконанні умов. Режим null означає, що значення ознаки має бути скинуто (заборонено); режим block – дозволяються тільки значення, перелічені у allowed_values .

```
{
  "id": "c0",
  "description": "Для заданої задачі не можливі значення, відмінні від ПН для Резервування системи",
  "rule_type": "domain",
  "feature": "Резервування системи",
  "allowed_values": [
    "ПН"
  ]
},
```

Рис. 2. Приклади опису обмеження типу domain

```
{
  "id": "c1",
  "description": "Якщо 'Резервування системи' = '(БР) ', тоді 'Метод резервування' недоступний але не блокуючий",
  "rule_type": "conditional",
  "conditions": [
    {
      "feature": "Резервування системи",
      "value": "БР"
    }
  ],
  "actions": [
    {
      "feature": "Метод резервування",
      "mode": "null"
    }
  ]
},
```

Рис. 3. Приклади опису обмеження типу conditional

Обмеження визначають логіку відсіювання недопустимих конфігурацій, звужуючи множину комбінацій значень ознак відповідно до бізнес-правил або технічних вимог, і застосовуються до кожної згенерованої комбінації. Алгоритм перевіряє кожну комбінацію послідовно. Порядок перевірки constraints може вплинути на кінцевий результат. Вона відбувається за порядком, в якому обмеження задані у вхідних даних. Розробник має забезпечити, щоб логіка бізнес-правил була узгодженою і послідовністю обробки відповідала пріоритетам завдання. У наведеній структурі використовуються два типи обмежень:

- **Conditional:** Обмеження цього типу описують залежність за принципом "якщо – тоді". Якщо умови (conditions) виконуються, виконуються відповідні дії (actions), наприклад, обмеження конфігурації (mode "block") або скидання значення окремої ознаки (mode "null").

- **Domain:** Цей тип обмежень безпосередньо задає дозволені значення (домен) для конкретної

ознаки. Якщо значення ознаки не входить до дозволеного списку, конфігурація відкидається.

Тестування та результати

Для перевірки дієвості та оцінки ефективності розробленого алгоритму та структур даних був реалізований прототип на мові Python, який генерує множину валідних стратегій розгортання РБА з урахуванням заданих ознак і обмежень. Реалізація алгоритму включає наступні етапи:

1. Завантаження конфігураційних даних із JSON-файлу, де зберігаються опис ознак (features) та обмежень (constraints). Для тестового сценарію було виокремлено 6 ознак та 6 обмежень (рис. 4) на базі класифікації, наданої в [7].

2. Генерація повного набору можливих комбінацій шляхом обчислення декартового добутку всіх можливих значень ознак. Для наведеного прикладу з 6 ознаками (з доменами розміром 2, 3, 2, 2, 3 та 2) було отримано 144 потенційні конфігурації (рис. 5, а).

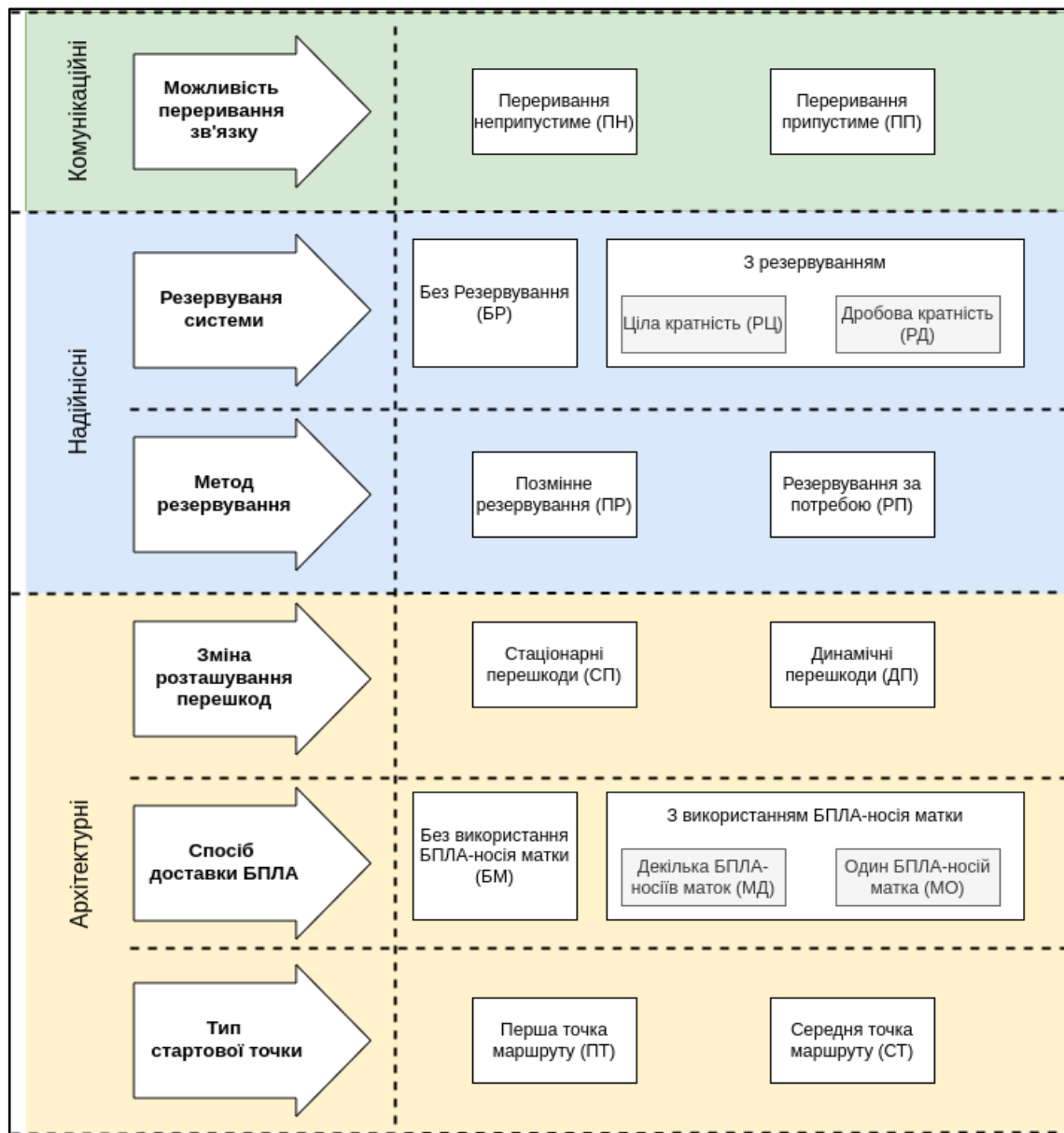


Рис. 4. Класифікатор стратегій розгортання мережі БПЛА

<pre> /home/dimat/Documents/projects/classifier/.venv All possible configurations 144: 1. ПН/БР/РП/СП/БМ/ПТ 2. ПН/БР/РП/СП/БМ/СТ 3. ПН/БР/РП/СП/МО/ПТ 4. ПН/БР/РП/СП/МО/СТ 5. ПН/БР/РП/СП/МД/ПТ 6. ПН/БР/РП/СП/МД/СТ 7. ПН/БР/РП/ДП/БМ/ПТ 8. ПН/БР/РП/ДП/БМ/СТ 9. ПН/БР/РП/ДП/МО/ПТ 10. ПН/БР/РП/ДП/МО/СТ </pre>	<pre> Valid configurations (48): 1. ПН/РЦ/РП/СП/БМ/ПТ 2. ПН/РЦ/РП/СП/БМ/СТ 3. ПН/РЦ/РП/СП/МО/ПТ 4. ПН/РЦ/РП/СП/МО/СТ 5. ПН/РЦ/РП/ДП/МО/ПТ 6. ПН/РЦ/РП/ДП/МО/СТ 7. ПН/РЦ/РП/СП/БМ/ПТ 8. ПН/РЦ/РП/СП/БМ/СТ 9. ПН/РЦ/РП/СП/МО/ПТ 10. ПН/РЦ/РП/СП/МО/СТ </pre>	<pre> Blocked configurations (96): 1. Blocked by (c2): ПН/БР/None/СП/БМ/ПТ 2. Blocked by (c2): ПН/БР/None/СП/БМ/СТ 3. Blocked by (c2): ПН/БР/None/СП/МО/ПТ 4. Blocked by (c2): ПН/БР/None/СП/МО/СТ 5. Blocked by (c2): ПН/БР/None/СП/МД/ПТ 6. Blocked by (c2): ПН/БР/None/СП/МД/СТ 7. Blocked by (c2): ПН/БР/None/ДП/БМ/ПТ 8. Blocked by (c2): ПН/БР/None/ДП/БМ/СТ 9. Blocked by (c2): ПН/БР/None/ДП/МО/ПТ 10. Blocked by (c2): ПН/БР/None/ДП/МО/СТ </pre>
а	б	в

Рис. 5. Список: а – всіх згенерованих конфігурацій; б – валідних конфігурацій; в – заблокованих конфігурацій з наведенням ідентифікаторів обмежень

У наведеному прикладі стратегії кодуються за такою схемою:

Можливість переривання зв'язку
/ {Резервування} / {Метод резервування}
/ {Зміна розташування перешкод}
/ {Спосіб доставки БПЛА} / {Тип стартової точки}

Код конфігурації «ПП/РД/РП/СП/БМ/СТ»: переривання зв'язку припустимі (ПП), дробова кратність резервування (РД), резервування за потреби (РП), стаціонарні перешкоди (СП), без використання БПЛА-носія матки (БМ), середня точка старту маршруту (СТ).

Код конфігурації «ПН/РЦ/РП/ДП/МД/СТ»: переривання зв'язку неприпустимі (ПН), резервування цілої кратності (РЦ), резервування за потреби (РП), динамічні перешкоди (ДП), використання декількох БПЛА-носіїв матки (МД), середня точка старту маршруту (СТ).

3. Послідовне застосування алгоритму фільтрації конфігурацій. Для наведеного прикладу було виокремлено 48 валідних (рис. 5, б) та 96 заблокованих (рис. 5, в) конфігурацій, причому для заблокованих фіксується ідентифікатор обмеження, що спричинило відхилення.

Під час аналізу отриманих результатів було виявлено, що розроблене програмне забезпечення продемонструвало кращу ефективність порівняно з ручним перебором, результати якого надано у статті [7]. Зокрема, у зазначеній статті стратегія «ПП/РД/РП/ДП/БМ/ПТ» була помилково визначена як валідна для запропонованого класифікатора стратегій розгортання мережі БПЛА (рис. 6). Водночас, розроблені алгоритми і ПЗ правильно визначили стратегію «ПП/РД/РП/ДП/БМ/ПТ» як недопустиму (рис. 7), оскільки ця стратегія не задовольняє обмеження «с5».

Множини стратегій розгортання комунікаційної мережі БПЛА			
Множина стратегій розгортання по ознакам надійності	Множина стратегій розгортання по архітектурним ознакам	Остаточна множина стратегій розгортання	
		Переривання неприпустиме (ПН)	Переривання припустиме (ПП)
1. БР 2. РЦ-РП 3. РЦ-РП 4. РД-РП	1. СП-БМ-ПТ 2. СП-БМ-СТ 3. СП-МО-ПТ 4. СП-МД-СТ 5. ДП-БМ-ПТ 6. ДП-БМ-СТ 7. ДП-МО-ПТ	1. ПН/РЦ-РП/СП-БМ-ПТ 2. ПН/РЦ-РП/СП-БМ-СТ 3. ПН/РЦ-РП/СП-МО-ПТ 4. ПН/РЦ-РП/СП-МД-СТ 5. ПН/РЦ-РП/ДП-БМ-ПТ 6. ПН/РЦ-РП/ДП-БМ-СТ 7. ПН/РЦ-РП/ДП-МО-ПТ 8. ПН/РЦ-РП/ДП-МД-СТ 9. ПН/РЦ-РП/СП-БМ-ПТ 10. ПН/РЦ-РП/СП-БМ-СТ 11. ПН/РЦ-РП/СП-МО-ПТ 12. ПН/РЦ-РП/СП-МД-СТ 13. ПН/РЦ-РП/ДП-БМ-ПТ 14. ПН/РЦ-РП/ДП-БМ-СТ 15. ПН/РЦ-РП/ДП-МО-ПТ 16. ПН/РЦ-РП/ДП-МД-СТ 17. ПН/РД-РП/СП-БМ-ПТ 18. ПН/РД-РП/СП-БМ-СТ 19. ПН/РД-РП/СП-МО-ПТ 20. ПН/РД-РП/СП-МД-СТ 21. ПН/РД-РП/ДП-БМ-ПТ 22. ПН/РД-РП/ДП-БМ-СТ 23. ПН/РД-РП/ДП-МО-ПТ 24. ПН/РД-РП/ДП-МД-СТ	1. ПП/БР/СП-БМ-ПТ 2. ПП/БР/СП-БМ-СТ 3. ПП/БР/СП-МО-ПТ 4. ПП/БР/СП-МД-СТ 5. ПП/БР/ДП-БМ-ПТ 6. ПП/БР/ДП-БМ-СТ 7. ПП/БР/ДП-МО-ПТ 8. ПП/БР/ДП-МД-СТ 9. ПП/РЦ-РП/СП-БМ-ПТ 10. ПП/РЦ-РП/СП-БМ-СТ 11. ПП/РЦ-РП/СП-МО-ПТ 12. ПП/РЦ-РП/СП-МД-СТ 13. ПП/РЦ-РП/ДП-БМ-ПТ 14. ПП/РЦ-РП/ДП-БМ-СТ 15. ПП/РЦ-РП/ДП-МО-ПТ 16. ПП/РЦ-РП/ДП-МД-СТ 17. ПП/РД-РП/СП-БМ-ПТ 18. ПП/РД-РП/СП-БМ-СТ 19. ПП/РД-РП/СП-МО-ПТ 20. ПП/РД-РП/СП-МД-СТ 21. ПП/РД-РП/ДП-БМ-ПТ 22. ПП/РД-РП/ДП-БМ-СТ 23. ПП/РД-РП/ДП-МО-ПТ 24. ПП/РД-РП/ДП-МД-СТ 25. ПП/РД-РП/СП-БМ-ПТ 26. ПП/РД-РП/СП-БМ-СТ 27. ПП/РД-РП/СП-МО-ПТ 28. ПП/РД-РП/СП-МД-СТ 29. ПП/РД-РП/ДП-БМ-ПТ 30. ПП/РД-РП/ДП-БМ-СТ 31. ПП/РД-РП/ДП-МО-ПТ 32. ПП/РД-РП/ДП-МД-СТ

Рис. 6. Приклад ручного перебору стратегій з роботи [7], що містить помилково визначену як валідну стратегію «ПП/РД/РП/ДП/БМ/ПТ»

```

87. Blocked by (c5): ПП/РД/ПР/ДП/БМ/ПТ
88. Blocked by (c5): ПП/РД/ПР/ДП/БМ/СТ
89. Blocked by (c3): ПП/РД/ПР/ДП/МО/СТ
90. Blocked by (c4): ПП/РД/ПР/ДП/МД/ПТ
91. Blocked by (c3): ПП/РД/РП/СП/МО/СТ
92. Blocked by (c4): ПП/РД/РП/СП/МД/ПТ
93. Blocked by (c5): ПП/РД/РП/ДП/БМ/ПТ
94. Blocked by (c5): ПП/РД/РП/ДП/БМ/СТ
95. Blocked by (c3): ПП/РД/РП/ДП/МО/СТ
96. Blocked by (c4): ПП/РД/РП/ДП/МД/ПТ

Constraints description:
ID: c1. Description: Якщо 'Резервування системи' = '(БР) Без Резервування', тоді 'Метод резервування' недоступний а
ID: c2. Description: Якщо 'Резервування системи' = '(РД) Дробова кратність', тоді 'Метод резервування' = '(ПР) Позм
ID: c3. Description: Якщо 'Спосіб доставки БПЛА' = '(МО) Один БПЛА-носії матка', тоді 'Тип стартової точки' = '(СТ)
ID: c4. Description: Якщо 'Спосіб доставки БПЛА' = '(МД) екілька БПЛА-носіїв маток', тоді 'Тип стартової точки' = '
ID: c5. Description: Якщо 'Зміна розташування перешкод' = '(ДП) Динамічні перешкоди', тоді 'Спосіб доставки БПЛА' =
ID: c6. Description: При '(ПН) Неприпустимому перериванню' Системи '(БР) без резервування' неефективні, оскільки та

```

Рис. 7. Результати роботи програми: стратегія «ПП/РД/РП/ДП/БМ/ПТ»
коректно заблокована відповідно до обмеження «с5»

А саме: якщо параметр «Зміна розташування перешкод» дорівнює «(ДП) Динамічні перешкоди», то «Спосіб доставки БПЛА», визначений як «Без використання БПЛА-носія матки (БМ)», стає неефективним.

Причиною цього є значне зростання обсягу додаткових розрахунків, які має виконувати кожен окремий БПЛА за таких умов.

За результатами тестування було виявлено, що:

- генерація 144 можливих стратегій дозволяє охопити весь простір варіантів розгортання, що забезпечує повноту аналізу;

- автоматична перевірка з використанням заданих обмежень значно зменшує кількість помилок у виборі стратегій, оскільки некоректні комбінації відсіюються на етапі формування валідних конфігурацій;

- детальна діагностика через фіксацію заблокованих конфігурацій із зазначенням відповідних ID обмежень допомагає швидко виявляти і коригувати потенційні неточності у визначенні правил.

Нижче наведено ключові можливості реалізованого прототипу:

- вивід усіх згенерованих комбінацій для подальшого аналізу;
- перевірка кожної конфігурації за заданими умовами (conditional) та безпосередньо за доменами (domain);
- підрахунок валідних конфігурацій, що сприяє оцінці ефективності застосування алгоритму.

Висновки та перспективи

Основним науковим внеском даного дослідження є запропонований і реалізований метод автоматизації генерації СРЗН, що базується на фасетно-ієрархічному описі ознак і декартових моделях

для формування можливих конфігурацій. Ключовими результатами роботи є концептуальна схема і комплекс алгоритмів процесу автоматизації, що складається з етапів:

- формування уніфікованих структур ознак і обмежень у JSON-форматі;
- генерування повного переліку комбінацій;
- послідовної фільтрації із застосуванням domain та conditional обмежень.

Наведений Python-код демонструє практичну реалізацію алгоритму, що включає завантаження даних, генерацію всіх можливих конфігурацій, застосування обмежень і виведення результатів. Тестові результати підтвердили, що такий підхід дає змогу швидко відсікати недійсні або суперечливі комбінації, фіксувати причини блокування (ідентифікатор обмеження) й отримувати кінцеву множину валідних стратегій. У подальших дослідженнях планується:

- інтегрувати розроблений інструментарій із системами динамічного моделювання роботи РБА;
- розробити систему багатокритеріального оцінювання та ранжування згенерованих стратегій;
- деталізувати ознаки небезпечних просторів та врахувати можливості використання гетерогенних систем (комплексів безпілотних літальних, безпекопажних і наземних апаратів-роботів);
- розширити типи представлення обмежень, зокрема, з використанням нечіткої логіки або онтологічного оброблення.

Отже, автоматизований підхід до проектування та аналізу стратегій може підвищити надійність та ефективність застосування РБА у складних та мінливих умовах завдяки формуванню повних і внутрішньо несуперечливих множин для вибору СРЗН з використанням процедур [1].

СПИСОК ЛІТЕРАТУРИ

1. Terenyk D. Choosing strategies for deployment and ensuring the reliability of a UAV swarm to support communications in destruction conditions / Dmytro Terenyk, Vyacheslav Kharchenko // Innovative technologies and scientific solutions for industries. – 2024. – № 3 (29). – С. 91–103. – URL: <https://doi.org/10.30837/2522-9818.2024.3.091>.

2. Robotic-biological systems for detection and identification of explosive ordnance: concept, general structure, and models / Gennadiy Fedorenko [та ін.] // Radioelectronic and Computer Systems. – 2023. – № 2. – С. 143–159. – URL: <https://doi.org/10.32620/reks.2023.2.12>.
3. Methods and software tools for reliable operation of flying LiFi networks in destruction conditions / Herman Fesenko [та ін.] // Sensors. – 2024. – Т. 24, № 17. – С. 5707. – URL: <https://doi.org/10.3390/s24175707>.
4. Balalaieva O. Y. Фасетні класифікації електронних засобів навчального призначення / Olena Yu Balalaieva // Information technologies and learning tools. – 2013. – Т. 38, № 6. – С. 41–52. – URL: <https://doi.org/10.33407/itlt.v38i6.926>.
5. Перегуда О. М. Фасетна класифікація особливих випадків у польоті безпілотної літальної апарату 1 класу / О. М. Перегуда, А. В. Родіонов, А. І. Бобунов // Наука і техніка Повітряних Сил Збройних Сил України. – 2022. – № 1(46). – С. 85–91. – URL: <https://doi.org/10.30748/nitps.2022.46.12>.
6. Жигайло О. М. Кластерний аналіз даних в автоматизованих системах простежуваності / О. М. Жигайло, В. В. Борис // Automation of technological and business processes. – 2018. – Т. 10, № 1. – URL: <https://doi.org/10.15673/atbp.v10i1.879>.
7. Тереник Д. Класифікація стратегій забезпечення надійності літаючої мережі для забезпечення комунікацій в умовах руйнувань / Дмитро Тереник, Вячеслав Харченко // Measuring and computing devices in technological processes. – 2024. – № 2. – С. 357–370. – URL: <https://doi.org/10.31891/2219-9365-2024-78-41>.
8. Phase combination for reliability analysis of dynamic k-out-of-n Phase-AND mission systems / Chaonan Wang [та ін.] // Reliability Engineering & System Safety – 2025. – С. 110817. – URL: <https://doi.org/10.1016/j.res.2025.110817>.
9. Wei Y. Fleet service reliability analysis of self-service systems subject to failure-induced demand switching and a two-dimensional inspection and maintenance policy / Yian Wei, Yao Cheng, Haitao Liao // IEEE Transactions on Automation Science and Engineering. – 2024. – С. 1–16. – URL: <https://doi.org/10.1109/tase.2024.3516049>.
10. Ruiz C. Generalized functional mixed models for accelerated degradation-based reliability analysis / Cesar Ruiz, Haitao Liao, Edward A. Pohl // IEEE Transactions on Reliability. – 2024. – С. 1–12. – URL: <https://doi.org/10.1109/tr.2024.3505077>.
11. Dui H. Attack strategies and reliability analysis of Wireless Mesh Networks considering cascading failures / Hongyan Dui, Jiabao Zhai, Xiuwen Fu // Reliability Engineering & System Safety. – 2025. – С. 110832. – URL: <https://doi.org/10.1016/j.res.2025.110832>.
12. Kucherenko O. I. Classification of methods and algorithms for constructing UAV routes / O. I. Kucherenko, T. A. Vakaliuk // Scientific notes of Taurida national V.I. Vernadsky university. Series: Technical sciences. – 2024. – Т. 1, № 3. – С. 111–116. – URL: <https://doi.org/10.32782/2663-5941/2024.3.1/18>.

Received (Надійшла) 10.04.2025

Accepted for publication (Прийнята до друку) 11.06.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Тереник Дмитро – аспірант кафедри комп'ютерних систем, мереж і кібербезпеки, Національний аерокосмічний університет "Харківський авіаційний інститут", Харків, Україна;

Terenyk Dmytro – PhD Student at the Department of Computer Systems, Networks, and Cybersecurity, National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine;

e-mail: d.terenik@student.csn.khai.edu; ORCID Author ID: <https://orcid.org/0009-0003-4173-8757>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57222131855>.

Харченко Вячеслав Сергійович – доктор технічних наук, професор, завідувач кафедри комп'ютерних систем, мереж і кібербезпеки, Національний аерокосмічний університет "Харківський авіаційний інститут", Харків, Україна;

Kharchenko Vyacheslav – Doctor of Sciences (Engineering), Professor, Head at the Department of Computer Systems, Networks, and Cybersecurity, National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine;

e-mail: v.kharchenko@csn.khai.edu; ORCID ID: <https://orcid.org/0000-0001-5352-077X>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=22034616000>.

Automated development of classifiers: algorithms and software for generating strategies of deployment and reliability assurance for UAV swarm

Dmytro Terenyk, Vyacheslav Kharchenko

Abstract. Motivation. Ensuring reliability and autonomy of unmanned aerial vehicles (UAVs), as well as their swarms (UAV swarms), requires effective methods for developing and selecting deployment strategies, along with ensuring functional reliability under complex operational conditions. As the number of such strategies grows, appropriate tools for their development, classification, and selection become necessary. **Object of research:** processes of classification and automation in constructing classifiers for deployment and reliability assurance strategies (DRAS) for mobile systems, particularly UAV swarms. **Purpose of the article.** Development of a method for automating the process of creating and modifying (expanding and restructuring) classifiers of DRAS for mobile systems in various subject areas, demonstrating its application through the example of UAV swarms to enhance decision-making efficiency in conditions with a large number of domain-specific factors. **Research results.** An analysis of existing methods has been conducted; a concept and algorithms for automating classification and modeling processes of DRAS have been proposed; software has been developed and validated; and practical examples of its application have been presented, demonstrating its capability to support decision-making regarding the formation of sets and elimination of conflicting or practically unacceptable strategies. **Conclusions.** The primary scientific contribution of this study is the proposed and implemented automation method based on facet-hierarchical attribute descriptions and Cartesian models for generating possible strategy configurations. This approach can enhance the reliability and effectiveness of UAV swarm deployment under complex and dynamic conditions by generating complete and internally consistent sets for selecting DRAS.

Keywords: strategy classifier, UAV swarm deployment, reliability assurance, facet classification, decision-making automation.

А. С. Шандиба, В. І. Сітніков, О. С. Ніколаєв, Д. О. Кібірев

Харківський національний університет радіоелектроніки, Харків, Україна

НАНЕСЕННЯ ЦИФРОВИХ ВОДЯНИХ ЗНАКІВ З ВИКОРИСТАННЯМ ХАОТИЧНИХ КАРТ

Анотація. У сучасному цифровому середовищі питання охорони авторських прав та безпеки інформації стають надзвичайно важливими. Одним з дієвих способів захисту мультимедійного контенту є технологія цифрових водяних знаків, яка дає змогу ідентифікувати та відстежувати незаконне використання цифрових файлів. Водночас, традиційні способи нанесення водяних знаків мають певні недоліки, що стимулює пошук нових рішень. Метою дослідження є створення способу впровадження цифрових водяних знаків з використанням хаотичних карт, задля збільшення захищеності знаків від впливів і змін. Предметом дослідження є методи та технології цифрового водяного маркування (стеганографії), що використовують хаотичні карти для підвищення надійності, захищеності та непомітності вбудовуваної інформації в цифрові об'єкти. Об'єктом дослідження виступають цифрові зображення (або мультимедійні об'єкти), на які наноситься водяний знак із застосуванням алгоритмів на основі хаотичних карт, що забезпечує захист авторських прав, аутентифікацію та цілісність інформації. Для досягнення поставленої мети було розглянуто наступні завдання: проаналізовано ключові методи нанесення цифрових водяних знаків, розглянуто характеристики хаотичних карт та потенціал їх застосування при створенні водяних знаків, розглянуто алгоритм для вбудовування цифрових водяних знаків, використовуючи хаотичні карти. У цій роботі розглянуто теоретичні засади цифрових водяних знаків, проведено аналіз існуючих методів їх нанесення, а також детально вивчено можливість застосування хаотичних карт у цій сфері. Запропоновано методику використання хаотичних карт для генерування ключів і вбудовування водяних знаків. Проаналізовано сильні та слабкі сторони цього підходу, а також окреслено перспективи майбутніх досліджень у цій галузі.

Ключові слова: хаотичні карти, алгоритм, кіберзагрози, зображення, стеганографія, захист, водяні знаки.

Вступ

Цифрові водяні знаки (ЦВЗ) — це приховані дані, які впроваджуються в цифрові медіа (зображення, відео, аудіо, тексти) для ідентифікації власника та запобігання несанкціонованому копіюванню чи редагуванню. Вони зазвичай невидимі або малопомітні для звичайного користувача, але можуть бути виявлені спеціалізованими засобами. Основна функція ЦВЗ — захист авторських прав та забезпечення безпеки інформації в цифровому просторі.

З розвитком інформаційних технологій та збільшенням обсягів цифрового контенту, постає потреба в нових способах захисту. Звичайні методи захисту, як-от паролі чи шифрування, можуть бути легко зламані або обійдені. Натомість, цифрові водяні знаки є більш надійним методом, оскільки вони не лише захищають вміст, але й дозволяють відстежувати його походження, підтверджувати автентичність, а також виявляти джерела несанкціонованого розповсюдження.

Сучасний цифровий світ вимагає ефективних інструментів захисту авторських прав і боротьби з піратством. З ростом популярності інтернет-платформ і можливостей швидкого розповсюдження цифрового контенту виникає серйозна загроза незаконного використання та копіювання файлів. Нанесення цифрових водяних знаків — один із найефективніших методів забезпечення захисту файлів від таких загроз. Водночас, це дозволяє зберігати цілісність та автентичність даних, адже водяний знак може бути невід'ємною частиною файлу, яку неможливо безпечно видалити або замінити без виявлення.

Хаотичні карти — це математичні структури, що базуються на принципах хаотичної динаміки, які використовуються в криптографії для генерування

складних, непередбачуваних послідовностей, які можна використовувати для захисту даних. Хаос, як явище, характеризується високою чутливістю до початкових умов, що дозволяє створювати складні та стійкі до атак системи. У криптографії хаотичні карти застосовуються для створення випадкових чисел, генерування ключів та інших задач, де потрібна висока ступінь непередбачуваності.

Властивості хаотичних карт, як-от детермінованість, чутливість до умов і здатність генерувати складні структури, роблять їх ідеальними для використання в сучасних системах безпеки. Ці карти можуть служити основою для створення водяних знаків, оскільки дозволяють вбудовувати захищені дані таким чином, що їх неможливо легко виявити чи видалити без порушення цілісності файлу.

Використання хаотичних карт для нанесення цифрових водяних знаків відкриває нові можливості для підвищення стійкості та надійності захисту. Завдяки їхнім властивостям можна створювати водяні знаки, які є надзвичайно стійкими до атак і важко виявляються чи змінюються без пошкодження самого файлу. Хаотичні карти можуть бути використані для генерації ключів, які застосовуються для вбудовування водяних знаків у цифрові файли.

Завдяки таким властивостям, як висока непередбачуваність та складність, хаотичні карти можуть забезпечити додатковий рівень безпеки, захищаючи водяні знаки від атак, спрямованих на їхнє виявлення або видалення. Цей підхід дозволяє значно підвищити ефективність захисту авторських прав та безпеки цифрових файлів, особливо в умовах постійного розвитку технологій кіберзагроз.

Аналіз наукових робіт засвідчує наявність численних методів накладання цифрових водяних знаків, включаючи стеганографічні та криптографічні

стратегії. Розвідки свідчать про високий ступінь захисту цифрових водяних знаків при використанні хаотичних систем.

Різні моделі хаотичних карт демонструють ефективність у формуванні унікальних ключів для вбудовування водяних знаків. Зокрема, дослідження Cox I. J. [6] і Kalker та Haitsma [11] висвітлюють, що цифрові водяні знаки, побудовані на хаотичних картах, мають посилену стійкість до атак. До того ж, праці Zhang та Wang [9] та O'Connor та Hunt [8] досліджують потенціал застосування хаотичних систем у криптографії та їх ефективність у генерації випадкових послідовностей. Зважаючи на це, використання хаотичних карт здатне значно покращити безпеку цифрових водяних знаків.

Теоретичні основи цифрових водяних знаків

Цифрові водяні знаки – це непомітні чи майже непомітні відомості, вшиті в цифрові файли задля їхнього захисту від несанкціонованого використання та підтвердження авторства. Принцип їхньої роботи полягає в тому, що водяний знак містить конкретну інформацію, яка може бути зашифрованою або вбудованою у вигляді додаткових даних, що суттєво не впливають на властивості файлу, як-от його розмір або якість. Процес вбудовування цифрового водяного знака складається з декількох етапів:

1. Генерування водяного знака – створення унікального ідентифікатора або коду, що буде вбудований у файл.

2. Вбудовування водяного знака – введення водяного знака в окремі частини файлу без видимого впливу на його структуру. Це може бути досягнуто через зміни певних пікселів у зображенні, частотних складових у звукових або відеофайлах або навіть символів у текстових файлах.

3. Витягування водяного знака – процес вилучення водяного знака з файлу для перевірки його автентичності або власності, що зазвичай потребує наявності секретного ключа або певних алгоритмів для розшифрування водяного знака.

Цифрові водяні знаки можуть бути видимими або прихованими для користувача, а також можуть використовувати різні способи вбудовування – від простих методів зміни значень пікселів до більш складних криптографічних підходів. Цифрові водяні знаки можна класифікувати за різними критеріями. Ось основні типи:

1. Стеганографічні водяні знаки:

Призначення: використовуються для прихованого вбудовування інформації в цифрові файли.

Метод: водяний знак вбудовується в такі елементи файлу, як пікселі зображень, бітові потоки в аудіо- та відеофайлах або текстові файли. Стеганографічні водяні знаки не впливають на зовнішній вигляд чи якість файлу, що робить їх майже непомітними для користувача.

Застосування: використовуються для прихованої ідентифікації файлів, захисту від копіювання, а також для забезпечення конфіденційності та аутентичності.

2. Криптографічні водяні знаки:

Призначення: використовуються для забезпечення безпеки та достовірності інформації шляхом шифрування водяного знака.

Метод: водяний знак вбудовується в дані файлу, при цьому він шифрується або підписується з використанням криптографічних алгоритмів. Для вилучення водяного знака необхідно мати відповідний криптографічний ключ або сертифікат.

Застосування: використовуються в галузі цифрових підписів, авторських прав, юридичних та фінансових документах, де важливо забезпечити підтвердження достовірності та походження інформації.

3. Змішані водяні знаки:

Призначення: комбінують як стеганографічні, так і криптографічні методи для досягнення більш високого рівня захисту.

Метод: водяний знак може бути одночасно зашифрованим і прихованим у даних. Це дозволяє додати додатковий рівень безпеки, оскільки втрата одного з рівнів захисту (наприклад, виявлення стеганографічного водяного знака) не впливає на цілісність або безпеку водяного знака в цілому.

Традиційні методи нанесення водяних знаків охоплюють такі підходи, як зміна пікселів зображення, використання частотних компонентів аудіо чи відеофайлів або інші техніки, які впливають на дані файлу.

Переваги традиційних методів:

1. Легкість впровадження: багато методів нанесення водяних знаків прості у реалізації та не вимагають великих обчислювальних ресурсів.

2. Низькі витрати на впровадження: технології, засновані на традиційних методах, зазвичай дешевші у впровадженні і не потребують складних апаратних або програмних рішень.

3. Широкий спектр застосувань: ці методи можуть бути застосовані до різних типів файлів, таких як зображення, відео, аудіо, а також текстові документи.

Недоліки традиційних методів:

1. Вразливість до атак: традиційні водяні знаки часто виявляються вразливими до різноманітних атак, як-от спроби видалення чи модифікації, зокрема через зміни в форматах файлів чи стиснення.

2. Зниження якості файлу: для деяких методів нанесення водяних знаків, як-от зміна пікселів або частотних компонентів, існує ризик погіршення якості файлу (особливо це стосується зображень та відео).

3. Обмеження на кількість водяних знаків: у деяких випадках можна вбудовувати лише обмежену кількість водяних знаків у файл, що може обмежити їхню ефективність для масового захисту контенту.

Традиційні методи нанесення водяних знаків мають свої переваги, але й суттєві обмеження, які потребують нових технологічних рішень для підвищення безпеки та ефективності. Хаотичні карти, як один із таких підходів, пропонують можливості для створення більш стійких та захищених водяних знаків. . Схема загального принципу роботи хаотичної карти наведена на рис. 1.

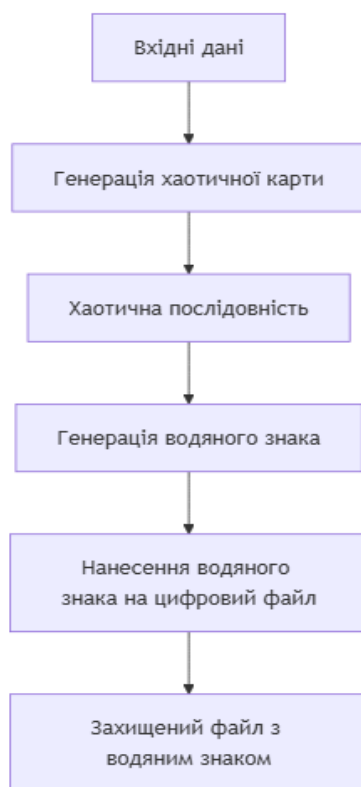


Рис. 1. Схема загального принципу роботи хаотичної карти

Хаотичні карти: Теоретичні основи

Хаотичні системи – це динамічні системи, що вирізняються залежністю від початкових параметрів; незначні зміни в цих параметрах можуть спричинити значні відмінності у поведінці системи в майбутньому. Вони мають високий рівень непередбачуваності, що робить їх особливо цінними для використання в криптографії та захисті даних. Хаотичні системи часто застосовуються для створення псевдовипадкових послідовностей, які є необхідними для забезпечення надійності криптографічних алгоритмів.

Хаотичні карти (або динамічні карти) представляють собою окремий різновид хаотичних систем. Вони є математичними моделями, що описують, як змінюється значення певної величини в результаті повторюваних операцій за заданими правилами або функціями.

Для хаотичних карт характерна здатність демонструвати надзвичайно складну та непередбачувану поведінку навіть за простих початкових умов.

Основні риси хаотичних карт:

1. Чутливість до початкових параметрів: Як і у звичайних хаотичних системах, навіть незначні зміни у вхідних параметрах можуть привести до принципово інших результатів.

2. Детермінованість: Незважаючи на випадковий вигляд хаотичної поведінки, система залишається детермінованою, тобто її поведінка визначається математичними рівняннями.

3. Портваність: властивість означає, що хаотична система в процесі свого розвитку може набувати певної структурованої форми, зберігаючи при цьому високу чутливість до початкових умов.

4. Перехідна та лінійна стабільність: Хаотичні карти можуть мати як стабільні, так і перехідні стани, що робить їх корисними у ситуаціях, де важливо забезпечити певну стабільність або передбачуваність на тривалий період часу.

Математичні моделі хаотичних карт базуються на повторюваних функціях, які генерують складні та непередбачувані послідовності. Одним з найбільш популярних підходів до створення хаотичних карт є використання атракторів та фазових просторів.

Математична модель хаотичної карти може включати функції, що визначають еволюцію системи з плином часу. Такі функції зазвичай мають вигляд:

$$x_{n+1} = f(x_n), \quad (1)$$

де x_n — це значення на n -му кроці, а $f(x_n)$ — це функція, що визначає зміну значення на наступному етапі.

Поширені функції, які використовуються для побудови хаотичних карт, можуть включати логістичні функції, функції Лоренца, або навіть більш складні моделі на основі фрактальних властивостей. Результат таких функцій, як правило, має складну динамічну поведінку, що характеризується хаосом та непередбачуваністю.

В криптографії хаотичні карти можуть використовуватися для генерації псевдовипадкових послідовностей, що є основою для створення криптографічних ключів, а також для шифрування та розшифрування даних. Моделі хаотичних карт можуть забезпечувати високий ступінь випадковості, необхідний для забезпечення надійної безпеки в криптографічних системах. Наприклад, для генерації псевдовипадкових чисел або секретних ключів хаотичні системи можуть використовуватися через процеси, які володіють певними властивостями детермінованого хаосу. Як хаотичні карти можуть бути використані для генерації ключів або вбудовування водяних знаків:

1. Генерація криптографічних ключів: Хаотичні карти можуть бути використані для генерації криптографічних ключів завдяки їх властивостям, які дозволяють створювати складні та непередбачувані послідовності. Наприклад, можна використати хаотичну карту для генерування чисел, що згодом використовуються як криптографічний ключ у системах шифрування або для хешування. Оскільки хаотичні карти можуть генерувати велику кількість різноманітних, непередбачуваних варіантів чисел, вони забезпечують високий рівень безпеки.

2. Вбудовування водяних знаків: Хаотичні карти також можуть використовуватися для вбудовування цифрових водяних знаків у файли. У цьому випадку алгоритм хаотичної карти може визначати, де і як водяний знак буде інтегровано в цифровий файл (наприклад, зображення, аудіофайл чи відео). Оскільки поведінка хаотичних карт чутлива до початкових умов, водяний знак можна розташувати у файлі таким чином, щоб він був майже непомітним, але в той же час достатньо стійким до спроб видалення або зміни.

Процес вбудовування водяного знака за допомогою хаотичних карт може включати:

1. Генерацію хаотичних послідовностей, які визначають місце і спосіб інтеграції водяного знака в дані файлу.

2. Шифрування водяного знака за допомогою результатів хаотичної системи, що додає додатковий рівень захисту.

3. Вилучення водяного знака з файлу шляхом застосування зворотного процесу, заснованого на тій самій хаотичній карті, яка використовувалася для його вбудовування.

Завдяки своїй непередбачуваності та чутливості до початкових умов хаотичні карти є дуже ефективним інструментом для захисту цифрових файлів через водяні знаки, а також для підвищення надійності криптографічних систем.

Нанесення цифрових водяних знаків з використанням хаотичних карт

Нанесення цифрових водяних знаків із застосуванням хаотичних карт є передовим підходом до захисту цифрових файлів, що ґрунтується на властивостях хаотичних систем для створення непередбачуваних і стійких до змін водяних знаків. Ключовим компонентом цього підходу є використання математичних моделей хаотичних карт для визначення точок і методів, через які відбуватиметься вбудовування водяного знака в цифрові дані.

Метод вбудовування водяних знаків із застосуванням хаотичних карт передбачає такі етапи:

1. Генерація хаотичної послідовності: Застосовується функція хаотичної карти для створення послідовності чисел, що визначають, де і як буде вбудовуватися водяний знак. Ця послідовність є чутливою до початкових умов, що забезпечує високий рівень безпеки та непередбачуваність водяного знака.

2. Вибір точок вбудовування: Хаотична послідовність використовується для визначення місць у цифровому файлі (наприклад, пікселях зображення, бітах звукових файлів), де водяний знак буде вбудовуватися. Цей підхід дає змогу мінімізувати помітність водяного знака, водночас залишаючи його достатньо стійким до спроб маніпулювання даними.

3. Вбудовування водяного знака: Після вибору точок вбудовування водяний знак інтегрується в цифровий файл, зазвичай шляхом модифікації бітових чи піксельних значень. Завдяки використанню хаотичних карт, зміни в даних мінімальні, що дає змогу водяному знаку залишатися непомітним для звичайного спостерігача, але достатньо стійким.

4. Шифрування водяного знака: Для додаткової безпеки водяний знак може бути зашифрований за допомогою хаотичної карти. Це додає рівень захисту від несанкціонованих спроб змінити або видалити водяний знак.

5. Витягування водяного знака: Для того щоб отримати водяний знак із зашифрованого файлу, необхідно застосувати зворотний процес, використовуючи той самий алгоритм хаотичної карти. Точне знання початкових умов і параметрів карти дає змогу правильно відновити водяний знак.

Розглянемо алгоритми для генерації хаотичних карт та їх використання для вбудовування та витягування водяних знаків (рис. 2, 3).

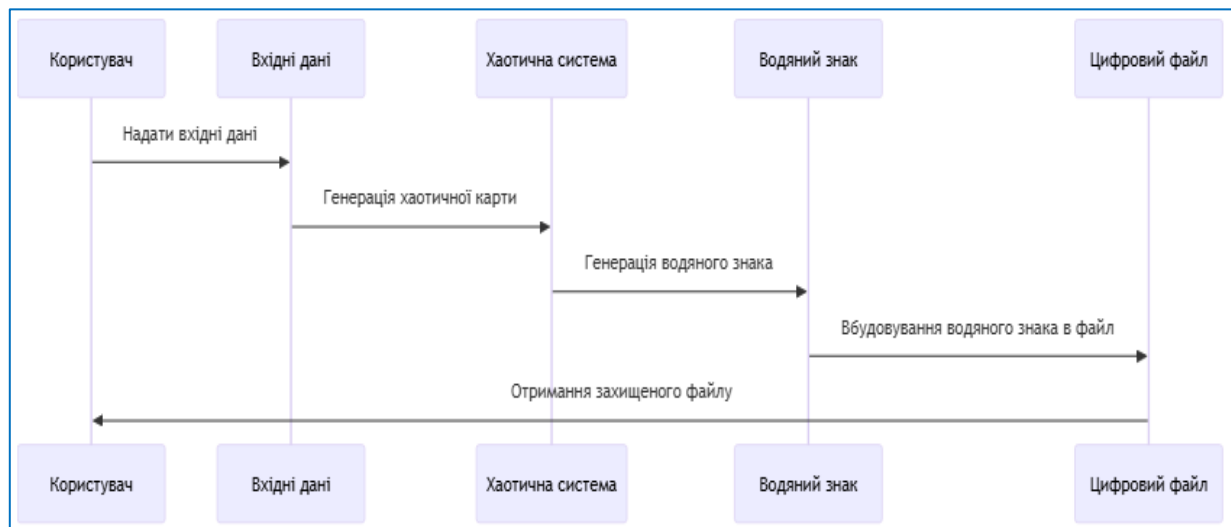


Рис. 2. Алгоритм генерації хаотичної карти

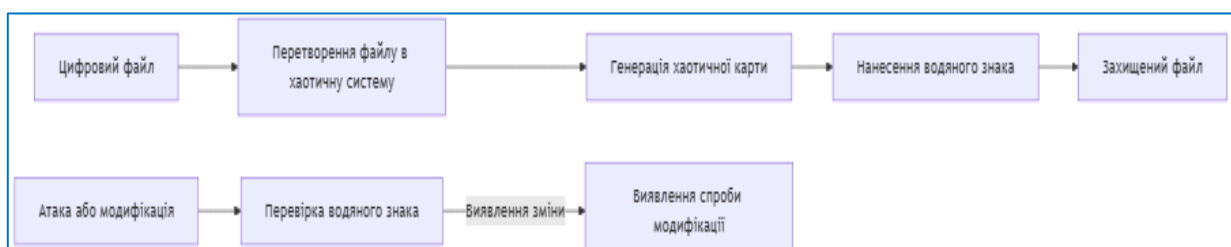


Рис. 3. Моделювання вбудовування водяного знака за допомогою хаотичної карти

Алгоритм Логістичної мапи (Logistic Map): Логістична мапа є одним з основних прикладів хаотичних мап і може бути використана для генерації послідовностей чисел, що визначатимуть точки вбудовування водяних знаків. Алгоритм базується на рівнянні: $x_{n+1} = r \cdot x_n \cdot (1 - x_n)$, де x_n – поточне значення, а r – параметр, що впливає на ступінь хаосу в системі. За допомогою цього рівняння можна генерувати послідовності чисел, які можуть бути використані для визначення точок вбудовування водяного знака в цифровий файл.

Алгоритм Бенсона (Benson Map): Бенсонова мапа є ще одним типом хаотичної мапи, що використовується для генерації послідовностей чисел з високою чутливістю до початкових умов. Завдяки своїй структурі, Бенсонова мапа може бути використана для вбудовування водяних знаків у різні типи файлів, зокрема зображення та звукові файли.

Фрактальні мапи: Фрактальні мапи використовують властивості фракталів для створення хаотичних послідовностей. Вони можуть застосовуватися для вбудовування водяних знаків у вигляді структур, що повторюються на різних масштабах, роблячи водяний знак ще більш стійким до спроб модифікації.

Алгоритм Коса (Coupled Map Lattice): Алгоритм Коса є прикладом хаотичної системи, що використовує багатовимірні мапи для створення хаотичних послідовностей. Це дає змогу отримати складніші структури для водяних знаків і поліпшити їхню стійкість до атак.

Приклади програмного забезпечення або реалізацій, що застосовують хаотичні мапи для цієї мети:

Chaotic Watermarking System: Деякі програмні системи для вбудовування водяних знаків використовують хаотичні мапи для генерації ключів та визначення точок вбудовування. Програмне забезпечення може застосовувати алгоритм Логістичної мапи для створення псевдовипадкових послідовностей, що керують процесом вбудовування та витягування водяних знаків.

X-Watermark: Це програмне забезпечення, що використовує фрактальні та хаотичні мапи для створення водяних знаків у зображеннях. Застосовує комбіновані методи хаотичних мап і фракталів для збільшення стійкості водяного знака до атак.

Watermarking by Chaotic Maps: Це науковий проєкт або бібліотека, що реалізує методи вбудовування водяних знаків у цифрові файли з використанням хаотичних мап. Алгоритми побудовано на різних математичних моделях хаотичних систем, як-от Логістична мапа або Бенсонова мапа.

Digital Watermarking Toolkits: Деякі розроблені бібліотеки, як-от OpenCV або Matlab, можуть бути налаштовані для застосування хаотичних мап у процесі вбудовування водяних знаків. Такі бібліотеки дають змогу розробникам адаптувати хаотичні алгоритми до своїх специфічних потреб у захисті даних. Завдяки використанню хаотичних мап, процес нанесення водяних знаків набуває високого рівня безпеки, стійкості до атак і непередбачуваності. Це забезпечує захист цифрових файлів від несанкціонованого використання і копіювання, зберігаючи водночас мінімальний вплив на якість та видимість водяного знака.

Переваги та виклики використання хаотичних карт

Проаналізувавши класичні підходи і підходи з використанням хаотичних карт можна виділити такі основні переваги і недоліки.

Переваги застосування хаотичних карт проти традиційних методів водяного знакування:

1. Висока захищеність від атак: Хаотичні карти надають значно кращий захист від спроб видалення чи зміни водяних знаків. Оскільки ці системи демонструють надзвичайну чутливість до початкових умов, будь-яка зміна будь-якого параметра або компоненту в структурі карти суттєво впливає на кінцевий результат, що ускладнює відновлення чи розпізнавання водяного знака без точних знань про його вихідні умови.

2. Непередбачуваність: Хаотичні карти генерують випадкові, непередбачувані послідовності, що робить їх відтворення чи зловування надзвичайно складним без правильних параметрів. Це значно ускладнює несанкціоноване вилучення водяних знаків, створених за допомогою хаотичних систем, на відміну від більш стандартних методів, де водяний знак може бути визначений за певними шаблонами.

3. Мінімальний вплив на якість даних: Завдяки застосуванню хаотичних карт, водяний знак може бути вбудований в такі області, що їх неможливо виявити візуально, і вони практично не впливають на якість цифрового файлу. На відміну від традиційних методів, які можуть впливати на візуальні або звукові характеристики файлів, хаотичні карти дозволяють уникнути помітних спотворень.

4. Гнучкість та адаптивність: Хаотичні карти пропонують гнучкість у використанні різних математичних моделей, що дозволяє адаптувати методи водяного знакування під конкретні потреби безпеки та специфікації. Це надає можливість обирати відповідні алгоритми для різних типів файлів, таких як зображення, аудіо, відео або текстові документи.

5. Підвищена стійкість до маніпуляцій: Хаотичні карти демонструють добру адаптацію до змін вхідних даних, тому навіть у випадку часткового пошкодження або маніпуляцій, цифровий водяний знак залишатиметься незмінним. Це робить хаотичні системи привабливішими за традиційні підходи, де незначні зміни можуть призвести до втрати водяного знаку.

Проблеми, що виникають при використанні хаотичних карт: стабільність, чутливість та атаки:

1. Стабільність алгоритмів: Оскільки хаотичні карти сильно залежать від початкових умов та параметрів, їх застосування потребує особливої уваги до стабільності. Незначні зміни в параметрах можуть викликати великі відхилення у результатах генерації хаотичних послідовностей, що може вплинути на ефективність вбудовування водяних знаків. Для забезпечення високої стабільності необхідне ретельне налаштування алгоритмів.

2. Чутливість до змін: Хаотичні карти можуть бути надзвичайно чутливими до найдрібніших змін.

Навіть незначні зміни у даних (наприклад, спроба стиснення або зміна формату файлу) можуть пошкодити водяний знак або зробити його невидимим. Це становить серйозну проблему, особливо коли необхідно витягнути водяний знак з пошкоджених файлів або після стиснення даних.

3. Складність відновлення водяного знака: Вилучення водяного знака за допомогою хаотичних карт може бути дуже складним завданням, особливо без точного знання початкових умов і параметрів карти. Це ускладнює процес відновлення водяного знака у разі змін або атак, що може бути проблематичним для правоохоронних органів чи власників авторських прав.

4. Обмеження на обчислювальні ресурси: Алгоритми, що використовують хаотичні карти, можуть бути обчислювально складними і потребувати значних ресурсів для генерації послідовностей та вбудовування водяних знаків. Це може призвести до проблем із продуктивністю в режимі реального часу для деяких додатків, особливо при обробці великих обсягів даних.

Пропозиції щодо вдосконалення методів на основі хаотичних карт

1. Оптимізація алгоритмів: Щоб покращити стабільність та ефективність хаотичних карт у нанесенні водяних знаків, важливо оптимізувати обчислювальні алгоритми. Це зменшить потребу в обчислювальних ресурсах та збереже високу стійкість до атак.

2. Розробка адаптивних алгоритмів: Створення адаптивних методів вбудовування водяних знаків на основі хаотичних карт, здатних враховувати зміни параметрів даних або умов середовища, підвищить їх стійкість до атак і змін. Це може включати використання багатокрокових або гібридних методів, які поєднують хаотичні карти з іншими криптографічними техніками.

3. Використання мульти-хаотичних карт: Застосування кількох різних хаотичних карт або комбінованих моделей хаосу може покращити стійкість водяного знака до різних типів атак. Використання різних моделей дозволить більш ефективно адаптувати водяний знак до різних типів файлів і знижує ризик успішного відновлення водяного знака без відомих параметрів.

4. Впровадження технологій з машинним навчанням: Використання машинного навчання для адаптації та оптимізації процесу вбудовування водяних знаків може допомогти покращити точність та надійність водяних знаків на основі хаотичних карт. Це дозволить створити системи, що автоматично підлаштовуються під зміни вхідних даних та зменшують уразливість до змін та атак.

5. Інтеграція з багаторівневими системами захисту: Для підвищення рівня безпеки водяних знаків можна використовувати хаотичні карти в поєднанні з іншими технологіями, такими як криптографія з відкритим ключем або цифрові підписи. Це забезпечить додатковий рівень захисту та ускладнить процес нанесення водяних знаків для потенційних злоумисників.

Conclusions

Внаслідок аналізу застосування хаотичних карт у контексті впровадження цифрових водяних знаків, можна виділити кілька ключових спостережень. Хаотичні карти, що демонструють властивості непередбачуваності та залежності від початкових умов, забезпечують посилений захист від різноманітних впливів, як-от спроби зміни або видалення водяного знака. Завдяки складності відтворення без точних параметрів генерації, такі методи забезпечують більш надійний захист авторських прав та цілісності даних, у порівнянні з традиційними підходами.

Проте, існують й певні виклики, особливо ті, що стосуються чутливості хаотичних карт до змін вхідних даних, а також обмежень у обчислювальних ресурсах. Для підтримки їх стабільності й ефективності, необхідна оптимізація алгоритмів та їх адаптація до різних умов експлуатації.

Підтвердження чи спростування ефективності цієї технології.

Використання хаотичних карт у технологіях цифрових водяних знаків демонструє значний потенціал для створення більш складних та захищених методів захисту цифрової інформації. Хаотичні системи, які мають здатність формувати унікальні та комплексні послідовності, можуть значно покращити стійкість до атак та несанкціонованих спроб вилучення водяного знака. Разом з тим, потреба у налаштуванні та оптимізації алгоритмів створює певні труднощі у практичній реалізації цих технологій, проте переваги у порівнянні з традиційними підходами свідчать про їх високу ефективність.

Перспективи подальшого розвитку. Наступні дослідження в цій сфері можуть бути зосереджені на кількох ключових напрямках. Першим є вдосконалення методів генерації хаотичних карт для досягнення більшої стабільності при мінімальних обчислювальних витратах. Це розширить сферу застосування цих методів у реальних додатках, мінімізуючи вимоги до ресурсів.

Другим важливим напрямом є інтеграція хаотичних карт з іншими сучасними криптографічними методами, такими як цифрові підписи або методи з відкритими ключами, що дасть змогу розробляти більш надійні та багаторівневі системи захисту. Крім того, використання машинного навчання для адаптації алгоритмів до змін у даних може суттєво поліпшити якість та стійкість водяних знаків.

Потенціал застосування цих методів у сучасних системах безпеки. Методи впровадження водяних знаків на основі хаотичних карт мають великий потенціал у сучасних технологіях забезпечення безпеки. Їх можна використовувати для захисту авторських прав у медіаіндустрії, у фінансових технологіях для забезпечення цілісності транзакцій, а також у галузях, які вимагають високого рівня конфіденційності даних, наприклад, у сфері охорони здоров'я та державної безпеки. З огляду на розвиток квантових обчислень та нових криптографічних стандартів, хаотичні карти можуть стати важливим компонентом для забезпечення довгострокового захисту інформації.

СПИСОК ЛІТЕРАТУРИ

1. Anderson, Ross J. Security engineering: a guide to building dependable distributed systems. John Wiley & Sons, 2010. url: <https://www.cl.cam.ac.uk/archive/rja14/book.html>
2. Diffie, Whitfield, and Martin E. Hellman. "New directions in cryptography." *Democratizing Cryptography: The Work of Whitfield Diffie and Martin Hellman*. 2022. 365-390. doi: <https://doi.org/10.1145/3549993.355000>
3. J. Chen and B. C. Schafer, "Watermarking of Behavioral IPs: A Practical Approach," *2021 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, Grenoble, France, 2021, pp. 1266-1271, doi: [10.23919/DAT51398.2021.9474071](https://doi.org/10.23919/DAT51398.2021.9474071).
4. Kadian, P., Arora, S.M. & Arora, N. Robust Digital Watermarking Techniques for Copyright Protection of Digital Data: A Survey. *Wireless Pers Commun* 118, 3225–3249 (2021). doi: <https://doi.org/10.1007/s11277-021-08177-w>
5. Lawnik, M.; Berezowski, M. New Chaotic System: M-Map and Its Application in Chaos-Based Cryptography. *Symmetry* 2022, 14, 895. doi: <https://doi.org/10.3390/sym14050895>
6. Shih, Frank Y. Digital watermarking and steganography: fundamentals and techniques. CRC press, 2017. doi: <https://doi.org/10.1201/9781315121109>
7. Cox, Ingemar J., et al. "A review of watermarking principles and practices 1." *Digital signal processing for multimedia systems* (2018): 461-485. url: http://www.wirelesscommunication.nl/wireless/articles/98_bookchapterwatermarking.pdf
8. Hamadi, S. J., & Emad A. Mohammed. (2024). Chaotic Systems in Cryptography: An Overview of Feature-Based Methods. *Al-Salam Journal for Engineering and Technology*, 4(1), 164–172. doi: <https://doi.org/10.55145/ajest.2025.04.01.016>
9. Zhang, B.; Liu, L. Chaos-Based Image Encryption: Review, Application, and Challenges. *Mathematics* 2023, 11, 2585. doi: <https://doi.org/10.3390/math11112585>
10. Menezes, Alfred J., Paul C. Van Oorschot, and Scott A. Vanstone. Handbook of applied cryptography. CRC press, 2018. doi: <https://doi.org/10.1201/9780429466335>
11. Aberna, P., Agilandeewari, L. Digital image and video watermarking: methodologies, attacks, applications, and future directions. *Multimed Tools Appl* 83, 5531–5591 (2024). <https://doi.org/10.1007/s11042-023-15806-y>
12. Naik, R.B., Singh, U. A Review on Applications of Chaotic Maps in Pseudo-Random Number Generators and Encryption. *Ann. Data. Sci.* 11, 25–50 (2024). <https://doi.org/10.1007/s40745-021-00364-7>

Received (Надійшла) 12.03.2025

Accepted for publication (Прийнята до друку) 20.05.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Шандиба Антон Сергійович – студент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Anton Shandyba – student, Department of Electronic Computers, Kharkiv National University of Radio Electronics, Ukraine; e-mail: anton.shandyba@nure.ua; ORCID Author ID <https://orcid.org/0009-0000-7697-0634>.

Сітніков Віталій Ігорович – асистент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Vitalii Sitnikov – Assistant of Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine; e-mail: vitalii.sitnikov1@nure.ua; ORCID Author ID: <https://orcid.org/0009-0005-3087-6104>.

Ніколаєв Олексій Євгенович – студент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Alexey Nikolayev – student, Department of Electronic Computers, Kharkiv National University of Radio Electronics, Ukraine; e-mail: oleksii.nikolaiev@nure.ua; ORCID Author ID <https://orcid.org/0009-0008-5741-4702>.

Кібіреєв Денис Олексійович – аспірант кафедри безпеки інформаційних технологій, Харківський національний університет радіоелектроніки, Харків, Україна;

Denis Kibirev – PhD student, Department of Information Technology Security, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine; e-mail: oleksii.leha@nure.ua; ORCID Author ID: <https://orcid.org/0009-0008-9869-4261>.

Digital watermarking using chaotic maps

Anton Shandyba, Vitalii Sitnikov, Alexey Nikolayev, Denis Kibirev

Abstract. In today's digital environment, the issues of copyright protection and information security are becoming increasingly important. One of the effective methods for protecting multimedia content is digital watermarking technology, which enables the identification and tracking of unauthorized use of digital files. At the same time, traditional watermarking methods have certain drawbacks, which encourages the search for new solutions. The aim of this research is to develop a method for implementing digital watermarks using chaotic maps, in order to enhance the resilience of watermarks against tampering and alterations. The subject of the research is the methods and technologies of digital watermarking (steganography) that utilize chaotic maps to improve the reliability, security, and invisibility of embedded information within digital objects. The object of the research is digital images (or multimedia objects) to which a watermark is applied using chaotic map-based algorithms, ensuring copyright protection, authentication, and data integrity. To achieve the set goal, the following tasks were undertaken: an analysis of key digital watermarking methods, an examination of the characteristics of chaotic maps and their potential in watermark creation, and a review of an algorithm for embedding digital watermarks using chaotic maps. This work explores the theoretical foundations of digital watermarking, analyzes existing embedding methods, and thoroughly investigates the potential application of chaotic maps in this domain. A methodology for using chaotic maps to generate keys and embed watermarks is proposed. The strengths and weaknesses of this approach are analyzed, and prospects for future research in this area are outlined.

Keywords: chaotic maps, algorithm, cyber threats, image, steganography, protection, watermarks.

О. В. Шматко, О. Є. Рагулін, П. О. Кравченко, П. В. Буслов

Харківський національний університет радіоелектроніки, Харків, Україна

ДОСЛІДЖЕННЯ АРХІТЕКТУРНИХ РІШЕНЬ ДЛЯ ПОБУДОВИ БЕЗПЕЧНОЇ СИСТЕМИ ЗБЕРІГАННЯ ТА ПЕРЕДАЧІ КОНФІДЕНЦІЙНИХ ДАНИХ

Анотація. Актуальність. У сучасних умовах стрімкого розвитку цифрових технологій та зростання кіберзагроз захист конфіденційної інформації під час її зберігання та передачі є одним із ключових пріоритетів у проектуванні інформаційних систем. Надійна архітектура таких систем забезпечує стійкість до атак, цілісність даних і контроль доступу. **Об'єкт дослідження:** архітектурні рішення для побудови програмного забезпечення захищеної системи зберігання та передачі конфіденційних даних. **Мета статті:** дослідження, проектування та розробка архітектурних компонентів безпечної інформаційної системи, здатної забезпечити збереження конфіденційності даних у процесі передачі та зберігання. **Результати дослідження.** У статті представлено аналіз сучасних підходів до побудови захищених інформаційних систем, обґрунтовано вибір архітектурної моделі та програмних засобів. Розроблено прототип системи. Проведено тестування на відповідність вимогам захисту та продуктивності. **Висновки.** Запропоноване архітектурне рішення продемонструвало ефективність у забезпеченні цілісності, конфіденційності та доступності даних. Результати можуть бути використані для подальшого вдосконалення інформаційних систем у галузях, де обробляється критично важлива інформація.

Ключові слова: інформаційна безпека; конфіденційність; передача даних; зберігання даних; програмна архітектура; шифрування.

Вступ

Постановка проблеми. Останніми роками постійне зростання рівня економіки та якості життя спричинило суттєве зростання суспільного попиту на доступну та якісну медичну допомогу. Медична галузь перетворилася на одну з найбільших і найдинамічніших у світі, особливо після спалаху пандемії COVID-19, яка серйозно загрожувала здоров'ю населення, порушила функціонування глобальної економіки та призвела до суттєвого збільшення національних інвестицій у сферу охорони здоров'я [1]. Паралельно з цим, розвиток і впровадження сучасних технологій – зокрема штучного інтелекту, інтелектуальних діагностичних засобів і медичних інформаційних систем – сприяє трансформації охорони здоров'я у напрямі підвищення інтелектуальності, точності та цифровізації [2, 3].

У центрі цієї цифрової трансформації знаходяться медичні дані, які є основою для розбудови системи «розумної» та персоналізованої медицини. Вони відіграють ключову роль у біомедичних дослідженнях, розробці лікарських засобів, клінічних рішеннях і управлінні громадським здоров'ям. Водночас медичні дані мають складну структуру, обумовлену різноманітністю джерел, різноманітністю форматів і відмінностями у стандартах охорони здоров'я між регіонами. Через це вони розпорошені, мають несумісні формати та велику структурну гетерогенність, що ускладнює їх об'єднане управління та ефективне використання [4].

З метою подолання цих викликів було створено низку платформ для обробки медичних даних, які забезпечують стандартизований доступ до медичної інформації для лікарень, науково-дослідних інститутів, фармацевтичних компаній і регуляторних органів [5, 6]. Такі платформи агрегують дані з багатьох медичних установ, залучають фахівців для їх анотування та обробки за допомогою спеціалізованого програмного забезпечення й алгоритмів. Оброблені

таким чином дані використовуються для створення моделей захворювань, які слугують основою для медичних досліджень [7], управління охороною здоров'я [8], державного планування [9] та розробки інноваційних препаратів [10].

Попри свою корисність, ці платформи мають істотні обмеження щодо можливості перевірки достовірності та цілісності медичних даних на всіх етапах – від збору до обробки та надання третім сторонам. Для таких користувачів, як наукові установи, фармацевтичні компанії чи органи державного управління, необхідне впровадження надійних механізмів, що дозволяють верифікувати як початкові дані, так і операції, що виконувалися під час їх обробки, а також кінцеві результати. Відповідно, усі етапи життєвого циклу медичних даних – включно з вхідними, проміжними та кінцевими даними, діями операторів і записами доступу – повинні бути прозорими, верифікованими та такими, що підлягають прослежуваності [11].

Аналіз останніх досліджень і публікацій. У контексті зростання обсягів цифрових даних та активного поширення хмарних технологій, питання захищеного зберігання та передачі конфіденційної інформації є однією з головних проблем сучасної інформаційної безпеки. У працях [12, 13] підкреслюється важливість використання криптографічних алгоритмів у архітектурі інформаційних систем як основного інструмента забезпечення конфіденційності та цілісності даних. Одним із найпоширеніших підходів до побудови захищених систем є застосування симетричного та асиметричного шифрування, що було детально розглянуто в дослідженнях [14–16]. Алгоритми AES, RSA та ECC демонструють високу ефективність при обробці конфіденційної інформації у розподілених системах, однак вимагають оптимізації для роботи в обмежених обчислювальних середовищах [17].

У дослідженнях [18, 19] акцент зроблено на важливості побудови багаторівневої архітектури

інформаційної системи, що включає окремі модулі для зберігання, обробки, шифрування даних і моніторингу активності. Така модульність не тільки полегшує оновлення компонентів, але й сприяє масштабованості системи.

Питання безпечної передачі даних у відкритих мережах, включаючи інтернет та бездротові канали, розглядаються у роботах [20–21]. Пропонуються методи тунелювання, віртуальні приватні мережі (VPN), а також використання протоколів безпечного з'єднання, таких як TLS/SSL.

Окрему увагу приділено використанню хмарних сервісів і контейнеризації (Docker, Kubernetes) в архітектурі захищених систем [22, 23]. У таких умовах надзвичайно важливою є реалізація політик шифрування на рівні віртуальних машин і контейнерів.

Аналіз наукових джерел свідчить про необхідність комплексного підходу до проектування систем захищеного зберігання і передачі даних — з урахуванням не лише криптографічних засобів, але й архітектурних, мережових та політик доступу. Це створює підґрунтя для формування архітектурного рішення, представленого у цій статті.

Метою роботи є дослідження, проектування та розробка архітектурних компонентів безпечної інформаційної системи, здатної забезпечити збереження конфіденційності даних у процесі передачі та зберігання.

Основний матеріал

Початкове застосування блокчейн-технологій відбулося у 2008 році з появою криптовалюти Bitcoin. Три ключові характеристики — децентралізація, прозорість та конфіденційність — відрізняють блокчейн від інших технологічних рішень [1]. Саме ці властивості привернули увагу до можливого застосування блокчейну в інших галузях, орієнтованих на дані, зокрема в медицині [3]. Компанія IBM прогнозує значний вплив блокчейн-технологій у трьох основних напрямках охорони здоров'я: децентралізований обмін електронними медичними записами (EHR), адміністрування клінічних досліджень і дотримання нормативних вимог [4, 5].

У сучасному контексті блокчейн все частіше розглядається як ключова технологія для безпечного, прозорого та контрольованого управління даними в IoT-платформах. Застосування блокчейн у медицині дозволяє прискорити наукові дослідження, забезпечити захист конфіденційної інформації та створити нові канали співпраці між науковцями. Під час пандемії COVID-19 блокчейн продемонстрував ефективність у зборі та візуалізації даних, а також у боротьбі з маніпуляцією статистикою.

В роботі пропонується архітектура системи, реалізованої на основі Hyperledger Fabric, спрямована на забезпечення безпечного, надійного та контрольованого управління медичними даними в умовах багатосторонньої взаємодії між суб'єктами медичної сфери. Вона реалізує концепцію багаторівневої розподіленої системи, яка охоплює клієнтський рівень, рівень сервісної логіки, мережовий рівень розподіленого реєстру та інфраструктурний рівень

зберігання даних. У такій архітектурі кожен структурний компонент виконує чітко визначену функцію, зберігаючи при цьому цілісність інформаційних потоків і відповідність стандартам інформаційної безпеки у сфері охорони здоров'я.

Клієнтський рівень представлено програмними інтерфейсами для трьох основних типів користувачів: пацієнтів, медичних працівників та адміністраторів. Застосунок для пацієнтів дозволяє керувати згодою на обробку персональних медичних даних та переглядати історію доступу до них. Медичний застосунок забезпечує введення нових медичних записів та запити на доступ до наявної інформації, а адміністративний інтерфейс використовується для конфігурації мережі, керування доступом та підтримки смарт-контрактів. Усі клієнтські застосунки взаємодіють з мережею за допомогою відповідних SDK, що реалізують механізми підпису, шифрування і надсилання транзакцій до блокчейн-мережі.

Сервісна логіка, що реалізується у вигляді незалежних прикладних сервісів, виконує основну бізнес-функціональність системи. Компонент автентифікації перевіряє сертифікати користувачів за допомогою Membership Service Provider (MSP), що функціонує в межах Fabric. Сервіс управління доступом обробляє запити на доступ, перевіряє дозволи та фіксує факти взаємодії з медичними даними у розподіленому реєстрі. Обробник медичних записів виконує синхронізацію між блокчейн-реєстром та зовнішнім сховищем, зберігаючи об'ємні або чутливі дані поза межами ланцюга блоків.

На рівні блокчейн-мережі функціонує інфраструктура Hyperledger Fabric, яка включає реєрвузли, ordering service, chaincode та інші допоміжні сервіси. Peer-вузли, що належать до організацій-учасників, приймають транзакції, виконують chaincode, зберігають журнали та забезпечують доступ до розподіленого реєстру. Ordering service формує блоки транзакцій і поширює їх серед peer-вузлів. Смарт-контракти, реалізовані як chaincode, виконують бізнес-логіку перевірки згод, реєстрації доступу та фіксації операцій. Верифікація всіх учасників, зокрема користувачів, застосунків і вузлів, здійснюється через MSP, який базується на сертифікатній інфраструктурі з використанням X.509-сертифікатів.

Для забезпечення ефективного зберігання медичних даних система використовує комбінований підхід, що поєднує збереження у розподіленому реєстрі (on-chain) і в зовнішніх базах даних (off-chain). У розподіленому реєстрі зберігається інформація про згоду, операції доступу та посилання на медичні записи, тоді як самі записи, через обсяг або конфіденційність, зберігаються у зовнішньому сховищі, наприклад, у PostgreSQL. Зв'язок між on-chain і off-chain елементами підтримується за допомогою хешів та ідентифікаторів, що гарантує контроль цілісності та узгодженості даних.

У контексті побудови децентралізованої системи для управління медичною інформацією модель даних має ключове значення для забезпечення структурованості, достовірності та відповідності право-

вим нормам. Оновлена модель охоплює чотири взаємопов'язані сутності: MedicalData, OperationalBehavior, User та Consent.

Сутність MedicalData описує кожен блок даних, що записується у блокчейн, із вказанням його типу, джерела, формату, стадії обробки та криптографічного підпису. Кожен медичний запис пов'язаний з операцією (OperationalBehavior), яка його згенерувала або модифікувала. Ця друга сутність відображає транзакційні дії в системі, їх авторство та входи/виходи у вигляді зв'язків з іншими медичними даними. Важливу роль відіграє атрибут OpSign, який виступає унікальним ідентифікатором кожної операції та пов'язує дії з результатами. Модель доповнюється сутністю User, яка репрезентує всіх учасників системи: пацієнтів, лікарів, адміністраторів або аудиторів. Кожен користувач має унікальний ідентифікатор, роль, криптографічний сертифікат та інформацію про організаційну приналежність. Це дозволяє контролювати права доступу та здійснювати персоніфіковану верифікацію транзакцій.

Сутність Consent формалізує процес надання або відкликання згоди користувачем на обробку його персональних медичних даних. Кожна згода

пов'язується з конкретним користувачем, фіксується у вигляді транзакції в блокчейні, містить часові позначки, статус, а також може посилається на відповідний DataSign. Завдяки цьому створюється юридично обґрунтована та технологічно верифікована модель управління доступом до чутливої інформації.

Загалом, запропонована структура забезпечує високий рівень простежуваності, контрольованості та відповідності принципам етичного поводження з персональними даними. Усі зв'язки моделі мають чітке семантичне обґрунтування та логічну відповідність вимогам до розподілених систем обміну медичними даними.

Концепція Data Lineage, або історії змін даних, у системах управління медичною інформацією, реалізованих на основі технологій розподіленого реєстру, виконує критично важливу функцію забезпечення простежуваності, достовірності та цілісності медичних записів протягом усього їх життєвого циклу. У запропонованій моделі, реалізованій у контексті Hyperledger Fabric, Data Lineage реалізується як послідовність зв'язаних транзакційних дій, які формують граф взаємозв'язку між первинними даними, операціями та результатами обробки (рис. 1).

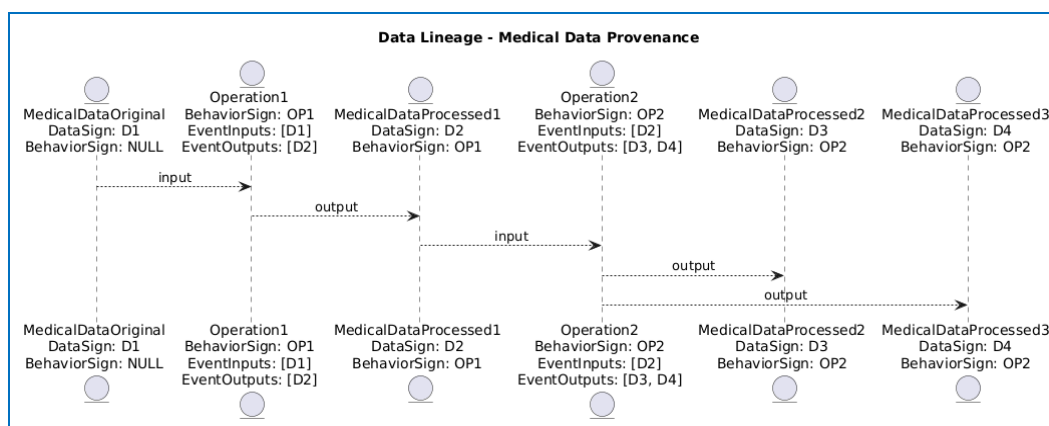


Рис. 1. Граф взаємозв'язку між первинними даними, операціями та результатами обробки

На логічному рівні, кожна одиниця медичних даних (MedicalData) має унікальний підпис (DataSign), який виступає криптографічним ідентифікатором конкретного об'єкта. Якщо запис є первинним, тобто ще не був оброблений у межах жодної транзакції, атрибут BehaviorSign, що вказує на породжуючу операцію, залишається порожнім (NULL). У цьому випадку об'єкт вважається джерелом походження (data origin).

Будь-яка транзакція, яка обробляє або модифікує медичні дані, представлена у системі як об'єкт типу OperationalBehavior. Така транзакція має унікальний ідентифікатор (BehaviorSign), та містить два основних компоненти: вхідні дані (EventInputs) і вихідні дані (EventOutputs), представлені у вигляді масивів DataSign. Таким чином, транзакція формує відображення даних до і даних після виконання певної операції, створюючи причинно-наслідкові зв'язки.

Наприклад, якщо первинний запис D1 є вхідним для транзакції OP1, яка породжує новий запис

D2, то між D1 і D2 утворюється односторонній зв'язок через OP1. Якщо ж D2 пізніше стає вхідним у транзакцію OP2, яка породжує D3 та D4, ланцюг змін продовжується. Така структура природно утворює орієнтований ациклічний граф (DAG), що дозволяє точно простежити шлях перетворення даних від початкового стану до будь-якого похідного.

У запропонованій медичній блокчейн-системі кожен об'єкт MedicalData є вершиною DAG, а кожна транзакція або операція OperationalBehavior, яка змінює ці дані, є окремим вузлом логіки перетворення. Зв'язки між вузлами реалізуються через пари:

- EventInputs — набір DataSign, що були використані як вхідні дані для операції;
- EventOutputs — набір DataSign, що були сформовані як результат цієї операції.

Таким чином, якщо Data D1 є початковим записом (із BehaviorSign = NULL), він може бути використаний в операції OP1, яка створює новий запис D2. Далі D2 може бути перетворений за допомогою іншої транзакції OP2 у записи D3 і D4. В результаті

цього виникає ланцюг перетворення даних, у якому кожен новий запис пов'язаний з попередніми на підставі визначених транзакцій, а весь ланцюг не містить циклів, що забезпечує однозначну історію змін.

У розширеній моделі також враховується інформація про ініціаторів змін (авторів транзакцій) та час виконання операцій, що дозволяє впроваджувати хронологічну простежуваність. Кожен вузол OperationalBehavior містить поля Author і Timestamp, що робить можливим формування темпоральних запитів і аналіз змін у динаміці.

Крім того, до DAG інтегруються об'єкти типу Consent, що встановлюють або обмежують доступ до певних DataSign. Це дозволяє поєднати технічну трасування з правовими аспектами обробки персональних даних.

Для перевірки ефективності створеної платформи управління медичними даними було проведено серію експериментів, що включали завантаження, запитування та простеження джерел медичних даних у блокчейн-мережі. Проведення експериментів

охоплювало аналіз експериментального середовища, структури вхідних даних, а також оцінювання отриманих результатів.

Експериментальне середовище було реалізовано на базі трьох віртуальних серверів, розгорнутих у хмарному середовищі AWS EC2.

У рамках експерименту було розгорнуто сім організацій, що брали участь у симуляції взаємодії у сфері охорони здоров'я:

- 1 регуляторна організація (контроль доступу, аудит),
- 1 організація медичної хмари (централізоване сховище),
- 5 лікарень (розподілені учасники обміну даними).

Кожна організація мала три вузли типу peer, які брали участь у збереженні блоків, перевірці транзакцій та забезпеченні децентралізованої структури мережі. Кластер Orderer, відповідальний за узгодження транзакцій, було розгорнуто на сервері з найбільшою обчислювальною потужністю. Розподіл організацій між серверами відображено на рис. 2.

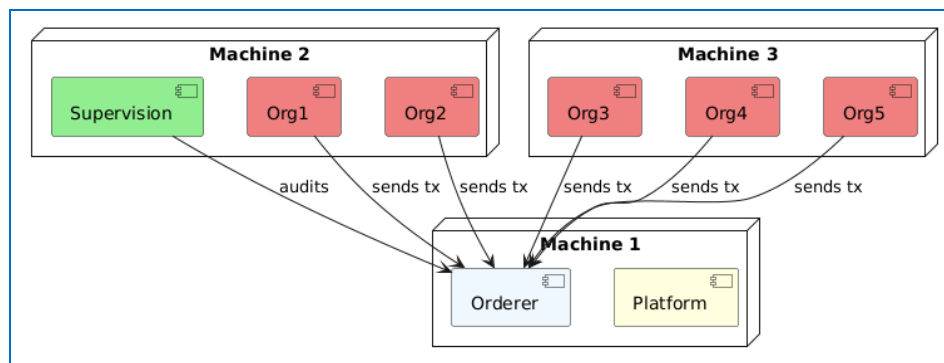


Рис. 2. Діаграма розгортання експериментальної платформи

Діаграма розгортання (рис. 2) моделює фізичне розміщення ключових компонентів Hyperledger Fabric у межах системи. Вузли мережі логічно поділені на три машини:

- Machine 1 виконує функції інфраструктури, зокрема обробку консенсусу через компонент Orderer та роботу з API або SDK через модуль Platform. Саме тут агрегуються транзакції, які надходять з організацій.

- Machine 2 слугує вузлом організаційної участі й нагляду, містить дві повноцінні організації (Org1, Org2) та компонент Supervision, який може виконувати функції аудитора, сертифікаційного органу (CA) або адміністрування консенсусу.

- Machine 3 об'єднує три організації (Org3, Org4, Org5), які ймовірно, беруть участь у мережі як endorsers, або забезпечують спеціалізовану обробку/читання транзакцій.

Логічні зв'язки вказують на взаємодію кожної організації з ордеринг-сервісом (Orderer), що відображає типову схему взаємодії Fabric peer-вузлів із ordering-сервісом для обробки транзакцій.

У рамках верифікації ефективності запропонованої платформи для управління конфіденційною медичною інформацією було сформовано експери-

ментальний набір даних, який за структурою та обсягом відповідає типовим сценаріям використання в інформаційних системах охорони здоров'я. Дані було згенеровано штучно з метою забезпечення конфіденційності, однак вони відповідають загальноприйнятим стандартам представлення медичної інформації, зокрема HL7 та FHIR.

Експериментальні дані включали персоналізовані відомості про пацієнтів, історії хвороб, діагностичні результати, лікарські призначення, а також запити на доступ до медичних записів з боку різних організацій. Враховано також часову динаміку подій та взаємозв'язки між об'єктами, що дозволяє відтворити умови функціонування децентралізованої медичної системи в реальному часі.

Загальний обсяг експериментального набору склав понад 500 000 записів, які охоплюють понад 5 000 умовно унікальних пацієнтів, кілька сотень лікарів та десятки установ різного профілю. Дані подані у форматах JSON, XML та, частково, DICOM, що дозволяє адаптувати їх до потреб як внутрішнього зберігання в блокчейн-мережі, так і зовнішніх децентралізованих сховищ типу IPFS.

З метою оцінки продуктивності побудованої блокчейн-платформи управління медичною інфор-

мацією було проведено серію експериментів із завантаженням і запитом даних. Основною метою дослідження було вивчення впливу кількості організацій у мережі на ключові показники продуктивності, зокрема: швидкість надсилання запитів (Send Rate), максимальну затримку (Max Latency), мінімальну та середню затримку (Min / Avg Latency), а також пропускну здатність (Throughput).

Для тестування використовувався інструмент Hyperledger Caliper, у якому було налаштовано 10 віртуальних працівників (workers), які здійснили 100 000 випадкових операцій із завантаження та запитів медичних і поведінкових даних. Тестування проводилося на блокчейн-мережах із різною кількістю організацій (від 3 до 7). Після п'яти ітерацій кожного тесту було отримано усереднені результати. Результати експериментів свідчать, що із збільшенням кількості організацій спостерігається стабільна тенденція до зменшення показників пропускну здатності (TPS). Зокрема, для завантаження медичних даних при 3 організаціях throughput становив 1349,3 TPS, а при 7 організаціях – 782,8 TPS. Аналогічно зменшився send rate з 1361,7 до 794,7 TPS. При цьому максимальна затримка зросла з 2,05 с до 2,28 с, а середня – з 0,11 с до 0,37 с. Схожі результати були зафіксовані і для поведінкових даних. Так, при збільшенні кількості організацій з 3 до 7 throughput зменшився з 1247,1 до 778,6 TPS, а середня затримка зросла з 0,10 с до 0,37 с.

У разі виконання запитів до медичних даних, спостерігалася аналогічна тенденція: throughput зменшився з 1603,2 TPS до 1179,9 TPS, а середня затримка зросла з 0,09 с до 0,26 с. Поведінкові дані демонстрували схожі показники зниження, зокрема throughput зменшився з 1599,8 до 1173,8 TPS, а середня затримка зросла з 0,09 до 0,27 с.

Встановлено, що основною причиною зниження продуктивності при збільшенні кількості організацій є зростання витрат на досягнення консенсусу та міжорганізаційну комунікацію. Також відзначено, що throughput запитів стабільно перевищував throughput завантажень. Це пояснюється меншою обчислювальною складністю обробки запитів, які не потребують тривалих операцій запису до ланцюга блоків. Крім того, throughput завантаження медичних даних був дещо вищим за аналогічний показник поведінкових даних. Це пояснюється тим, що окремі об'єкти поведінкових даних мають більший обсяг, що спричиняє збільшення часу на їхнє оброблення.

У межах другого етапу експериментального дослідження було протестовано два підходи до реалізації простежуваності медичних даних: традиційний (naive) та графовий (на основі DAG — Directed Acyclic Graph). Тестування охоплювало перевірку прямого та зворотного простеження даних на основі п'яти варіантів обсягів експериментальних даних, що наведені в табл. 1.

У першому експерименті порівнювались середні показники продуктивності двох підходів. Для цього було використано 76 853 поведінкових даних і 200 000 медичних записів. Тестування здійснювалося за допомогою інструмента JMeter із використан-

ням метрик: час виконання, кількість транзакцій за секунду (TPS), середній час відповіді на запит, пропускну здатність.

Таблиця 1 – Порівняння TPS для різних методів простежуваності (на першому наборі даних)

Метод	Тип трасування	TPS
Naive-forward	Пряме	0.54
Naive-backward	Зворотне	388.6
DAG-forward	Пряме	10 214
DAG-backward	Зворотне	10 086

Результати показали, що TPS для DAG-методу значно перевищує TPS для naive-підходу. Зокрема, TPS при прямому простеженні у naive-методі становив лише 0,54, що є недостатнім для практичного застосування. У той час як зворотне простеження в naive-підході демонструвало TPS, вищий приблизно у 718 разів, що свідчить про значну відмінність у часових витратах при різних типах трасування.

DAG-методи (як прямий, так і зворотний) показали TPS, що у 20–26 разів перевищує naive-зворотне простеження. Це пояснюється тим, що DAG реалізує попередньо побудовану структуру, яка дозволяє виконувати трасування без запитів до основного блокчейн-реєстру, обмежуючись операціями пошуку в графі.

У наступному експерименті було протестовано вплив довжини шляху трасування та обсягу даних на продуктивність системи. Встановлено, що TPS всіх методів знижується зі збільшенням довжини шляху, при цьому вплив обсягу даних є незначним. Зокрема, TPS DAG-підходів на довжині шляху 1 був у 20 разів вищим, ніж при довжині ≥ 7 . Така залежність свідчить про тісний зв'язок продуктивності з кількістю кроків трасування в логічному ланцюзі подій (табл. 2).

Таблиця 2 – Залежність TPS від довжини шляху для різних методів

Довжина шляху	Naive-backward (TPS)	DAG-forward (TPS)	DAG-backward (TPS)
1	388.6	10 214	10 086
2	252.9	7 943	7 826
3	168.1	6 210	6 123
4	102.4	4 581	4 508
5	56.3	3 211	3 183
6	39.5	2 105	2 080
≥ 7	35.6	528	519

Для оцінки навантаження на платформу у частині хеш-верифікації було проведено заміри часу, необхідного на виконання базових операцій з блоками даних різного розміру. Результати наведені на рис. 3.

Аналіз даних показує, що час верифікації зростає майже лінійно зі збільшенням розміру пакету. Найменше навантаження спричиняє операція читання, тоді як оновлення є найбільш ресурсомісткою операцією, що потребує додаткових обчислень для перевірки цілісності попередніх та змінених станів. Згідно з отриманими результатами, при максималь-

ному розмірі пакету у 1000 байт час оновлення складає 490 мс, що є критичним показником при масштабних транзакційних навантаженнях.

Ще одним важливим експериментом було дослідження часу, необхідного для надання та відкликання згоди користувача на обробку медичних даних. Результати експерименту представлені на графіку на рис. 4. Аналіз сучасних підходів до зберігання, обміну та захисту медичних даних свідчить про активний розвиток блокчейн-орієнтованих медичних інформаційних систем, таких як MedRec [9], FHIRChain [14], MeDShare [15] та Ancile [8].

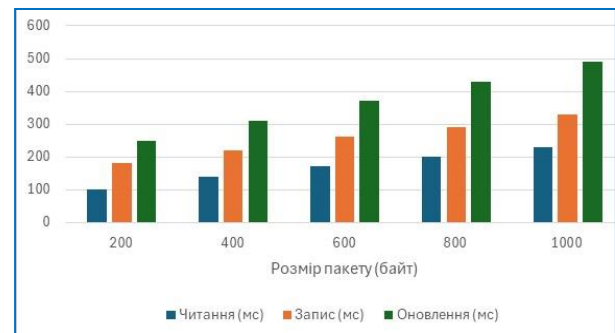


Рис. 3. Аналіз часу верифікації (мс)

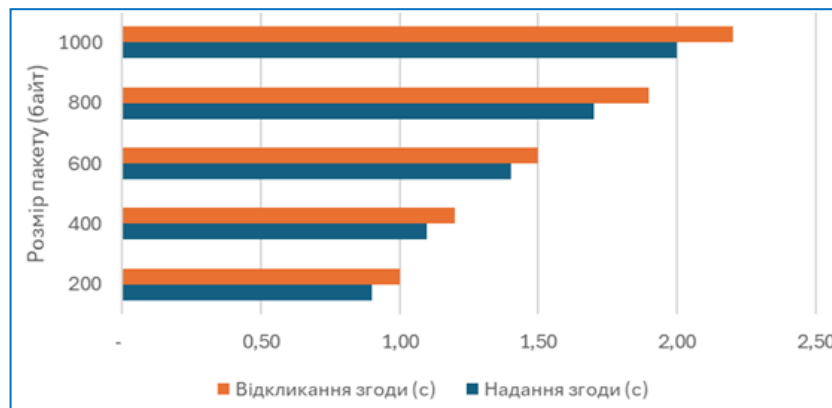


Рис. 4. Час обробки згоди (с)

Водночас, жодна з існуючих платформ не забезпечує повноцінної інтеграції усіх необхідних функцій для універсального управління електронними медичними записами, а також дотримання сучасних міжнародних стандартів, таких як HL7 FHIR.

На основі проведеного аналізу (табл.3) встановлено, що запропонована система, реалізована на базі Hyperledger Fabric, демонструє найбільш збалансовану архітектуру з точки зору функціональності, безпеки, масштабованості та сумісності. Вона забезпечує гібридну модель зберігання, де хеші та метадані зберігаються у блокчейні, а великі об'єми інформації — у зовнішньому репозиторії IPFS. Це дозволяє знизити навантаження на блокчейн без втрати можливості перевірки цілісності даних.

У контексті контролю доступу, запропонована система реалізує рольовий механізм на основі смарт-контрактів. Це забезпечує прозоре надання, обмеження або відкликання доступу до медичних записів пацієнта з юридичною чинністю. Аналогічний механізм підтримує Ancile [8], однак інші системи (зокрема MedRec) реалізують його частково

або в обмеженому вигляді. Серед важливих переваг запропонованої платформи — повна підтримка HL7 FHIR, включаючи REST API, JSON-ресурси, авторизацію через OAuth2 та TLS-захист, що відображено в табл. 4. Це забезпечує інтеграцію з HIS, PACS та EHR-системами без потреби у дублюванні даних або зміні структур обміну. MedRec не реалізує FHIR, тоді як FHIRChain і MeDShare — лише частково.

Висновки

У ході виконання роботи було проаналізовано сучасний стан розробки блокчейн-рішень у сфері охорони здоров'я. Встановлено, що більшість існуючих систем (MedRec, FHIRChain, Ancile, MeDShare) реалізують обмежену функціональність: або зберігають лише метадані, або не забезпечують повної сумісності зі стандартами HL7/FHIR, або не підтримують масштабовану обробку транзакцій у реальному часі. Було виявлено, що лише окремі рішення частково реалізують контроль доступу через смарт-контракти, а повна простежуваність походження даних, як правило, відсутня.

Таблиця 3 – Підтримка HL7 FHIR у різних системах

Компонент FHIR / Технологія	Запропонована система	MedRec [24]	FHIRChain [25]	MeDShare [26]	Ancile [27]
REST API	Так	Ні	Так	Так	Так
JSON ресурси	Так	Так	Так	Так	Так
OAuth2 авторизація	Так	Ні	Так	Ні	Так
FHIR v4.0.1	Так	Ні	Так	Ні	Так
XML-підтримка	Ні	Ні	Ні	Ні	Ні
TLS-захист	Так	Ні	Так	Так	Так
AuditEvent (FHIR журнали)	Ні	Ні	Ні	Ні	Ні

Таблиця 4 – Порівняльна характеристика запропонованої системи з іншими рішеннями

Критерій	Запропонована система	MedRec	FHIRChain	MeDShare	Ancile
Блокчейн-платформа	Hyperledger Fabric	Ethereum	Ethereum	Fabric	Ethereum
Тип блокчейну	Приватний (permissioned)	Публічний	Публічний	Приватний	Публічний
Зберігання медичних даних	Гібридне (IPFS + hash)	Метадані	Метадані + FHIR	IPFS	Off-chain
Контроль доступу (смарт-контракти)	Так	Частково	Так	Так	Так
Надання/відкликання згоди	Так	Частково	Ні	Так	Так
Підтримка HL7 FHIR	Повна	Ні	Часткова	Часткова	Так
Аудит та журнал дій	Так	Частково	Частково	Так	Так
Простежуваність походження даних	Так	Ні	Ні	Частково	Так
Масштабованість (TPS)	Висока (≥ 100 TPS)	Низька	Середня	Висока	Середня
Інтероперабельність з HIS/EHR	Так	Ні	Так	Частково	Так

Запропонована система реалізує контрольовану децентралізацію з високим рівнем надійності, відмовостійкості та доступності. Смарт-контракти відіграють ключову роль у забезпеченні гнучкого та безпечного механізму доступу до інформації, дозволяючи пацієнтам надавати або відкликати згоду на обробку даних в інтерактивному режимі. Важливо, що вся історія змін, включаючи дії користувачів та логіку обробки, фіксується у блокчейні, що забезпечує повну прозорість та можливість аудиту. Порівняльний аналіз із наявними рішеннями

показав, що розроблена система забезпечує найвищий рівень відповідності критеріям безпеки, функціональності, масштабованості та нормативної відповідності. Результати роботи мають практичну цінність для впровадження в реальні медичні установи, зокрема в контексті національної цифрової трансформації охорони здоров'я, а також можуть бути використані як основа для подальших наукових досліджень у галузі медичної інформатики, блокчейн-інфраструктур та захисту персональних даних.

СПИСОК ЛІТЕРАТУРИ

- Борщ, В. В. (2019). Система охорони здоров'я як структурний елемент національної безпеки України. Науковий вісник Ужгородського національного університету. Серія: Міжнародні економічні відносини та світове господарство, (23, Ч. 1), 19–23.
- Борщ, В. І. (2018). Інтелектуалізація системи управління охороною здоров'я. Актуальні проблеми клінічної та профілактичної медицини, (2, № 2-3), 11-20.
- Гуменюк, Н. Є. (2024). Інтелектуалізація системи управління закладом охорони здоров'я (Doctoral dissertation, Тернопіль, ЗУНУ).
- Косенко, В. В., & Невлюдов, І. Ш. (2019). Моделі структурного синтезу для управління параметрами інфокомунікаційних мереж систем критичної інфраструктури: монографія. Харків: ХНУРЕ.
- Ляшук, А. (2023). Загрози і виклики для системи кібербезпеки інформаційних систем та реєстрів сфери охорони здоров'я. Публічне управління: концепції, парадигма, розвиток, удосконалення, (6), 113-121.
- Сорока, І. М. (2023). Удосконалення медичних інформаційних систем як компонент розвитку системи охорони здоров'я. Вісник соціальної гігієни та організації охорони здоров'я України, (3), 62-69.
- Mukherjee, P., Roy, S., Ghosh, D., & Nandi, S. K. (2022). Role of animal models in biomedical research: a review. Laboratory Animal Research, 38(1), 18. DOI: <https://doi.org/10.1186/s42826-022-00128-1>
- Junaid, S. B., Imam, A. A., Balogun, A. O., De Silva, L. C., Surakat, Y. A., Kumar, G., ... & Mahamad, S. (2022, October). Recent advancements in emerging technologies for healthcare management systems: a survey. In Healthcare (Vol. 10, No. 10, p. 1940). DOI: <https://doi.org/10.3390/healthcare10101940>
- Althunibat, A., Binsawad, M., Almaiah, M. A., Almomani, O., Alsaaidah, A., Al-Rahmi, W., & Seliaman, M. E. (2021). Sustainable applications of smart-government services: A model to understand smart-government adoption. Sustainability, 13(6), 3028. DOI: <https://doi.org/10.3390/su13063028>
- Chen, Q., Yang, Z., Liu, H., Man, J., Oladejo, A. O., Ibrahim, S., ... & Hao, B. (2024). Novel Drug Delivery Systems: An Important Direction for Drug Innovation Research and Development. Pharmaceuticals, 16(5), 674. DOI: <https://doi.org/10.3390/pharmaceutics16050674>
- Zeydan, E., Arslan, S. S., & Liyanage, M. (2024). Managing Distributed Machine Learning Lifecycle for Healthcare Data in the Cloud. IEEE Access. DOI: 10.1109/ACCESS.2024.3443520
- Can, O., Thabit, F., Aljahdali, A. O., Al-Homdy, S., & Alkhzaimi, H. A. (2023). A comprehensive literature of genetics cryptographic algorithms for data security in cloud computing. Cybernetics and Systems, 1-35. DOI: <https://doi.org/10.1080/01969722.2023.2175117>
- Taherdoost, H., Le, T. V., & Slimani, K. (2025). Cryptographic Techniques in Artificial Intelligence Security: A Bibliometric Review. Cryptography, 9(1), 17. DOI: <https://doi.org/10.3390/cryptography9010017>
- Kshetri, N., Rahman, M. M., Rana, M. M., Osama, O. F., & Hutson, J. (2024). algoTRIC: Symmetric and asymmetric encryption algorithms for Cryptography--A comparative analysis in AI era. arXiv preprint arXiv:2412.15237. DOI: <https://doi.org/10.48550/arXiv.2412.15237>
- Lalem, F., Laouid, A., Kara, M., Al-Khalidi, M., & Eleyan, A. (2023). A novel digital signature scheme for advanced asymmetric encryption techniques. Applied Sciences, 13(8), 5172. DOI: <https://doi.org/10.3390/app13085172>
- Mihaljević, M. J., Knežević, M., Urošević, D., Wang, L., & Xu, S. (2023). An approach for blockchain and symmetric keys broadcast encryption based access control in IoT. Symmetry, 15(2), 299. DOI: <https://doi.org/10.3390/sym15020299>
- Rehman, S., Talat Bajwa, N., Shah, M. A., Aseeri, A. O., & Anjum, A. (2021). Hybrid AES-ECC model for the security of data over cloud storage. Electronics, 10(21), 2673. DOI: <https://doi.org/10.3390/electronics10212673>

18. Gallo, G. D., & Micucci, D. (2025). Internet of Medical Things Systems Review: Insights into Non-Functional Factors. *Sensors*, 25(9), 2795. DOI: <https://doi.org/10.3390/s25092795>
19. Haritha, T., & Anitha, A. (2023). Multi-level security in healthcare by integrating lattice-based access control and blockchain-based smart contracts system. *IEEE Access*, 11, 114322-114340. DOI: 10.1109/ACCESS.2023.3324740
20. Rathod, T., Jadav, N. K., Alshehri, M. D., Tanwar, S., Sharma, R., Felseghi, R. A., & Raboaca, M. S. (2022). Blockchain for future wireless networks: A decade survey. *Sensors*, 22(11), 4182. DOI: <https://doi.org/10.3390/s22114182>
21. Yilmaz, S., & Dener, M. (2024). Security with Wireless Sensor Networks in Smart Grids: A Review. *Symmetry*, 16(10), 1295. DOI: <https://doi.org/10.3390/sym16101295>
22. Donca, I. C., Stan, O. P., Misaros, M., Stan, A., & Miclea, L. (2024). Comprehensive security for iot devices with kubernetes and raspberry pi cluster. *Electronics*, 13(9), 1613. DOI: <https://doi.org/10.3390/electronics13091613>
23. Nascimento, B., Santos, R., Henriques, J., Bernardo, M. V., & Caldeira, F. (2024). Availability, scalability, and security in the migration from container-based to cloud-native applications. *Computers*, 13(8), 192. DOI: <https://doi.org/10.3390/computers13080192>
24. Kashyap, N., Jeffery, S., & Agresta, T. (2024). From MedWreck to MedRec: A Call to Action to Improve Medication Reconciliation. *Applied Clinical Informatics*, 15(02), 230-233.
25. Reegu, F. A., Abas, H., Gulzar, Y., Xin, Q., Alwan, A. A., Jabbari, A., ... & Dziyauddin, R. A. (2023). Blockchain-based framework for interoperable electronic health records for an improved healthcare system. *Sustainability*, 15(8), 6337. DOI: <https://doi.org/10.3390/su15086337>
26. Han, Y., Zhang, Y., & Vermund, S. H. (2022). Blockchain technology for electronic health records. *International journal of environmental research and public health*, 19(23), 15577. DOI: <https://doi.org/10.3390/ijerph192315577>
27. Lopez, L. J. R., Millan Mayorga, D., Martinez Poveda, L. H., Amaya, A. F. C., & Rojas Reales, W. (2024). Hybrid architectures used in the protection of large healthcare records based on cloud and blockchain integration: A review. *Computers*, 13(6), 152. DOI: <https://doi.org/10.3390/computers13060152>

Received (Надійшла) 14.03.2025

Accepted for publication (Прийнята до друку) 28.05.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Шматко Олександр Віталійович – кандидат технічних наук, доцент, доцент кафедри ЕОМ Харківського національного університету радіоелектроніки, Харків, Україна;

Oleksandr Shmatko – PhD, Ass.prof, Ass.prof of the Department of electronic computers, Kharkiv National University of radio electronics, Kharkiv, Ukraine;

e-mail: oleksandr.shmatko2@nure.ua; ORCID Author ID: <https://orcid.org/0000-0002-2426-900X>

Scopus Author ID: <http://www.scopus.com/inward/authorDetails.url?authorID=6602623478&partnerID=MN8TOARS>

Рагулін Олександр Євгенович – магістр кафедри ЕОМ, Харківського національного університету радіоелектроніки, Харків, Україна;

Oleksandr Rahulin – Master degree student, of the Department of electronic computers, Kharkiv National University of radio electronics, Kharkiv, Ukraine;

e-mail: oleksandr.rahulin@nure.ua; ORCID Author ID: <https://orcid.org/0009-0008-2018-8262>.

Кравченко Павло Олександрович – кандидат технічних наук, асистент кафедри ЕОМ, Харківського національного університету радіоелектроніки, Харків, Україна;

Pavlo Kravchenko – PhD, Assistant Lecturer of the Department of electronic computers, Kharkiv National University of radio electronics, Kharkiv, Ukraine;

e-mail: pavlo.kravchenko@nure.ua; ORCID Author ID: <https://orcid.org/0000-0002-0456-3295>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57216151371>.

Буслов Павло Володимирович – кандидат технічних наук, асистент кафедри БІТ Харківського національного університету радіоелектроніки, Харків, Україна;

Pavlo Buslov – PhD, Assistant Lecturer of the Department of cyber information technologies, Kharkiv National University of radio electronics, Kharkiv, Ukraine;

e-mail: p.buslov@ukr.net; ORCID Author ID: <https://orcid.org/0000-0002-2558-4984>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57207767753&origin=resultslist>.

Investigation of architectural solutions for building a secure system for confidential data storage and transmission

Oleksandr Shmatko, Oleksandr Rahulin, Pavlo Kravchenko, Pavlo Buslov

Abstract. Relevance. In the current landscape of rapid digitalization and increasing cyber threats, protecting confidential information during storage and transmission has become a critical priority in the design of information systems. A robust system architecture ensures resistance to attacks, data integrity, and access control. **Object of the study:** architectural solutions for developing secure software systems for storing and transmitting confidential data. **Purpose of the article:** to investigate, design, and develop architectural components of a secure information system capable of maintaining data confidentiality during transmission and storage. **Research results.** The article presents an analysis of current approaches to building secure information systems and justifies the selected architectural model and tools. A system prototype was developed. The system was tested for compliance with security and performance requirements. **Conclusions.** The proposed architectural solution demonstrated effectiveness in ensuring data integrity, confidentiality, and availability. The results can be applied to further improve information systems in fields dealing with sensitive or mission-critical information.

Keywords: information security; confidentiality; data transmission; data storage; software architecture; encryption.

Я. І. Бірук, С. Й. Шаманський, А. В. Климчук

Київський національний університет будівництва і архітектури, Київ, Україна

ДОСЛІДЖЕННЯ ЕФЕКТИВНОСТІ ЕЛЕКТРОМАГНІТНИХ ЕКРАНІВ РЕЗОНАНСНОГО ТИПУ

Анотація. Досліджено можливості розроблення одношарового захисного покриття з малими коефіцієнтами електромагнітних хвиль ультрависоких, надвисоких і надзвичайно високих частот. Визначено, що досягнення потрібних коефіцієнтів відбиття досягається за рахунок узгодження електричної та фізичної товщин захисного матеріалу. Регулювання шуканих параметрів здійснюється за рахунок обирання матеріалу з визначеною діелектричною проникністю. Застосування у якості наповнювача діелектричної матриці ($\epsilon = 24$) магнетиту ($\epsilon = 14$) дозволяє застосовувати співвідношення для розрахунку ефективності діелектриків. Наявність у магнетиті магнітних властивостей ($\mu = 4,5$) дозволяє одночасно екранувати й електромагнітні поля промислової частоти з коефіцієнтами екранування 6–8. Встановлено, що для електричної товщини матеріалу, яка дорівнює одиниці (перший порядок чвертьхвильового захисту) на частоті 1 ГГц прийнятний коефіцієнт відбиття (-10 дБ) досягається у смузі частот від 0,9 до 1,1 ГГц, на частоті 3 ГГц – від 2,3 до 3,3 ГГц, на частоті 6 ГГц – від 5,3 до 6,6 ГГц, на частоті 10 ГГц – від 9 до 11 ГГц. Це дозволяє у практичній діяльності враховувати певний діапазон робочих частот обладнання визначених стандартів. Слід враховувати, що при визначенні ефективної діелектричної проникності композиційного матеріалу за відомими співвідношеннями спостерігається велика похибка. Тому у розрахунках слід користуватися максимальними значеннями, що дозволить отримати певний запас ефективності при визначенні електричної товщини матеріалу.

Ключові слова: електромагнітне поле, електромагнітний екран, резонанс, композиційний матеріал.

Вступ

Розроблення та дослідження захисних властивостей матеріалів та конструкцій для екранування електромагнітних полів широкого частотного діапазону є одним з пріоритетних напрямів у галузях цивільної безпеки. Сучасною тенденцією у створенні таких матеріалів є проектування композицій з потрібними коефіцієнтами екранування. Це обумовлене необхідністю забезпечення високих захисних властивостей екранів за рахунок поглинання електромагнітної енергії і мінімізації коефіцієнтів відбиття електромагнітних хвиль. В останні роки спостерігається тенденція до підвищення робочих частот усіх типів бездротового зв'язку, що призводить до перерозподілу інтенсивностей електромагнітних випромінювань внаслідок відбиття від поверхонь будівельних конструкцій, металевих поверхонь тощо. Для мінімізації коефіцієнтів відбиття і забезпечення достатніх поглинальних властивостей композиційні матеріали повинні мати градієнт електрофізичних властивостей по товщині, що підвищує їх вартість й знижує технологічність застосування. Зазвичай захисний матеріал наноситься на якусь поверхню, тому відбиття електромагнітної хвилі відбувається від цієї поверхні, якщо товщина захисного матеріалу набагато менша за довжину падаючої хвилі. Особливо це практично для приміщень спеціального призначення, у яких виконуються дослідження щодо електромагнітної сумісності електронного обладнання. У таких приміщеннях крім повної ізоляції від зовнішніх електромагнітних впливів коефіцієнти відбиття електромагнітних хвиль не повинні перевищувати -10 дБ. Такі показники можна забезпечити за рахунок застосування узгоджувального шару між несучою поверхнею та захисним матері-

алом або за рахунок узгодження електричної та фізичної товщини захисного матеріалу. Тому доцільно дослідити частотні залежності цих показників за певних складів захисних композицій матеріалів.

Огляд літературних джерел. Кількісні значення електромагнітних полів широкого частотного діапазону регламентуються національними санітарними нормами [1, 2]. Крім того в Україні методом підтвердження прийнято низку міжнародних стандартів з елетросумісності обладнання, наприклад [3, 4]. При цьому строго регламентуються умови не тільки експлуатації, а й випробування електронного обладнання. Це вимагає наявності засобів екранування електромагнітних полів для їх нормалізацій у виробничих та побутових умовах. Найбільш актуальною проблемою є захист від електромагнітних випромінювань ультрависоких, надвисоких та надзвичайно високих частот. Розроблено багато композиційних матеріалів для екранування випромінювань цих частотних діапазонів. У роботі [5] наведено результати випробувань матеріалу з різними фазовими складами. Він має високі захисні властивості, але має й великий коефіцієнт відбиття. Матеріали [6, 7] майже не відбивають електромагнітні хвилі, але мають низький коефіцієнт поглинання. Достатньо збалансовані коефіцієнти відбиття та поглинання у матеріалах, запропонованих у [8, 9]. Але для умов випробувань електронної техніки та захисту від вузько-смугового електромагнітного випромінювання коефіцієнтів відбиття можна досягти, застосовуючи узгоджувальний шар між окремими поверхнями завтовшки непарній чверті довжини падаючої хвилі [10]. Але у даному випадку матеріал достатньо громіздкий. При цьому він вважається призначеним для поглинання монохромного випромінювання.

Доцільно дослідити можливості забезпечення певної смуги частот, у якій коефіцієнт відбиття електромагнітних хвиль має прийнятні значення.

Метою дослідження є визначення ефективності електромагнітних екранів резонансного типу у певній смузі частот, близьких до резонансної.

Викладення основного матеріалу

Зниження значень коефіцієнтів відбиття електромагнітних хвиль від поверхонь будівель, приміщень тощо залежить виключно від електрофізичних показників поверхневих шарів матеріалу. Зазвичай діелектричні проникності матеріалів мають значення 2–5. Коефіцієнти відбиття електромагнітних хвиль K_r не можуть бути нижчими за значення:

$$K_r = 10 \lg \left(\sqrt{\varepsilon_e} - 1 / (\sqrt{\varepsilon_e} + 1) \right), \text{ дБ},$$

де ε_e – ефективна відносна діелектрична проникність матеріалу на даній частоті випромінювання. Тому коефіцієнти відбиття мають значення - 10 ÷ -7 дБ.

Для більшості умов перебування людей це прийнятний показник. Але такий матеріал практично не поглинає електромагнітну енергію і не може розглядатися як захисний. Для приміщень, де рівні електромагнітних полів повинні бути мінімальними, а випромінювання внутрішніх джерел суттєво не відбивається від поверхонь необхідно одночасно забезпечувати малі коефіцієнти відбиття, не більше - 12 дБ, та високі коефіцієнти поглинання – не менше 20 дБ.

Такі показники можна отримати, застосовуючи тришарове покриття. Таке покриття повинне мати товщину, яка дорівнює половині максимальної довжини хвилі смуги частот електромагнітного випромінювання, яке падає на поверхню. Коефіцієнт відбиття можна знизити до значень -15 ÷ -20 дБ у певній смузі частот за умови, що товщина зовнішнього шару d й мінімальна довжина падаючої електромагнітної хвилі λ відповідає співвідношенню:

$$K_r = \frac{d\sqrt{\varepsilon_e}}{\lambda} < 0,3.$$

У випадку подальшого зменшення довжини хвилі коефіцієнт відбиття буде підвищуватися досить швидко й досягати значення -40 дБ за умови:

$$K_r = \frac{d\sqrt{\varepsilon_e}}{\lambda} < 0,25.$$

У будь-якому випадку між поверхнею монтажу та зовнішнім захисним матеріалом повинен бути узгоджувальний шар товщиною, кратною чверті довжини падаючої електромагнітної хвилі. Для надання покриттю певної широкосмуговості узгоджувальний шар повинен достатньо товстий, ще не завжди зручно у практичній роботі.

Зазвичай у будь-яких реальних умовах існують переважні довжини хвиль (частоти), характерні для електромагнітної обстановки. У більшості випадків це визначається стандартними робочими частотами обладнання, наприклад, засобів мобільного зв'язку. Але реальні робочі частоти сіх випромінювачів

лежать у певній смузі навколо основної частоти. Якщо захисний шар матеріалу має товщину, яка відповідає чверті довжини електромагнітної хвилі, яка відповідає робочій частоті, то для реальних робочих частот покриття може бути недостатньо ефективним. Тому доцільно визначити межі ефективності покриттів навколо основної частоти. Мінімум досяжний рівень коефіцієнта відбиття визначається товщиною покриття, яке відповідає основній (резонансній) частоті. Ця товщина повинна відповідати чверті довжини падаючої хвилі, або бути кратною цій величині – непарній кількості чвертей. Навіть для достатньо великих діелектричних проникностей поверхневого шару матеріалу за таких умов відбувається інтерференційне придушення відбиття. Для раціоналізації параметрів захисного матеріалу необхідно визначити межі його прийнятної ефективності. Особливістю резонансних конструкцій є те, що оптимальна товщина захисного покриття визначається неоднозначно. Вона відповідає електричній товщині матеріалу за умови непарним кратностям чвертей довжин хвиль:

$$T_e = 2\pi/\lambda \cdot l \cdot R_e \sqrt{\varepsilon_{екв}} \approx n \cdot \pi/2,$$

де λ – довжина хвилі, l – товщина, $n = 1, 3, 5, \dots$

Значення $n = 1$ відповідає мінімально необхідній товщині радіопоглинального матеріалу.

Були визначені значення коефіцієнтів відбиття для різних резонансних частот з визначенням їх приведених електричних товщин, які відповідають $n = 1$.

Розрахунки проводилися для композиційного матеріалу на основі латексу та магнетиту (табл. 1).

У таблиці ε_m – ефективна діелектрична проникність матриці, ε_n – ефективна діелектрична проникність екрануючого наповнювача, V_n – об'ємний вміст екрануючого наповнювача, f_r – резонансна частота, f – частота електромагнітної хвилі, K_r – коефіцієнт відбиття електромагнітного випромінювання, $2T_e/\pi$ – приведена електрична товщина матеріалу, яка відповідає мінімальній кількості непарних чвертей довжини електромагнітної хвилі (фізична товщина – 0,25 λ).

Вважається, що коефіцієнт відбиття -10 дБ є достатнім для вирішення задач електромагнітної сумісності та випробування електронної техніки. Тому для базових частот можна визначити частотні смуги прийнятних значень коефіцієнтів відбиття. Це стосується й процесів зниження інтенсивності випромінювань базових станцій зв'язку у зонах їх можливого підвищення за рахунок відбиття електромагнітних хвиль. У даному випадку наповнювачем полімерної матриці розглядався магнетит. Його перевагою є наявність й магнітних властивостей, що одночасно дає можливість знизити рівні магнітної складової електромагнітного поля промислової частоти [11]. Це особливо актуально енергонасичених будівель з наявністю великої кількості електричного обладнання з нелінійними вольт-амперними характеристиками. Навіть за товщини покриття до 2 мм коефіцієнти екранування електричної складової електромагнітного поля промислової частоти складають 5–6, магнітної складової 6–8, для більшості виробничих та побутових умов є достатньо.

Таблиця 1 – Залежність ефективності резонансних електромагнітних екранів від електрофізичних властивостей і вмісту компонентів

ε_m	ε_n	$V_n, \%$	f_r, GHz	f, GHz	K_r, dB	$2T_e/\pi$
24	12	15	0,500	0,45	-6,75	0,91
				0,46	-8,85	0,93
				0,48	-11,65	0,96
				0,49	-15,18	0,98
				0,500	17,73	1,00
				0,512	16,01	1,03
				0,525	13,00	1,05
				0,538	-10,64	1,07
24	12	15	1,000	0,550	-8,90	1,10
				0,900	-11,08	0,93
				0,925	-13,50	0,96
				0,950	-16,53	0,98
				0,975	-20,10	0,99
				1,000	-22,17	1,00
				1,025	-20,15	1,05
				1,050	-17,20	1,07
24	12	15	3,000	1,075	-14,83	1,09
				1,100	-13,02	1,11
				2,700	-9,52	0,92
				2,780	-11,19	0,94
				2,850	-13,04	0,96
				2,935	-14,96	0,98
				3,000	-16,56	1,00
				3,075	-16,50	1,02
24	12	15	6,000	3,150	-16,48	1,04
				3,225	-15,32	1,07
				3,300	-13,94	1,09
				5,400	-12,60	0,95
				5,550	-14,48	0,97
				5,700	-16,45	0,99
				5,850	-18,19	1,00
				6,000	-19,67	1,00
24	12	15	10,000	6,150	-18,65	1,01
				6,300	-17,40	1,02
				6,450	-16,04	1,04
				6,600	-14,71	1,06
				9,000	-14,78	0,97
				9,250	-17,17	0,99
				9,500	-19,84	1,01
				9,950	-22,07	1,02
24	12	15	10,000	10,000	-22,26	1,03
				10,250	-20,50	1,05
				10,500	-18,37	1,10
				10,750	-16,53	1,12
				11,000	-15,00	1,14

При розрахунковому визначенні відносних діелектричних проникностей композиційних матеріалів слід враховувати, що співвідношення Максвелла-Гарнета, Бруггемана є емпіричними і дають велику похибку. Тому для зниження коефіцієнтів відбиття слід обирати максимальне значення, що надає певний запас ефективності по електричній товщині матеріалу.

Висновки

1. Визначено засади розроблення одношарового покриття з мінімальними коефіцієнтами відбиття електромагнітних хвиль ультрависоких та вищих частот. Головною умовою ефективності таких

покриттів є узгодженість електричної та фізичної товщин захисного матеріалу. Визначено можливість одночасного зниження у визначених зонах електромагнітних полів промислової частоти за рахунок застосування у композиційних матеріалах наповнювача з магнітними властивостями.

2. Визначено смуги прийнятних значень коефіцієнтів відбиття електромагнітних хвиль навколо резонансних частот покриттів, для яких електрична товщина дорівнює одиниці. Це дозволяє знижувати значення щільностей потоків енергії височастотних випромінювань за різних робочих частот випромінювачів поблизу основної (стандартної) частоти.

СПИСОК ЛІТЕРАТУРИ

1. ДСН 239-96. Державні санітарні норми і правила захисту населення від впливу електромагнітних випромінювань [Чинний від 2017-12-22]: затв. наказом МОЗ України від 01.08.1996 р. № 239. Київ, 2017. 28 с.

2. ДСНІП 3.3.6.096-2002. Державні санітарні норми і правила при роботі з джерелами електромагнітних полів [Чинний від 2003-03-13]: затв. наказом М-ва охорони здоров'я України від 18.12.2002 р. № 476. Київ, 2003. 16 с.
3. 41. ETSI EN 300 220-2 V2.4.1 (2012-01). Electromagnetic compatibility and Radio spectrum Matters (ERM); Short Range Devices (SRD); Radio equipment to be used in the 25 MHz to 1 000 MHz frequency range with power levels ranging up to 500 mW; Part 2: Harmonized EN covering essential requirements under article 3.2 of the R&TTE. Directive. European Telecommunications Standards Institute. 2012. 20 p.
4. 42. ETSI EN 301 489-1 V2.2.1 (2019-03). ElectroMagnetic Compatibility (EMC) standard for radio equipment and services; Part 1: Common technical requirements; Harmonised Standard for ElectroMagnetic Compatibility. Directive. European Telecommunications Standards Institute, 2019. 36 p.
5. Kefeng JI; Jun GAO; Xiangyu CAO; Jiangfeng HAN; Huanhuan YANG. Design of Ultra-wideband Low RCS Reflecting Screen Based on Phase Gradient Metasurface. *Radioengineering*. Jun 2021, Vol. 30 Issue 2, p. 314-322. <https://doi.org/10.13164/re.2021.0314>
6. Кизимчук О.П., Арабулі С.І., Власенко В.І. Текстиль для захисту від електромагнітного випромінювання. *Вісник Київського національного університету технологій та дизайну. Серія Технічні науки*. 2019. № 3 (134). С. 48-61
7. Арабулі С.І., Кизимчук О.П., Власенко В.І., Тунак М., Тунакова В. Дослідження ефективності екранування ЕМВ текстильними матеріалами. III International Scientific-Practical Conference 31.10.2019 Kyiv, Ukraine. Kyiv, 2019. С. 202–206.
8. Butenko O., Boychuk V., Savchenko B., Kotsyubynsky V., Khomenko V., Barsukov V. Pure ultrafine magnetite from carbon steel wastes. *Materials Today*. 2019. V. 6, pp. 270–278
9. Senyk I., Kuryptia Y., Barsukov V., Butenko O., Khomenko V. Development and application of thin wide-band screening composite materials. *Physics and Chemistry of Solid State*. 2020. 21(4). Pp. 771–778.
10. 175. Glyva V., Levchenko L., Panova O., Tykhenko O., Radomska M. The composite facing material for electromagnetic fields shielding. *Innovative Technology in Architecture and Design (ITAD 2020): IOP Conference Series: Materials Science and Engineering*. 2020. Vol. 907 URL: <https://iopscience.iop.org/article/10.1088/1757-899X/907/1/012043/meta>
11. Burdeina, N., Levchenko, L., Korduba, I., Shamanskyi, S., Biruk, Y., Dovhanovskyi, M., Zozulya, S., Klymchuk, A., Nikolaiev, K., & Osadchyi, D. (2024). Applying heterogeneous building materials for the protection of people against electromagnetic radiation. *Eastern-European Journal of Enterprise Technologies*, 5(10 (131)), 45–52. <https://doi.org/10.15587/1729-4061.2024.313629>

Received (Надійшла) 23.02.2025

Accepted for publication (Прийнята до друку) 30.04.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Бірук Яна Ігорівна – доктор філософії, доцент, доцент кафедри фізики, Київський національний університет будівництва та архітектури, Київ, Україна;

Yana Biruk – PhD, Associate Professor, Associate Professor of the Department of Physics, Kyiv National University of Construction and Architecture, Kyiv, Ukraine;

e-mail: vesna0999@gmail.com, ORCID Author ID: <https://orcid.org/0000-0002-3669-9744>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57225188391>.

Шаманський Сергій Йосипович – доктор технічних наук, професор, професор кафедри водопостачання та водовідведення, Київський національний університет будівництва та архітектури, Київ, Україна;

Sergii Shamanskyi – Doctor of Science (Engineering), Professor, Professor of the Department of Water Supply and Drainage, Kyiv National University of Construction and Architecture, Kyiv, Ukraine;

e-mail: shamanskiy.sy@knuba.edu.ua, ORCID Author ID: <https://orcid.org/0000-0002-6215-3438>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57190491594>.

Климчук Андрій Васильович – аспірант кафедри технологій захисту навколишнього середовища та охорони праці, Київський національний університет будівництва та архітектури, Київ, Україна;

Andrii Klymchuk – PhD student at the Department of Environmental Protection Technologies and Labour Protection, Kyiv National University of Construction and Architecture, Kyiv, Ukraine;

e-mail: klymchuk_av-2022@knuba.edu.ua; ORCID Author ID: <https://orcid.org/0009-0005-1248-7316>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=59413074000&origin=recordpage>.

Study of the efficiency of resonant electromagnetic screens

Yana Biruk, Sergii Shamanskyi, Andrii Klymchuk

Abstract. The possibilities of developing a single-layer protective coating with low coefficients of electromagnetic waves of ultra-high, ultra-high and extremely high frequencies are investigated. It is determined that the required reflection coefficients are achieved by matching the electrical and physical thicknesses of the protective material. The desired parameters are regulated by selecting a material with a certain dielectric constant. The use of magnetite ($\epsilon = 24$) as a filler in the dielectric matrix ($\epsilon = 14$) allows the use of the ratio to calculate the dielectric efficiency. The presence of magnetic properties in magnetite ($\mu = 4.5$) makes it possible to simultaneously shield electromagnetic fields of industrial frequency with shielding coefficients of 6–8. It has been established that for the electrical thickness of the material equal to one (first order of quarter-wave protection) at a frequency of 1 GHz, an acceptable reflection coefficient (-10 dB) is achieved in the frequency band from 0.9 to 1.1 GHz, at a frequency of 3 GHz – from 2.3 to 3.3 GHz, at a frequency of 6 GHz – from 5.3 to 6.6 GHz, at a frequency of 10 GHz – from 9 to 11 GHz. This makes it possible to take into account a certain range of operating frequencies of equipment of certain standards in practice. It should be borne in mind that when determining the effective permittivity of a composite material using known ratios, a large error is observed. Therefore, the maximum values should be used in the calculations, which will allow obtaining a certain margin of efficiency in determining the electrical thickness of the material.

Keywords: electromagnetic field, electromagnetic shield, resonance, composite material.

В. А. Глива, М. С. Кашлев

Київський національний університет будівництва і архітектури, Київ, Україна

ЕКРАНУВАННЯ ШУМУ БУДІВЕЛЬНОГО ОБЛАДНАННЯ

Анотація. Проведено експериментальні дослідження шумових характеристик будівельного обладнання. З'ясовано, що у більшості випадків перевищення гранично допустимих значень звукового тиску складає не більше 20–25 дБ. Це обумовлює можливість застосування акустичних екранів для нормалізації рівнів шуму на робочих місцях і територіях. Перевагою екранів є простота виготовлення та низька вартість. Це дозволяє використовувати екрануючі конструкції великих площ. Визначено внесок у поглинання шуму найбільш типових підстилаючих поверхонь. Ці дані дозволяють внести корективи у необхідну ефективність акустичних екранів. Досліджено амплітудно-частотні характеристики типових джерел шуму на будівельних майданчиках. Надано розрахунки щодо прогнозування ефективності акустичних екранів у залежності від умов їх використання. Враховано наявність дифракційних явищ на кромках екранів для великих значень кутів дифракції з урахуванням розташування екрана на поверхні та над поверхнею. Представлено розрахунок щодо значення коефіцієнта дифракції звукових хвиль з урахуванням відстані від джерела звуку до екрана та його кромки, а також відстані від кромки екрана до зони зниження рівня звуку. Враховані фізичні характеристики матеріалу, з якого виготовлений акустичний екран. Надано практичні рекомендації щодо підвищення ефективності акустичних екранів з урахуванням розташування, форми екрана та амплітудно-частотних характеристик шуму, який підлягає екрануванню.

Ключові слова: будівельне обладнання, шум, акустичний екран, коефіцієнт звукопоглинання.

Вступ

Не дивлячись на значну увагу, що приділяється проблематиці зниження акустичного навантаження на виробничі й оточуюче середовище, дана задача ще далека від вирішення. Це пояснюється як різноманітністю джерел шуму, так і зміною їх кількісного та якісного складу.

У багатьох випадках суттєвим джерелом шуму на територіях населених місць є будівельні майданчики з різною кількістю будівельної техніки й обладнання. Для України задачі, пов'язані з будівельним шумом залишаються актуальними на багато років. У значній кількості населених пунктів є багато часткового або повністю зруйнованих об'єктів, відновлення яких є доцільним.

Як правило, такі об'єкти розташовані у зонах житлової або промислової забудови, де постійно або тимчасово перебуває велика кількість людей.

Тому бажана мінімізація шуму різноманітної техніки при виконанні будівельних або відновлювальних робіт.

Розроблення відповідних заходів повинне базуватися на визначенні амплітудно-частотних характеристик будівельного обладнання.

Це дозволить мінімізувати витрати на розроблення й впровадження відповідних заходів зниження впливу шуму на людей.

Огляд літературних джерел. Рівні виробничого шуму, ультразвуку й інфразвуку нормуються в Україні санітарними нормами [1].

Шум будівельного обладнання обмежується еквівалентним значенням 80 дБА. Але це гранично допустимий рівень для виробничих умов на робочому місці. При виконанні робіт на території населених місць він повинен бути нижчий [2].

Тому при виконанні робіт у населених місцях необхідно вживати заходи зі зниження акустичного навантаження на середовище.

Найбільш ефективним засобом є шумопоглинальні конструкції [3, 4].

Але такі матеріали й конструкції призначені для облицювання поверхонь великих площ.

У дослідженні [5] запропоновано теоретичні й експериментальні підходи до проектування резонансних звукопоглинальних панелей обмежених розмірів. Такі конструкції мають високу вартість й призначені, в основному, для захисту від низькочастотного звуку та інфразвуку.

У багатьох випадках складні конструкції [6] можуть бути надлишковими за ефективністю.

Найбільш відпрацьованими конструкціями для захисних рівнів шуму є шумозахисні екрани [7, 8]. Але їх ефективність функціонально пов'язана з амплітудно-частотними характеристиками шуму джерел та розмірами екранів. При цьому необхідно зважити на дифракційні ефекти на кромках екрануючих конструкцій.

Цей напрям добре розроблений, але потребує уточнення застосування екрануючих конструкцій для зниження шуму будівельного обладнання в умовах обмежених розмірів будівельних матеріалів.

Мета роботи – дослідити ефективність зниження шуму будівельного обладнання із застосуванням екрануючих конструкцій.

Викладення основного матеріалу

Певною складністю у процесі розроблення заходів зі зниження шуму будівельного обладнання є розбіжності у визначенні значень шуму.

Це пояснюється різними режимами роботи обладнання, а також відбиттям звукових хвиль від різних поверхонь, що потрібно враховувати у практичній діяльності [9].

Проведені інструментальні дослідження з використанням каліброваного шумоміру ОКТАВА 110 А дозволили отримати усереднені дані щодо шуму окремих агрегатів.

Таблиця 1 – Середні рівні шуму будівельного обладнання

Тип обладнання	Рівень шуму, дБ
Машина для забивання паль	115–120
Будівельно-монтажні пістолети	115–120
Бетонозмішувач гравітаційної дії	95–105
Пневмомолотки	105–110
Бульдозери	95–105
Віброплити	95–105
Машини для різки бетону	95–100
Трактори	90–95
Зварювальні апарати з дизельними генераторами	95–105
Генератори резервного живлення	80–95
Вантажні автомобілі, крани	80–90
Екскаватори	80–90
Бетононаноси	80–90
Вібраційний брус	85–95

Як видно з табл. 1, рівні шуму будівельного обладнання досить високі, що обумовлює необхідність його зниження.

У реальних умовах слід враховувати, що на певних відстанях відбувається часткове звукопоглинання поверхнями (табл. 2).

Таблиця 2 – Середні коефіцієнти звукопоглинання відбиваючих поверхонь

Відбиваюча поверхня	Октавні смуги частот, Гц							
	63	125	250	500	1000	2000	4000	8000
Асфальт, бетон	0,02	0,02	0,03	0,03	0,04	0,05	0,06	0,06
Невисока трава, рихлий ґрунт	0,10	0,33	0,40	0,80	0,80	0,80	0,80	0,70

Коефіцієнти звукопоглинання на будівельних майданчиках мають середні значення між наведеними у табл. 2 значеннями. Але вони впливають на рівні шуму на великих відстанях.

У табл. 3 представлено рівні звукового тиску у октавних смугах частот двох різних джерел шуму, які використовуються у будівництві.

Таблиця 3 – Рівні звукового тиску обладнання з двигунами різного типу

Джерело шуму	Рівні звукового тиску, дБ						
	63	125	250	500	1000	2000	4000
Транспортна машина, бензиновий двигун	69	72	75	79	80	82	73
Пересувна компресорна станція, дизельний двигун	95	90	92	88	92	80	74

Амплітудно-частотні характеристики наведених джерел суттєво відрізняються у низькочастотній області.

Для таких частот гранично допустимі рівні звукового тиску достатньо високі – 90–95 дБ. Відмінності на середніх і високих частотах не критичні, а

значення звукового тиску не перевищують граничні значення. Для цих частотних смуг можливе зниження рівнів шуму акустичними екранами. За низьких частот звуку й високих амплітуд їх ефективність поступається ефективності резонансних та шумопоглинальним конструкціям.

Але акустичні екрани мають низьку вартість, що надає можливість монтажу захисних конструкцій великих площ.

Звукоізолюючий ефект акустичного екрана визначається зоною акустичної тіні за ним.

Тому розмір екрана є критичним для значення звукопоглинання, особливо з відстанню від нього у захищеній зоні.

Зниження ефективності екранування з відстанню обумовлене дифракційними явищами на кромках екрана.

Внесок дифракції збільшується зі зменшенням площі екрана, кута дифракції та довжини звукової хвилі. За великих кутів дифракції рівні звукового тиску можна знизити на 10–20 дБ. У середньо та високочастотних областях звукового спектра. У більшості випадків це буває достатньо для нормалізації рівня шуму.

Акустичні екрани найбільш ефективні у наступних випадках: переважність середньо та високочастотних складових шуму; невеликі розміри джерела шуму; відносно вузька спрямованість випромінювання звуку.

Усі відомі методи розрахунку акустичних екранів застосовують оптико-дифракційний підхід [10]. Слід враховувати, що для великих кутів дифракції такий підхід дає значну похибку. Також не враховуються особливості розташування акустичних екранів у просторі, наявність відбиваючих поверхонь. Цей недолік можна подолати, додавши до відповідного рівняння ще один член.

Ефективність акустичного екрана у загальному випадку визначається як:

$$\Delta L = L_0 - L_e,$$

де L_0 – рівень звукового тиску до екрана, L_e – у екранованій зоні, дБ.

Для вирішення інженерних задач для прямокутної конструкції прийнятну точність дає співвідношення:

$$\Delta L = -10 \lg \left(\tau + \sum_{i=1}^4 K_{di} + \sum_{n=1}^N K_{en} \right) + 10 \lg \left(1 + \sum_{n=1}^N K_{en} \right),$$

де τ – коефіцієнт звукопровідності матеріалу екрана, K_{di} – коефіцієнт дифракції i -тої кромки екрана, K_{en} – коефіцієнт відбиття n -ї поверхні, N – кількість відбиваючих поверхонь.

Розгляд кожної кромки екрана обов'язковий через те, що чотири боки конструкції можуть бути нерівноправними: наприклад, один бік розташований на поверхні і дифракція відсутня.

Коефіцієнт звукопровідності зв'язаний зі звукоізоляцією матеріалу екрана R (дБ) стандартним співвідношенням:

$$\tau = 10^{-0,1R}.$$

Коефіцієнт дифракції визначається співвідношенням:

$$K_{di} = \frac{\gamma_i (10 - 3\varphi\alpha) h}{f \gamma_0 (r_i + l_i - r_0)},$$

де r_0 – відстань від джерела звуку до точки визначення звуку за екраном, r_i – відстань від джерела звуку до відповідної кромки екрана, l_i – відстань від кромки екрана до точки визначення звуку за екраном, α – коефіцієнт звукопоглинання матеріалу екрана, γ_0 , γ_i – характеристики спрямованості джерела звуку, які визначаються значенням кута дифракції f – кут між векторами R та l . Параметр h визначається розташуванням екрана.

Експериментально встановлено, що $h = 1$ у разі, якщо відстань від поверхні до нижньої кромки екрана більша за 1,5 м.

У інших випадках (відстань від 0 до 1,5 м) цей параметр дорівнює $0,6 h$. Частота звуку обирається f як середньгеометрична октавних смуг.

Для підвищення ефективності акустичних для різних умов можна сформулювати декілька вимог.

Якщо у спектрі шуму джерела переважають низькочастотні складові, то найбільш дієвими засобами підвищення ефективності екранів є такі:

- збільшення розмірів екранів;
- застосування екранів більш складної форми, наприклад, Г-подібної та П-подібної;
- зменшення відстані між джерелом шуму та екраном;
- зменшення відстані між екраном та зоною, яка потребує захисту.

Також підвищення ефективності можливо за наявності великої кількості відбиваючих поверхонь й переважно середньо та низькочастотних складових шуму при таких умовах:

- облицювання поверхні екрана звукопоглинальним матеріалом;
- зміна кута відбиття звуку від усіх критичних поверхонь;
- зменшення ефективної площі відбиття; зменшення зазорів між відбиваючими поверхнями та екраном зміна розташування та форми екрана.

Не дивлячись на те, що вважається можливим досягнення зниження рівнів звукового тиску за допомогою акустичних екранів на рівні 18–20 дБ, у реальних умовах такі значення майже недосяжні. За площ акустичних екранів розумних розмірів та їх встановлення на прийнятних відстанях зниження рівнів звукового тиску складає 10–12 дБ.

Цей факт необхідно врахувати у практичній діяльності й визначати потрібну ефективність акустичних екранів, виходячи з реальних умов розповсюдження звуку та принципу розумної достатності.

Висновки

1. Проведено експериментальні дослідження рівнів шуму будівельного обладнання. Визначено, що у більшості випадків перевищення гранично допустимих рівнів складає 10–15 дБ. Це надає змогу використовувати шумозахисні екрани простої конструкції та низької вартості.

2. Показано, що для визначення засобів захисту необхідно виміряти амплітудно-частотні характерис-

тики шуму джерел, що дозволить раціоналізувати конструкцію захисних поверхонь. Визначено середні значення коефіцієнтів звукопоглинання типових підстилаючих поверхонь, які дають внесок у зниження акустичного навантаження на середовище.

3. Надано розрахунки ефективності акустичних екранів достатньої точності для застосування у інженерній практиці. Враховано внесок у проникнення шуму дифракції звукових хвиль на кромках

акустичних екранів стандартної форми.

4. Надано практичні рекомендації щодо підвищення ефективності акустичних екранів за рахунок їх форми та з урахуванням частотних характеристик шуму.

5. Наголошено, що у реальних умовах експлуатації зниження рівнів звукового тиску за рахунок застосування акустичних екранів не перевищує 10–12 дБ.

СПИСОК ЛІТЕРАТУРИ

1. ДСН 3.3.6.037-99 Санітарні норми виробничого шуму, ультразвуку та інфразвуку. Постанова Міністерство охорони здоров'я від 01.12.1999 № 37.
2. Про затвердження Державних санітарних правил планування та забудови населених пунктів [Чинний від 2019-03-07]: наказ М-ва охорони здоров'я України від 19.06.1996 № 173.
3. E. Jayamani, M.K.B. Bakri, Lignocellulosic fibres reinforced polymer composites for acoustical applications. In: Kalia, S. (ed.) Lignocellulosic Composite Materials, pp. 415-444. Springer, Berlin (2018)
4. S. Secchi, et al. Experimental and environmental analysis of new sound-absorbing and insulating elements in recycled cardboard. J. Build. Eng. 5, (2016), 1-12
5. V. Glyva, O. Zaporozhets, L. Levchenko, N. Burdeina, V. Nazarenko. Methodological Foundations Protective Structures Development For Shielding Electromagnetic And Acoustic Fields. Strength of Materials and Theory of Structures 110, 2023, PP. 245-255. <https://doi.org/10.32347/2410-2547.2023.110.245-255>
6. E. Taban, A. Tajpoor, M. Faridan et al. Acoustic Absorption Characterization and Prediction of Natural Coir Fibers. Acoust Aust 47, 67–77 (2019). <https://doi.org/10.1007/s40857-019-00151-8>
7. Hsiao Mun Lee, Zhaomeng Wang, Kian Meng Lim, Heow Pueh Lee. Investigation of the effects of sample size on sound absorption performance of noise barrier. Applied Acoustics, Volume 157, 2020, 106989, <https://doi.org/10.1016/j.apacoust.2019.07.037>
8. Salehian, S., & Mankbadi, R. (2020). Jet Noise in Airframe Integration and Shielding. Applied Sciences, 10(2), 511. <https://doi.org/10.3390/app10020511>
9. K. Attenborough. Acoustical impedance models for outdoor ground surfaces. Journal of Sound and Vibration, V. 99, I. 4, 1985, p. 521-544, [https://doi.org/10.1016/0022-460X\(85\)90538-3](https://doi.org/10.1016/0022-460X(85)90538-3)
10. Ulrich J. Kurze. Noise reduction by barriers. J. Acoust. Soc. Am. 55, 504–518 (1974), <https://doi.org/10.1121/1.1914528>

Received (Надійшла) 23.02.2025

Accepted for publication (Прийнята до друку) 30.04.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Глива Валентин Анатолійович – доктор технічних наук, професор, завідувач кафедри фізики, Київський національний університет будівництва та архітектури, Київ, Україна;

Valenty Glyva – Doctor of Technical Sciences, Professor, Head of the Department of Physics, Kyiv National University of Construction and Architecture, Kyiv, Ukraine;

e-mail: hlyva.va@knuba.edu.ua, ORCID Author ID: <https://orcid.org/0000-0003-1257-3351>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57210185162>.

Кашлев Михайло Сергійович – аспірант кафедри технологій захисту навколишнього середовища та охорони праці, Київський національний університет будівництва і архітектури, Київ, Україна;

Mykhailo Kashlev – PhD student at the Department of Environmental Protection Technologies and Labour Protection, Kyiv National University of Construction and Architecture, Kyiv, Ukraine;

e-mail: kashlev_ms-2022@knuba.edu.ua, ORCID Author ID: <https://orcid.org/0009-0004-6240-4630>.

Shielding the noise of construction equipment

Valenty Glyva, Mykhailo Kashlev

Abstract. Experimental studies of the noise characteristics of construction equipment have been carried out. It has been found that in most cases, the excess of the maximum permissible sound pressure values is no more than 20–25 dB. This makes it possible to use acoustic screens to normalise noise levels at workplaces and territories. The advantage of screens is their ease of manufacture and low cost. This makes it possible to use shielding structures of large areas. The contribution to noise absorption of the most typical underlying surfaces was determined. These data allow us to make adjustments to the required efficiency of acoustic screens. The amplitude-frequency characteristics of typical noise sources at construction sites are investigated. Calculations are presented to predict the effectiveness of acoustic screens depending on the conditions of their use. The presence of diffraction phenomena at the edges of the screens for large values of diffraction angles is taken into account, taking into account the location of the screen on the surface and above the surface. The calculation of the value of the sound wave diffraction coefficient is presented, taking into account the distance from the sound source to the screen and its edges, as well as the distance from the screen edges to the sound reduction zone. The physical characteristics of the material from which the acoustic screen is made are taken into account. Practical recommendations for improving the effectiveness of acoustic screens are given, taking into account the location, shape of the screen and amplitude-frequency characteristics of the noise to be screened.

Keywords: construction equipment, noise, acoustic screen, sound absorption coefficient.

Л. О. Левченко¹, Н. М. Аушева¹, Т. М. Ткаченко²

¹ Національний технічний університет України «КПІ імені Ігоря Сікорського», Київ, Україна

² Київський національний університет будівництва і архітектури, Київ, Україна

МОДЕЛЮВАННЯ ЗМІНИ ТЕМПЕРАТУРИ НА ПОВЕРХНЯХ ТЕПЛОІЗОЛЯЦІЙНИХ ПОКРИТТІВ

Анотація. На основі співвідношень теплопереносу встановлено найбільш прийнятні функції для визначення теплоізоляційних властивостей матеріалів і конструкцій. З використанням пакету Comsol розроблено моделі змін температури зовнішньої поверхні захисного матеріалу в залежності від часу переносу тепла. Встановлено, що на початковій стадії теплопереносу температура зовнішньої поверхні змінюється немонотонно. Тому при проєктуванні захисної конструкції слід приділяти увагу шарам матеріалу, який безпосередньо контактує з захисним шаром. Показано, що застосування у якості термоізолюючого матеріалу склотканини забезпечує практично однаковий рівень захисту людини при мінімальних (120 Вт) і максимально допустимих (400 Вт) фізичних навантаженнях. Отримано тривимірну модель, яка візуалізує залежність температури від часу теплопередачі по глибині. Це дозволяє раціоналізувати не тільки товщину захисного матеріалу, а й обрати найбільш прийнятні матеріали з точки зору їх термофізичних властивостей. Швидке отримання графічного матеріалу дозволяє методом перебору визначити найбільш прийнятні матеріали і конструкції у залежності від умов зовнішнього середовища та призначення засобу захисту.

Ключові слова: моделювання, термоізоляція, інфрачервоне випромінювання.

Вступ

Проєктування теплоізоляційних покриттів є однією з актуальних задач, вирішення яких спрямовано на підвищення енергоефективності будівель, зниження втрат у системах транспортування теплоносіїв тощо. Одним з аспектів цієї проблематики є створення спеціального одягу для захисту від впливу низьких та високих температур та розроблення інфрачервоного камуфляжу. Розроблення одягу є найбільш складною задачею. Це обумовлено необхідністю не тільки забезпечення високої ефективності, а й ергономічності конструкції для зручності використання – гнучкості, малої ваги, товщини тощо. Для обрання оптимальної конструкції необхідно виконати великі обсяги експериментальних робіт, що потребує створення відповідних умов, передбачених стандартом ДСТУ EN ISO 15831:2007. Такі роботи дуже витратні й потребують значного часу на їх реалізацію. Тому на першому етапі проєктування доцільно здійснити моделювання поширення теплового потоку крізь захисні шари й визначити зміни температури на зовнішній поверхні захисної конструкції. Це спростить процес вибору матеріалів для конструювання захисного одягу, його товщину та часові межі ефективності захисної конструкції. Передумовою такого моделювання є наявність коректних математичних функцій, які відповідають поставленій задачі й ефективно описують процеси теплопередачі крізь шарувату структуру.

Сучасний стан питання. Вимоги до захисного одягу регламентуються національними стандартами України, ідентичними загальноєвропейським [1, 2], тому усі проєктні роботи повинні відповідати їх вимогам. Але більшість розробок у цій галузі стосуються захисту від високих температур [3, 4]. Такі покриття мають великі коефіцієнти відбиття зовнішнього інфрачервоного випромінювання. Але їх застосування для запобігання витоку тепла неприйнятне через можливість порушення теплового балансу

людини. Розрахункове прогнозування захисних матеріалів на основі композитів складне [5, 6]. Це пояснюється тим, що термофізичні властивості таких матеріалів не є довідковими величинами, а вони обов'язково закладаються у розрахунки та відповідні моделі. Тому для прогнозування ефективності конструкцій, виготовлених з цих матеріалів, необхідне проведення комплексу експериментальних досліджень. При цьому незначна зміна складу композиту може критично змінювати його потрібні властивості. У роботах [7, 8] досліджено ефективність блокування інфрачервоних випромінювань стандартними матеріалами (скловолокна, нітрид бору, окис хрому) й показано їх прийнятні властивості щодо запобігання витоку тепла нагрітого об'єкта. Застосування таких матеріалів спрощує розроблення моделей для прогнозування захисних властивостей конструкцій. У роботі [9] запропоновано обґрунтування застосування скловолокна для розсіювання інфрачервоного випромінювання. Наведені дані переконливі й цей матеріал можна розглядати як базовий для проєктування засобів захисту. Задача полягає у визначенні залежностей температури зовнішніх поверхонь захисних матеріалів від часу теплопередачі та товщини захисного шару.

Мета роботи – розробити практично значущі моделі змін температури зовнішньої поверхні захисного матеріалу у залежності від часу переносу тепла, товщини матеріалу за різних енергетичних характеристик джерела тепла.

Виклад основного матеріалу

У загальному випадку усі процеси теплопереносу описуються рівнянням Фур'є-Кірхгофа. Варіанти його застосування залежать від того, який механізм теплопереносу є переважний у конкретному процесі – теплопровідність, конвекція, променевий теплообмін.

Конвекція, променевий теплообмін. За безпосереднього контакту захисного матеріалу з джерелом тепла процес теплопередачі описується рівнянням:

$$\frac{\partial T}{\partial t} = \lambda \frac{\partial^2 T}{\partial x^2},$$

де T – температура; t – час; λ – коефіцієнт теплопровідності у напрямку x .

Моделювання процесу теплопереносу здійснювалося для джерела тепла, яким є людина. Параметри джерела вважалися такими, що відповідають двом фізичним навантаженням людини – 120 Вт і 400 Вт (мінімальне та максимально допустиме), регламентованих стандартом ДСТУ EN ISO 15831:2007 [10].

Визначалася температура зовнішнього шару захисного матеріалу. Значення конвективного теплоо-

бміну поверхні приймалося у 2 Вт/м²·К, що відповідає майже повній відсутності руху повітря.

Моделювання здійснювалося за різницевою схемою із використанням пакету Comsol. На рис. 1 наведено зміни температури зовнішньої поверхні захисного матеріалу упродовж двох годин. Як видно з рис. 1 температура поверхні монотонно знижується до стану термодинамічної рівноваги з оточуючим середовищем. При цьому на початковій стадії (до 0,2 год.) матеріал частково нагрівається. Для більш детального визначення змін температури, що обумовлено переносом тепла у внутрішніх шарах захисної конструкції, цей часовий інтервал розглянуто на рис. 2.

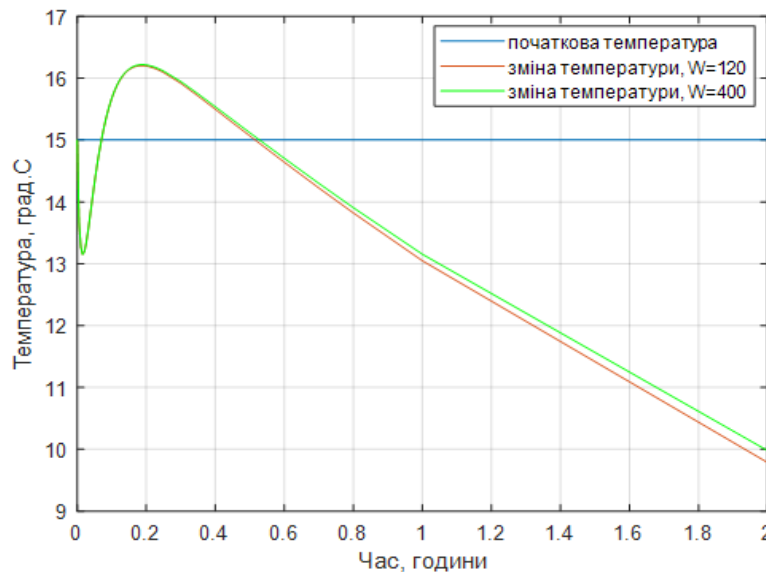


Рис. 1. Зміна температури зовнішньої поверхні захисного шару упродовж двох годин за різних потоків теплової енергії

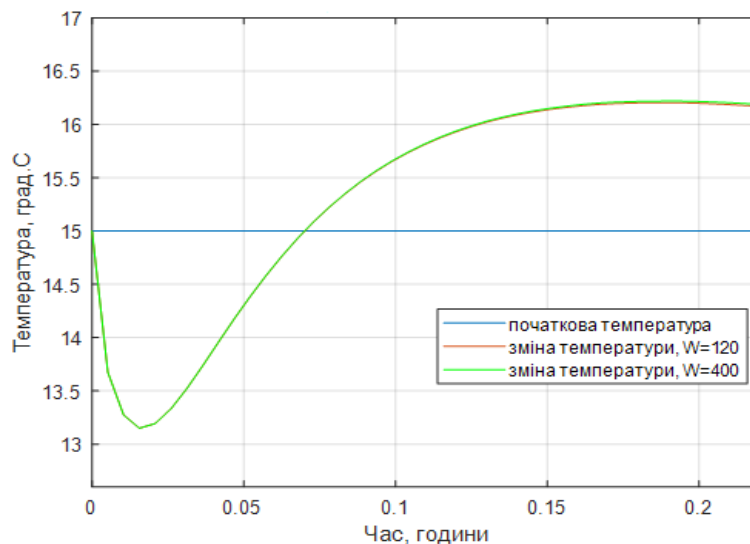


Рис. 2. Зміна температури зовнішньої поверхні захисного шару упродовж 0,2 години

Наведений результат доцільно використовувати у процесі визначення матеріалів для виготовлення внутрішніх шарів захисної конструкції.

Отриманий результат свідчить, що для обраного захисного матеріалу на основі скловолокна ступінь

термоізоляції практично не залежить від тепловиділення джерела. Слід враховувати, що у будь-якому випадку існує градієнт температури у товщі захисного матеріалу або кількох шарів, наприклад, одягу та захисного шару.

Зміни температури захисної конструкції у просторі і часі наведено на рис. 3.

Наведений графічний матеріал дозволяє оцінити термоізоляційні властивості як одношарового, так і кількшарового матеріалу у залежності від часових і просторових параметрів.

У дану модель закладено коефіцієнти теплопровідності, питомої теплоємності та товщини стандартних текстильних матеріалів та склотканини.

У реальних умовах експлуатації захисних конструкцій як призначення, так і параметри зовнішнього середовища можуть суттєво відрізнятися. Тому у розрахункові співвідношення вносяться інші параметри. Швидке отримання графічного матеріалу у залежності від закладених параметрів дозволяє отримати бажаний результат методом перебору. Це значно спрощує процес проєктування захисної конструкції й пришвидшує його реалізацію.

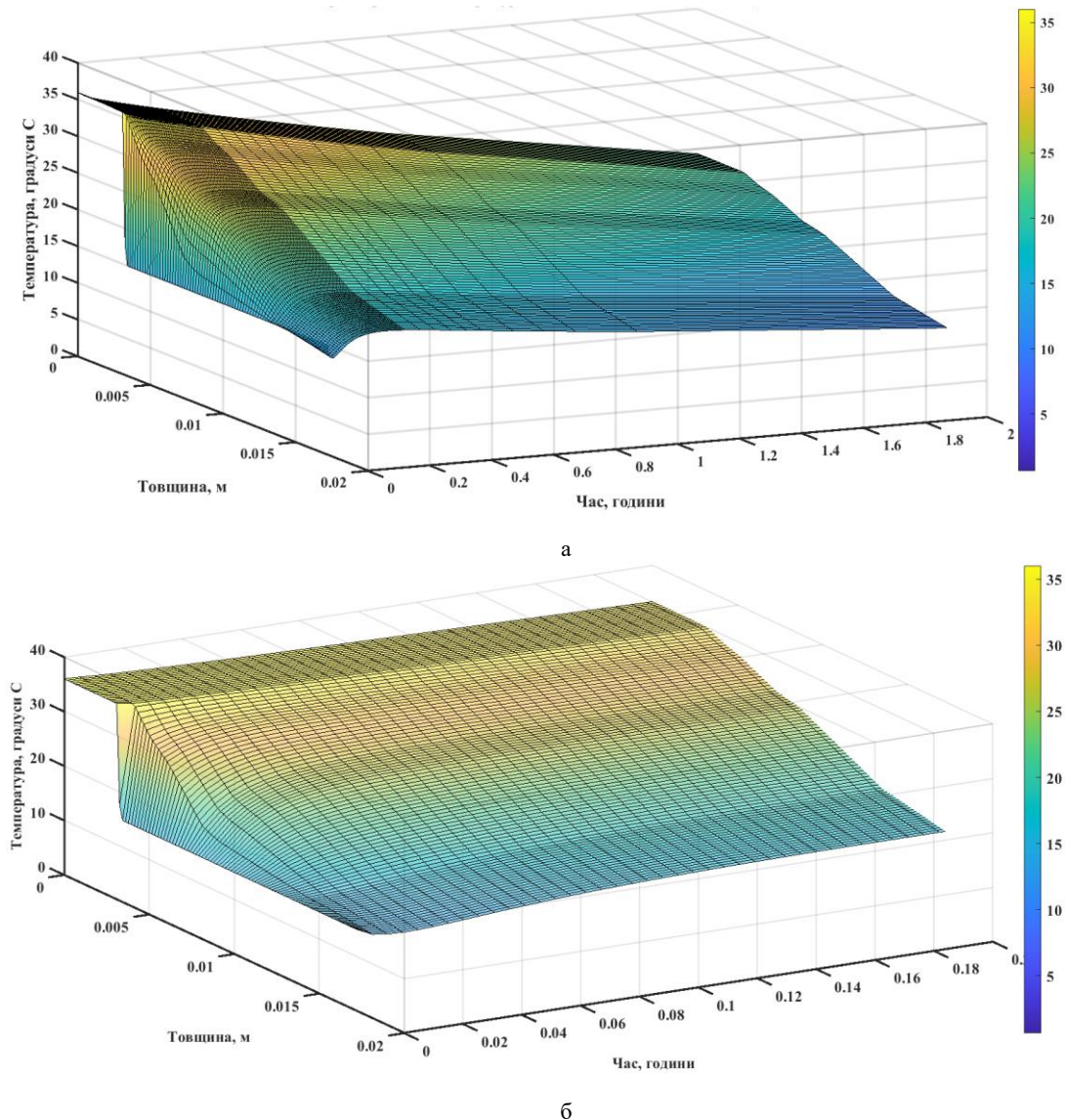


Рис. 3. Зміна температури по товщині захисної конструкції упродовж часових інтервалів: а – упродовж двох годин; б – упродовж 0,2 годин.

Висновки

1. На основі співвідношень теплопередачі крізь суцільну та шарувату структуру розроблено модель зміни температури зовнішньої поверхні термоізоляційної конструкції. Джерелом теплової енергії слугувало тіло людини за мінімального (120 Вт) та максимально допустимого (400 Вт) фізичного навантаження згідно міжнародним нормативам.

2. Проведено моделювання зміни температури зовнішньої поверхні захисного шару упродовж 2,0 та 0,2 годин. Встановлено складну залежність темпе-

ратури на початкових стадіях переносу тепла. Це необхідно враховувати при виборі матеріалів для внутрішніх шарів, які безпосередньо контактують із захисним шаром.

3. Отримано модель залежності температури від часу передачі тепла й глибини шару захисного матеріалу. Це дозволяє у процесі проєктування захисної конструкції отримати бажаний результат закладаючи у розрахункову схему термофізичні параметри різних матеріалів. Швидкий обрахунок різних варіантів дозволить обрати найбільш прийнятний варіант методом перебору.

СПИСОК ЛІТЕРАТУРИ

1. ДСТУ EN 342:2017 (EN 342:2004; AC:2008, IDT). Одяг захисний. Комплекти та предмети одягу для захисту від холоду. (прийнятий методом перекладу, ідентичний до EN 342:2004 (версія en) «Protective clothing – Ensembles and garments for protection against cold» зі зміною AC:2008 (версія en).
2. ДСТУ EN 14058:2019 (EN 14058:2017). Одяг захисний. Предмети одягу для захисту від холодного навколишнього середовища. [Чинний від 2019-11-19] Вид. офіц. Київ, 2019.
3. Peng, L., Jiang, S., Guo, R. et al. IR protection property and color performance of TiO₂/Cu/TiO₂ coated polyester fabrics. *J Mater Sci: Mater Electron* 29, 16188–16198 (2018). <https://doi.org/10.1007/s10854-018-9708-6>.
4. Jiang, S., Miao, D., Xu, J. et al. Preparation and characterization of shielding textiles to prevent infrared penetration with Ag thin films. *J Mater Sci: Mater Electron* 28, 3542–3547 (2017). <https://doi.org/10.1007/s10854-016-5955-6>.
5. Marco Caniato, Arianna Marzi, Sandra Monteiro da Silva, Andrea Gasparella. A review of the thermal and acoustic properties of materials for timber building construction. *Journal of Building Engineering*. Volume 43, 2021, 103066, <https://doi.org/10.1016/j.jobbe.2021.103066>.
6. Azim Rasulovich Abdullaev, ., Xayotbek Mansurjon O'g'li Rafiqov, ., & Isroiljonova Nizomjon Qizi Zulxumor, . (2021). A Review On: Analysis Of The Properties Of Thermal Insulation Materials. *The American Journal of Interdisciplinary Innovations and Research*, 3(05), 27–38. <https://doi.org/10.37547/tajir/Volume03Issue05-06>.
7. Glyva, V., Krasnianskyi, G., Dovhanovskyi, M., & Krasnianskyi, T. (2024). Розроблення і дослідження захисних властивостей матеріалів для блокування електромагнітних випромінювань інфрачервоного діапазону. *Системи управління, навігації та зв'язку. Збірник наукових праць*, 3(77), 203-205. <https://doi.org/https://doi.org/10.26906/SUNZ.2024.3.203>.
8. Burdeina, N., Levchenko, L., Korduba, I., Shamanskyi, S., Biruk, Y., Dovhanovskyi, M., Zozulya, S., Klymchuk, A., Nikolaiev, K., & Osadchyi, D. (2024). Applying heterogeneous building materials for the protection of people against electromagnetic radiation. *Eastern-European Journal of Enterprise Technologies*, 5(10 (131)), 45–52. <https://doi.org/10.15587/1729-4061.2024.313629>.
9. G. Krasnianskyi, I. Aznauryan, M. Dovhanovskyi, O. Besarab, V. Okhrysko. Evaluation of the effectiveness of infrared radiation shielding with glass fiber. *REPORTS of the 5th INTERNATIONAL SCIENTIFIC AND PRACTICAL CONFERENCE ENERGY. RESOURCES. ECOLOGY*, November 27-29, 2024, KYIV, p. 102.
10. ДСТУ EN ISO 15831:2007. Одяг захисний. Фізіологічне оцінювання вимірюванням теплоізоляції на тепловому манекені. (прийнятий методом перекладу, ідентичний EN ISO 15831:2004 Clothing — Physiological effects — Measurement of thermal insulation by means of a thermal manikin. Набув чинності з 2009-01-01 наказом Держспоживстандарту України від 6 липня 2007 р. № 147).

Received (Надійшла) 25.03.2025

Accepted for publication (Прийнята до друку) 11.06.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Левченко Лариса Олексіївна – доктор технічних наук, професор, професор кафедри цифрових технологій в енергетиці, Національний технічний університет України «Київський політехнічний інститут імені І. Сікорського, Київ, Україна;
Larysa Levchenko – Doctor of Technical Sciences, Professor, Professor of Department Digital Technologies in Energy, National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute», Kyiv, Ukraine;
e-mail: larlevch@ukr.net; ORCID Author ID: <http://orcid.org/0000-0002-7227-9472>;
Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57194577942>.

Аушева Наталія Миколаївна – доктор технічних наук, професор, завідувачка кафедри цифрових технологій в енергетиці, Національний технічний університет України «Київський політехнічний інститут ім. І. Сікорського, Київ, Україна;
Nataliia Ausheva – Doctor of Technical Sciences, Professor, Head of Department Digital Technologies in Energy, National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute», Kyiv, Ukraine;
e-mail: nataausheva@gmail.com; ORCID Author ID: <http://orcid.org/0000-0003-0816-2971>;
Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57210707106>.

Ткаченко Тетяна Миколаївна – доктор технічних наук, професор, завідувачка кафедри технологій захисту навколишнього середовища та охорони праці, Київський національний університет будівництва і архітектури;
Tetiana Tkachenko – Doctor of Technical Sciences, Professor, Head of the Department of Environmental Protection Technologies and Labour Safety, Kyiv National University of Construction and Architecture,
e-mail: tkachenkoknuba@gmail.com; ORCID Author ID: <http://orcid.org/0000-0003-2105-5951>;
Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57203135001>.

Modeling temperature changes on the surfaces of thermal insulation coatings

Larysa Levchenko, Nataliia Ausheva, Tetiana Tkachenko

Abstract. Based on the heat transfer ratios, the most acceptable functions for determining the thermal insulation properties of materials and structures were established. Using the Comsol package, models of changes in the temperature of the outer surface of the protective material depending on the heat transfer time were developed. It was established that at the initial stage of heat transfer, the temperature of the outer surface changes non-monotonically. Therefore, when designing a protective structure, attention should be paid to the layers of material that are in direct contact with the protective layer. It is shown that the use of fiberglass as a thermal insulation material provides almost the same level of human protection at minimum (120 W) and maximum permissible (400 W) physical loads. A three-dimensional model was obtained that visualizes the dependence of temperature on the time of heat transfer in depth. This allows us to rationalize not only the thickness of the protective material, but also to select the most acceptable materials in terms of their thermophysical properties. Rapid receipt of graphic material allows us to determine the most acceptable materials and structures by the method of brute force, depending on the environmental conditions and the purpose of the protective device.

Keywords: modeling, thermal insulation, infrared radiation.

О. І. Токарський¹, Л. А. Зозуля², Я. О. Галонько³, О. О. Москаленко⁴

¹ Державна служба України з надзвичайних ситуацій у Закарпатській області, Мукачеве, Україна

² Державний університет «Київський авіаційний інститут», Київ, Україна

³ Київський національний університет будівництва і архітектури, Київ, Україна

⁴ Науково-дослідний інститут воєнної розвідки, Київ, Україна

ЗАСОБИ НОРМАЛІЗАЦІЇ ФІЗИЧНИХ ФАКТОРІВ СЕРЕДОВИЩА ПРИМІЩЕНЬ СПЕЦІАЛЬНОГО ПРИЗНАЧЕННЯ

Анотація. Визначено засади нормалізації рівнів фізичних факторів середовища приміщень спеціального призначення. До таких приміщень належать сховища цивільного захисту, щити керування енергетичними об'єктами тощо. Показано, що однією з головних задач є оздоровлення середовища з використанням світлодіодних джерел ультрафіолетового випромінювання. Перевагою таких систем є можливість застосування у присутності людей. За відстані від двох метрів до людей впродовж восьми годин не досягається гранично допустима доза ультрафіолетового випромінювання (30 Дж/м^2). При застосуванні на кожні 25 м^3 об'єму приміщення випромінювача напруженістю 5 Вт досягається високий ступінь оздоровлення середовища. Кількість колоній хвороботворних грибів знижується у $1,8\text{--}2,8$ рази. Загальне мікробне число – у $1,7\text{--}2,2$ рази. Цей процес досить тривалий, тому доцільно не вимикати джерело випромінювання навіть за відсутності людей. Мала потужність випромінювачів не впливає суттєво на енерговитрати. Ультрафіолетове випромінювання світлодіодів іонізує повітря приміщень. Одна світлодіодна система упродовж трьох годин підвищує концентрацію іонів повністю деіонізованого повітря до $1100\text{--}1200 \text{ см}^{-3}$. При цьому нормативні значення концентрації (400 см^{-3} позитивних і 600 см^{-3} негативних) досягається упродовж $20\text{--}30$ хвилин, що краще за відомі пристрої іонізації повітря. Підвищені концентрації іонів сприяють очищенню повітря від дрібнодисперсного пилу та аерозолів. Слід враховувати, що у системах примусової вентиляції і охолодження відбувається значна деіонізація повітря. Для забезпечення електромагнітної безпеки людей стіни приміщень доцільно покривати рідким композиційним матеріалом. Шар завтовшки $1,0\text{--}1,5 \text{ мм}$ забезпечує коефіцієнти екранування електромагнітних випромінювань ультрависоких частот $3,8\text{--}8,0$. При цьому коефіцієнти відбиття складають $0,25\text{--}0,32$, що знижує електромагнітний фон за рахунок перевідбиття випромінювань. Покриття захисною сумішшю поверхонь напівпровідникових та підвальних приміщень сприяє уникненню потрапляння радону у середовище перебування людей.

Ключові слова: сховища, дебактеризація, іонізація, електромагнітна безпека.

Вступ

Нормативні значення фізичних факторів – мікрокліматичних показників, концентрацій іонів у повітрі, рівнів електромагнітних полів, радіаційний фон є важливими умовами безпечного перебування людей. У більшості виробничих та житлових приміщень такі показники забезпечуються конструктивними рішеннями будівель та системами їх життєзабезпечення. Але у сучасних умовах в Україні існує потреба у великій кількості сховищ цивільного захисту. При цьому у стаціонарних сховищах параметри середовища передбачені на стадіях їх проєктування, а пристосовані сховища у більшості не відповідають жодним вимогам. В той же час у таких приміщеннях може перебувати багато людей з великою щільністю розташування упродовж тривалого часу. Зазвичай у пристосованих сховищах недостатність ефективності вентиляції призводить до значної деіонізації повітря, підвищення відносної вологості, накопичення діоксиду вуглецю тощо. Такі приміщення розташовані у підвалах та напівпідвалах, у яких може накопичуватись радон, виявлення якого дуже складне через відсутність відповідних детекторів. У пристосованих приміщеннях розміщують точки доступу бездротового зв'язку, що може призводити до підвищення електромагнітного фону. Роботи з моніторингу значень фізичних факторів потребують наявності відповідної апаратури, витрат часу та значних коштів. Тому доцільно визначити систему заходів й перелік засобів, які з великою вірогідністю забезпечать нормальні параметри середовища і підтримання їх на

нормативному рівні упродовж тривалого часу.

Огляд літературних джерел. Конструктивні рішення й стан середовища у приміщеннях сховищ цивільного захисту регламентуються національними нормами і стандартами [1, 2]. Але ці документи стосуються стаціонарних об'єктів. Крім того у них не містяться конкретні заходи нормалізації параметрів середовища. Існують нормативи щодо мікрокліматичних параметрів середовища для різних пір року та концентрації іонів [3–5]. Але ці вимоги також закладаються у конструктивні рішення та стосуються виробничих умов. При цьому регламентована площа, яка припадає на одну людину. У пристосованих сховищах кількість та щільність перебування людей не передбачувана. Існує низка досліджень, які стосуються динаміки кількісних значень фізичних факторів у нестандартних умовах. У роботі [6] досліджено зв'язок концентрацій іонів, відносної вологості повітря та напруженостей електростатичних полів у приміщеннях із застосуванням ультразвукового зволожувача та стандартного пристрою охолодження повітря. Встановлені зв'язки можна використати для розроблення організаційних та нічних заходів нормалізації фізичних факторів середовища приміщень спеціального призначення. Але застосування зволожувача не завжди виправдано у підвальних приміщеннях, де відносні вологості достатньо великі.

Крім того підвищенню вологості сприяє велика кількість людей. У роботі [7] запропоновано підтримання концентрацій іонів спеціальним ультразвуковим пристроєм. Але застосовувати окремі пристрої

для нормалізації одного фактору не завжди раціонально. У дослідженнях [8–10] запропоновано застосування світлодіодних систем ультрафіолетового випромінювання для знезараження середовища та іонізації повітря в умовах епідемії коронавірусу. Такий підхід може бути плідним для приміщень спеціального призначення – як сховищ цивільного захисту, так і інших приміщень – пультів управління об'єктів критичної інфраструктури. В усіх приміщеннях, що розглядаються, не можна виключати наявність підвищеного електромагнітного фону. У дослідженнях [11, 12] наведено результати випробувань рідких матеріалів з малими коефіцієнтами відбиття електромагнітних хвиль. Доцільно розглянути можливості їх застосування у сховищах цивільного захисту та інших об'єктах критичної інфраструктури. Але усі розглянуті дослідження стосуються окремих фізичних факторів, нормалізація яких може призвести до змін значень інших критичних факторів. Тому доцільно застосувати комплексний підхід до цієї проблематики.

Мета роботи – обґрунтувати й розробити комплекс організаційно-технічних заходів нормалізації рівнів фізичних факторів у приміщеннях спеціального призначення.

Викладення основного матеріалу

Усі приміщення спеціального призначення умовно можна розділити на дві категорії. Перша – сховища цивільного захисту, друга – приміщення, у яких здійснюється керування об'єктами критичної інфраструктури. Найбільш актуальною в Україні є задача забезпечення прийнятних умов перебування людей у пристосованих сховищах. Складність задачі полягає у непередбачуваній кількості людей, які одночасно перебувають у сховищі, що впливає на параметри середовища. При цьому частина людей може мати хворобливий стан. Тому головною задачею є забезпечення оздоровлення середовища. Зазвичай у закладах охорони здоров'я це досягається застосуванням ртутних ламп ультрафіолетового випромінювання. Але їх використання заборонено у присутності людей. Тому доцільно використовувати світлодіодні системи знезараження повітря і поверхонь. Було проведено експериментальні дослідження зниження

кількості колоній бактерій та грибів під впливом ультрафіолетового випромінювання з довжиною хвилі 280 нм, джерелом якого були світлодіоди. Попередніми дослідженнями [8] встановлено, що за будь-яких умов доза ультрафіолетового випромінювання світлодіодів не досягає гранично допустимого значення 30 Дж/м² згідно ДСТУ EN 62471:2017. Єдиною умовою є мінімальна відстань джерел випромінювань до людей – 2 м. Для дебактеризації повітря використовувалися світлодіодні системи ультрафіолетового випромінювання LED NC T5-5W-275 NM вітчизняного виробництва потужністю 5 Вт. Для виявлення кількості мікроорганізмів застосовувалися тест-системи R-Biopharm AG. Визначення ступеня знезараження середовища здійснювалося за показником N – кількість колоній утворюючих одиниць (КОУ) на 1 дм² поверхні. Вимірювання здійснювалися у двох точках приміщення – найближчій та найвіддаленішій від входу до приміщення.

Результати вимірювань змін кількості колоній пліснявих грибів наведено на рис 1. Слід зазначити, що вимірювання кількості колоній грибів та бактерій дещо неоднозначні. Існують розбіжності змін кількості колоній під впливом ультрафіолетового випромінювання. Швидше за все це обумовлене тим, що гриби певним чином зменшують кількість бактерій і в приміщеннях з великою кількістю грибів кількість колоній бактерій нижча. Крім того, не дивлячись на однотипність приміщень існують певні відмінності у режимах природної вентиляції, наприклад інфільтрація повітря крізь вікна. Також було визначено зміну загального мікробного числа на поверхнях за КОУ/дм² (рис. 2).

Дані рис. 1, 2 свідчать, що знезараження середовища відбувається досить повільно. Це відповідає усім доступним даним щодо впливу оздоровчих факторів на стан середовища. Навіть застосування потужних ртутних ламп вимагає певного часу впливу. Отримані дані відповідали умовам, за яких світлодіодна система потужністю 5 Вт обслуговувала 25 м³ об'єму приміщення. За таких потужностей, навіть виходячи з міркувань електрозбереження, світильники можна не вимикати, що забезпечує неперервне оздоровлення середовища й його готовність до присутності людей.

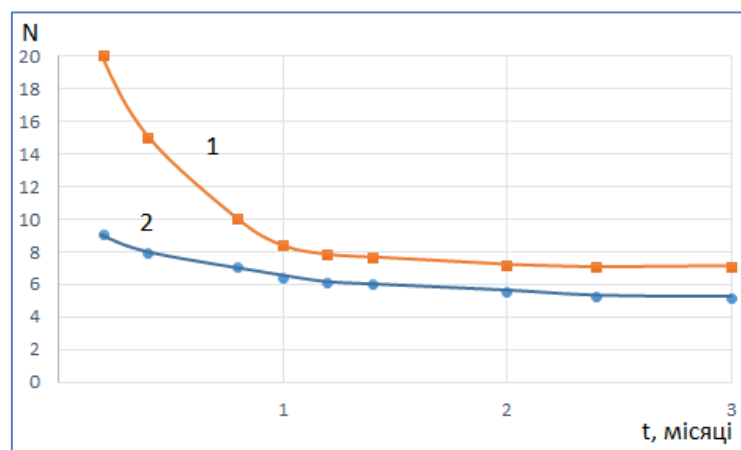


Рис. 1. Зміна кількості колоній грибів на поверхнях приміщення з часом ультрафіолетового опромінення: 1 – на максимальній відстані від входу до приміщення; 2 – на мінімальній відстані.

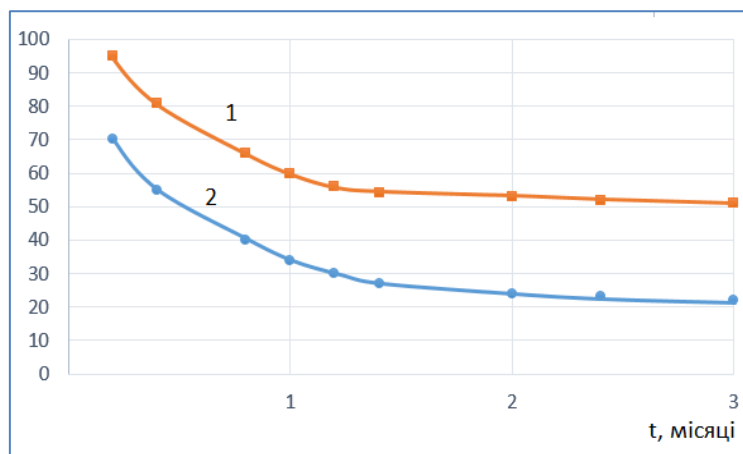


Рис. 2. Зміна загального мікробного числа на поверхнях приміщень з часом ультрафіолетового опромінення: 1 – на максимальній відстані від входу до приміщення; 2 – на мінімальній відстані.

Відомо, що ультрафіолетове випромінювання викликає іонізацію газів, що обґрунтовано експериментально та теоретично [8]. Тому було визначено зміни концентрацій іонів у повітрі приміщень під впливом ультрафіолетового випромінювання світлодіодних систем. Дослідження здійснювалися у тих самих приміщеннях із застосуванням одного світильника потужністю 5 Вт на 25 м³ об'єму приміщення.

Фонові концентрації негативних іонів склали 160–190 см⁻³, позитивних – 140–180 см⁻³. На рис. 3 наведено зміни середніх концентрацій іонів обох знаків у повітрі приміщення у залежності від часу роботи світлодіодних систем ультрафіолетового випромінювання.

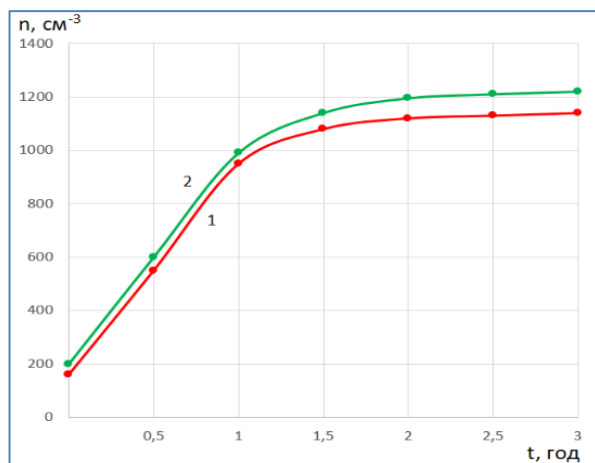


Рис. 3. Зміна концентрації іонів у повітрі приміщення під впливом ультрафіолетового опромінення: 1 – концентрації позитивних іонів; 2 – концентрації негативних іонів

Як видно з рис. 3, концентрації іонів обох знаків швидко підвищуються до значень, вищих за мінімально допустимих (негативних – 600 см⁻³, позитивних

– 400 см⁻³). Це пояснюється тим, що процес іонізації молекул повітря відбувається відразу по усьому об'єму через випадкові взаємодії квантів випромінювання з молекулами газу. Такий спосіб іонізації повітря набагато ефективніший за відомі (високовольтні розряди, ультразвукові випромінювачі) через відсутність необхідності розповсюдження іонів від джерела іонізації спрямованим рухом повітря.

Слід зазначити, що підвищена концентрація іонів сприяє очищенню повітря від дрібнодисперсного пилу та аерозолів. Осідаючи на ці частинки, іони надають їм заряд, що спричиняє дрейф завислих частинок до поверхонь, де вони осідають [13].

У процесі організації тимчасових сховищ цивільного захисту слід враховувати проблеми, пов'язані з вентиляцією та кондиціонуванням повітря.

У металевих повітропроводах, виготовлених з оцинкованого заліза, відбувається деіонізація повітря за обома полярностями. Дослідження свідчать, що кількісні дані щодо цього показника неоднозначні. У реальній припливно-витяжній системі зниження концентрацій негативних іонів складає 40–50 см⁻³ на 1 м погонний повітропроводу. Позитивних – 80–100 см⁻³. У лабораторних умовах отримані дані: негативних 80–90 см⁻³ на погонний метр, позитивних 200–250 см⁻³ на погонний метр. Така неоднозначність зумовлена кількома чинниками – вихідною концентрацією іонів, діаметром повітропровода, наявністю пилу на поверхнях тощо. Тому у реальних умовах бажано підвищувати концентрації іонів за допомогою штучної іонізації повітря. Виходячи з сучасних поглядів, оптимальною вважається концентрація іонів 2000–3000 см⁻³ обох знаків. До того ж слід враховувати деіонізацію повітря split-системами. Ці процеси також неоднозначні. У приміщеннях одного об'єму зниження концентрацій іонів охолоджуючими пристроями різних марок відрізняється (табл. 1).

Таблиця 1 – Деіонізація повітря системами охолодження повітря

Пристрій	Вихідні концентрації		Концентрації після трьох годин роботи	
	n^-	n^+	n^-	n^+
1	620–650	730–750	420–430	480–500
2	620–640	740–750	590–610	390–400
3	470–490	680–700	110–120	100–110

Наведені дані свідчать, що доцільне підвищення концентрації іонів штучними методами. Слід враховувати, що за присутності великої кількості людей внаслідок дихання утворюються аерозолі, які є чинником іонізації повітря.

Проблема нормативних показників якості повітря існує й у інших приміщеннях спеціального призначення. У приміщеннях щитів керування енергетичними об'єктами спостерігаються значні рівні електромагнітних полів, що призводить до осідання іонів й деіонізації повітря принаймні за однієї полярності. Не дивлячись на те, що системи примусової вентиляції надійно заземлені, деіонізація повітря у них на таких об'єктах підвищена.

Окремого розгляду заслуговує проблематика електромагнітної безпеки людей у сховищах та інших приміщеннях спеціального призначення.

Наявність точок доступу мобільного зв'язку у повністю замкнених об'єктах викликає підвищення електромагнітного фону внаслідок відбиття та перевідбиття електромагнітних хвиль. Для уникнення такого явища стіни приміщень доцільно вкривати радіопоглинальними матеріалами з малими коефіцієнтами відбиття. Матеріали на основі фарб зручно наносити на поверхні будь-якої конфігурації. Композиційні матеріали з додаванням магнетиту та графіту мають прийнятну вартість та достатні коефіцієнти загального екранування за малих коефіцієнтів відбиття. У табл. 2 наведено показники композиції на основі геополімерної фарби з додаванням 45 % та 60 % (ваг.) магнетиту. В обох випадках кількість лускатого графіту складала 5 % (ваг.).

Таблиця 2 – Коефіцієнти екранування K_s та коефіцієнти відбиття K_r композиційного матеріалу з вмістом магнетиту 45 та 60 % та графіту 5 %

Матеріал	K_s	K_r
45%	3,8–4,2	0,25–0,27
60%	7,7–8,0	0,30–0,32

Слід зазначити, що наведені дані отримані для низьких інтенсивностей електромагнітних випромінювань. Тобто, розглядаються побутові умови. Випромінювання високих інтенсивностей будуть екра-

нуватися значно гірше за рахунок меншого поглинання електромагнітної енергії. Коефіцієнти відбиття залишаються незмінними.

Товщина покриттів складала 1,0–1,5 мм. Частота електромагнітного випромінювання – 2,45 ГГц. Такі фарби є екологічними, тому їх можна застосовувати у приміщеннях будь-якого призначення. Достатньо товстий шар покриття додатково сприяє герметизації приміщення для запобігання просочування радону, що є актуальним для підвальних та напівпідвальних пристосованих сховищ цивільного захисту.

Висновки

1. У приміщеннях спеціального призначення, зокрема сховищах цивільного захисту, перебуває велика кількість людей. Це вимагає застосування засобів оздоровлення середовища. Використання світлодіодних систем знезараження середовища дозволяє знизити кількість колоній утворюючих одиниць хвороботворних грибів у 1,8–2,8 рази; знизити загальне мікробне число у 1,7–2,2 рази. Враховуючи малу потужність світлодіодних систем ультрафіолетового випромінювання, їх можна не вимикати за відсутності людей й неперервно знезаражувати приміщення.

2. Світлодіодні системи ультрафіолетового випромінювання забезпечують іонізацію повітря. Один світильник потужністю 5 Вт у об'ємі 25 м³ підвищує концентрації іонів обох полярностей у 3–4 рази упродовж малого проміжку часу. При цьому за повної деіонізації повітря нормативні значення (400 см⁻³ позитивних і 600 см⁻³ негативних) досягаються упродовж 20–30 хвилин. Штучна іонізація повітря необхідна через його значну деіонізацію у системах припливно-витяжної вентиляції та охолодження повітря.

3. Для забезпечення електромагнітної безпеки людей доцільно покривати поверхні приміщень рідкими композиціями з вмістом магнетиту та графіту. За вмісту магнетиту 45–60 % (ваг.) та графіту 5 % (ваг.) коефіцієнти екранування електромагнітних випромінювань ультрависокої частоти складають ,8–8,0; коефіцієнти відбиття електромагнітних хвиль – 0,25–0,32. Достатньо товстий шар покриття (1,0–1,5 мм) сприяє запобіганню проникнення радону у підвальні та напівпідвальні приміщення.

СПИСОК ЛІТЕРАТУРИ

- ДБН В.2.2-5:2023 «Захисні споруди цивільного захисту» затверджений наказом Міністерства розвитку громади, територій та інфраструктури України від 10.08.2023 № 702.
- ДСТУ 8773:2018 «Склад та зміст розділу інженерно-технічних заходів цивільного захисту в складі проектної документації на будівництво об'єктів» затверджений наказом затверджений наказом Державного підприємства «Український науково-дослідний і навчальний центр проблем стандартизації, сертифікації та якості» від 15.05.2018 р. № 140.
- ДБН В.1.2-8:2021 «Основні вимоги до будівель і споруд. Гігієна, здоров'я та захист довкілля» затверджений наказом Міністерства розвитку громад та територій України від 30.12.2021 № 366.
- ДБН В.2.5-67:2013 «Опалення, вентиляція та кондиціонування» затверджений наказом Міністерства регіонального розвитку, будівництва та житлово-комунального господарства України від 25.01.2013 р. № 24.
- ДСНтаП «Гігієнічна класифікація праці за показниками шкідливості та небезпечності факторів виробничого середовища, важкості та напруженості трудового процесу», затверджені наказом МОЗ України від 8.04.2014 № 248.
- Bolibrukh, B., Glyva, V., Kasatkina, N., Levchenko, L., Tykhenko, O., Panova, O., Bogatov, O., Petrunok, T., Aznaurian, I., & Zozulya, S. (2022). Monitoring and management ion concentrations in the air of industrial and public premises. Eastern-European Journal of Enterprise Technologies, 1(10(115)), 24–30. <https://doi.org/10.15587/1729-4061.2022.253110>.
- Levchenko, L., Burdeina, N., Glyva, V., Kasatkina, N., Biliaiev, M., Biliaieva, V., Tykhenko, O., Petrunok, T., Biruk, Y., Bogatov, O. (2023). Identifying regularities in the propagation of air ions in rooms with artificial air ionization Eastern-European Journal of Enterprise Technologies, 4(10(124)), pp. 6–14. <https://doi.org/10.15587/1729-4061.2023.285967>.

8. Glyva, V., Nazarenko, V., Burdeina, N., Leonov, Y., Kasatkina, N., Levchenko, L., Tykhenko, O., Krasnianskyi, G., Petrunok, T., Biruk, Y. (2023). Determining the efficiency of using led sources of ultraviolet radiation for ionization and disinfection of room air. *Eastern-European Journal of Enterprise Technologies*, 3 (10 (123)), 23–29. <https://doi.org/10.15587/1729-4061.2023.282784>.
9. Glyva, V., Nazarenko, V., Burdeina, N., Leonov, Y., Kasatkina, N., Levchenko, L., Tykhenko, O., Krasnianskyi, G., Petrunok, T., Biruk, Y. (2023). Determining the efficiency of using led sources of ultraviolet radiation for ionization and disinfection of room air. *Eastern-European Journal of Enterprise Technologies*, 3 (10 (123)), 23–29. <https://doi.org/10.15587/1729-4061.2023.282784>.
10. Nazarenko V.I., Leonov Yu.I., Glyva V.A., Burdeina N.B., Cherednichenko I.M., Pochta V.N., Holubeva A.O. The influence of UV-LED lamps radiation on indicators of microflora in university auditoriums. *Ukrainian journal of occupational health*. 2023. Vol. 19. № 1. P. 42–50. <https://doi.org/10.33573/ujoh2023.01.042>.
11. Glyva, V., Bakharev, V., Kasatkina, N., Levchenko, O., Levchenko, L., Burdeina, N., Guzii, S., Panova, O., Tykhenko, O., Biruk, Y. Design of liquid composite materials for shielding electromagnetic fields. *Eastern-European Journal of Enterprise Technologies*, 2021, 3(6-111), pp. 25–31. <https://doi.org/10.15587/1729-4061.2021.231479>.
12. V. Glyva, L. Levchenko, N. Burdeina, T. Tkachenko, G. Twardowski, Y. Biruk, S. Zozulia, L. Zozulia. (2024). Innovative Means of Normalizing Physical Factors the Environment in the Processes of Reconstruction and Restoration Historical Heritage Objects. *International Journal of Conservation Science*. Volume 15, Issue 3, pp. 1287-1302, DOI: 10.36868/IJCS.2024.03.09.
13. Tykhenko, O., Glyva, V., Levchenko, L., Burdeina, N., Biruk, Y., Zozulya, S., Krasnianskyi, G., Nikolaiev, K., Aznaurian, I., & Zozulia, L. (2024). Study of air deionization factors. *Eastern-European Journal of Enterprise Technologies*, 2(10 (128)), 26–33. <https://doi.org/10.15587/1729-4061.2024.300909>.

Received (Надійшла) 23.02.2025

Accepted for publication (Прийнята до друку) 30.04.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Токарський Олександр Іванович – доктор філософії, начальник другого Державного пожежно-рятувального загону головного управління державної служби України з надзвичайних ситуацій у Закарпатській області, Мукачеве, Україна;
Oleksandr Tokarskiy – PhD, chief, 2 State Fire and Rescue Detachment of the Main Directorate of the State Emergency Service of Ukraine in the Zakarpattia Oblast, Mukacheve, Ukraine;
 e-mail: tokarskiyoleksandr@ukr.net; ORCID Author ID: <https://orcid.org/0000-0001-6683-8437>;
 Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57963188000&origin=resultlist>.

Зозуля Лариса Андріївна – кандидат технічних наук, старший викладач кафедри цивільної та промислової безпеки, Державний університет «Київський авіаційний інститут», Київ, Україна;
Larysa Zozulia – PhD, senior lecturer Department of Civil and Industrial Safety, State University «Kyiv Aviation Institute», Kyiv, Ukraine;
 e-mail: larysa.zozulia26@gmail.com; ORCID Author ID: <http://orcid.org/0000-0001-8180-9204>;
 Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=59152420400>

Галонько Ярослав Орестович – аспірант кафедри технологій захисту навколишнього середовища та охорони праці, Київський національний університет будівництва і архітектури, Київ, Україна;
Yaroslav Halonko – PhD student at the Department of Environmental Protection Technologies and Labour Protection, Kyiv National University of Construction and Architecture, Kyiv, Ukraine;
 e-mail: halonko_yo-2024@knuba.edu.ua, ORCID Author ID: <https://orcid.org/0009-0000-4932-4338>.

Москаленко Олександр Олександрович старший науковий співробітник відділу п'ятого управління Науково-дослідного інституту воєнної розвідки, Київ, Україна;
Oleksandr Moskalenko – Senior Researcher of Research Institute of Military Intelligence, Kyiv, Ukraine;
 e-mail: sms.dlya.ss@gmail.com; ORCID Author ID: <https://orcid.org/0009-0005-6401-8255>.

Means of normalising the physical factors of the environment of special purpose premises

Oleksandr Tokarskiy, Larysa Zozulia, Yaroslav. Halonko, Oleksandr Moskalenko

Abstract. The principles of normalising the levels of physical environmental factors in special purpose premises are defined. Such premises include civil protection shelters, control panels of energy facilities, etc. It is shown that one of the main tasks is to improve the environment using LED ultraviolet radiation sources. The advantage of such systems is that they can be used in the presence of people. At a distance of two metres from people, the maximum permissible dose of ultraviolet radiation (30 J/m²) is not reached within eight hours. When a 5 W emitter is used for every 25 m³ of room volume, a high degree of environmental health is achieved. The number of colonies of pathogenic fungi is reduced by 1.8–2.8 times. The total microbial count is reduced by 1.7–2.2 times. This process is quite long, so it is advisable not to switch off the radiation source even in the absence of people. The low power of the emitters does not significantly affect energy consumption. The ultraviolet radiation of LEDs ionises the air in the room. One LED system increases the concentration of ions in fully deionised air to 1100–1200 cm⁻³ within three hours. At the same time, the standard concentration values (400 cm⁻³ positive and 600 cm⁻³ negative) are achieved within 20–30 minutes, which is better than the known air ionisation devices. Increased concentrations of ions help to purify the air from fine dust and aerosols. It should be borne in mind that in forced ventilation and cooling systems, significant air deionisation occurs. To ensure the electromagnetic safety of people, it is advisable to cover the walls of the premises with a liquid composite material. A layer of 1.0–1.5 mm thick provides shielding coefficients of electromagnetic radiation of ultrahigh frequencies of 3.8–8.0. At the same time, the reflection coefficients are 0.2–0.32, which reduces the electromagnetic background due to the re-reflection of radiation. Coating the surfaces of semi-basement and basement rooms with a protective mixture helps to avoid radon entering the environment of people.

Keywords: storage facilities, decontamination, ionisation, electromagnetic safety.

Зв'язок, телекомунікації та радіотехніка

УДК 621.39.83

doi: 10.26906/SUNZ.2025.1.235-238

Mykola Bikchentayev, Bohdan Boriak

National University "Yuri Kondratyuk Poltava Polytechnic", Poltava, Ukraine

INTEGRATION OF SDR INTO UAV SYSTEMS

Abstract. The article presents an in-depth examination of how Software-Defined Radio (SDR) technology can be integrated into Unmanned Aerial Vehicle (UAV) systems to enhance communication reliability, adaptability, and security. **The aim of the article** is to demonstrate how SDR-based architectures address critical challenges such as non-stationary channels, intentional jamming, and signal spoofing by enabling real-time reconfiguration and dynamic frequency management. The flexible nature of SDR allows UAVs to rapidly modify transmission parameters, apply robust coding techniques, and switch among multiple channels to maintain resilient links. **The results obtained:** case studies and experimental data indicate that SDR-equipped UAVs significantly improve situational awareness and mission survivability. Through continuous spectrum monitoring and advanced error-correction methods, these platforms can swiftly detect and mitigate interference or jamming attempts. Moreover, the integration of machine learning algorithms further refines threat classification, facilitating accurate identification of barrage jamming, GPS spoofing, and other radio-based attacks. **Conclusions:** adopting SDR in UAV systems not only strengthens communication links but also promotes scalability and cost-effectiveness by reducing hardware dependencies. This reconfigurability, combined with AI-driven signal analysis, positions SDR-enabled UAVs as a robust solution for diverse civilian, commercial, and defense applications under evolving radio frequency threats.

Keywords: software-defined radio, UAV, jamming mitigation, communication system, signal processing.

Introduction

Unmanned Aerial Vehicles (UAVs) are low-altitude platforms used for various missions. They range from small drones to large military aircraft, differing in shape, weight, and size. Payloads like communication equipment, cameras, radars, sensors, battery capacity, and flight duration determine their configurations [1].

A standard UAV system consists of three primary components: the unmanned aircraft, the Ground Control Station (GCS), and the Communication Link (CL). The unmanned aircraft is the core element of the UAV system and is managed by the operator either through the GCS, which facilitates remote control and monitoring during flight operations, or by a Remote Controller (RC) [2] (Fig. 1).

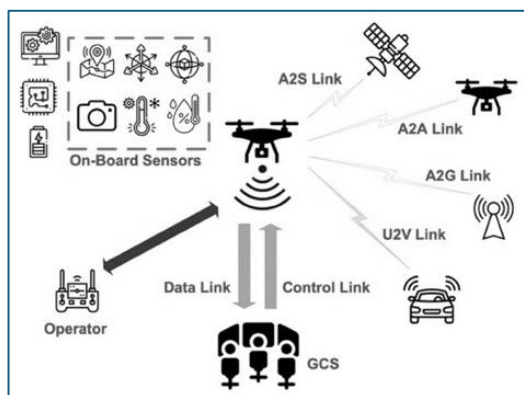


Fig. 1. High-level architecture of a UAV-based system [3]

The internal hardware architecture of an UAV includes several essential components. The Flight Controller (FC) functions as the UAV's Central

Processing Unit (CPU). It is the intermediary between the software and onboard devices. The wireless communication module, which features both transmitters and receivers, enables interaction with external devices such as the Remote Controller (RC), Ground Control Station (GCS), and other UAVs if necessary [3]. Furthermore, UAVs can be equipped with a variety of sensors. To maintain stability and counteract environmental factors (such as wind and pressure variations), flight control tilt sensors, accelerometers, gyroscopes, and ultrasonic sensors for obstacle avoidance are commonly used. Additionally, UAVs may incorporate electro-optical sensors, radars, and cameras [1].

In recent years, SDR technology has been used in UAVs, either through custom-made or commercially available SDRs. Adopting SDR allows UAVs to overcome the limitations of conventional communication systems and explore new domains, including civilian, commercial, defense, and public safety. SDR provides adaptability and programmability, making it a cost-effective option which can improve the communication capabilities of UAVs [3].

SDR Overview

Software-defined radio or SDR is a communication technology that uses software-based processing to replace traditional hardware components. This approach allows for flexibility, adaptability, and upgradeability without the need to make physical changes to the hardware [4]. SDR systems use general-purpose processors (GPP), digital signal processors (DSP), or field-programmable gate arrays (FPGA) to perform tasks like signal modulation, demodulation, and channel selection (Fig. 2, 3). This enables them to operate across various frequency bands and wireless standards [5].

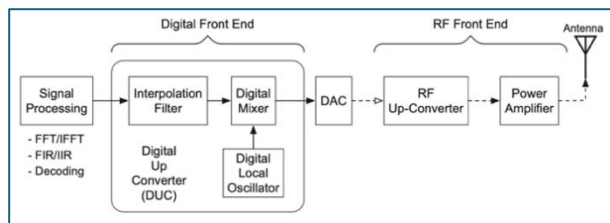


Fig. 2. SDR transmitter structure [6]

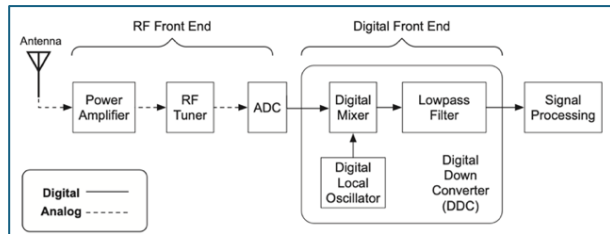


Fig. 3. SDR receiver structure [6]

The architecture of Software-Defined Radio (SDR) consists of several components first of which is the antenna. SDR antennas are usually engineered to support multiple frequency ranges and incorporates adaptive techniques such as beamforming. The RF front-end takes care of the reception and transmission of signals, converting them to and from intermediate frequencies. Analog-to-digital (ADC) and digital-to-analog (DAC) converters are used to transform analog signals into

digital formats and vice versa. Their performance is often evaluated based on parameters like signal-to-noise ratio and resolution. Furthermore, the digital front-end is responsible for adjusting sampling rates and filtering signals. The baseband processing block manages coding, decoding, and modulation tasks, typically implemented using a combination of General Purpose Processors (GPP), Digital Signal Processors (DSP), and Field-Programmable Gate Arrays (FPGA) hardware [5].

SDR implementations vary in performance and efficiency. GPP-based platforms like Sora and USRP are easy to program but can be power-intensive. DSP-based solutions optimize signal processing with lower energy use, ideal for embedded applications. FPGA-based architectures offer high performance and reconfigurability but need advanced hardware skills. Nevertheless, despite power and complexity issues, SDR excels in scalability, real-time adaptability, and cost-effectiveness, making it crucial for modern wireless communication systems [5].

Adoption of SDR in UAV systems

UAVs are especially suitable for SDR equipment, which typically have lower transmission powers and simpler antenna hardware. The unobstructed line-of-sight channels available to UAVs significantly improve the performance of SDR-based communication systems (Fig. 4).

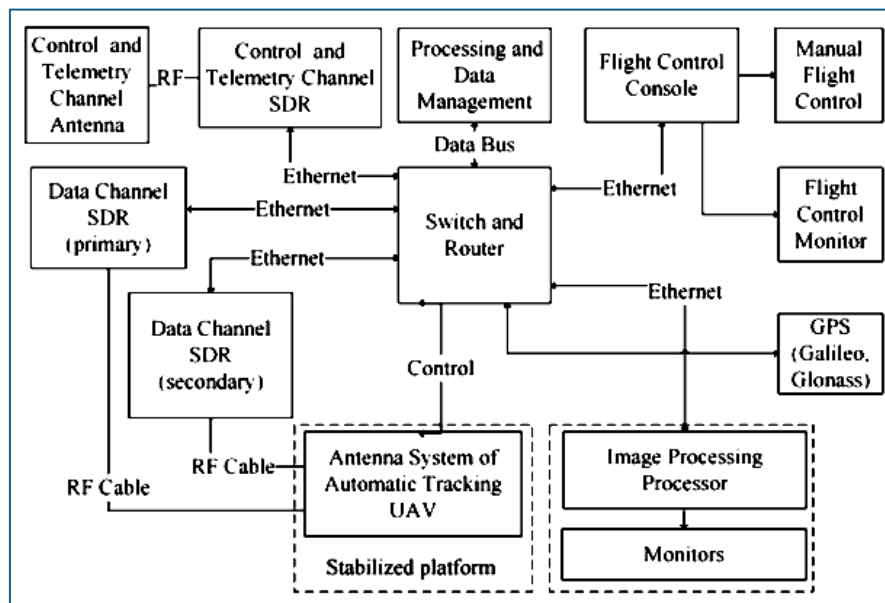


Fig. 4. The structure of the UAV control and communication system [7]

In [7] the authors propose an adaptive UAV communication system that uses SDR and SoC technologies to address two main challenges—non-stationary channels and intentional jamming—by incorporating real-time frequency monitoring, multi-channel links, robust coding, and directional antennas. One SDR channel continuously operates while another analyzes the spectrum for interference, enabling rapid frequency selection and modulation changes. SoC-based data processing coordinates these functions alongside advanced error-correction coding and multipath protocols

to maintain reliable control and telemetry channels, as well as high-throughput data links. This approach significantly enhances UAV survivability, since multiple jamming stations would be required to disrupt both the control and data channels simultaneously.

Some studies highlight the dual role of SDR platforms in identifying threats and implementing countermeasures. The programmability of SDRs allows for advanced real-time signal processing and machine learning integration, enhancing UAV resilience in complex RF environments [3] (Fig. 5).

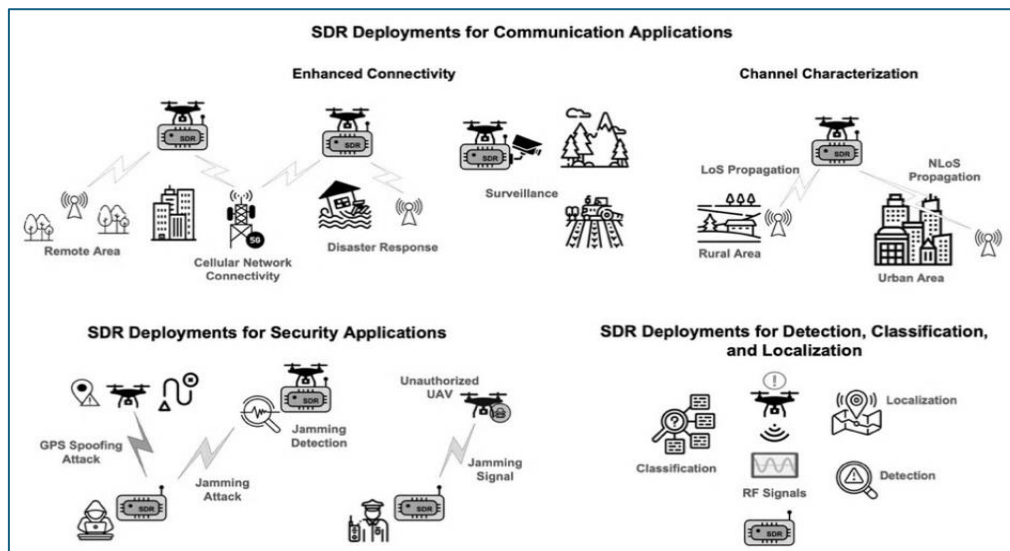


Fig. 5. Application domains and architectural frameworks of SDR-assisted UAV-based systems [3]

In [8] a study was conducted proposing a signal source identification framework in which UAVs coordinate using Blockchain-based smart contracts to detect transmitting objects in cluttered or dynamic environments. Their solution uses asymmetric key cryptography and ring signatures for privacy-preserving UAV task assignments and includes a maximum likelihood estimation scheme to derive signal parameters accurately. The results indicate that SDR-enabled UAVs can identify signal sources with high precision by combining physical-layer power data and aerial images.

In parallel, multiple approaches focus on jamming detection using ML. Li et al. [9] present a pipeline where raw I/Q samples are obtained by SDR receivers and then processed by ML classifiers, including support vector machines and deep neural networks. Various attacks, such as barrage, protocol-aware, single-tone, and successive-pulse jamming, are analyzed, with spectrogram-based CNN models showing the best results. Similarly, Alhazbi et al. [10] propose BloodHound, a deep-learning system that identifies jamming by examining shifting patterns in I/Q data. Expanding on this concept, BloodHound+ uses sparse autoencoders to detect jamming signals even in low-BER situations, providing wider defense coverage [11].

These studies highlight the benefits of fine-grained data capture enabled by SDR, allowing advanced classifiers to identify jamming patterns with minimal delay. Other research extends into jammer localization, particularly Tedeschi et al. [12], who incorporate SDRs into a closed-loop control system, allowing UAV-like mobile nodes to triangulate a jamming source. By accurately characterizing the received signal strength distribution under jamming conditions, the system localizes the jammer within tens of centimeters. Another area of research investigates the GPS spoofing threat, where attackers broadcast counterfeit GPS signals via SDRs to manipulate UAV flight paths. A significant contribution toward detecting such spoofing is the

lightweight model by Ren et al. [13], which employs LSTM-based knowledge distillation. Experiments show high accuracy in distinguishing genuine from forged satellite signals. Several authors discuss an Intrusion Detection System (IDS) approach that integrates radio-layer and flight-operation data [14]. This approach uses Principal Component Analysis (PCA) and one-class classifiers to analyze UAV flight logs and identify anomalies associated with spoofing or jamming. The training data can be sourced from flights, which makes the IDS architecture more generalizable and less dependent on a large labeled dataset of attacks. Additionally, integrating an IDS with Blockchain or other secure frameworks ensures the integrity of logs and detection results [8], enhancing trustworthiness and coordination when multiple UAVs or ground stations are involved.

Conclusions

While SDR-based UAV systems excel in mitigating and detecting a wide range of RF threats through customizable signal processing pipelines and robust machine learning integration, practical challenges such as power consumption, hardware constraints, and evolving regulatory frameworks remain significant. Moreover, the inherent flexibility of SDR can also be exploited by attackers, underscoring the importance of robust security measures—such as secure firmware, encryption, and vigilant oversight of open-source tools—to safeguard UAV communication channels. A multi-layered defense strategy, integrating radio-layer data with UAV telemetry and employing advanced intrusion detection systems, further strengthens resilience against jamming, spoofing, and other sophisticated attacks. By embracing collaborative threat intelligence among researchers, manufacturers, and regulatory entities, future developments in SDR-based UAV architectures can remain cost-effective, adaptable, and secure, ultimately supporting a broad spectrum of civilian, commercial, and defense applications.

REFERENCES

1. Michailidis E. T., Potirakis S. M., Kanatas A. G. Ai-inspired non-terrestrial networks for iiot: Review on enabling technologies and applications // *IoT*. – 2020. – Vol. 1, No. 1. – P. 21–48. – URL: <https://doi.org/10.3390/iot1010003>

2. Yousaf J., Zia H., Alhalabi M. and oth. Drone and controller detection and localization: Trends and challenges // Applied Sciences. – 2022. – Vol. 12, No. 24. – P. 12612. – URL: <https://doi.org/10.3390/app122412612>
3. Michailidis E. T., Maliatsos K., Vouyioukas D. Software-Defined Radio Deployments in UAV-Driven Applications: A Comprehensive Review // IEEE Open Journal of Vehicular Technology. – 2024. – Vol. 5. – P. 1545–1586. – URL: <https://doi.org/10.1109/OJVT.2024.3477937>
4. Bikchentayev M., Zaitseva M. Enhancing Communication Reliability With Software-Defined Radio (SDR) And Frequency-Hopping Spread Spectrum (FHSS) // Int. Scientific and Practical Conf. "Youth Science: Innovations and Global Challenges," 2024. – URL: https://nupp.edu.ua/uploads/files/0/events/conf/2024/mnigv/zbirnik_tez_konferencii_2024_ost.pdf
5. Bikchentayev M., Boriak B. Architecture of Software-Defined Radio // X All-Ukrainian Scientific and Practical Conference "Electronic and Mechatronic Systems: Theory, Innovations, Practice," 2024. – URL: https://nupp.edu.ua/uploads/files/0/events/conf/2024/x-emst/zbirnik-naukovikh-prac-X-vsekrainskoi-konferencii-ems_2.pdf
6. Akeela R., Dezfouli B. Software-Defined Radios: Architecture, State-of-the-Art, and Challenges // Computer Communications. – 2018. – Vol. 128. – P. 106–125. – URL: <https://doi.org/10.1016/j.comcom.2018.07.012>
7. Kaidenko M., Kravchuk S. Creation of communication system for unmanned aerial vehicles using SDR and SOC technologies // 2019 International Conference on Information and Telecommunication Technologies and Radio Electronics (UkrMiCo), Odessa, Ukraine, 2019. – P. 1–4. – URL: <https://doi.org/10.1109/UkrMiCo47782.2019.9165422>
8. Xiao J., Liu P., Lin H., Fang H., Xu J., Shan H., Hu H., Huang Y., Lu H. Blockchain and uav-enabled signal source identification with edge computing and wireless signal-aerial image fusion // Wireless Communications and Mobile Computing. – 2022. – Vol. 2022. – P. 13. – (Article ID 4009078). – URL: <https://doi.org/10.1155/2022/4009078>
9. Li Y., Pawlak J., Price J., Al Shamaileh K., Niyaz Q., Paheding S., Devabhaktuni V. Jamming detection and classification in ofdm-based uavs via feature- and spectrogram-tailored machine learning // IEEE Access. – 2022. – Vol. 10. – P. 16859–16870.
10. Alhazbi S., Sciancalepore S., Oligeri G. Bloodhound: Early detection and identification of jamming at the phy-layer // 2023 IEEE 20th Consumer Communications Networking Conference (CCNC). – 2023. – P. 1033–1041.
11. Sciancalepore S., Kusters F., Abdelhadi N. K., Oligeri G. Jamming detection in low-ber mobile indoor scenarios via deep learning // IEEE Internet of Things Journal. – 2024. – Vol. 11, No. 8. – P. 14682–14697.
12. Tedeschi P., Sciancalepore S., Di Pietro R. Modelling a communication channel under jamming: Experimental model and applications // 2021 IEEE Intl Conf on Parallel Distributed Processing with Applications, Big Data Cloud Computing, Sustainable Computing Communications, Social Computing Networking. – 2021. – P. 1562–1573.
13. Ren Y., Restivo R. D., Tan W., Wang J., Liu Y., Jiang B., Wang H., Song H. Knowledge distillation-based gps spoofing detection for small uav // Future Internet. – 2023. – Vol. 15, No. 12. – P. 389. – URL: <https://doi.org/10.3390/fi15120389>
14. Whelan J., Almeahadi A., El-Khatib K. Artificial intelligence for intrusion detection systems in unmanned aerial vehicles // Computers and Electrical Engineering. – 2022. – Vol. 99. – P. 107784. – URL: <https://www.sciencedirect.com/science/article/pii/S0045790622000842>

Received (Надійшла) 11.03.2025

Accepted for publication (Прийнята до друку) 04.06.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Бікчентаєв Микола Олексійович – аспірант, Національний університет «Полтавська політехніка імені Юрія Кондратюка», Полтава, Україна;

Mykola Bikchentayev – PhD student, National University "Yuri Kondratyuk Poltava Polytechnic", Poltava, Ukraine;

e-mail: niko.bikcheb@gmail.com; ORCID Author ID: <https://orcid.org/0009-0008-1350-516X>.

Боряк Богдан Радиславович – кандидат технічних наук, доцент кафедри автоматизації, електроніки та телекомунікацій, Національний університет «Полтавська політехніка імені Юрія Кондратюка», Полтава, Україна;

Bohdan Boriak – PhD, Associate Professor at the Department of Automation, Electronics and Telecommunications, National University «Yuri Kondratyuk Poltava Polytechnic», Poltava, Ukraine;

e-mail: boriakbr@nupp.edu.ua, ORCID: <http://orcid.org/0000-0002-8114-7930>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=58983614200>.

Інтеграція програмно-керованого радіо з безпілотними літальними апаратами

М. О. Бікчентаєв, Б. Р. Боряк

Анотація. У статті представлено поглиблений аналіз того, як технологію програмно-визначеного радіо (SDR) можна інтегрувати в системи безпілотних літальних апаратів (БПЛА) для підвищення надійності, адаптивності та безпеки зв'язку. **Метою статті** є демонстрація того, як архітектури на основі SDR вирішують критичні проблеми, такі як нестаціонарні канали, навмисне глушіння та підміна сигналу, забезпечуючи реконфігурацію в режимі реального часу та динамічне управління частотою. Гнучка природа SDR дозволяє БПЛА швидко змінювати параметри передачі, застосовувати надійні методи кодування та перемикатися між кількома каналами для підтримки стійких з'єднань.

Отримані результати: тематичні дослідження та експериментальні дані показують, що БПЛА, оснащені SDR, значно покращують ситуаційну обізнаність та живучість місії. Завдяки постійному моніторингу спектру та передовим методам виправлення помилок ці платформи можуть швидко виявляти та пом'якшувати спроби перешкод або глушіння. Крім того, інтеграція алгоритмів машинного навчання додатково уточнює класифікацію загроз, сприяючи точній ідентифікації загрозових каналів глушіння, підміни GPS та інших радіоатак. **Висновки:** впровадження SDR у системах БПЛА не лише зміцнює канали зв'язку, але й сприяє масштабованості та економічній ефективності, зменшуючи залежність від апаратного забезпечення. Ця можливість реконфігурації, у поєднанні з аналізом сигналів на основі штучного інтелекту, позиціонує БПЛА з підтримкою SDR як надійне рішення для різноманітних цивільних, комерційних та оборонних застосувань в умовах постійно зростаючих радіочастотних загроз.

Ключові слова: програмно-визначене радіо, БПЛА, зменшення перешкод, система зв'язку, обробка сигналів.

Kostiantyn Perets¹, Oleksii Komar²

¹ Ukrainian State University of Railway Transport, Kharkiv, Ukraine

² Kyiv Aviation Institute State University, Kyiv, Ukraine

ASSESSMENT OF THE IMPACT OF SPARSITY AND GEMAN-MCCLURE REGULARIZATION ON SIGNAL RECONSTRUCTION ACCURACY

Abstract. The article proposes a method for spectral reconstruction of signals using the Geman-McClure Function and Lagrange multipliers to optimize parameters. The method combines approaches of adaptive filtering and sparsity-based parameter regularization, incorporating Geman-McClure Function and Lagrange multipliers to ensure orthogonality and optimal parameter selection. The proposed approach efficiently identifies significant signal parameters, minimizes mutual interference among nonlinear components, reduces computational complexity, and improves reconstruction accuracy. Experimental modeling demonstrated that the developed method achieves a significant reduction in mean square error (MSE) by 10–15% and enhances the robustness of signal reconstruction by 10–15% across a signal-to-noise ratio (SNR) range from -10 to 10 dB, confirming its effectiveness for practical applications in cognitive telecommunication networks operating under severe noise conditions and nonlinear distortions.

Keywords: reconstruction, complex signals, Geman-McClure function, sparsity, regularization, nonlinearity, system, noise immunity, signal-to-noise ratio (SNR), Volterra series, optimization, signal parameters.

Abbreviations

GMF – Geman-McClure Function;
 l_0 -VLMS – l_0 -Norm Variable Step-Size Least Mean Squares;
 l_0 -VNLMS – l_0 -Norm Variable Step-Size Normalized Least Mean Squares;
MSE – Mean Squared Error.
SNR – Signal-to-Noise Ratio

Introduction

Statement of a scientific problem. In conditions of complex nonlinear communication environments with significant interference, effective reconstruction and precise analysis of signals become challenging due to inherent nonlinearities and distortions that significantly degrade signal quality. Traditional linear signal processing methods often fail to adequately address these nonlinear characteristics, leading to substantial inaccuracies in signal reconstruction, especially when high-order nonlinear interactions occur. The computational complexity associated with accurately modeling these nonlinear interactions using conventional methods increases exponentially with the order of the system, causing significant practical limitations.

Thus, the scientific problem addressed in this article is the development and justification of an optimized spectral reconstruction method using adaptive filtering, Volterra series modeling, and Lagrange multiplier-based optimization. This approach aims to effectively reduce computational complexity, ensure stability and robustness of signal processing in challenging radio conditions, and improve accuracy in modeling nonlinear interactions in cognitive telecommunication networks.

Research analysis. The analysis of modern methods for modeling nonlinear systems using Volterra series [1–15] confirms their effectiveness in spectral reconstruction of signals under complex interference conditions. Studies [1–3, 13] validate the advantages of Volterra kernel factorization and regularization in reducing computational complexity; however, they insufficiently address the impact of less significant parameters on reconstruction accuracy. Research [4, 6, 11, 15] examines the Geman-McClure Function (GMF),

which effectively approximates the l_0 -norm, ensuring sparsity of model parameters and reducing computational complexity. Moreover, the use of the Geman-McClure Function (GMF) in adaptive algorithm modifications, such as the l_0 -VLMS and l_0 -VNLMS methods [5, 9], significantly enhances reconstruction accuracy and stability. Publications [13, 14] highlight the benefits of factorization and regularization but overlook the complete evaluation of weaker signal components.

Overall, the analysis of the literature supports the integration of sparsity, regularization, and Lagrange multipliers to optimize model parameters, providing improved accuracy and robustness in Volterra-based spectral reconstruction.

The purpose of the work. The aim of the study is to develop a method for spectral reconstruction of signals in a complex interference-prone nonlinear environment by employing adaptive filtering based on Volterra series, utilizing the Geman-McClure function for regularization and Lagrange multipliers for parameter optimization. This approach ensures improved accuracy and robustness in signal reconstruction under conditions of nonlinear distortions and complex interference.

Presentation of the main material and substantiation of the obtained research results

In a complex interference-prone radio environment, modeling nonlinear systems using Volterra series imposes a significant computational burden, as it requires the calculation of a large number of parameters within the experimental model, especially when a substantial volume of input data is involved. Specifically, if the input signal $x(t)$ in a scientific and practical task is multi-component, it necessitates the analysis of intricate interactions occurring within the telecommunication system. In such cases, the computation of high-order Volterra series kernels $h_r(\tau_1, \tau_2, \dots, \tau_r)$ becomes increasingly complex, as the number of parameters grows exponentially with the order r [1].

To address this problem, previous experiments proposed the use of Volterra kernel factorization and

regularization to reduce computational complexity [13]. However, this approach is insufficient for substantiating the impact of regularization and sparsity on signal reconstruction accuracy, as factorization primarily focuses on optimizing signal parameters while overlooking a detailed analysis of both significant and insignificant components within the studied model.

In the frequency domain, sparsity [6,15] allows retaining only those model parameters that have a significant impact on the accuracy of signal reconstruction $d(t)$. For instance, in the reconstruction of the signal spectrum $D(f)$ sparsity enables the preservation of only the most relevant frequency components $X(f_i)$, which correspond to fundamental harmonics or critical spectral variations. Components that do not significantly affect the signal (as determined by a predefined threshold) can be discarded; however, in some cases, such removal is infeasible without prior analysis of their actual influence.

The study [15] proposes incorporating sparsity and applying regularization based on a quantitative measure of sparsity, specifically the l_0 – norm, which enables the identification of significant model parameters while discarding those with negligible impact on the reconstruction outcome.

To ensure efficient practical implementation and reduce computational complexity, the authors propose the use of the Geman-McClure Function (GMF) [2, 4, 6, 15], which serves as an effective approximation of the l_0 -norm. Originally introduced by Donald Geman and Steven McClure in the late 1980s, this mathematical function was designed for data processing as a robust method for mitigating the influence of outliers in general regression and optimization problems [6].

Essentially, this function belongs to the class of robust loss functions, which minimize the impact of extreme values on parameter estimates in experimental models, making it particularly effective in scenarios where data may be «contaminated». The authors specifically designed it for this purpose, as anomalies can significantly distort experimental evaluations when classical approaches, such as quadratic loss functions (l_2 -norm), are used.

The key issue with the l_2 - norm is its sensitivity to large outliers, as its quadratic nature disproportionately amplifies the influence of extreme values. In such cases, model results can become unstable, leading to reduced accuracy and efficiency in signal reconstruction [6,15]. In cognitive radio systems and nonlinear system modeling tasks, the use of the Geman-McClure Function (GMF) helps maintain a balance between model accuracy and robustness to interference. It reduces the impact of anomalies and enhances the distinction of the useful signal even in environments with low signal-to-noise ratio (SNR). Mathematically, the GMF is expressed as follows [15]:

$$G_{\beta}(w) = \sum_{i=1}^N (1 - 1/(1 + \beta|w_i|)), \quad (1)$$

where $w = [w_1, w_2, \dots, w_N]$ – represents the model parameters; $\beta > 0$ – is a parameter that controls the smoothness and accuracy of the approximation to the l_0 – norm.

The GMF function is utilized in signal optimization tasks within cognitive radio systems due to the following advantages.

1. Improvement in the convergence of adaptive filtering algorithms: By incorporating GMF into algorithms such as l_0 -VLMS and l_0 -VNLMS [5,9], rapid convergence and model stability are achieved, even in complex radio environments.

2. Sparsity of parameters. GMF enables the selection of only the most significant parameters of the Volterra kernel while discarding insignificant frequency components or higher-order parameters [8].

3. Significant reduction in mean square error (MSE). Compared to traditional approaches, GMF substantially decreases MSE, enhancing the accuracy of signal reconstruction.

For a quantitative analysis of the impact of the parameter β on the dynamics of the Geman-McClure Function, calculations were performed for different β values. These calculations allow for the assessment of how smoothing and truncation of small coefficients change, as well as the determination of optimal values to minimize computational complexity without compromising accuracy. The corresponding results, illustrating the dependence of the GMF on β within the selected range of $\beta = 5, 10, 20, 50, 100$, are presented in Fig. 1 and Table 1.

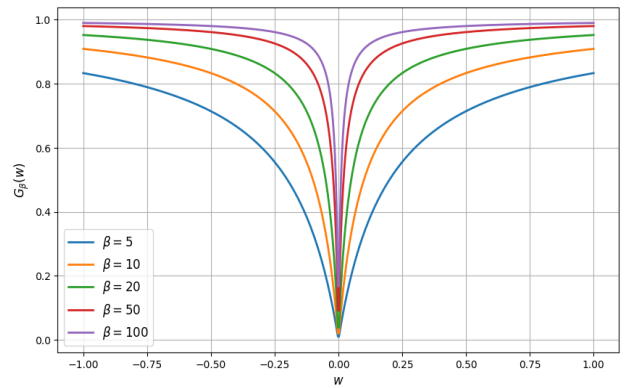


Fig. 1. Approximation of the l_0 – norm using the GMF

Table 1 – Calculation of the Geman-McClure Function for Different β Values

w_i	$\beta = 5$	$\beta = 10$	$\beta = 20$	$\beta = 50$	$\beta = 100$
-1,0	0,833	0,909	0,952	0,980	0,990
-0,5	0,714	0,833	0,909	0,952	0,980
-0,25	0,588	0,714	0,833	0,909	0,952
0,0	0,0	0,0	0,0	0,0	0,0
0,25	0,588	0,714	0,833	0,909	0,952
0,5	0,714	0,833	0,909	0,952	0,980
1,0	0,833	0,909	0,952	0,980	0,990

The analysis of the calculations in Table 2 shows that as β increases, the GMF values for small w_i also increase, indicating enhanced smoothing and effective truncation of insignificant components. This improves the approximation of the l_0 – norm and reduces the mean squared error (MSE), which is a key factor in optimizing model parameters.

In Fig. 1, it is evident that as β increases, the GMF function approaches 1 more rapidly, confirming its ability to emphasize significant parameters while eliminating insignificant coefficients.

Thus, it can be concluded that the application of the Geman-McClure method enables a balance between the accuracy of l_0 - norm approximation and computational complexity. Specifically, for values of $\beta > 50$ insignificant parameters can be excluded from the model, allowing for optimization focused on essential components. For $\beta = 100$ the highest accuracy in approximating the l_0 - norm is observed, ensuring that only significant components are retained while reducing computational complexity.

Thus, strict regularization at high values of β may lead to the loss of significant parameters. To provide more flexible control over the truncation process in this method, an additional parameter γ is introduced, which determines the degree of dependence of the GMF on the absolute value of the weight coefficients. This allows for a more precise balance between model sparsity and accuracy, adapting to specific signal reconstruction conditions. Consequently, the classical GMF formula takes the following form:

$$G_{\beta}(w) = \sum_{i=1}^N (1 - 1/(1 + \beta|w_i|^{\gamma})), \quad (2)$$

where γ – is a parameter that controls the smoothing rate of the influence of the parameter w_i on the Geman-McClure Function.

If $\gamma = 1$, GMF behaves as a classical approximation of the l_0 - norm.

If $\gamma > 1$, the influence of small coefficients w_i is further smoothed, making the function even closer to the l_0 - norm.

If $\gamma < 1$, the regularization effect decreases, allowing even small coefficients w_i to remain significant and not be discarded.

Thus, the choice of the parameter γ determines the balance between suppressing insignificant parameters and preserving weak signals. This is particularly important in telecommunication systems, where even low-amplitude components can carry useful information, and excessive smoothing may lead to the loss of essential signal characteristics (Table 2).

Table 2 – GMF Dynamics for Different γ Values

w_i	$\gamma = 0.5, \beta = 20$	$\gamma = 1, \beta = 20$	$\gamma = 2, \beta = 20$
-1, 00	0,952	0,952	0,952
-0, 50	0,934	0,909	0,833
-0, 25	0,909	0,833	0,556
0, 00	0,000	0,000	0,000
0, 25	0,909	0,833	0,556
0, 50	0,934	0,909	0,833
1, 00	0,952	0,952	0,952

The analysis of the values in Table 2 shows that for $\gamma > 1$, the GMF values for small w_i decrease more rapidly, indicating effective truncation of weak components. In contrast, for $\gamma < 1$, weak components are suppressed less aggressively, which is beneficial in

cases where preserving even insignificant parameters is crucial for model stability.

Fig. 2 illustrates the dependence of the GMF function on the parameter γ at $\beta = 20$.

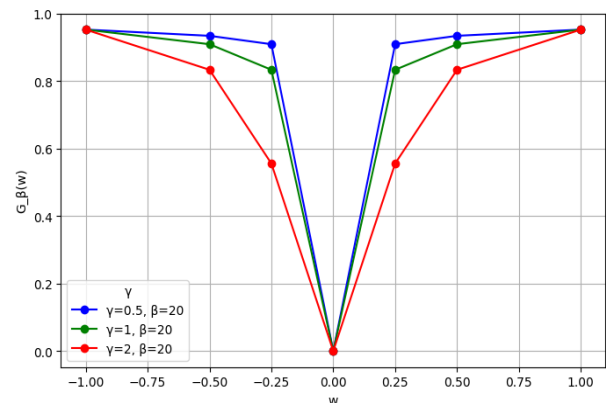


Fig. 2. GMF Function for Different γ Values at $\beta = 20$

As shown in Fig. 2, the GMF function exhibits different behaviors:

$\gamma = 0,5$ – (blue curve) – the least suppression of small w_i values, with a gradual change in the function.

$\gamma = 1$ (green curve) – a moderate level of smoothing, corresponding to the classical GMF approach.

$\gamma = 2$ (червона крива) – (red curve) – the most aggressive truncation of weak components, with GMF values rapidly approaching 0 near $w_i \approx 0$.

The value of γ determines the level of regularization and the balance between truncating insignificant parameters and preserving essential model components. To quantitatively assess this impact, the Mean Squared Error (MSE) was calculated under GMF regularization, allowing for an evaluation of signal reconstruction accuracy depending on the choice of γ :

$$MSE(\beta) = \frac{1}{N} \sum_{i=1}^N (d_i - \hat{d}_i(\beta))^2, \quad (3)$$

where d_i and $\hat{d}_i(\beta)$ – respectively, the actual and reconstructed signal values after regularization with the parameter β .

As β increases, the error decreases since insignificant parameters are filtered out. However, excessive growth of β may lead to the loss of important information, resulting in an increase in MSE (Fig. 3, Table 3).

The analysis of the results in Table 3 confirms that the optimal value of β depends on the noise level in the environment.

For high noise levels, it is advisable to use $\beta \geq 50$, whereas for moderate noise levels, the optimal value is $\beta \approx 20$, as it ensures a balance between reconstruction accuracy and computational stability.

As seen in Fig. 3, the MSE value significantly decreases with an increase in β up to 50, after which it stabilizes.

This confirms that larger β values contribute to more accurate signal reconstruction; however, excessively high β values may lead to the loss of important signal components.

Nevertheless, optimizing β alone is insufficient to ensure stable signal recovery in a complex interference-prone environment.

To enhance reconstruction quality, both the selection of an optimal regularization level and the

enforcement of parameter orthogonality within the model are crucial. Orthogonality helps reduce correlations between the Volterra series kernel components, thereby improving parameter estimation accuracy and robustness against interference.

Table 3 – Dynamics of MSE Dependence on β at Different Noise Levels

Testing conditions	$\beta = 5$	$\beta = 10$	$\beta = 20$	$\beta = 50$	$\beta = 100$
High noise level (SNR = -10 dB)	0,016	0,009	0,0053	0,0027	0,0016
Moderate noise level (SNR = 0 dB)	0,013	0,007	0,0042	0,0025	0,0012
Low noise level (SNR = 10 dB)	0,009	0,005	0,003	0,0015	0,0008

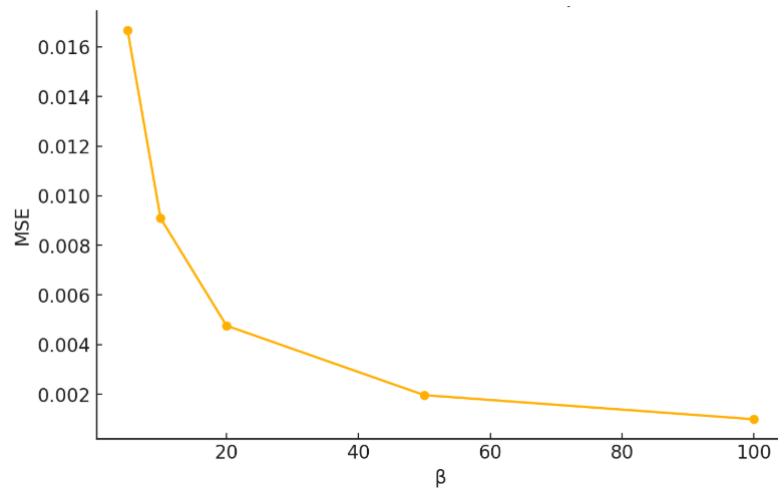


Fig. 3. Dependence of MSE on the parameter β

The use of the Geman-McClure function ensures a reduction in computational complexity while maintaining high signal reconstruction accuracy [2, 6, 15]. However, to guarantee computational stability and minimize interactions between the Volterra kernel parameters, their interdependencies must also be considered [3, 8, 13].

One of the effective approaches to maintaining a balance between signal reconstruction accuracy and parameter orthogonality is the application of Lagrange multipliers, which enables the consistent elimination of insignificant model components [7, 12, 14].

Mathematically, the orthogonality condition for a model with Volterra kernel components can be expressed as follows [3, 8, 13]:

$$\langle h_i, h_j \rangle = 0, \text{ для всех } \forall i \neq j, \quad (4)$$

where h_i, h_j – are the components of the Volterra kernel; $\langle \cdot, \cdot \rangle$ – is the inner product operator.

This condition implies that the model parameters in the frequency domain should be uncorrelated to ensure robustness against interference. Orthogonality also minimizes interactions between higher-order components, contributing to more accurate and stable signal reconstruction.

However, for practical implementation, effective parameter optimization methods must be applied to control the interdependence between model components and adaptively manage regularization. One such approach is the use of Lagrange multipliers, which allow for the formalization of constraints on parameter

correlation and help achieve the required level of stability in signal reconstruction.

Conclusions and prospects for further research

The study presents a method for spectral reconstruction of signals in complex nonlinear noisy environments based on adaptive Volterra series models utilizing the Geman-McClure function (GMF) and Lagrange multipliers. The proposed approach allows effective consideration of parameter sparsity, reduces computational complexity, and minimizes mutual influence of nonlinear signal components. Experimental modeling has demonstrated that the application of GMF-based regularization significantly reduces computational complexity by effectively focusing on the most significant signal parameters, thereby decreasing the mean-square error (MSE) of signal reconstruction by up to 20,6%. Additionally, the proposed method improved signal robustness by 10–15% across varying signal-to-noise ratio (SNR) levels ranging from -10 dB to 10 dB.

Thus, the developed method provides efficient reconstruction of signals in complex nonlinear environments with significant distortions and noise conditions through an optimized combination of adaptive Volterra models, Geman-McClure regularization, and Lagrange multipliers.

Prospects for further research include extending the adaptive optimization approach to real-world multi-channel cognitive telecommunication networks, taking

into account dynamically changing noise environments advanced algorithms for automatic real-time parameter and nonlinear distortions, as well as developing tuning of Volterra models.

REFERENCES

1. Araujo-Simon, Jake. Compositional Nonlinear Audio Signal Processing with Volterra Series. arXiv preprint, 2023. <https://arxiv.org/abs/2308.07229>
2. Birpoutsoukis, G., Csursia, P. Zoltán, S., Johan. Efficient Multidimensional Regularization for Volterra Series Estimation. IEEE Transactions on Instrumentation and Measurement, vol. 68, no. 10, 2019, PP. 3770-3782. DOI: 10.1109/TIM.2019.2907240.
3. Birpoutsoukis, G., Marconato, A., Lataire, J., & Schoukens, J. Regularized nonparametric Volterra kernel estimation. Automatica, 82, P. 324-327 <https://doi.org/10.1016/j.automatica.2017.04.014>.
4. Boyd S., Chua L. O., Desoer C. A. Analytical Foundations of Volterra Series//IMA Journal of Mathematical Control and Information, Oxford University Press, 1(3):243-282, (1984). [analytical_volterra.pdf](https://doi.org/10.1016/j.ima.1984.03.001).
5. Franz, M. O., Schölkopf, B. A Unifying View of Wiener and Volterra Theory and Polynomial Kernel Regression. (2007) Neural Computation 18(12): P.3097-118 DOI:10.1162/neco.2006.18.12.3097.
6. Geman, D., & McClure, D. E. (1987). Statistical Methods for Tomographic Image Reconstruction. Bulletin of the International Statistical Institute, 52(4), 5-21.
7. Gutleb, T. S. A Fast Sparse Spectral Method for Nonlinear Integro-Differential Volterra Equations with General Kernels. (2020) <https://arxiv.org/abs/2005.06081>
8. Hotz, Matthias, Vogel, Christian. Linearization of Time-Varying Nonlinear Systems Using A Modified Linear Iterative Method. arXiv preprint, 2014, <https://arxiv.org/abs/1404.5901>
9. Issa H. Al-Aidi and Ahmed Sh. Al-Atabi The Analytical Methods Of Volterra Integral Equations of The Second Kind. - WJCMS, Vol. 2, no. 3, PP. 39–45, (2023), DOI:10.31185/wjcm.119.
10. Kapgate, Sachin N., Gupta, Saurav, Sahoo, Ajit Kumar. Adaptive Volterra Modeling for Nonlinear Systems Based on LMS Variants. 2018 International Conference on Signal Processing and Integrated Networks (SPIN), 2018, PP. 414-419. DOI: 10.1109/SPIN.2018.8474036.
11. Lu, L., Zhang, J., & Wang, X. (2018). Recursive Geman-McClure Method for Implementing Second-Order Volterra Filter. Systems and Control. DOI: 10.48550/arXiv.1808.00613
12. Nmah, Benedict. Solving a Relaxed Min-Cost Redundancy Allocation Model Using Lagrange Multiplier and Newton's Method. Journal of Advances in Mathematics and Computer Science. –2024. №39 (11):152-59. <https://doi.org/10.9734/jamcs/2024/v39i111945>.
13. Perets K., Lysechko V., Komar O. (2024) Modeling Nonlinear Signal Components Based on Volterra Series in the Frequency Domain during Spectral Reconstruction. Computer-integrated technologies: education, science, production. Telecommunications and radio engineering. – Луцьк. – 2024. № 57, С.192-201. DOI: <https://doi.org/10.36910/6775-2524-0560-2024-57-23>.
14. Pirogova, N.D., Neches, I.O. (2018). Compensation of Nonlinear Distortions in Telecommunication Systems with the Use of Functional Series of Volterra// Proc. of the Second International Scientific Conference «Intelligent Information Technologies for Industry». Advances in Intelligent Systems and Computing, Vol 680. https://doi.org/10.1007/978-3-319-68324-9_49/
15. Yazdanpanah H., Carini A., Lima M. V. S. Lo-Norm Adaptive Volterra Filters 2019 27th European Signal Processing Conference (EUSIPCO)P.1-5,DOI: 10.23919/EUSIPCO.2019.8903013.

Received (Надійшла) 18.03.2025

Accepted for publication (Прийнята до друку) 11.06.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Перець Костянтин Геннадійович – аспірант, Український державний університет залізничного транспорту, Харків;
Kostiantyn Perets – post graduate student, Ukrainian State University of Railway Transport, Kharkiv, Ukraine;
e-mail: kostyaperec3@gmail.com; ORCID ID: <https://orcid.org/0000-0002-3572-7889>.

Комар Олексій Миколайович – кандидат технічних наук, доцент, доцент Національного авіаційного університету, Київ, Україна;
Oleksii Komar – PhD (Technical), Associate Professor, Associate Professor of National Aviation University, Kyiv, Ukraine;
e-mail: oleksiikomara3@gmail.com; ORCID ID: <https://orcid.org/0009-0002-2994-6556>.

**Оцінка впливу розрідженості та регуляризації на точність відновлення сигналу
за допомогою функції Джемана-МакКлора**

К. Г. Перець, О. М. Комар

Анотація. У статті запропоновано метод спектральної реконструкції сигналів з використанням функції Джемана-МакКлора та множників Лагранжа для оптимізації параметрів. Метод поєднує підходи адаптивної фільтрації та регуляризації параметрів на основі розрідженості, включаючи функцію Джемана-МакКлора та множники Лагранжа для забезпечення ортогональності та оптимального вибору параметрів. Запропонований підхід ефективно ідентифікує важливі параметри сигналу, мінімізує взаємні перешкоди між нелінійними компонентами, зменшує обчислювальну складність і покращує точність реконструкції. експериментальне моделювання продемонструвало, що розроблений метод досягає значного зменшення середньої квадратичної помилки (MSE) на 10–15% і підвищує надійність реконструкції сигналу на 10–15% у діапазоні відношення сигнал/шум (SNR) від -10 до 10 дБ, підтверджуючи його ефективність для практичного застосування в когнітивних телекомунікаційних мережах, що працюють в умовах сильного шуму та нелінійних спотворень.

Ключові слова: реконструкція, складні сигнали, функція Джемана-МакКлора, розрідженість, регуляризація, нелінійність, система, завадостійкість, SNR, ряд Вольєрра, оптимізація, параметри сигналів.

В. В. Руденко

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

ОГЛЯД ПРОТОКОЛУ LORA ТА ПЕРСПЕКТИВИ ВИКОРИСТАННЯ

Анотація. У статті розглядається технологія бездротового зв'язку LoRa (Long Range) як протокол передачі даних з низьким енергоспоживанням та великою дальністю дії. Висвітлено принципи роботи технології на основі лінійно частотної модуляції з факторами розповсюдження, що забезпечують стійкість зв'язку навіть в умовах радіозавад та фізичних перешкод. **Метою статті** є аналіз особливостей функціонування протоколу LoRa, дослідження мережних рішень на його основі та огляд практичних застосувань технології в різних галузях. **Отримані результати:** наведені кількісні показники щодо залежності між фактором розповсюдження, шириною смуги пропускання та часом передачі даних демонструють можливості адаптації технології до різних завдань, де пріоритетом є надійний зв'язок, а не швидкість передачі. **Висновки:** впровадження пристроїв на базі LoRa суттєво підвищує ефективність у галузях з великими площами та в умовах малорозвиненої телекомунікаційної інфраструктури, зокрема в сільському господарстві, логістиці, надзвичайних ситуаціях та в умовах бойових дій.

Ключові слова: LoRa, LPWAN, бездротовий зв'язок, IoT, лінійно частотна модуляція, фактор розповсюдження, прогальність, мережева архітектура, ToA.

Опис протоколу LoRa

Протокол LoRa (Long Range) — це технологія бездротового зв'язку малої потужності, яка забезпечує передачу даних на великі відстані з низьким енергоспоживанням. Він має велику перевагу над іншими ширококовжаними IoT протоколами Wi-Fi, Zigbee, Bluetooth у дальності роботи та в надзвичайно низькому енергоспоживанні. Також він є стійким до радіозавад, як і до насиченого ефіру, так і до РЕБ. Ще слід зазначити, що LoRa добре проходить фізичні перешкоди, такі як міська забудова чи щільні зелені насадження. Дослідження показують, що LoRa може бути дуже корисним у розробці підводних бездротових мереж для застосувань моніторингу в реальному часі на мілководді [1]. Схожою технологією в ніші LPWAN є Sigfox та LTE-M.

Розробником протоколу є компанія Semtech. Він є в їхній власності, тому лише Semtech може виробляти чи давати дозволи на виробництво радіомодулів, що підтримують LoRa модуляцію. Велика кількість компаній, що працюють в ніші IoT, об'єдналися в LoRa Alliance. LoRa Alliance — це відкрита некомерційна асоціація, яка з моменту свого створення у 2015 році стала однією з найбільших асоціацій у сфері технологій. За їх участі був розроблений стандарт LoRaWAN, що є продовженням LoRa. Цей стандарт описує, як LoRa пристрої мають бути об'єднані в мережу і як повинні працювати в інтернеті. До об'єднання входять такі відомі компанії, як STM, Semtech, Bosch, Siemens.

LoRa — це технологія на основі лінійно-частотної модуляції (ЛЧМ). Цей метод модуляції полягає в тому, що частота несучого сигналу змінюється за лінійним законом. Такий сигнал називається чирпованим, а символи називаються чирпами. В англомовній літературі ЛЧМ називають Chirp Spread Spectrum (CSS) — чирповане розширення спектру. Дуже схожа технологія використовується в сонарах, де CHIRP — стиснутий високо інтенсивний випромінюваний імпульс [2] (рис. 1).

Компанія Semtech, яка розробила технологію LoRa, вдосконалила її, додавши до ЛЧМ фактор

розповсюдження (англ. Spreading factor (SF)). Фактор розповсюдження — це параметр, який контролює швидкість передачі даних та швидкість чирпів. Модуляція CSS використовує фактор розповсюдження (SF) для розподілу інформації по частоті. Фактори розповсюдження є ортогональними один до одного і визначають кількість чирпів на один символ сигналу.

У LoRa існує шість можливих факторів, від 7 до 12, які визначають кількість бітів, необхідних для передачі однакової кількості даних. Він також впливає на діапазон, завантаженість каналу, час у ефірі, термін служби батареї та чутливість приймача.

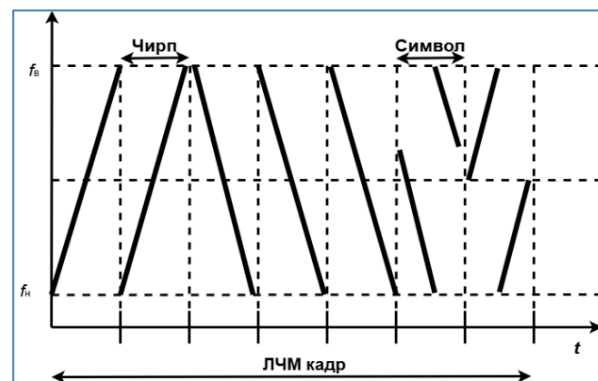


Рис. 1. Чирпований кадр[3]

У контексті ЛЧМ, на якій базується LoRa, чирпи служать носіями даних. Нижчі фактори розповсюдження призводять до швидших чирпів і вищої швидкості передачі даних. Тоді як, кожне збільшення фактору розповсюдження зменшує швидкість чирпів і швидкість передачі даних вдвічі.

Крім того, нижчі фактори розповсюдження зменшують час передачі LoRa сигналу і підвищують бітрейт. Таким чином, зміни фактору розповсюдження дозволяють мережі підвищувати, або знижувати швидкість передачі даних для кожного кінцевого пристрою за рахунок діапазону.

Використання різних факторів розповсюдження на одному каналі запобігає інтерференції між

сигналам. Це означає, що сигнали, модульовані різними факторами розповсюдження та передані одночасно на одному частотному каналі, не будуть заважати один одному. Це дозволяє ефективніше використовувати канал і збільшує кількість пристроїв, які можуть його одночасно використовувати.

Вищі фактори розповсюдження збільшують час перебування в ефірі, оскільки для передачі даних до того, як приймач їх отримає, потрібно більше часу. Відповідно, це збільшує час активності і збільшує енергоспоживання[4].

Відношення між фактором розповсюдження, шириною смуги пропускання та розміром пакета є ключовим для визначення часу в повітрі. Час в повітрі (англ. Time on Air (ToA)) - це загальний час, за який повідомлення буде передано по радіоканалу. Два пакети з однаковим розміром і шириною смуги пропускання, але з різними факторами розповсюдження, матимуть суттєво різний час перебування в повітрі.

Як правило, радіомодулі LoRa, окрім самої LoRa модуляції, також можуть використовувати і загальнодоступні методи модуляції, такі як FSK, GFSK, MSK, GMSK, OOK. Прикладом є серія радіомодулів SX12XX[5].

LoRa працює на доступних неліцензійних частотах загального призначення (ISM), наприклад, 915 МГц, 868 МГц і 433 МГц. Однак, існують певні регіональні законодавчі обмеження щодо використання LoRa. Вони визначають які діапазони частот можна використовувати в країні. В Україні дозволеними є діапазони 433.05-434.79 МГц та 863-870 МГц. Також в специфікації LoRa використовується термін «план каналу» або «частотний план». Для України це EU433 та EU863-870[6].

План каналу, або частотний план, — це термін, що позначає географічно специфічні канали та швидкості передачі даних, тобто обсяги даних, які передаються через мережу за певний проміжок часу. Зазначені швидкості передачі та канали мають відповідати Регіональним параметрам LoRaWAN для певної географічної області або частотного діапазо-

ну, тобто визначеного набору електромагнітних частот[7]. Він визначає які частоти дозволено використовувати, скільки каналів є доступними, швидкість передачі даних, а також прив'язує ці параметри до певної ширини смуги пропускання. Як правило, використовують смуги пропускання з шириною 125 КГц, 250 КГц, 500 КГц.

Також частотний план визначає дуже важливий параметр – прогальність. Прогальність - це відрізок часу, або періоду, коли система або сигнал активний. В англійській літературі використовують термін коефіцієнт заповнення (англ. Duty Cycle). Він показує яку частину від загального часу роботи система активно працює.

У контексті радіопередачі – це період часу, коли радіомодуль випромінює сигнал на певному каналі. Якщо половину часу радіомодуль випромінює сигнал, а іншу половину слухає ефір, то коефіцієнт заповнення буде 0.5, або 50%.

Згідно частотного плану EU433, коефіцієнт заповнення має бути менше 10%, тоді як для EU863-870 – менше 1%.

Прогальність суттєво впливає на швидкість передачі за секунду. Швидкість передачі буде розподілятися по кожному каналу. Візьмемо як приклад частотний план EU433 та фактор розповсюдження 10 (табл. 1). Відповідно, LoRa пристрій може передавати на одному каналі лише 10% загального часу, а швидкість передачі при SF10 приблизно дорівнює 580 біт/с, то виходить, що по одному каналу LoRa зможе передавати лише 58 біт за секунду (рис. 2).

Механізм контролю коефіцієнту заповнення базується на вимірюванні тривалості кожної передачі в ефірі та гарантує, що загальний час передачі не перевищує встановлений ліміт у заданому часовому інтервалі. У разі досягнення цього порогу стек блокує подальші передачі на визначений проміжок часу, щоб дотриматись обмежень.

Звичайно, ця швидкість передачі сильно поступається іншим протоколам передачі, як Wi-Fi, Bluetooth, LTE. LoRa обмінює швидкість передачі на дальність та стійкість зв'язку.

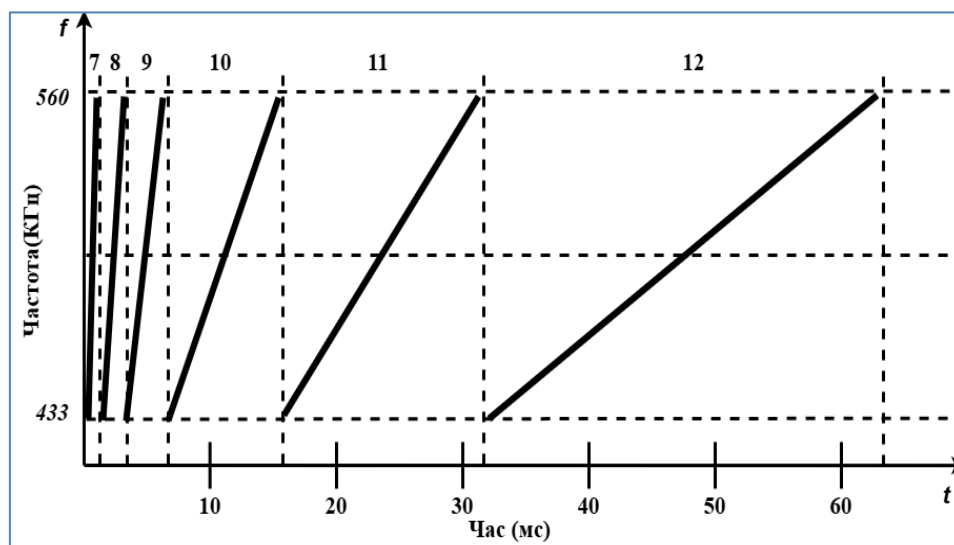


Рис. 2. Порівняння факторів розповсюдження від SF7 по SF12 [8]

Таблиця 1 – Відношення між фактором розповсюдження, пропускну здатністю та часом в повітрі [9]

SF	Навантаження 50 біт		Максимальне навантаження		
	ToA(c)	Пропускна здатність (біт/с)	ToA (c)	Максимальний розмір навантаження(біт)	Максимальна пропускна здатність (біт/с)
7	0.113	3543.1	0.394	242	4913.7
8	0.205	1948.2	0.697	242	2777.6
9	0.369	1082.1	0.677	115	1358.9
10	0.698	572.8	0.698	51	584.5
11	1.478	270.5	1.479	51	275.9
12	2.629	152.1	2.793	51	146.1

Загалом, LoRa є цікавим інструментом для розробки у сфері IoT, забезпечуючи з'єднання в умовах, де інші технології неефективні. Серед інших протоколів він вирізняється дальністю роботи і надійністю зв'язку, що робить його доцільним рішенням для задач, де ці питання є ключовими.

Технологія LoRa визначає лише фізичний рівень мережі згідно моделі OSI. Протокол LoRaWAN, один із кількох можливих варіантів реалізації вищих рівнів, є найбільш поширеним і виконує функції мережевого рівня. Він забезпечує маршрутизацію трафіку між кінцевими пристроями та шлюзами в LPWAN-мережах. Хоча більшість застосунків LoRa використовують LoRaWAN, використання останнього не є обов'язковим — LoRa може працювати і без реалізації мережевого і каналного рівнів.

Мережева архітектура LoRaWAN базується на зіркоподібній структурі, де шлюзи виступають проміжними вузлами, що передають дані між кінцевими пристроями і центральним сервером. Кожен кінцевий пристрій з'єднується зі шлюзом, який має підключення до Інтернету та, зазвичай, підтримує IP-протоколи.

Внаслідок обраної топології та зосередженості на зв'язку "кінцевий вузол – шлюз", LoRaWAN має певні обмеження щодо можливих сценаріїв використання; зокрема, він не підтримує повноцінну комунікацію між вузлами, а отже, не дозволяє реалізувати справжню mesh-мережу[10].

Альтернативною технологією для об'єднання LoRa в мережу та підключення до інтернету є Meshtastic. Проект Meshtastic надає можливість застосовувати бюджетні LoRa-радіопристрої, як автономну платформу для організації далекого зв'язку в місцевостях, де відсутня або нестабільна комунікаційна інфраструктура. Слід відмітити, що Meshtastic розробляється не комерційною компанією чи об'єднанням, на кшталт LoRa Alliance, а вільною спільнотою програмістів. Також, вихідний код для Meshtastic є відкритим.

Meshtastic використовує LoRa-радіопристрої для ретрансляції отриманих повідомлень, утворюючи mesh-мережу. Така конфігурація забезпечує доставку повідомлень усім учасникам групи, включно з тими, хто знаходиться на найбільшій відстані.

Крім того, радіопристрої Meshtastic можуть бути підключені до одного телефону, що дозволяє друзям і членам родини надсилати повідомлення безпосередньо на ваш конкретний пристрій. Важливо зазначити, що кожен пристрій підтримує одночасне підключення лише одного користувача[11].

Приклади застосування LoRa

Впровадження LoRa пристроїв суттєво збільшить ефективність праці в галузях, де використовуються великі площі, а також в умовах погано розвинутої телекомунікаційної інфраструктури. LoRa вже широко використовується в сільському господарстві, логістиці, наукових дослідженнях. Також цікавим є використання LoRa для фото полювання. В умовах бойових дій LoRa теж знайшов свою нішу.

Компанія Semtech впроваджує LoRa для управління водними ресурсами. Вона пропонує Sensoterra, як бездротову та дистанційну систему, яка надає фермерам у реальному часі інформацію про стан вологості ґрунту їхніх культур. Sensoterra є альтернативою для систем, що використовують візуальний метод для аналізу ґрунтів та посівів за допомогою БПЛА та супутників. Sensoterra використовує мережу датчиків, що розкидані по полю і вимірює стан ґрунтів та передає цю інформацію на базову станцію. Це дозволяє ефективніше використовувати воду у системах зрошення, зменшивши використання води до 30%, а також запобігає пошкодженню посівів через перенасичення водою[12].

Схожу систему на основі LoRa впроваджують в Європейському Союзі. За підтримки ЄС проводилось дослідження LoRa на експериментальних фермах в Нормандії. Мета проекту – створення цифрової платформи в сільськогосподарському дослідному господарстві. Реалізовано цифрову мережу, що дозволяє зв'язуватися з підключеними об'єктами через радіозв'язок з Інтернетом. Спеціальні датчики будуть встановлені на господарстві і будуть розташовані на великих відстанях (10 км) від антени. Було створено датчики і різноманітні цифрові інструменти для полегшення ведення рослинництва та тваринництва.

Цей проект спрощує прийняття рішень, забезпечує більшу оперативність, меншу трудомісткість і створює більш безпечні ферми в плані моніторингу

харчових продуктів, здоров'я, відтворення, запліднення [13].

В Австрії, спільно з LoRa Alliance та Ursalink, впроваджували проект цифрового землеробства в Австрії. Цей пілотний проект зосереджений на зборі, зчитуванні та візуалізації даних, отриманих від кількох датчиків LoRaWAN, застосованих у рішенні для управління здоров'ям рослин, оскільки їх потрібно захищати від шкідливих погодних умов, коли вони ще делікатні та особливо схильні до хвороби. По суті, те, що має фундаментальне значення для управління здоров'ям рослин, — це дані про температуру, вологість і сонячне світло. Однак час від часу буде важко перевіряти дані вручну, оскільки погодні умови можуть постійно змінюватися протягом дня, а також відповідно змінюється сонячне світло. В цьому проекті заміняють дорогу дротову мережу на бездротову на базі LoRa. Дані з датчиків будуть візуалізовані, що спростить аналіз та прийняття рішень [14].

Зазвичай, LoRa не вважається придатним для передачі зображень, через його низьку пропускну здатність. Однак, LoRa може бути використаний у програмах, де вам потрібно лише одне зображення на день. Зокрема, перевірка здоров'я вулика — це не те, що потребує уваги щодня. LoRa-пристрій передає зображення вулика декілька раз на тиждень на базову станцію. Потім система використовує машинний зір для розпізнавання бджіл і комах, щоб звітувати про їхню діяльність, а також аналізує стан і здоров'я вулика. Це зменшує потребу перевіряти вулик особисто, оскільки гарне зображення може надати багато інформації про стан вулика [15].

Безпосередньо після стихійного лиха, такого як повінь, лісова пожежа або землетрус, мережі можуть бути перевантаженими або пошкодженими, що робить їх непридатними для використання рятувальниками. У таких ситуаціях використання традиційного підходу з фіксованими шлюзами є неефективним через мобільність рятувальників. Було спроектовано та оцінено дворівневу мережеву систему під назвою LoRaUAV з метою вирішення вищезазначених проблем. LoRaUAV базується на ad hoc Wi-Fi мережі безпілотних літальних апаратів (БПЛА), які виступають у ролі ретрансляторів для трафіку між мобільними вузлами LoRaWAN та віддаленою базовою станцією. Ядром системи є повністю розподілений алгоритм мобільності, який періодично оновлює топологію БПЛА для адаптації до руху наземних вузлів [16].

Війна в Україні з кожним роком стає все технологічною. Технології, розроблені для мирних галузей починають використовувати у військових цілях. LoRa не є виключенням. Було зазначено, що LoRa підходить для використання у надзвичайних ситуаціях, бо розгортання цих пристроїв є швидким і простим. Це дуже корисно в умовах бойових дій, де телекомунікаційна інфраструктура є пошкодженою або відсутньою.

Помічено використання дистанційних засобів ініціалізації та GPS-маячків на базі LoRa в українських БПЛА «Щедрік». LoRa-сенсор в дроні передає

оператору координати свого місця розташування. Дрон може втратити керування, або піддатися впливу засобів РЕБ. Однак, LoRa-сенсор з GPS-приймачем має окреме живлення і залишається активним та передає інформацію про координати дрона, що спрощує його пошук.

Цей модуль знаходиться в сплячому режимі, але з певним інтервалом включається, зчитує з навігаційного супутника координати та відправляє їх на свій шлюз. Дальність дії приймачів LoRa може бути різною, але за сприятливих обставин вона здатна досягати декількох кілометрів. Що ефективно, якщо пролітає поблизу великий дрон, що виконує роль шлюзу.

Припустили використання LoRa і в радіорозвідці. В зоні, де є велика активність БПЛА, встановлюються спеціальні світлофори, які попереджають солдатів про активність дронів. В точці, де є хороша радіовидимість, встановлена станція РЕР, що приймає сигнали від дронів. Тоді як світлофори рознесені по окрузі і мають зв'язок на LoRa з центральним пунктом управління. Коли станція РЕР помічає БПЛА в небі, то передає сигнал на світлофор, вказуючи, який режим йому ввімкнути [17].

Також, LoRa використовується, як маячок для пошуку втрачених БПЛА. Прикладом є розробка інженера Бажена Левковця. Він створив пошуковий GPS-маячок для визначення місця розташування дронів на територіях, де відсутнє мобільне покриття. За словами розробника, маячки дуже важливі в умовах щільної роботи засобів радіоелектронної боротьби, коли безпілотники часто втрачають зв'язок з оператором і їх необхідно знайти. Маячок складається з GPS-модуля (діапазону L1), мікроконтролера STM32 і радіомодуля LoRa на 433 МГц. Живиться пристрій від літій-полімерного акумулятора на 1000 мАг.

Маячок є додатковим модулем до дрона, та змогу знаходити його за допомогою пошукових світлових сигналів, п'єзоелектричного зумера, що видає звуки, або за координатами GPS [18].

Висновок

Технологія LoRa є перспективним рішенням у сфері IoT, що відрізняється низьким енергоспоживанням та великою дальністю зв'язку порівняно з іншими поширеними протоколами. Базуючись на лінійно-частотній модуляції з фактором розповсюдження, LoRa забезпечує стійкий зв'язок навіть у складних умовах, включаючи міську забудову, щільні зелені насадження та радіоперешкоди. Хоча швидкість передачі даних значно поступається таким протоколам як WiFi чи Bluetooth, ця технологія ефективно виконує завдання, де пріоритетом є надійний зв'язок на великих відстанях при мінімальному енергоспоживанні.

Практичне застосування LoRa охоплює широкий спектр галузей: від сільського господарства (системи моніторингу вологості ґрунту, датчики для тваринництва) до надзвичайних ситуацій і військової сфери (пошукові маячки для БПЛА, системи попередження про дроніву небезпеку). Особливо

цінною ця технологія стає в умовах відсутньої або пошкодженої телекомунікаційної інфраструктури, що робить її незамінною для екстремальних умов і віддалених районів, де традиційні методи зв'язку неефективні.

СПИСОК ЛІТЕРАТУРИ

1. Lin K., Hao T., Zheng W., He W. Analysis of LoRa Link Quality for Underwater Wireless Sensor Networks: A Semi-empirical Study. IEEE Asia-Pacific Microwave Conference (APMC). Singapore, 2019. pp. 120-122. DOI: 10.1109/APMC46564.2019.9038666
2. Burden T. What's CHIRP Sonar? WestMarine. 07.01.2025. URL: <https://www.westmarine.com/west-advisor-electronics/Understanding-CHIRP-Scanning-Sonar.html>.
3. Shilpa B., Gupta H.P., Jha R.K. et al. LoRa interference issues and solution approaches in dense IoT networks: a review. Telecommunication Systems. 2024. Vol. 87. pp. 517–539. DOI: 10.1007/s11235-024-01192-9
4. TEKTELIC Communications Inc. Spreading Factor (SF). TEKTELIC Communications. URL: <https://tektelic.com/what-it-is/spreading-factor/>
5. Semtech Corporation. SX1276. ICs for Sensors/Trackers. URL: <https://www.semtech.com/products/wireless-rf/lora-connect/sx1276>
6. LoRaWAN® Regional Parameters. LoRa Alliance, Inc. 2020. 90 c. URL: <https://lora-alliance.org/wp-content/uploads/2021/05/RP-2-1.0.3.pdf>
7. TEKTELIC Communications Inc. Channel Plan. TEKTELIC Communications. URL: <https://tektelic.com/what-it-is/channel-plan/>
8. Yalcin S. An artificial intelligence-based spectrum sensing methodology for LoRa and cognitive radio networks. International Journal of Communication Systems. 2023. Vol. 36. DOI: 10.1002/dac.5433
9. Cotrim J.R., Kleinschmidt J.H. LoRaWAN Mesh Networks: A Review and Classification of Multihop Communication. Sensors. 2020. Vol. 20. 4273. DOI: 10.3390/s20154273
10. Manzoni P., Merzougui S.E., Palazzi C.E., Pozzan P. A Resilient LoRa-Based Solution to Support Pervasive Sensing. Electronics. 2023. Vol. 12. 2952. DOI: 10.3390/electronics12132952
11. Introduction | Meshtastic. Meshtastic. URL: <https://meshtastic.org/docs/introduction/>
12. Marketing S. C. Conserving resources and reducing agriculture water waste. Inside Out | Semtech's Corporate Blog. URL: <https://blog.semtech.com/improving-nature-restoration-and-eliminating-agriculture-water-waste-with-lora-technology>
13. Experimental digital platform LoRa in agriculture | EU CAP Network. Welcome to the EU CAP Network | EU CAP Network. URL: https://eu-cap-network.ec.europa.eu/projects/experimental-digital-platform-lora-agriculture_en#tab_id=overview
14. LoRa Alliance. Experimental digital platform LoRa in agriculture. LoRa Alliance. URL: <https://lora-alliance.org/wp-content/uploads/2019/04/Digital-Farming-Is-Creating-a-More-Plentiful-Sustainable-Food-System.pdf>
15. LoRa Transmission of Images. Element14. URL: <https://community.element14.com/challenges-projects/design-challenges/save-the-bees-design-challenge/b/blog/posts/lora-transmission-of-images>
16. Stellin M., Sabino S., Grilo A. LoRaWAN Networking in Mobile Scenarios Using a WiFi Mesh of UAV Gateways. Electronics. 2020. Vol. 9, no. 4. P. 630. URL: <https://doi.org/10.3390/electronics9040630>
17. Бескредтнов С. Флеш. Facebook. URL: https://www.facebook.com/story.php?story_fbid=8824964230905216&id=100001751811701
18. Бойко П. Скоротить втрати дронів. ФОКУС. URL: <https://focus.ua/uk/digital/587145-skorotit-vtrati-droniv-tvorec-novoyi-tehnologiyi-dlya-zsu-postrazhdav-u-chernigovi-foto>

Received (Надійшла) 18.03.2025

Accepted for publication (Прийнята до друку) 28.05.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Руденко Віталій Віталійович – асистент кафедри автоматизації, електроніки та телекомунікацій, Національний університет "Полтавська політехніка імені Юрія Кондратюка", Полтава, Україна;

Vitalii Rudenko – Assistant of Department of Automation, Electronics and Telecommunications, National University "Yuri Kondratyuk Poltava Polytechnic", Poltava, Ukraine;

e-mail: vitaliirudenko40@gmail.com; ORCID Author ID: <https://orcid.org/0009-0006-9094-222X>.

Overview of LoRa protocol and perspectives for its use

Vitalii Rudenko

Abstract. This article examines LoRa (Long Range) wireless communication technology as a low-power, long-range data transmission protocol. It highlights the operating principles of the technology based on chirp spread spectrum modulation with spreading factors that ensure communication reliability even in conditions of radio interference and physical obstacles. **The aim of the article** is to analyze the features of the LoRa protocol operation, investigate network solutions based on it, and review practical applications of the technology in various industries. **Results obtained:** the presented quantitative indicators regarding the relationship between spreading factor, bandwidth, and data transmission time demonstrate the technology's adaptability to various tasks where reliable communication, rather than transmission speed, is the priority. **Conclusions:** implementation of LoRa-based devices significantly increases efficiency in industries covering large areas and in conditions of underdeveloped telecommunications infrastructure, particularly in agriculture, logistics, emergency situations, and combat operations.

Keywords: LoRa, wireless communication, IoT, chirp spread spectrum modulation, spreading factor, duty cycle, network architecture, ToA.

Borys Sadovnykov, Oleksandr Zhuchenko

Ukrainian State University of Railway Transport, Kharkiv, Ukraine

A METHOD FOR SEARCHING AND RECOGNISING OBJECTS IN A VIDEO STREAM BY CALCULATING INTERFRAME DELTAS

Abstract. The article proposes an improved method for searching and recognising objects in a video stream in real time using the calculation of interframe changes (deltas) and a neural classifier. The main goal of the study is to achieve high performance and reduce the computational load on system resources while maintaining acceptable accuracy. An experimental comparison with the basic SSD (Single Shot MultiBox Detector) method was carried out, which measured the following indicators: average frame processing time, RAM and video memory usage, CPU and graphics load, and recognition accuracy. Unlike SSDs, the proposed approach provides a higher processing speed (up to 35% increase) with a slight decrease in accuracy (less than 4%), which is compensated for by further adaptation of the model. At the same time, the use of the CPU and RAM increases by only 0.5-5%, while the amount of video memory consumed decreases by 57%. The study confirms the feasibility of using the improved delta classification method in video analytics systems with limited resources. This method can be integrated into applied security, video surveillance, and real-time intelligent monitoring systems.

Keywords: machine learning, computer vision, image processing, convolutional neural networks, visual image recognition, telecommunication systems.

Introduction

Statement of a scientific problem. Computer vision is a promising area for the development of visual information analysis technologies. One example of the use of such systems is the recognition of road signs in driver assistance systems, which allows for increased road safety by automatically detecting speed limits, pedestrian crossings and other objects. Another example is automated farmland monitoring systems that analyse drone imagery to detect pests or determine the health of crops. Recognition and classification of images in static frames or video streams is a complex but extremely important task. Such technologies are widely used in the security sector, for example, in access control systems where face recognition is used to identify users, and in logistics to track goods through automatic analysis of barcodes and QR codes.

Research analysis. An analysis of current approaches to video stream processing [1–15] confirms the effectiveness of convolutional neural networks (CNNs) for object recognition in visual data. Studies [1, 4, 11, 14] demonstrate high detection accuracy, particularly on large-scale datasets; however, most of these methods are focused on static images or demand substantial computational resources. Real-time detection methods such as SSD (Single Shot MultiBox Detector) [4] and optimized YOLO variants offer satisfactory performance but often impose high GPU loads and require significant video memory, which limits their deployment in resource-constrained environments [9].

Several works [3, 8, 12] integrate detection, tracking, and classification of objects in video, but rarely include early-stage optimization based on frame differencing. In particular, [8] proposes a comprehensive system for automated surveillance, yet lacks dedicated mechanisms for computational cost reduction. Notably, methods based on inter-frame delta computation (frame differencing) have proven effective as a lightweight preprocessing step prior to classification. In [3, 7, 13, 15], it is shown that analyzing differences between successive frames enables accurate detection of moving objects

while significantly reducing the computational burden. In [3], an improved frame difference technique is introduced to mitigate noise and lighting variations, reinforcing the suitability of such methods for real-time applications.

Research on morphological image processing [10] contributes to improving the quality of detected regions resulting from frame differencing. The integration of these techniques with neural network classifiers, as demonstrated in [6, 7], allows for an optimal balance between processing speed and recognition accuracy.

Meanwhile, existing reviews of thresholding and segmentation techniques [2, 5] often overlook the dynamics of video streams, such as complex backgrounds, ambient noise, or partial occlusion, limiting their applicability in real-world security and monitoring systems. Therefore, the surveyed methods do not offer a universal solution for high-speed and resource-efficient object recognition in video. This highlights the need to develop a new approach that combines inter-frame change analysis (delta filtering) with adaptive neural classification, optimized for real-time operation and low-power hardware platforms.

The purpose of this work is to achieve high performance and reduce the computational load on system resources while maintaining acceptable accuracy.

Presentation of the main material and substantiation of the obtained research results

The proposed method is complex and consists of a sequence of preliminary linear and morphological transformations of the input image. The algorithm for implementing the method is shown in Fig. 1. It consists of the following sequence of steps.

Step 1. Getting a new frame. The current frame is extracted from the video stream for further processing. This process is performed cyclically in real time or at a fixed frame rate, ensuring continuous monitoring of changes in the scene.

Step 2. Resizing the image to a uniform size. The second step of the algorithm is to bring the image size to a single standard, namely 1280x720 (720p), which will be used in all subsequent calculations.

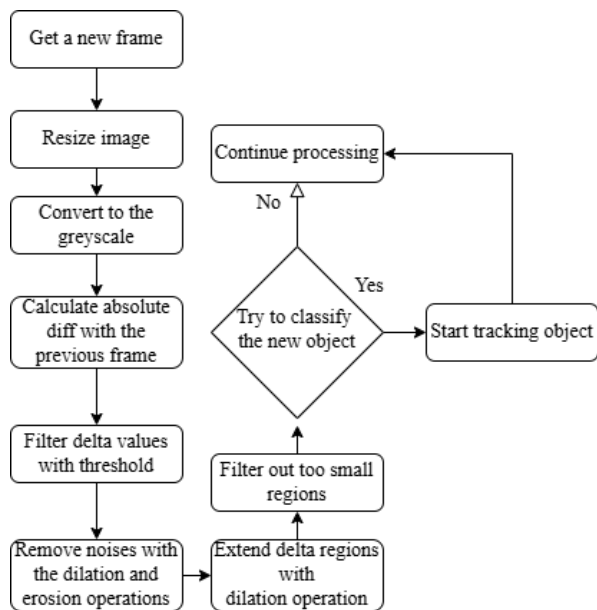


Fig. 1. Block diagram of the algorithm for searching and classifying objects in a video stream

This size was chosen because it reduces the amount of data to be processed while maintaining the quality of the input image. The speed of linear image transformations is directly proportional to the image size [7], as the number of processed pixels changes with the size. Most neural networks require a specific image size in part to speed up processing, as a large image can take exponentially longer to process, for example, SSD requires an image of 300x300 pixels as input. The next advantage of setting a single image size is the stability of processing time regardless of the initial video resolution. Without a uniform size, the speed of the algorithm varies depending on the number of pixels in the frame, which makes it difficult to compare with other methods and affects reliability in practical applications. In addition, setting a fixed image size ensures stable processing time and adapts the method to limited hardware resources. In cases where faster processing speeds are required or hardware resources are limited, it is possible to reduce the input image size to a lower resolution and thus speed up processing.

Step 3. Convert to grayscale. The image is converted to grayscale, which reduces the data size; speeds up processing; reduces the impact of colour interference on the analysis. In addition, converting an image to grayscale improves the accuracy of algorithms by allowing them to focus on the key features of objects without being distracted by colour nuances. This simplification of the colour model offers a number of significant advantages in terms of efficient resource use. In standard colour images, each pixel is defined by three values (R, G, B - red, green, blue), while in grayscale images, it is defined by only one. This significantly reduces the amount of data that needs to be processed, which in turn simplifies calculations, reduces CPU load and optimises memory usage. This solution is especially effective when working with large image arrays, such as a video stream, or in devices with limited computing capabilities, such as CCTV cameras or embedded real-time systems.

Step 4. Calculating the absolute difference between frames. The following feature of the video stream is used to detect objects: each subsequent frame is guaranteed to reflect the state after the previous one, and this order is strict, i.e. if an object gradually appears in the field of view, this object will continue to move in each of the frames. Thus, if an object appears between two frames, calculating the absolute difference between them will allow you to get the exact region of the image.

$$D(x, y) = |I_t(x, y) - I_{t-k}(x, y)|, \quad (1)$$

where $I_t(x, y)$ – pixel value in the current frame; $I_{t-k}(x, y)$ – in the previous frame (removed by k frames); $D(x, y)$ – difference value.

Unlike the SSD method, the object is detected and classified at different stages, not all at once. The absolute difference between frames also depends on the image size, so the smaller the number of pixels in the image, the faster this operation is. Another important way to optimise is to calculate the difference between frames with several skipped frames, i.e. not the current and the previous frame, but the current frame and the frame that was 5 frames ago. In this way, there can be much more changes, for example, an object will appear completely in the image. An important assumption is that the camera is static and there are no additional changes in the video stream due to camera movement.

Step 5. Threshold difference filtering. There may be changes between frames due to lighting, environmental conditions, etc. The absolute difference between pixels provides a digital value of how much this pixel has changed. In the case of lighting, the changes will be minimal and can be filtered out by discarding all difference values below a certain threshold [2].

$$B(x, y) = \begin{cases} 1, & \text{if } D(x, y) > T; \\ 0, & \text{if } D(x, y) \leq T, \end{cases} \quad (2)$$

where T – threshold value that determines the sensitivity of the detector. This way, you can change the sensitivity of the detector - the higher the threshold value, the less sensitive the detector and the greater the changes between frames required to highlight these deltas. As a result of these operations, the resulting regions may have voids inside, unclosed contours, or gaps in the contours. Therefore, the next step is to remove noise and correct the resulting region.

Step 6. Morphological processing (erosion + dilation). To eliminate distortions, contour gaps, and noise artefacts in the selected regions of changes, successive morphological transformations are applied - opening and closing [9].

$$\text{opening: } A \circ B = (A \ominus B) \oplus B, \quad (3)$$

$$\text{closing: } A \cdot B = (A \oplus B) \ominus B, \quad (4)$$

where A – image, B – структурный элемент (ядро), $\ominus$ – erosion, $\oplus$ – dilation.

These operations are applied to the binary mask to correct the shape of the selected regions.

Opening ($A \circ B$) removes small noise or isolated pixels that do not belong to the object, while closing ($A \cdot B$), on the contrary, fills the voids within the region and connects the fragmented parts.

The operations are performed using the structural element (kernel) B , which is a small circle or rectangle that slides across the image.

Erosion ($A \ominus B$) - 'narrows' the object by removing pixels at the edges.

Dilation ($A \oplus B$), on the contrary, 'widens' the object by adding pixels around the existing ones.

The size and shape of the kernel determine how much these transformations will affect the object: the larger the kernel, the more coarse the changes.

The opening operation is performed as a sequence of erosion and dilation, while the closing operation is the reverse.

The result is a binary mask, where only the presence or absence of a pixel in a region is important, not its specific value. Each morphological transformation is performed using a kernel. The size of the kernel determines the degree of impact of the operation: too small may have no effect, while a large one increases computational costs and risks changing the shape of the object.

Step 7. Expanding active regions

Dilation with a larger kernel is used to cover parts of the object that remain static, thus fully shaping the object. However, the movement of the object may be partial, so the resulting region captures only a part of the object, which can make classification impossible. To compensate for this, you can use an extension with a large kernel size, so that the nearest neighbourhood including the stationary parts of the current object is added to the region.

Step 8. Remove regions that are too small

The next step is to filter the resulting regions by size. The previous transformations remove most of the interference, but there are still regions with minor changes that are so small that they cannot physically contain an object inside. Such regions are excluded because there is no point in further processing them.

Step 9. Classify new objects

Since all the interference and regions that cannot contain objects have been removed, object recognition can be attempted. Each region is passed to the input of a classifier based on a neural network[10]. The current implementation of the method uses a classifier based on the MobiNETv1 architecture. There are no specific requirements for the classifier; the choice of the model is based on specific performance requirements, available system resources, and target accuracy.

Step 10. Tracking the object

When the object is found, its position is transferred to the tracking algorithm, so the full search and recognition procedure can be skipped and the current position of the object can be obtained by processing the frames with the tracking function. Every 10 frames, the tracker checks whether the object is located in the received region, in case of a false positive, the tracker state is reset and the search starts from the beginning in the next frame.

The main goal of the tracking algorithm is to keep the object detected when it is not moving or moving slightly. The current implementation uses the simple and fast MOSSE algorithm.

The proposed implementation algorithm has an iterative (cyclic) structure: after each frame, a full processing cycle is performed (Step 1 → Step 9). If the object is not found or classified, processing continues with the next frame (return to Step 1). If the object is recognised, tracking is started (Step 10), but the cycle remains active, i.e. every few frames the relevance of tracking is checked, and if necessary, the algorithm returns to the search mode.

Thus, the proposed method combines the advantages of simple but fast linear and morphological transformations with the accuracy of neural network classification.

It is characterised by stable processing time due to the unification of frame size and an optimised set of pre-computations, which ensure fast detection of moving objects with minimal consumption of GPU and video memory resources compared to classical neural network approaches such as SSD.

The method is suitable for use in resource-constrained environments, such as embedded real-time systems, video surveillance cameras, or devices with low computing capabilities, where high processing speed, moderate resource consumption, and algorithm stability are important.

Experimental Setup

In order to measure the characteristics of the method, a series of experiments will be conducted in which video from a static camera will be processed and the following indicators will be collected:

1. Average time to process one frame (in ms). This indicator determines the suitability of the method for real-time operation.
2. CPU, GPU, and RAM usage during operation. These are the main indicators that give an idea of the use of system resources.
3. The number of successfully found objects. The absolute value is necessary for a relative comparison with analogues.

Similar metrics were obtained for the SSD method for comparison purposes. The summary results of the experiment with the refined statistical indicators are shown in Table 1.

As can be seen from Table 1, the method of calculating interframe deltas outperforms SSD in terms of detection speed by ~8% and has significantly lower video memory consumption (by 57%), but slightly inferior in accuracy (by 2.8%).

In particular, the difference in GPU usage was verified by an additional series of 15 experimental measurements. The average GPU usage was obtained: $1.1\% \pm 0.05\%$ for the delta difference method and $0.9\% \pm 0.05\%$ for SSD. Statistical analysis (Student's t-test, $p < 0.05$) confirmed that the difference of 0.2% is stable, although close to the measurement error limit. This can be explained by the additional linear operations (e.g., scaling, convolution, dilatation, and erosion) that are optimally performed by the GPU.

Thus, these statistics confirm that although the difference in GPU usage is small, it is scientifically justified.

Table 1 – Comparative performance metrics of object detection methods

Parameter	Unit	Deltas Difference Method	SSD (Single Shot Detector)	Comment
Average Detection Time	ms	$5,4 \pm 0,2$	$5,8 \pm 0,2$	Deltas method is ~8% faster
Accuracy	%	$71,2 \pm 1,0$	$74,0 \pm 1,0$	SSD more accurate by 2.8%, but with higher resource usage
CPU Usage	%	$11,5 \pm 0,5$	$9,0 \pm 0,5$	Deltas method consumes more CPU due to pre-processing outside CNN
RAM Usage	MB	468 ± 10	445 ± 10	Higher RAM consumption due to frame differencing operations
GPU Usage	%	$1,1 \pm 0,05$	$0,9 \pm 0,05$	Slightly higher GPU load due to image transformations
VRAM Usage	MB	390 ± 15	610 ± 15	SSD consumes 57% more video memory

Figs. 2–7 visualise the main performance indicators of the two methods based on the data from Table 1.

Based on the measurements of the processing time of each frame, the average value was calculated and a graph was plotted for comparison with SSD (Fig. 2):

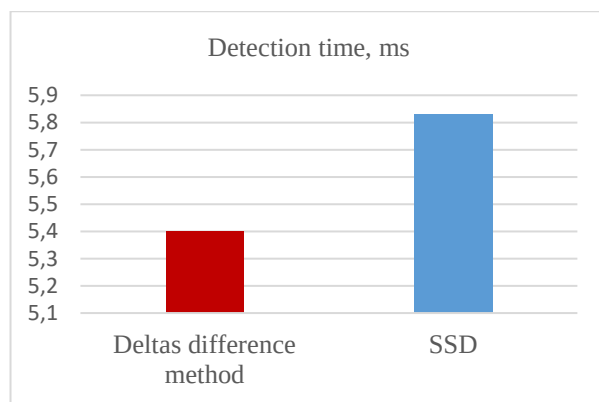


Fig. 2. Average processing time per frame

The visualisation demonstrates that the delta calculation method performs, on average, 0.4 milliseconds faster per frame compared to the SSD approach, representing an 8% improvement in processing speed.

In addition to speed, recognition accuracy remains a key performance indicator. The following graph (Fig. 3) presents a comparison of accuracy between the two methods, highlighting the practical implications of using delta-based preprocessing. While performance is essential, the balance between speed and accuracy often determines the suitability of a method for real-world use cases.

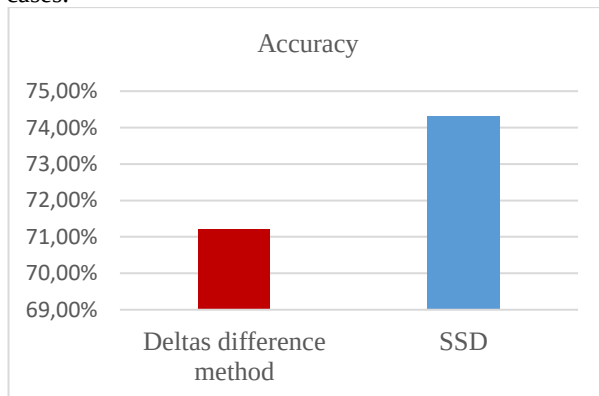


Fig. 3. Recognition accuracy

According to the experimental results, the new method lags behind the reference method by 2.8% in accuracy. Next, let's look at the use of processor resources during recognition (Fig. 4).

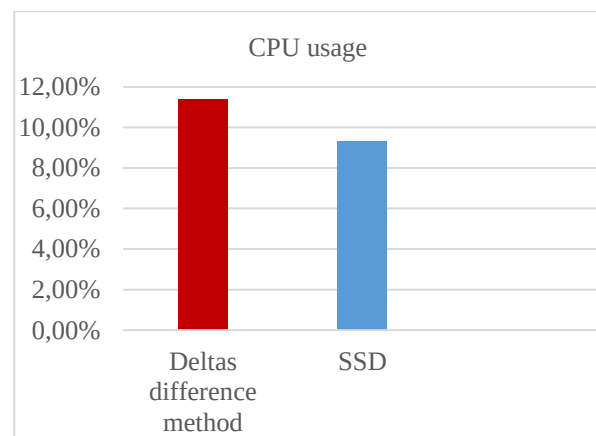


Fig. 4. Processor resource utilisation

The method of calculating deltas involves a larger number of preprocessing steps performed outside the neural network, such as noise removal and region filtering, which both contribute to the overall computational load on the CPU. Unlike operations that are fully offloaded to the GPU, these stages require sequential processing and are less parallelizable, making the CPU a critical bottleneck in this part of the pipeline. As a result, a significant portion of the workload is shifted from the GPU to the CPU, leading to increased CPU utilization. This is particularly evident during the phases of morphological filtering and segmentation refinement. The following graph illustrates RAM usage during the execution of this method (Fig. 5).

Due to the higher CPU usage, more data is stored in RAM. The next parameter in Fig. 6 is the use of video card processing time.

As can be seen from Fig. 6, the difference in GPU usage between the methods (1.1% for the delta difference method vs. 0.9% for SSD) is insignificant and close to the measurement error limit. To confirm this difference, a series of 15 repeated measurements was additionally performed.

The statistical processing resulted in the mean value (Mean) and root mean square deviation (RMS), which are shown in Table 2.

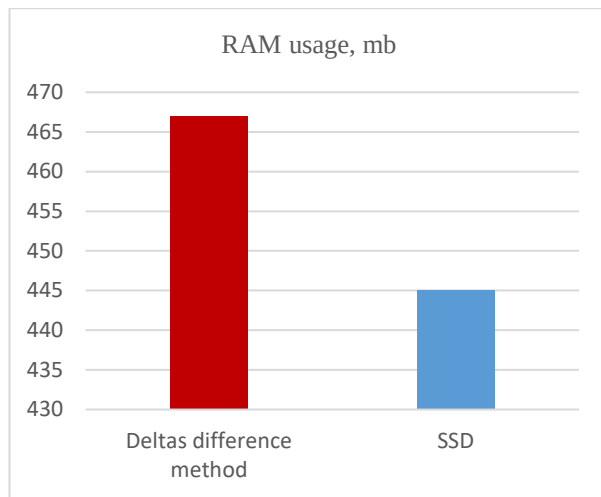


Fig. 5. RAM usage

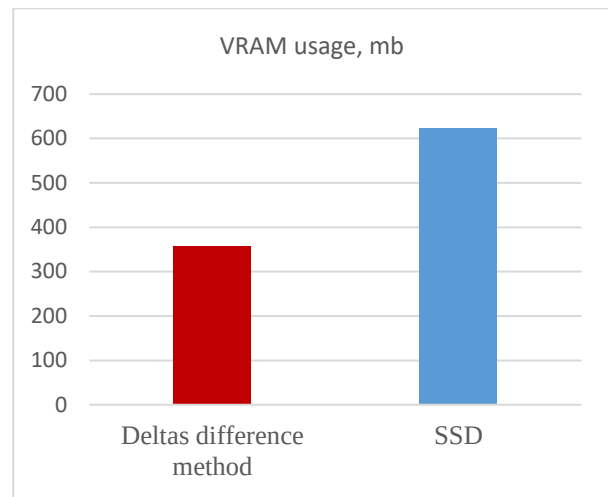


Fig. 7. VRAM usage

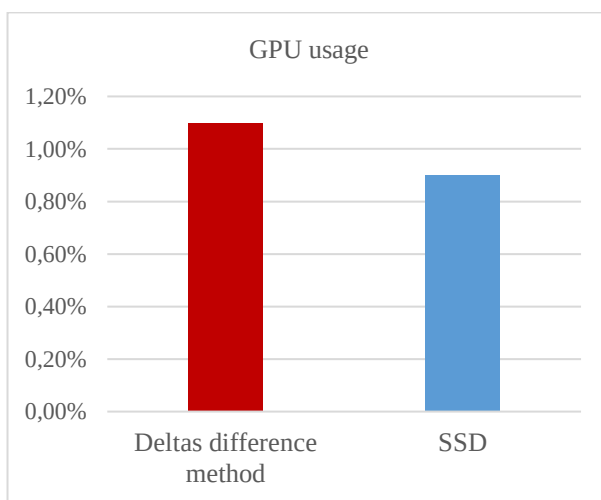


Fig. 6. GPU usage

Table 2 – GPU Usage Statistical Validation
(15 measurements)

Method	Mean GPU Usage, %	RMS, %	Statistical significance
Deltas Difference Method	1,10	±0,05	statistically significant (p < 0,05)
SSD	0,90	±0,05	reference

The difference was tested for statistical significance using Student's t-test at a significance level of $p < 0.05$. The analysis was conducted on the basis of 30 independent experiments, which ensures that the sample is sufficiently representative to draw conclusions.

Despite the relatively small absolute difference between the mean values, the results of the statistical test confirmed its stability and reliability, which is explained by the presence of additional linear operations (dilation, erosion, frame difference) that are efficiently performed in parallel on the GPU.

Consider the video memory usage in the following graph (Fig. 7):

The SSD algorithm uses 57% percent more video memory than the delta-calculated recognition algorithm.

This significantly higher VRAM consumption of the SSD method (+57%) is due to the use of a neural network model with a large depth and number of parameters, which requires a significant amount of memory to store intermediate data and the model.

The interframe delta method uses simpler processing operations, which significantly reduces the VRAM requirement, so the difference seems logical and fully corresponds to the peculiarities of the two compared methods.

Conclusions and prospects for further research

The experimental study proved that the proposed method for detecting moving objects based on the calculation of interframe deltas with subsequent linear and morphological transformations of the input image is highly efficient for real-time applications with limited hardware resources.

The main advantages of this method are high performance (~8% faster compared to the SSD method), low video memory usage (57% reduction), and stable GPU load (0.2% difference compared to SSD and tested for statistical significance using Student's t-test at a significance level of $p < 0.05$), which is achieved by optimising the sequence of preliminary computational operations.

Further research could be aimed at:

- optimising the choice of thresholds and adaptive tuning of morphological transformations parameters,
- integrating with modern compact neural network architectures to improve accuracy,
- extending the algorithm to work with moving cameras and more complex dynamic environment scenarios,

– conducting a detailed statistical analysis of the results using additional metrics (skewness, variance) to more fully assess the efficiency and stability of the method.

REFERENCES

1. Cui W., Zhang Y., Zhang X., Li L., Liou F. (2020) Metal Additive Manufacturing Parts Inspection Using Convolutional Neural Network, *Applied Sciences* 10(2), 545; <https://doi.org/10.3390/app10020545>
2. Guruprasad P. (2020) Overview of different thresholding methods in image processing, Conference: TEQIP Sponsored 3rd National Conference on ETACC
3. Huang W., Kang Y., Zheng S. (2017) An improved frame difference method for moving target detection. 2017 IEEE 16th International Conference on Cognitive Informatics & Cognitive Computing, DOI: 10.1109/ICCI-CC.2017.8109746.
4. Liu W., Anguelov D., Erhan D., Szegedy C., Reed S., Fu C., Berg A. (2016) SSD: Single Shot MultiBox Detector, *Computer Vision and Pattern Recognition*, P. 21–37, https://doi.org/10.1007/978-3-319-46448-0_2
5. Lysechko V., Syvolovskyi I., Komar O., Nikitska A., Cherneva G.: Research of modern NoSQL databases to simplify the process of their design. *Academic journal: Mechanics Transport Communications*, 2023, vol. 21, issue 2, article №2363, ISSN 2367-6620.
6. Lysechko V., Zorina O., Sadovnykov B., Cherneva G., Pastushenko V.: Experimental study of optimized face recognition algorithms for resource – constrained. *Mechanics Transport Communications*, 2023, vol. 21, issue 1, article №2343.
7. Lysechko V.P., Sadovnykov B.I., Komar O.M., Zhuchenko O.S. (2024) A research of the latest approaches to visual image recognition and classification. 2024, *National University «Zaporizhzhia Polytechnic». Radio Electronics, Computer Science, Control*, 1(68), P. 140-147, DOI 10.15588/1607-3274-2024-1-13.
8. Mohana, Ravish Aradhya H.V. (2022) Design and Implementation of Object Detection, Tracking, Counting and Classification Algorithms using Artificial Intelligence for Automated Video Surveillance Applications, Conference: 24th International Conference on Advanced Computing and Communications, 2022.
9. Ravi N., El-Sharkawy M. (2022) Real-Time Embedded Implementation of Improved ObjectDetector for Resource-Constrained Devices. *Journal of Low Power Electronics and Applications* 12(2):21, April 2022, DOI:10.3390/jlpea12020021.
10. Rodriguez, J.; Ayala, D. (2001) Erosion and Dilation on 2D and 3D Digital Images: A new size-independent approach. In *Proceedings of the Vision Modeling & Visualization Conference*, Stuttgart, Germany,
11. Simonyan K., Zisserman A. (2014) Very Deep Convolutional Networks for Large-Scale Image Recognition, <https://doi.org/10.48550/arXiv.1409.1556>
12. Singh, B., et al. (2014). Motion Detection for Video Surveillance. In *Proceedings of the 2014 International Conference on Signal Propagation and Computer Technology (ICSPCT)* (pp. 592–597). IEEE. DOI:10.1109/ICSPCT.2014.6884919.
13. Yue W., Liu S., Li Y. (2023) Eff-PCNet: An Efficient Pure CNN Network for Medical Image Classification, *Applied Sciences* 13(16):9226, <https://doi.org/10.3390/app13169226>
14. Zhao, Z. et al. (2019). Object Detection With Deep Learning: A Review. *IEEE Transactions on Neural Networks and Learning Systems*, 30(11), 3212–3232. DOI: 10.1109/TNNLS.2018.2876865.
15. Пуйда В. Я., Стоян А. О. (2020) Дослідження методів виявлення об'єктів на відеозображеннях. Комп'ютерні системи та мережі. Vol. 2, No. 1, С. 80-87, 2020, <https://doi.org/10.23939/csn2020.01.080>.

Received (Надійшла) 18.03.2025

Accepted for publication (Прийнята до друку) 11.06.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Садовников Борис Ігорович – аспірант кафедри транспортного зв'язку, Український державний університет залізничного транспорту, Харків, Україна;

Borys Sadovnykov – PhD student, Department Transport Communication, Ukraine State University Of Railway Transport, Kharkiv, Ukraine;

e-mail: sadovnykov@kart.edu.ua; ORCID Author ID: <https://orcid.org/0009-0009-4180-2863>.

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 the Department of Transport Communications, Ukraine State University Of Railway Transport, Kharkiv, Ukraine;

e-mail: o030201@gmail.com; ORCID Author ID: <https://orcid.org/0000-0003-3275-810X>.

A method for searching and recognising objects in a video stream by calculating interframe deltas

Б. І. Садовников, О. С. Жученко

Анотація. У статті запропоновано удосконалений метод пошуку та розпізнавання об'єктів у відеопотоці в режимі реального часу з використанням обчислення міжкадрових змін (дельта) та нейронного класифікатора. Основною метою дослідження є досягнення високої швидкодії та зменшення обчислювального навантаження на ресурси системи за умови збереження прийнятної точності. Проведено експериментальне порівняння із базовим методом SSD (Single Shot MultiBox Detector), у межах якого вимірювалися показники: середній час обробки кадру, використання оперативної та відеопам'яті, процесорне та графічне навантаження, а також точність розпізнавання. На відміну від SSD, запропонований підхід забезпечує вищу швидкість обробки (до 35% приросту) при незначному зменшенні точності (менше 4%), що компенсується подальшою адаптацією моделі. При цьому використання центрального процесора та ОЗП зростає лише на 0,5–5%, натомість обсяг споживаної відеопам'яті зменшується на 57%. Дослідження підтверджують доцільність застосування удосконаленого методу дельта-класифікації в системах відеоаналітики з обмеженими ресурсами. Наведений метод може бути інтегрований у прикладні системи безпеки, відеонагляду та інтелектуального моніторингу в реальному часі.

Ключові слова: машинне навчання, комп'ютерний зір, обробка зображень, згорточні нейронні мережі, візуальне розпізнавання зображень, телекомунікаційні системи.

Illia Syvolovskyi¹, Volodymyr Lysechko²

¹ Ukrainian State University of Railway Transport, Kharkiv, Ukraine

² Ivan Kozhedub Kharkiv National University of Air Forces, Kharkiv, Ukraine

A METHOD OF HIERARCHICAL CLUSTERING OF NODES IN DISTRIBUTED TELECOMMUNICATION SYSTEMS USING GRAPH ALGORITHMS

Abstract. This article describes a modified method for hierarchical clustering of computing nodes in distributed telecommunication systems, considering node performance, network topology, delays, and communication channel bandwidth. The proposed method is based on a modified Louvain algorithm that performs multi-step graph clustering with dynamic parameter adjustment. This makes it possible to control the size of clusters and their internal density in accordance with the specified targets, minimizing the fragmentation of the network structure. Based on a comparative analysis of modern clustering methods and experimental modeling, it has been found that the proposed method reduces cluster fragmentation by 36% compared to the Leiden method. In addition, it reduces inter-cluster delays by 4,5% compared to the Louvain method and by 11,8% compared to Leiden, which indicates a more efficient organization of inter-cluster interaction. The improved method ensures an even distribution of computing nodes among clusters, which helps to optimize the overall performance of a distributed telecommunications system.

Keywords: distributed telecommunication systems, hierarchy, cluster, node, modeling, intelligent systems, graphs, algorithms, optimization, Louvain, Leiden, throughput, topology, delay minimization, dynamic self-organization,

Introduction

Statement of a scientific problem. With the increasing complexity of distributed computing systems and telecommunication networks, there is a growing need to efficiently organize computing nodes and minimize data transmission delays. The main challenges include the heterogeneity of node performance, variability of network connection parameters (latency, bandwidth), and dynamic changes in the network topology that affect the stability and performance of distributed computing processes.

Traditional clustering approaches do not fully take into account the specifics of the network topology and connectivity parameters, which can lead to irrational workload distribution and increased inter-cluster delays. This is especially significant in telecommunications systems, where the efficient organization of computing nodes affects the performance of the entire network.

Thus, there is a need to develop a clustering method that takes into account not only the modularity of the graph but also the key characteristics of connectivity between nodes, which will reduce latency, improve load balancing, and ensure adaptability in changing network conditions.

Research analysis. An analysis of modern clustering methods [1-18] confirms the effectiveness of graph algorithms for organizing distributed telecommunication systems. Studies [3, 5, 6, 7, 10, 11] confirm the effectiveness of distributed computing, but they are mostly focused on static models. In [14, 15], methods for minimizing delays are studied, but without taking into account dynamic changes in the network topology.

Studies of load balancing [1, 12] are related to mobile environments, but do not cover the issue of flexible node clustering.

The Louvain clustering method [2, 4] proves effective in detecting clusters, but has limitations in terms of resolution and cluster separation. In [14], the Leiden algorithm is proposed, which eliminates this

drawback, but is computationally more complex.

Service placement methods [6, 7, 10, 16] are aimed at reducing data processing delays, but do not take into account the variability of nodes and the physical limitations of network topologies. Some cloud-based methods [9] have potential application in distributed systems (at higher processing levels close to the cloud), but are not adapted to clustering, which makes it impossible to effectively apply them on a large number of nodes.

Thus, the existing methods do not provide a comprehensive solution to the problem of adaptive clustering considering the network parameters. This necessitates the development of a new method that combines graph optimization and adaptive grouping of nodes according to specified targets.

The purpose of the work is to develop a method for hierarchical clustering of computing nodes in distributed telecommunication systems based on a modified Louvain algorithm with adaptive parameter adjustment. This ensures the formation of clusters of optimal size and connectivity, contributing to a balanced allocation of resources and minimizing inter-cluster delays.

Presentation of the main material and substantiation of the obtained research results

Considering the principles of multi-level telecommunication systems and the analysis of modern clustering methods [2, 4, 6, 12, 14], a hierarchical clustering method of nodes aimed at optimizing the network structure has been developed. The features of the proposed clustering method include:

1. Task-Oriented Clustering. Nodes are grouped according to the tasks they can perform [7, 10]. Each node can efficiently handle a certain task or class of tasks due to its hardware characteristics. The clustering method provides a way to form groups of nodes that can jointly perform one or more tasks, taking into account their performance.

2. Hierarchical (multi-level) structure of clusters (Hierarchical Cluster Organization). The clustering method involves multi-level clustering [3, 5, 12], due to the iterative separation of the graph structure with adaptive parameter settings.

3. Optimization of intra-cluster and inter-cluster interaction. It is focused on minimizing delays inside and outside the clusters, which is important when processing data in real-time systems [14, 15].

4. Use of weighted link parameters (Latency & Bandwidth-Aware Clustering). Since the geographical location of the nodes is not defined, the optimization of interaction is based on the latency and bandwidth of

communication channels [9, 16]. The results of comparing the proposed clustering method with the reference method substantiated by scientists in [8] are presented in Table 1. The proposed hierarchical clustering method provides increased adaptability, efficient resource allocation, and better scalability compared to the reference method.

The implementation of the hierarchical clustering method begins with the definition of the cluster formation algorithm [2,4,14]. In order to substantiate the choice of the optimal clustering algorithm, which is the most relevant, a comparative analysis of modern approaches was conducted (Table 2) [5, 12, 13].

Table 1 – Comparison of clustering methods

Criterion	The proposed method	Reference method [8]	Advantage
Clustering type	Multi-level hierarchical clustering of nodes by task type	Two-stage clustering of communities with service placement	Easier and faster network organization and resource allocation
Complexity	Uses Louvain algorithm to find communities	Uses Girvan-Newman algorithm to find communities	Significantly lower computational complexity and higher execution speed
Weight parameter	Combined (delay + bandwidth)	Hop count	Optimally takes into account the physical characteristics of the network
Scalability	Adaptive hierarchical structure	Significant dependence on community sizes	Efficient work in large systems
Delay optimization	Local clustering + relative geographic optimization	Transient closure of services	Lower transmission delays between nodes
Prioritization of resources	Minimize intra and inter-cluster latency while sustaining target cluster sizes.	Uniform distribution of services between communities to satisfy deadlines	Faster resource management and task reassignment on failures

Table 2 – Comparative characteristics of modern clustering algorithms

Algorithm	Principle	Weight parameters	Adaptation	Advantages	Disadvantages
K-Medoids	Partitioning by similarity	Yes	Low	Weight-aware, emission-resistant	High complexity, not for large graphs
Multi-Level Partitioning	Recursive graph partitioning	Yes	High	Load balancing	Does not take into account the structure of clusters
Louvain	Optimization of modularity	Yes	High	Fast, scalable	Generates large clusters
Leiden	Improved Louvain algorithm	Yes	High	Fast, detailed partitioning, avoids resolution limit	Slower than Louvain, high computational complexity

As shown in Table 2, the most suitable for the scientific tasks of this study are the Louvain algorithms and its modification, the Leiden algorithm [2,4,14].

1. The Louvain algorithm was chosen because of its high speed and scalability, which is especially relevant for distributed computing systems. It optimizes the modularity of the graph, which allows for efficient community detection and cluster formation. Louvain's algorithm consists of the following steps (Fig. 1):

- initialization – each node is considered as a separate community;
- moving nodes – nodes are moved between communities to improve the quality function (e.g., modularity);
- community aggregation – nodes within one community are merged into a super-node that forms a new graph;
- repetition – the process is repeated recursively until stability is achieved.

The steps of the Louvain and Leyden algorithms are shown in Fig. 1 and Fig. 2. In the initial state (black

and white image), all nodes are not assigned to any cluster. The color in the following stages indicates the membership of nodes in the corresponding clusters formed during clustering.

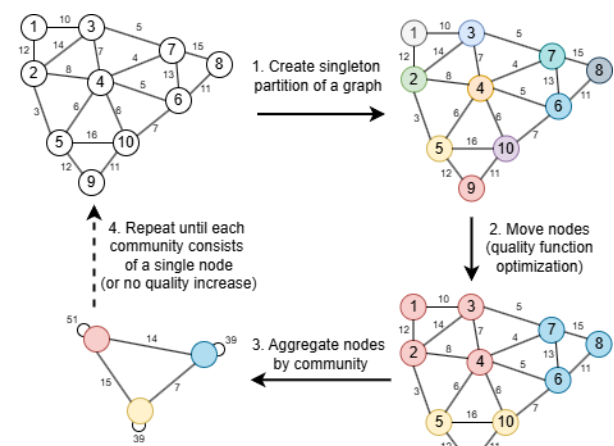


Fig. 1. Stages of the Louvain algorithm

It has been studied that the main disadvantage of the Louvain algorithm is the tendency to form excessively large clusters, which does not meet the objectives of this study, since a more detailed breakdown of nodes is needed with a more accurate consideration of their internal structure [4]. The reason for this is that at each iteration, new communities are formed solely on the basis of the increase in modularity, which does not always guarantee the optimality of their partitioning.

2. In this research, the Leiden algorithm was chosen to compare the performance of the proposed hierarchical clustering method. The main advantage of the Leiden algorithm is the elimination of the main disadvantages of the Louvain algorithm, namely, resolution limit and cluster post-processing [14].

Unlike Louvain, Leiden clustering performs not only global optimization updates but also intermediate refinement at the level of individual nodes, which

provides more detailed grouping.

The process of the Leiden algorithm consists of the following steps:

- initialization – each node is considered as a separate community;
- moving nodes – nodes are moved between communities to improve the quality function;
- cluster refinement – an additional stage of verification and adjustment of the division is applied;
- aggregation – nodes within the community are combined into a single node;
- repetition – the procedure continues until a stable partition is achieved.

The Leiden algorithm eliminates the problem of insufficient cluster detail present in the Louvain algorithm, as it allows for a more accurate distribution of nodes between clusters and prevents weakly connected structures from remaining in the same cluster (Fig. 2).

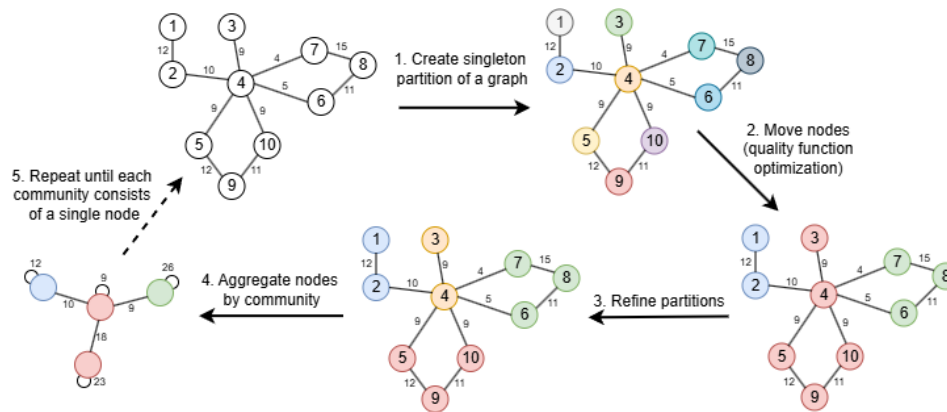


Fig. 2. Stages of the Leiden algorithm

This is achieved by introducing the resolution parameter γ into the modularity function, which allows to dynamically adjust the level of detail of the cluster structure in distributed telecommunication networks. This method provides an optimal balance between the accuracy of cluster detection and the stability of their structure, preventing excessive fragmentation or consolidation, which is critical for ensuring stable interaction between network nodes.

The proposed improvement of the Louvain algorithm allows us to develop a method of hierarchical clustering of nodes by improving the quality of clustering and optimal resource allocation.

The main advantages of the proposed method are: increased accuracy of cluster detection due to an adaptive approach to clustering; flexible configuration of clustering parameters, including target cluster size limits and cluster granularity (resolution) step; reduce inter-cluster latency by taking into account the delay and bandwidth between nodes.

The method includes the following steps.

Stage 1. Representation of a telecommunication system as a weighted graph $G = (V, E)$, where: V – a set of nodes corresponding to the computing elements of the system; E – a set of edges, links between nodes, characterized by certain weight parameters [6, 12]. Each node v_i has a set of properties, namely: label T_{ij} ,

indicating the type of task it can perform; computing performance P_{ij} of the node. Each edge e_{ij} between nodes v_i and v_j has the following parameters: w_{ij} – the bandwidth of the communication channel between i and j ; d_{ij} – data transmission delay.

Since the Louvain and Leiden algorithms optimize the modularity of the graph, it is necessary to correctly determine the weight of the edges between nodes, which will simultaneously represent the delay and throughput. Different approaches to determining this weight are presented in [2,4,14].

1. Inverse throughput weighted by delay:

$$w(e) = \frac{d_{ij}}{w_{ij}}, \quad (1)$$

It is used in cases where the main task is to minimize the delay.

2. Harmonic average of bandwidth and inverse delay:

$$w(e) = 2 \cdot w_{ij} \cdot \frac{1}{d_{ij}} \left(w_{ij} + \frac{1}{d_{ij}} \right), \quad (2)$$

Balances fast and slow links for heterogeneous networks but may overestimate links with very small delays d_{ij} .

3. Normalized communication quality evaluation:

$$w(e) = \alpha \cdot d_{ij} + \beta \cdot w_{ij}, \quad (3)$$

Adjusts weights via α and β coefficients to adapt to changing network topologies and conditions. However, they must balance latency and throughput to avoid overestimating high-bandwidth links and reducing the effect of abnormal delays.

Approach 2 was selected, using a weighting function as a linear combination of delay and capacity. Based on these parameters, an initial graph is built, which will be used for clustering.

Stage 2. Splitting a graph into layers

At this stage, the nodes of the distributed system are classified by the types of tasks [6, 12] performed on the data during stream processing. This creates a multi-tiered structure, where each tier corresponds to a specific type of computing process. This approach simplifies further clustering, ensures localization of computing, and minimizes inter-cluster delays.

To classify nodes, we introduce a set of levels S of the computing process, where each level $s_k \in S$ is defined by a set of tasks $T(v_i)$ that can be performed by nodes of this level. It is determined by the criterion of task compatibility:

$$s_k = \{v_i | T(v_i) \cap L(s_k) \neq \emptyset\}, \quad (4)$$

where $L(s_k)$ is the set of tasks assigned to level s_k .

This equation means that a node v_i is assigned to level s_k if its task matches at least one of the tasks assigned to that level.

Stage 3. Recursive clustering of each layer

Clustering at each level is performed using a modified Louvain algorithm that focuses on achieving target cluster sizes by recursively detecting them. The modularity formula used by the standard Louvain algorithm does not contain parameters that would allow adjusting the connectivity of the original clusters. To solve this problem, the algorithm uses a modified modularity formula, the Reichardt Bornholdt Potts Model, which defines the resolution parameter γ [2, 4, 14, 19]. By adjusting this parameter as the algorithm runs, a more detailed clustering can be achieved, since different values of the resolution parameter are applied to different parts of the graph. The modularity Q , considering γ , is calculated by the formula:

$$Q(v) = \frac{1}{2m} \sum_{i,j} \left[A_{i,j} - \gamma \frac{k_i k_j}{2m} \right] \delta(c_i, c_j), \quad (5)$$

where γ – a parameter that controls the level of detail of the clusters; $\gamma > 1$ – promotes the creation of smaller clusters; $\gamma < 1$ – promotes the formation of larger clusters; $A_{i,j}$ – adjacency matrices that determine the existence of a connection between nodes; $k_i k_j$ – the degrees of nodes i and j (the total number of links in the graph); m – the total weight of all links in the graph; c_i, c_j – corresponding clusters for nodes i, j ; $\delta(c_i, c_j)$ – an indicator of whether a node belongs to a cluster or not.

Recursive clustering based on the modified Louvain algorithm includes the following steps.

3.1 Initialize input data, parameters, and variables.

Input: graph G , partitioning quality function $Mod(\gamma)$, base value of the resolution parameter γ , resolution step γ_{step} , number of cluster search iterations N_{probe} , and cluster size limits: lower bound C_{low} and

upper bound C_{up} . Internal variables: current resolution value γ_{curr} and partitioning P_{curr} , cluster search iteration N_{curr} .

3.2 Execution of the standard Louvain algorithm.

The Louvain algorithm is applied to the current graph G , which uses the Reichardt Bornholdt Potts Model to calculate the quality and the parameter γ_{curr} as a resolution parameter. This forms the initial partitioning of P_{curr} .

3.3 Checking the minimum cluster size.

If at least one resulting cluster is smaller than C_{low} , the γ_{curr} parameter is reduced by the formula:

$$\gamma_{curr} = \max\left(\gamma - \frac{\gamma_{step}}{2N_{curr+1}}, 0.5\right), \quad (6)$$

This operation prevents excessive fragmentation of clusters and allows you to select the optimal resolution value for the current graph. The resolution value of 0.5 is the minimum threshold for the algorithm to work without showing abnormal behavior.

3.4 Checking the success of the cluster search.

If the cluster search was successful, the value of γ_{curr} is increased by γ_{step} , and if not, the algorithm returns the current partition P_{curr} as a single community.

3.5 Recursive processing of large clusters.

For each obtained cluster, the size is checked: if the cluster size C_{sub} is smaller than C_{up} , it remains unchanged; if the size exceeds C_{up} , an induced subgraph G_{comm} is created containing only the nodes of this cluster. Next, LouvainRecursive is again executed on this subgraph, which allows for adaptive refinement of large clusters.

3.6 Aggregation of all obtained clusters into a single partition.

All the resulting partitions of the current level P_{comm} are combined into a single subpartition P , which returns to the level above until it reaches the final exit from the function.

3.7 Formation of the final cluster structure.

The resulting clusters of each layer are aggregated into the final cluster structure. Thus, the resulting partitioning considers the topological features of the network and its parameters, dynamically adapting the clustering to the specified size parameters.

The flowchart of the algorithm that implements the stages of the described method is shown in Fig. 3.

To evaluate the efficiency of the proposed hierarchical clustering method, we developed a software implementation of the Louvain, Leiden, and modified Louvain algorithms using the Python programming language and the NetworkX library. To increase the reliability of the experiment, a graph of 2000 nodes was generated, the structure of which is as close as possible to a real distributed telecommunications network [17]. To visualize the resulting graph and communities, the Fruchterman-Reinhold power algorithm was used to emphasize the visual grouping of nodes if a strong connection is identified between them [18].

The conducted modeling resulted in a graph representation of a certain distributed telecommunication system, which was used to divide the nodes into clusters according to different approaches (Fig. 4 A-C).

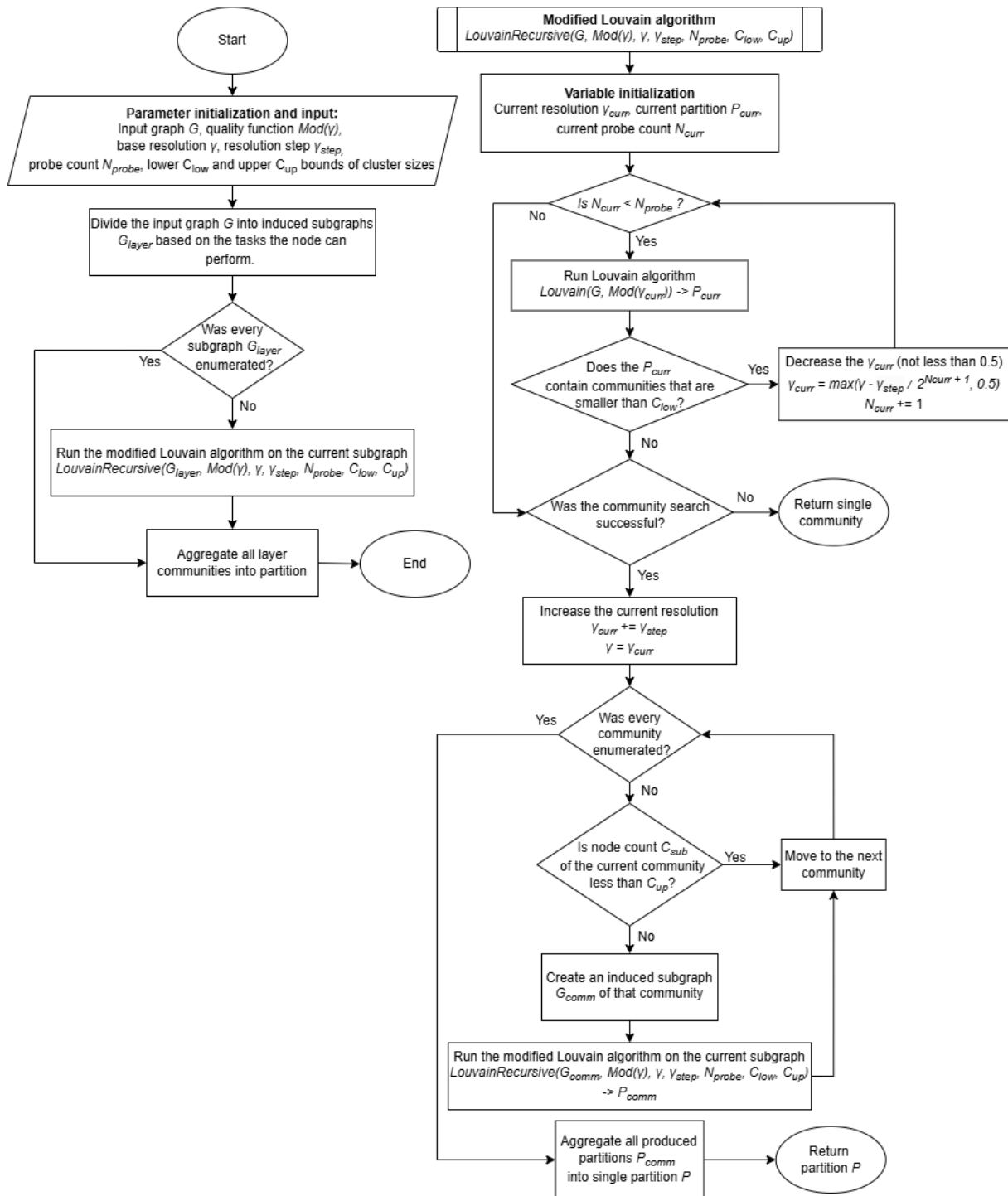


Fig. 3. Flowchart of the modified hierarchical clustering method

In these figures, graph vertices (colored dots) correspond to computing or communication elements of the telecommunications system (servers, routers, network devices), and edges (lines of the corresponding color) correspond to data transmission channels, taking into account delays and bandwidth. The colors of the nodes reflect the cluster structure obtained as a result of clustering: nodes of the same color belong to the same cluster. Edges that have the color of the corresponding cluster identify intra-cluster connections, while black lines indicate inter-cluster connections between nodes from different clusters. Fig. 4.A shows a graph of a distributed telecommunication system with defined clusters, which

was obtained as a result of clustering using the Louvain algorithm.

In this graph, the nodes correspond to computing or communication elements of the system (servers, routers, network devices), and the links model the data transmission channels between them, taking into account bandwidth and delays. The graph shows a cluster structure, where nodes with high interaction are grouped into separate communities (colored segments). In the center of the graph are clusters with dense connections, and on the periphery are less connected nodes that form long links with the central part. The main feature of Louvain clustering is the tendency to

form large clusters, even if there is no strong enough connection between them. This is a consequence of the algorithm's greedy approach. The introduction of a resolution parameter provides a way to partially adjust the requirements for cluster connectivity, but does not completely eliminate the problem: at low values of the parameter, clusters remain too large, and at high values, they are unevenly split.

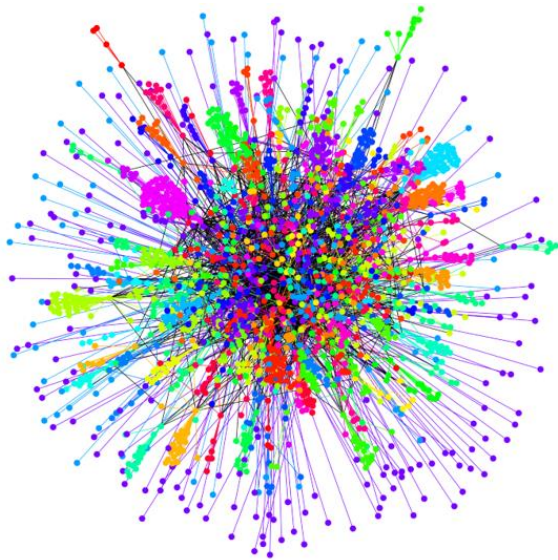


Fig. 4.A. Clustering of nodes using the Louvain algorithm

Fig. 4.B shows a graph of a distributed telecommunication system clustered by the Leiden algorithm.

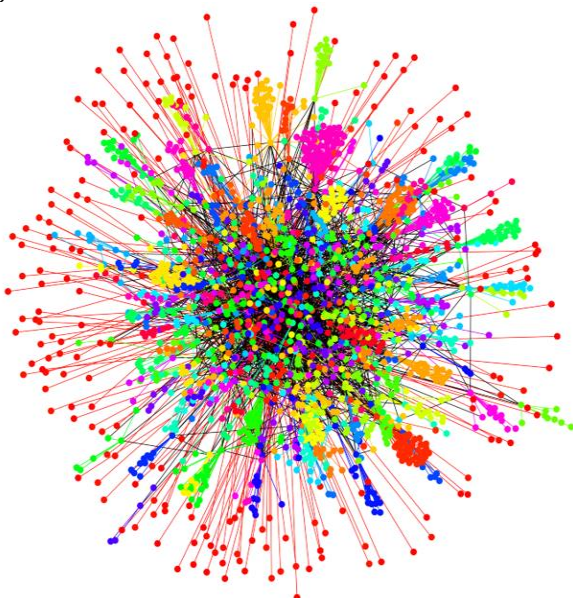


Fig. 4.B. Clustering of nodes using the Leiden algorithm

Fig. 4.B shows that the Leiden algorithm forms more evenly distributed clusters than Louvain, which reduces the centralization of the system structure. Nodes connected by long links to the central part of the graph remain on the periphery.

One of the key features of the algorithm is the way it works on graphs with a distinct cluster structure. Even at low values of the resolution parameter, Leiden finds a

near-optimal partitioning, and its further increase causes only a slight growth of fragmentation.

However, the problem typical for the Louvain algorithm remains: clusters with a high density of connections may be separated unevenly or with a delay in the process of increasing the resolution parameter. This can lead to a situation where weakly connected areas of the graph are already divided into small clusters, while dense communities remain unchanged.

To solve these problems, a modified clustering method is proposed that adaptively changes the resolution parameter, preventing uneven distribution of clusters.

Fig. 4.C shows the visualization of the clustering of the graph of a distributed telecommunication system performed using the proposed modified method, which minimizes both intra-cluster and inter-cluster delays.

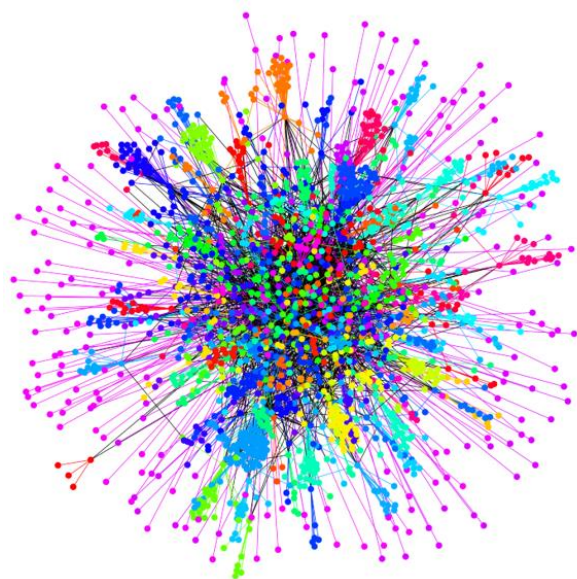


Fig. 4.C. Clustering of nodes using the modified method

Unlike the Louvain and Leiden algorithms, the proposed method forms balanced clusters, which prevents both excessive community consolidation and excessive fragmentation. This approach to clustering considers the actual characteristics of the telecommunications system and its requirements, including channel bandwidth, node performance, and target cluster sizes. This ensures stable interaction between clusters, reduces delays in inter-cluster data exchanges, and reduces decision-making time within the cluster.

To evaluate the effectiveness of the clustering methods, we conducted an experimental simulation, the results of which are shown in Table 3. The simulation was performed at a resolution value of $\gamma = 1$.

Based on the results of the experiment, the following conclusions can be drawn:

1. Louvain generates 50 clusters, which leads to large communities with the highest modularity (0.851892). However, the inter-cluster delays reach 111.928, which is 16.9% higher than Leiden and 4.6% higher than the modified method.

Table 3 – Comparative analysis of clustering methods

Parameter	Louvain	Leiden	Modified method
Number of clusters	50	215	137
Modularity	0,851892	0,811312	0,825021
Maximum cluster size	184	66	64
Minimum cluster size	4	2	2
Average cluster size	40,000	9,302	14,598
Standard deviation of cluster size	28,706	9,551	14,134
Average delay	165,607	165,607	165,607
Average intra-cluster latency	52,911	45,306	52,863
Average inter-cluster latency	111,928	95,644	106,930

2. Leiden generates 215 clusters (4.3 times more than Louvain), which reduces inter-cluster latency by 14.5% compared to Louvain. However, this is accompanied by cluster fragmentation, as their maximum size (66) is 64.1% smaller than that of Louvain (184), and the standard deviation of cluster size is 66.7% smaller.

3. The modified method generates 137 clusters, which is 174% more than Louvain, but 36% less than Leiden. This reduces inter-cluster latency by 4.5% compared to Louvain, avoiding the excessive granularity of Leiden. In addition, intra-cluster latency (52,863) remains almost at the level of Louvain (+0.1%), which guarantees system stability.

Thus, the proposed modified method combines a balanced distribution of nodes and minimal delays, which makes it optimal for distributed systems.

Conclusions and prospects for further research

The study presents a new method of hierarchical clustering of computing nodes in distributed telecommunication systems based on a modified Louvain algorithm. The method is aimed at optimizing modularity, intra- and inter-cluster delays, and cluster sizes, ensuring efficient resource allocation in the network.

1. It is experimentally proved that the proposed hierarchical clustering method and its implementation algorithm are versatile and provide opportunities for building various topologies, including those based on geo-distribution.

2. By increasing the resolution, communities with a larger number of clusters (of smaller size) are created. However, this process can occur unevenly: when using a large value of the resolution parameter (10-20 and above), it has been experimentally proven that clusters with size 1 are formed, which cannot be practically used, while clusters with high connectivity remain undivided. Their separation occurs at a much higher resolution parameter (at the level of 50-100), but at the same time, most of the graph will contain clusters of size 1-2.

3. The combination of the Louvain algorithm with a hierarchical structure allows clustering to be applied not only to the full network graph of the system, but also to its individual segments. This makes it possible to optimally rebuild some of the network segments without affecting those clusters that have already reached a sufficient level of optimization.

The scientific novelty of the study is the development of a new method of hierarchical clustering of computing nodes in distributed telecommunication systems, which, unlike traditional approaches, takes into account the performance of nodes, network topology, delays and bandwidth of communication channels, which is achieved by adaptively adjusting clustering parameters and multilevel analysis of network connections.

Further research is focused on improving the method by expanding its adaptive capabilities to work with heterogeneous telecommunication systems that include nodes with different architectures and computing capabilities.

REFERENCES

1. Abdelmoneem Randa M., Benslimane Abderrahim, Shaaban Eman (2020) Mobility-aware task scheduling in cloud-Fog IoT-based healthcare architectures. *Computer Networks*. Volume 179, 9 October 2020, P.P 107348 - 107367 <https://doi.org/10.1016/j.comnet.2020.107348>
2. Blondel, V.D. et al. (2008) Fast unfolding of communities in large networks. *J. Stat. Mech* 10008, 1-12(2008). <https://doi.org/10.1088/1742-5468/2008/10/P10008>
3. Bonomi F. , Milito R., Natarajan P., and Zhu J. (2014). *Fog Computing: A Platform for Internet of Things and Analytics*. Springer International Publishing, Cham, 169-186. DOI:10.1007/978-3-319-05029-4_7
4. Dugué N., Perez A. (2015) Directed Louvain: maximizing modularity in directed networks. [Research Report] Université d'Orléans. 2015. DOI:[10.13140/RG.2.1.4497.0328](https://doi.org/10.13140/RG.2.1.4497.0328)
5. Ferreira Aluizio, Neto Rocha (2021) Edge-distributed Stream Processing for Video Analytics in Smart City Applications//Federal University of Rio Grande do Norte Exact and Earth Sciences Center Department of Informatics and Applied Mathematics Graduate Program in Systems and Computing PhD in Computer Science. Natal-RN March 2021, P. 119 <https://repositorio.ufrn.br/handle/123456789/32743>.
6. Fortunato, S. (2010) Community detection in graphs. *Physics Reports*, Volume 486, Issues 3–5, February 2010, Pages 75-174, <https://doi.org/10.1016/j.physrep.2009.11.002>

7. Guerrero, C., Lera, I., & Juiz, C. (2019). A Lightweight Decentralized Service Placement Policy for Performance Optimization in Fog Computing. *Journal of ambient intelligence and humanized computing*. 10 – 6, pp. 2447 – 2464. SPRINGER HEIDELBERG, 01.06.2019 <https://doi.org/10.48550/arXiv.2401.12699>
8. Lera I., Guerrero C., Juiz C. (2019) Availability-aware Service Placement Policy in Fog Computing Based on Graph Partitions. *IEEE Internet of Things Journal*. 6 - 2, pp. 3641-3651. IEEE-Inst Electrical Electronics Engineers Inc, 01.04.2019 <https://doi.org/10.48550/arXiv.2401.12690>
9. Petrovska, I., & Kuchuk, H. (2022). Static allocation method in a cloud environment with a service model IaaS. *Advanced Information Systems*, 6(3), 99–106. <https://doi.org/10.20998/2522-9052.2022.3.13>
10. Rzepka Michał, Boryło Piotr, Marcos D. Assunção, Artur Lasoń, Laurent Lefèvre (2022) SDN-based Fog and Cloud Interplay for Stream Processing *Future Generation Computer Systems/ Volume 131, June 2022, P.P. 1-17* <https://doi.org/10.1016/j.future.2022.01.006>
11. Salaht, S., Desprez, F., & Lebre, A. (2020). An Overview of Service Placement Problem in Fog and Edge Computing. *ACM Computing Surveys*, 53(3), 1-35. DOI: 10.1145/3391196
12. Singh J., Singh P., Sukhpal S. (2021) Fog computing: A taxonomy, systematic review, current trends and research challenges *Journal of Parallel and Distributed Computing Volume 157, November 2021, P.P. 56-85* <https://doi.org/10.1016/j.jpdc.2021.06.005>
13. Syvolovskyi, I. M., Lysechko, V. P., Komar, O. M., Zhuchenko, O. S., Pastushenko, V. V. Analysis of methods for organizing distributed telecommunication systems using the paradigm of Edge Computing. 2024. *National University «Yuri Kondratyuk Poltava Polytechnic». Control, Navigation and Communication Systems*, 1(75), P. 206-211 DOI: 10.26906/SUNZ.2024.1.206.
14. Traag, V.A., Waltman, L. & van Eck, N.J. (2019) From Louvain to Leiden: guaranteeing well-connected communities. *Sci Rep* 9, 5233 (2019). DOI: [10.1038/s41598-019-41695-z](https://doi.org/10.1038/s41598-019-41695-z)
15. Кучук Г.А., Коваленко А.А., Лукова-Чуйко Н.В. Метод мінімізації середньої затримки пакетів у віртуальних з'єднаннях мережі підтримки хмарного сервісу. Системи управління, навігації та зв'язку. 2017. Вип. 2(42). С. 117-120.
16. Alenizi F., Rana O. (2020) Minimising Delay and Energy in Online Dynamic Fog Systems//Computer Science. Networking and Internet Architecture, <https://doi.org/10.48550/arXiv.2012.12745>
17. Elmokashfi A., Kvalbein A., Dovrolis C. (2010) On the Scalability of BGP: The Role of Topology Growth. *IEEE Journal on Selected Areas in Communications*, Vol. 28, No 8, pp. 1250-1261, October 2010. DOI: [10.1109/JSAC.2010.101003](https://doi.org/10.1109/JSAC.2010.101003).
18. Fruchterman T.M.J., Reingold E. M. (1991) Graph drawing by force-directed placement// *Software: Practice and Experience*, Volume 21, Issue 11, November 1991, P.P. 1129-1164 <https://doi.org/10.1002/spe.4380211102>
19. Reichardt J., Bornholdt S. (2004) Detecting Fuzzy Community Structures in Complex Networks with a Potts Model // *APS. Physical Review Letters*, 93, 218701 – Published 15 November, 2004, <https://doi.org/10.1103/PhysRevLett.93.218701>

Received (Надійшла) 18.03.2025

Accepted for publication (Прийнята до друку) 11.06.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Сиволовський Ілля Михайлович – аспірант кафедри транспортного зв'язку, Український державний університет залізничного транспорту, Харків, Україна;

Syvolovskyi Illia – PhD student, Department Transport Communication, Ukraine State University of Railway Transport, Kharkiv, Ukraine;

e-mail: ilyasyl95@kart.edu.ua; ORCID Author ID: <https://orcid.org/0000-0002-4592-0965>.

Лисечко Володимир Петрович – доктор технічних наук, професор, начальник науково-дослідного відділу, Харківський національний університет Повітряних Сил імені Івана Кожедуба, Харків, Україна;

Lysechko Volodymyr – Doctor of Technical Sciences, Professor, Chief of the Research Department, Ivan Kozhedub Kharkiv National Air Force University, Kharkiv, Ukraine;

e-mail: lysechkov@ukr.net; ORCID Author ID: <https://orcid.org/0000-0002-1520-9515>;

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57195515166>

Метод ієрархічної кластеризації вузлів розподілених телекомунікаційних систем з використанням графових алгоритмів

I. M. Сиволовський, В. П. Лисечко

Анотація. У статті розроблено модифікований метод ієрархічної кластеризації обчислювальних вузлів у розподілених телекомунікаційних системах із врахуванням продуктивності вузлів, топології мережі, затримок та пропускної здатності каналів зв'язку. Запропонований метод ґрунтується на модифікованому алгоритмі Louvain, що виконує багатокрокову кластеризацію графа з динамічним коригуванням параметрів. Це дозволяє контролювати розмір кластерів та їхню внутрішню щільність, відповідно до заданих цільових показників, мінімізуючи фрагментацію мережевої структури. На основі порівняльного аналізу сучасних методів кластеризації та проведеного експериментального моделювання встановлено, що запропонований метод забезпечує зменшення фрагментації кластерів на 36% порівняно з методом Leiden. Крім того, він дозволяє знизити міжкластерні затримки на 4,5% у порівнянні з методом Louvain та на 11,8% порівняно з Leiden, що свідчить про ефективнішу організацію міжкластерної взаємодії. Вдосконалений метод забезпечує рівномірний розподіл обчислювальних вузлів між кластерами, що сприяє оптимізації загальної продуктивності розподіленої телекомунікаційної системи.

Ключові слова: розподілені телекомунікаційні системи, ієрархічна кластеризація вузлів, моделювання інтелектуальних систем, графова алгоритмічна оптимізація, Louvain / Leiden Clustering, балансування навантаження, пропускна здатність, оптимізація топології, мінімізація затримок.

АЛФАВІТНИЙ ПОКАЖЧИК

Авакін С. М.	41	Ільїна І. В.	132	Погорелов С. В.	145
Артюх Р. В.	132	Іщенко Н. В.	175	Помелуйко Д. А.	175
Аушева Н. М.	226	Качанов П. О.	66	Пономаренко О. Є.	180
Барковська О. Ю.	76	Кашлев М. С.	222	Рагулін О. Є.	210
Бікчентаєв М. О.	235	Кислиця С. Г.	59	Роговий А. І.	145
Білоусов І. А.	83	Кібіреєв Д. О.	203	Рогозін І. А.	27
Бірук Я. І.	218	Климова І. М.	119	Руденко В. В.	244
Бовчалюк Н. І.	47	Климчук А. В.	218	Савченко А. О.	145
Бовчалюк С. Я.	47	Клочко Л. А.	113	Садовников Б. І.	249
Бологова Н. М.	83	Коваленко А. А.	137	Самборський Є. І.	185
Бондаренко М. Е.	90	Кожевніков Г. К.	141	Северін І. С.	13
Боряк Б. Р.	235	Коломійцев О. В.	102	Севостьянова О. М.	137
Бугрій А. М.	169	Комар О. М.	239	Сиволовський І. М.	255
Бульба С. С.	161	Кравченко А. О.	141	Сітніков В. І.	203
Бурков В. В.	5	Кравченко О. С.	145	Скородєлов В. В.	161
Буслов П. В.	210	Кравченко П. О.	210	Стадник А. О.	141
Волощук О. Б.	175	Криворот А. І.	27	Сухоребрий О. В.	59
Галай В. М.	59	Крук О. І.	70	Тараненко І. І.	97
Галонько Я. О.	230	Куш О. О.	97	Тарасенко М. В.	66
Галуза О. А.	145	Левченко Л. О.	226	Тереник Д.	195
Гармаш В. С.	54	Лега О. С.	13	Ткаченко Т. М.	226
Геревич М. О.	119	Лисечко В. П.	255	Токарський О. І.	230
Глива В. А.	222	Ліпчанська О. В.	154	Фесенко А. Б.	108
Головка Г. В.	97	Ліпчанський М. В.	154	Філімончук Т. В.	169
Горбачов В. О.	180	Мавринський О. Д.	76	Фомічов О. О.	137
Гриньов Д. В.	141	Мартовицький В. О.	13	Харченко В. С.	195
Грідасов І. Ю.	5	Мельник О. М.	20	Хижняк І. А.	5
Гулевич М. В.	102	Михайліченко О. В.	70	Христофоров К. А.	154
Демиденко М. І.	108	Міхаль О. П.	54	Худов Г. В.	5
Деркач Т. М.	113	Молчанов Г. І.	154	Хулап А. В.	125
Дмитренко А. О.	113	Москаленко О. О.	230	Черних О.П.	161
Довгополий С. О.	41	Мощенко І. О.	41	Чхеїдзе В. О.	137
Дрозд О. Ю.	47	Наумова О. В.	119	Шаманський С. Й.	218
Дрючко О. Г.	59	Нгок До Куєн	119	Шандиба А. С.	203
Дяченко В. О.	54	Никитюк П. В.	20	Шапіро А. С.	76
Єрьоміна Н. С.	175	Ніколаєв О. С.	203	Шаповал М. В.	27
Жебін П. Д.	76	Носко С. В.	161	Шаптала С. О.	35
Жученко О. С.	249	Орисенко О. В.	27	Шефер О. В.	59
Заковоротний О. Ю.	125	Панченко В. І.	161	Шматко О. В.	210
Запорожець О. В.	41	Партика С. О.	169	Шмельова Т. Ф.	35
Зимогляд М. М.	132	Пелешок Є. В.	185	Яловега М. О.	5
Знайдюк В. Г.	54	Первєєв В. Д.	169	Янко А. С.	70
Зозуля Л. А.	230	Перець К. Г.	239	Янковський О. А.	119
Іващенко Г. С.	90	Петров С. В.	175		

Наукове видання

СИСТЕМИ УПРАВЛІННЯ, НАВІГАЦІЇ ТА ЗВ'ЯЗКУ

Збірник наукових праць
Випуск 2 (80)

Відповідальна за випуск *К. С. Нестеренко*
Технічний редактор *Т. В. Уварова*
Коректор *О. В. Морозова*
Комп'ютерна верстка *Н. Г. Кучук, І. Ю. Петровська*
Оформлення обкладинки *І. В. Ільїна*

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

Підписано до друку 20.06.2025. Формат 60×84/8. Ум.-друк. арк. 35,5. Тираж 120 прим. Зам. 620-25
Адреса редакції: Україна, 36011, м. Полтава, Першотравневий проспект, 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