ISSN 2073-7394

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 4 (82)

Issue 4 (82)

Щоквартальне видання

Засноване у 2007 році

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

Засновник і видавець:
Національний університет
“Полтавська політехніка імені Юрія Кондратюка”

Телефон:
+38 (050) 302-20-71

E-mail редколегії:
kuchuk56@ukr.net

Інформаційний сайт:
http://journals.nupp.edu.ua/sunz

Quarterly

Founded in 2007

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

Founder and publisher:
National University
"Yuri Kondratyuk Poltava Polytechnic"

Phone:
+38 (050) 302-20-71

E-mail of the editorial board:
kuchuk56@ukr.net

Information site:
http://journals.nupp.edu.ua/sunz

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

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

*Затверджений до друку Вченою Радою Національного університету “*Полтавська політехніка імені Юрія Кондратюка*” (протокол від 2 грудня 2025 року № 16).*

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

З М І С Т

АВТОМОБІЛЬНИЙ, РІЧКОВИЙ, МОРСЬКИЙ ТА АВІАЦІЙНИЙ ТРАНСПОРТ

УПРАВЛІННЯ В СКЛАДНИХ СИСТЕМАХ

ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

ЦИВІЛЬНА БЕЗПЕКА

ЗВ'ЯЗОК, ТЕЛЕКОМУНІКАЦІЇ ТА РАДІОТЕХНІКА

Організації авторів

Державний університет «Київський авіаційний інститут» , Київ, Україна
Київський національний університет будівництва і архітектури, Київ, Україна
Київський фаховий коледж зв'язку, Київ, Україна
Міжрегіональна академія управління персоналом, Київ, Україна
Науково-дослідний інститут воєнної розвідки, Київ, Україна
Національна академія Служби безпеки України, Київ, 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 кораблебудування ім. адмірала Макарова, Миколаїв, Україна
Харківський національний університет міського господарства ім. О. М. Бекетова, Україна
Національний університет оборони України, Київ, Україна
Національний університет «Полтавська політехніка імені Юрія Кондратюка», Полтава. Україна
Одеський національний морський університет, Одеса, Україна
Triol Corporation, Харків, Україна
Харківський національний економічний університет імені Семена Кузнеця, Харків, Україна
Харківський національний університет радіоелектроніки, Харків, Україна
Харківський національний університет Повітряних Сил імені Івана Кожедуба, Харків, Україна
Українська державна льотна академія, Кропивницький
Університет сучасних технологій, Київ, Україна

Authors affiliation

State University "Kyiv Aviation Institute," Kyiv, Ukraine
Kyiv National University of Construction and Architecture, Kyiv, Ukraine
Kyiv Vocational College of Communication, Kyiv
Interregional Academy of Personnel Management, Kyiv, Ukraine
Defence Intelligence Research Institute, Ukraine
National Academy of the Security Service of Ukraine, Kyiv, 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
Admiral Makarov National University of Shipbuilding, Mykolaiv, Ukraine
Beketov National University of Urban Economy in Kharkiv, Ukraine
National Defence University of Ukraine, Kyiv
National University “Yuri Kondratyuk Poltava Polytechnic”, Poltava, Ukraine
Odesa National Maritime University, Odesa, Ukraine
Triol Corporation, Kharkiv, Ukraine
Simon Kuznets Kharkiv National University of Economics, Kharkiv, Ukraine
Kharkiv National University of Radio Electronics, Kharkiv, Ukraine
Ivan Kozhedub Kharkiv National Air Force University, Kharkiv, Ukraine
Ukrainian State Flight Academy, Kropyvnytskyi
University of Modern Technologies, Kyiv, Ukraine

Автомобільний, річковий, морський та авіаційний транспорт

УДК 656.61: 656.615:614.8 **doi: 10.26906/SUNZ.2025.4.5-11**

С. В. Волянський[1], Ю. В. Бичковський[2], О. М. Мельник[2], А. О. Волошин[2]

[1] Національний університет кораблебудування ім. адмірала Макарова, Миколаїв, Україна
[2] Одеський національний морський університет, Одеса, Україна

СИСТЕМНИЙ АНАЛІЗ ПРИЧИН НАВІГАЦІЙНИХ ІНЦИДЕНТІВ ПІД ЧАС ШВАРТУВАННЯ СУДЕН ТА ШЛЯХИ ВДОСКОНАЛЕННЯ МІЖНАРОДНОЇ НОРМАТИВНО-ПРАВОВОЇ БАЗИ

Анотація. Актуальність. У зв'язку зі зростанням розмірів торговельних суден і навантаженням на портову інфраструктуру, питання забезпечення безпеки під час процесу швартування набуває стратегічного значення. Часті інциденти в портах світу свідчать про недосконаліть існуючих практик і фрагментарність впровадження сучасних методів управління ризиками, зокрема Bridge Resource Management (BRM). **Об'єкт дослідження** - навігаційні інциденти під час швартування великотоннажних суден. **Мета** - системний аналіз чинників, що спричиняють інциденти під час швартування, та формування рекомендацій щодо вдосконалення міжнародної нормативної бази і процедур управління ризиками в портах. **Методологія.** У роботі застосовано порівняльний аналіз інцидентів, експертну оцінку ризиків, класифікацію причин на основі факторної моделі та аналіз міжнародних нормативних документів IMO (SOLAS, STCW, MSC.255(84) та ін.). **Результати.** Дослідження підтвердило, що ключовими причинами є перевищення швидкості підходу до причалу, неузгодженість дій буксирного супроводу, людський фактор, а також відсутність стандартизованих процедур BRM на підході до причалу. Порівняння практик у різних країнах засвідчило нерівномірність впровадження тренажерної підготовки та інтелектуальних систем підтримки рішень. **Висновки.** В роботі наведено рекомендації щодо гармонізації регіональних практик, сприяючих розвитку систем підтримки прийняття рішень та впровадженню тренажерної підготовки екіпажів суден. Доведено доцільність інтеграції BRM у порту та розробки адаптивних протоколів ризик-менеджменту із урахуванням реального стану інфраструктурного забезпечення.

Ключові слова: морський транспорт; швартування суден; керування судном; маневрування; інциденти; безпека навігації; міжнародні конвенції; людський фактор; портова інфраструктура; управління ризиками; управління ресурсами містку (BRM).

Вступ

Постановка проблеми. В умовах розвитку сучасного судноплавства безпечне маневрування суден у припортових водах набуває критичного значення як для збереження людських життів, так і для стабільності міжнародної логістики та захисту морського середовища. Попри наявність розвиненого правового інструментарію, передусім Конвенції SOLAS-74, Кодексу STCW-78 та пов'язаних з ними резолюцій IMO, щороку фіксуються серйозні інциденти під час швартування великотоннажних суден.

Особливе занепокоєння викликають випадки, коли аварії відбуваються в умовах, які мали б унеможливити такі події у світлу пору доби, за сприятливої погоди, при наявності повного складу штурманів на містку та використанні буксирів.

Аналітичний огляд гучних інцидентів останнього десятиліття (зокрема аварій YM Witness (2023), Milano Bridge (2021), Ever Given (2021), DALI (2024), Yuzhou Qi (2019) засвідчує, що причини таких подій мають комплексний характер. Як підсумок, у всіх випадках виявлено недостатню готовність портової інфраструктури, неефективну комунікацію між екіпажом і портовими службами, недосконале конструкторське виконання швартових елементів судна, а також відсутність чітких нормативних вимог щодо перевірки готовності причалу і усе перераховане разом створює критичну зону ризику на стику між судном та портом.

Аналіз останніх досліджень і публікацій. У науковій літературі питання безпеки швартування послідовно розглядаються крізь призму складності операцій у вузьких акваторіях, когнітивного навантаження екіпажу та появи нових технологій. Опитування операторів портів свідчать про обережний оптимізм щодо впровадження MASS, але лише за умов прозорих протоколів взаємодії та чітко визначеної відповідальності [1]. Для випробування автономних функцій необхідні реалістичні сценарії трафіку; підходи до їх екстракції й семплювання демонструють, що робастність досягається лише за широкого охоплення умов середовища і варіантів поведінки суден [2].

Традиційні флотські операції підтверджують, що аварійність при постановці суден з небезпечними вантажами значною мірою зумовлена поєднанням людського фактору, тиску графіка та метеоумов; регресійний аналіз дозволяє сформувати ієрархію ризик-факторів і транслювати її у практику контролю під час швартування [3].

Нові ризики породжує електрифіка-ція: GT-FFTA моделі для суден-електроходів вказують на складні ланцюги займання й поширення пожежі, що ускладнює аварійно-рятувальні дії та вимагає спеціальних протоколів моніторингу [4].

Систематичні огляди та бібліометрія фіксують зсув досліджень у бік людиноцентричних підходів [5]. Концепція Safety-II підкреслює, що надійність у складних маневрах забезпечується адаптивністю екіпажу та варіативністю процедур; це означає, що регламенти мають не лише забороняти відхилення, а й підтримувати успішні практики [6]. Функціонально-причинна аналітика (наприклад, TM-FRAM) показує, як моделювати взаємодію сенсорів, комунікацій і маневрових рішень у часі, що критично для операцій швартування [7]. Додаткову стабільність забезпечують високоточні прогнози офшорних вітрів та течій, особливо у режимі nowcasting, які можуть суттєво знизити ризики підходу до порту [8]. Водночас огляди вказують на дефіцит узгоджених метрик і бібліотек сценаріїв для тестування автономних навігаційних систем у портових маневрах [9].

Дослідження також демонструють, що управління ризиком має бути типоспецифічним: рішення, ефективні для пасажирських суден, не завжди підходять для танкерів чи балкерів [10]. Українські науковці конкретизують роль людського елементу: рівень безпеки судна, вплив стресу та організаційних факторів напряму визначають якість прийняття рішень при швартуванні [11–13]. У цьому контексті ситуаційна обізнаність (SA) визнається ключовим запобіжником: спільна картина середовища між мостиком, буксирами, лоцманом і берегом є обов'язковою умовою безпечного зближення [14]. У випадках браку об'єктивних даних застосовуються методики експертних оцінок, що дозволяють швидко актуалізувати параметри ризику [15]. Нарешті, проектно-орієнтовані моделі управління безпекою, акцентовані на людському чиннику, демонструють ефективність інтеграції навчання, тренажерних сесій і аудиту процедур у єдину систему ризик-менеджменту [16].

Таким чином, аналіз літератури по темі дослідження відображає системну природу аварій протягом процесу швартування і доводить, що їх подолання потребує міждисциплінарної інтеграції: сценарно-орієнтованих випробувань, людино-центричних методів Safety-II/FRAM, точних прогнозних сервісів та диференційованих регуляторних вимог.

Мета та завдання дослідження. Незважаючи на регулярні оновлення норм IMO, окремі аспекти безпеки під час процесу швартування судна залишаються нормативно неурегульованими або регулюються лише у формі необов'язкових рекомендацій. Особливо це підкреслюється в умовах зростання розмірів суден, інтенсифікації вантажообігу та кліматичних змін, що ускладнюють навігацію, постає об'єктивна необхідність перегляду підходів до регулювання безпеки на фінальних етапах маневрування а саме підходу до причалу, гальмування, взаємодії з буксирами та швартування.

Метою даної роботи є обґрунтування необхідності внесення конкретних змін до міжнародного нормативного середовища, зокрема до Конвенцій SOLAS-74 і STCW-78, а також супутніх кодексів з урахуванням практичного аналізу типових помилок, причин навігаційних аварій та неврахованих конструктивно-експлуатаційних чинників. Запропонований підхід поєднує аналіз морської практики, аналіз звітів про інциденти та практичні інженерні рішення щодо підвищення безпеки.

Аналіз сучасного стану проблеми. Проблема безпечного швартування великорозмірних суден та управління ризиками при їх маневруванні поблизу причалів набула особливої актуальності в останнє десятиліття. Цьому спряло ускладнення конструкцій суден, збільшення їх габаритів та вантажопідйомності, зміна кліматичних умов. Також нерівномірний розвиток портової інфраструктури, які спричиняють низку викликів, які вимагають перегляду підходів до планування та виконання швартових операцій.

Низка гучних аварій за участю крупнотоннажних суден демонструє наслідки недосконалої координації між судноводієм, буксирами та портовою службою. В табл. 1 узагальнено найбільш резонансні інциденти останніх років.

Таблиця 1 – **Приклади аварій суден під час підходу до порту або швартування**

Судно	Рік	Локація	Причина	Наслідки
YM Witness	2019	Колумбія	Некоректна оцінка інерції	Пробоїна в причалі, збитки в 5 млн дол.
Milano Bridge	2020	Пусан, Корея	Занадто висока швидкість	Руйнування 3 STS-кранів
DALI	2024	Балтімор, США	Втрата електроживлення	Обвал мосту Francis Scott Key
Yuzhou Qi	2016	Янцзи, Китай	Зрив швартовного плану	Контакт з береговою інфраструктурою
Durban collision	2017	ПАР	Порушення комунікації	Масштабні пошкодження терміналу

Приведені вище інциденти свідчать про системні вади в процесах управління ризиками на фінальній стадії маневрування та недостатню стандартизацію процесів між портами різних країн.

Швартування великотоннажних суден у портах з розвиненою інфраструктурою (ЄС, Північна Америка, Японія) супроводжується використанням інтелектуальних систем підтримки рішень, централізованого контролю швидкості та автоматичних систем позиціювання (MoorMaster, DockingAid, Virtual Fender тощо).

Водночас, у ряді портів країн, що розвиваються, інфраструктура не дозволяє повноцінно забезпечити рекомендації MSC.1/Circ.1175 щодо мінімізації залишкової енергії судна при контакті з причалом, табл.2.

Нормативно-правова база та її недоліки. Безпека швартування суден регулюється міжнародними документами, які охоплюють навігацію, підготовку екіпажу та управління ризиками. Зокрема, Конвенція SOLAS (Розділ V) зобов'язує судновласників і капітанів забезпечувати безпечну навігацію до повної зупинки судна в порту.

Таблиця 2 – **Порівняння систем швартування у різних регіонах**

Регіон / Країна	Автоматизація	Наявність прогнозних моделей	Використання датчиків інерції	Рівень аварійності (умовний)
Північна Європа	Висока	Так	Так	Низький
Китай	Середня	Частково	Обмежено	Середній
США	Висока	Так	Так	Низький
ПАР	Середня	Ні	Частково	Середній
Південна Америка	Низька	Ні	Ні	Високий

Конвенція STCW (Розділ A-V/2) визначає вимоги до підготовки капітанів і старших помічників для дій у прибережних районах і обмежених водах, а резолюція IMO MSC.255(84) включає положення концепції E-Navigation з використанням систем підтримки прийняття рішень під час маневрування суден.

Важливим інструментом регламентації є також ISM Code (A.741(18)), який визначає необхідність ідентифікації ризиків і впровадження процедур їх контролю в експлуатаційній практиці. Крім того, документ IMO MSC.947(23) містить рекомендації щодо управління маневруванням великих суден у вузьких акваторіях та під час постановки до причалу.

Однак аналіз цих документів свідчить, що більшість положень мають рамковий характер. Вони задають загальні принципи, але не формують уніфікованої моделі управління швартуванням. І це зумовлює необхідність врахування локального контексту, тобто особливостей конкретного порту, гідрологічних умов, типорозмірів суден та наявної інфраструктури. Відсутність детальної регламентації створює правові прогалини, які можуть ускладнювати застосування сучасних технологій (зокрема MASS) та перешкоджати створенню універсальних стандартів.

Таким чином проведений аналіз доводить необхідність створення інтелектуальних рішень підтримки судноводія у фазі швартування, які враховують не лише фізичні характеристики судна, але й оперативні ризики, параметри оточення та технічні можливості порту. Тому системний підхід до навчання, оновлення нормативів та уніфікація технологічних інтерфейсів є ключем до зниження аварійності.

Методологія дослідження. Методологічна основа даного дослідження ґрунтується на системному аналізі відомих інцидентів зіткнення суден з причальними інфраструктурними об’єктами у міжнародних портах. Враховуючи складність морських операцій і велику кількість чинників, що впливають на результат швартування, застосовано багатокритеріальний підхід з використанням експертного оцінювання, якісного та кількісного аналізу.

Критерії аналізу інцидентів. Для формалізованого вивчення кожного випадку інциденту було обрано п’ять ключових критеріїв, рис. 1.

Рис. 1. Критерії аналізу інцидентів

Наведені ключові критерії для аналізу морських інцидентів дозволяють системно оцінювати як технічні, так і організаційні аспекти подій. Формалізація таких критеріїв є важливою складовою розслідування, оскільки вона забезпечує об’єктивність оцінки та можливість порівняння між різними випадками, табл. 3. Критерії аналізу інцидентів здатні забезпечити багаторівневий підхід до оцінювання ризиків у процесі морської діяльності. Поєднання технічних параметрів (тип судна, метеоумови, стан причалу) та організаційних аспектів (ефективність буксирів, дії екіпажу й служб) дозволяє комплексно визначити першопричини аварійних ситуацій.

Таблиця 3 – **Критерії аналізу морських інцидентів**

№	Критерій	Зміст та значення
1	Тип судна та його тоннаж	Дозволяє врахувати маневрові характеристики, осадку та керованість судна, що впливає на його поведінку під час інциденту.
2	Метеоумови та гідрологічна ситуація	Враховуються вітер, хвилювання, течії, видимість та інші фактори, що безпосередньо впливають на безпеку маневрування.
3	Кількість та ефективність буксирного забезпечення	Аналізує залучені допоміжні засоби управління судном (буксири), їх адекватність і готовність.
4	Оцінка готовності причального фронту	Визначається відповідність технічного стану, габаритів та інфраструктури причалу для безпечного швартування.
5	Хронологія дій екіпажу, лоцмана та диспетчерської служби	Дозволяє реконструювати послідовність рішень, виявити можливі комунікаційні збої чи процедурні порушення.

Застосування таких критеріїв створює основу для систематизації висновків розслідувань та вироблення практичних рекомендацій щодо підвищення безпеки. Зокрема, аналіз хронології рішень екіпажу й комунікацій із береговими службами допомагає виявити людський фактор та його роль у розвитку подій, тоді як технічні критерії дають змогу оцінити, чи були створені належні умови для безпечного маневрування. У результаті формалізований підхід дозволяє не лише пояснювати причини конкретних інцидентів, а й прогнозувати ймовірність повторення подібних ситуацій, що є основою для превентивних заходів у сфері морської безпеки.

На основі аналізу десяти випадків аварійного швартування у портах різних країн було виокремлено домінуючі групи причин, які наведені у табл. 4.

Таблиця 4 – **Класифікація основних причин інцидентів**

№	Група причин	Сутність проблеми
1	Надмірна швидкість підходу	Судно перевищувало безпечну швидкість у момент швартування
2	Недостатня потужність буксирів	Наявні буксири не мали достатньої тягової сили для маневрування великим судном
3	Людський фактор	Помилки екіпажу, неправильні рішення лоцмана, відсутність злагодженості дій
4	Відсутність готовності причалу	Пошкоджені або застарілі кріплення, слабка інфраструктура, відсутність сигнального супроводу
5	Невраховані течії / вітер	Відсутність гідрологічної компенсації при плануванні траєкторії підходу

Представлена класифікація є основою для побудови моделі профілактики інцидентів, а також лягає в основу подальших рекомендацій з регуляторного вдосконалення норм IMO.

Метод експертної оцінки. У зв'язку з фрагментарністю офіційних звітів деяких портів, було застосовано також метод експертного опитування (Delphi-метод) серед морських лоцманів, капітанів з досвідом понад 10 років, представників портових адміністрацій та інженерів-механіків з судноремонтних підприємств. Оцінювання проводилось за шкалою важливості (1-5 балів) для кожного чинника, з наступним агрегуванням результатів у вигляді рейтингової матриці. Отримані дані використовувались для верифікації класифікації інцидентів та побудови карти ризиків.

Результати аналізу інцидентів та регіональні відмінності. В ході дослідження було розглянуто п'ять прикладів зіткнень суден із причальною інфраструктурою, які отримали широке висвітлення у міжнародних морських звітах (табл. 5). Кожен з випадків демонструє унікальні виклики щодо процесу керування судном, технічного оснащення і координації між портом та екіпажом.

Таблиця 5 – **Аналіз реальних інцидентів**

Судно	Рік	Місце інциденту	Ключові причини	Наслідки
YM Witness	2019	Port of Keelung, TW	Людський фактор, помилкове реагування на попутний вітер	Пошкодження причалу, розрив швартов
Milano Bridge	2020	Port of Busan, KR	Надмірна швидкість підходу, відсутність координації	Знищення контейнерного крана, мільйонні збитки
DALI	2024	Baltimore, US	Втрата керування при втраті електроживлення	Обвал моста Francis Scott Key
Yuzhou Qi	2021	Port of Shanghai, CN	Невраховані течії, помилки у взаємодії з буксирами	Часткове пошкодження палубних споруд
Durban incident	2022	Port of Durban, SA	Потужний боковий вітер, неузгоджені дії екіпажу та порту	Пошкодження причальної балки

Дані інциденти дозволяють виявити загальні закономірності, особливо щодо швидкості наближення, метеоумов та якості комунікацій між мостиком і портовою службою.

У ході аналізу підходів до організації швартування і забезпечення навігаційної безпеки в портах різних географічних регіонів були виявлені суттєві відмінності, що свідчать про варіативність практик залежно від локального правового поля, технічної інфраструктури та наявних ресурсів. Зокрема, дослідження виявило ключові розбіжності за трьома напрямами:

1. Інфраструктурна готовність причалів, яка варіюється від сучасних, технічно оснащених об'єктів із сенсорними системами контролю стану конструкцій до обмежено обладнаних швартувальних комплексів, що підвищує ймовірність аварій за несприятливих умов.

2. Політика використання буксирів для швартування, що охоплює як вимоги до обов'язковості їх застосування, так і характеристики флоту: потужність, льодовий клас, тоннаж і система зв'язку з портовими диспетчерами. У деяких регіонах використання буксирів є суворо регламентованим, в інших - залишається на розсуд капітана.

3. Ступінь залучення лоцманів до процесу прийняття рішень, зокрема роль лоцмана не лише як навігаційного радника, а як активного учасника управління маневрами. У частині портів лоцмани мають розширені повноваження щодо остаточного схвалення траєкторії швартування, тоді як в інших - виконують лише консультативну функцію, табл. 6.

Результати аналізу вказують на тісний зв'язок між ефективністю швартовної інфраструктури, стандартами підготовки екіпажу та ризиком інцидентів. Так у країнах із високим рівнем автоматизації причалів та жорсткими вимогами до буксирного забезпечення аварії майже не трапляються. Натомість регіони, де відсутній централізований контроль, демонструють підвищену вірогідність інцидентів навіть за сприятливих умов.

Таблиця 6 – **Регіональні відмінності у практиках підготовки причалів**

Регіон	Мінімальна кількість буксирів	Автоматизована підтримка причалу	Частка аварій (%)*
Західна Європа	2-3	Так (більшість портів)	1.2
Східна Азія	1-2	Частково	3.5
Близький Схід	1	Обмежено	4.7
США / Канада	2	Залежно від штату / провінції	2.1
Африка (загалом)	0-1	Рідко	>6.0

**Частка аварій - орієнтовна кількість випадків пошкодження причалів на 1000 заходів у період 2020-2023 рр.*

Представлений огляд міжнародних актів у табл. 7, свідчить, що безпека швартування розглядається комплексно: від технічних вимог до суден (SOLAS) до підготовки персоналу (STCW) та організації управління ризиками (ISM Code).

Особливої уваги заслуговують рекомендації MSC.947(23), які враховують зростання розмірів сучасних контейнеровозів та необхідність залучення більш потужних буксирів і технологій автоматизації.

Таблиця 7 – **Міжнародно-правові акти, які регулюють безпеку швартування**

Документ	Основні положення	Вплив на безпеку швартування
SOLAS (Міжнародна конвенція з охорони людського життя на морі)	Вимоги до конструкції, навігаційного обладнання, справності систем керування та швартування	Забезпечує технічну готовність суден до безпечного маневрування
STCW (Конвенція про підготовку і дипломування моряків)	Стандарти підготовки екіпажу, оцінка умов, використання буксирів, прийняття рішень у складних умовах	Підвищує професійну компетентність екіпажу під час швартування
MSC.255(84) - Кодекс для лоцманських служб	Визначає повноваження та взаємодію лоцманів із капітанами й портами	Забезпечує якісну координацію дій під час швартування
MSC.947(23) - Рекомендації щодо швартування великих суден	Норми використання буксирів, систем автоматичного позиціювання, контроль навантаження швартовних систем	Знижує ризики аварій при швартуванні великих суден
ISM Code (A.741(18)) - Кодекс безпеки для управління суднами	Вимагає впровадження систем управління ризиками та коригувальних дій	Формує системний підхід до управління ризиками швартування

Водночас актуальним завданням залишається адаптація цих міжнародних норм до національних правових систем. Наприклад, в Україні їх імплементація потребує врахування специфіки портової інфраструктури, рівня технічного забезпечення й підготовки персоналу. Таким чином, ефективне застосування цих документів дозволяє знизити ймовірність інцидентів, забезпечити стабільність портових операцій і сприяти розвитку безпечного мореплавства.

Обговорення результатів. Аналіз інцидентів, розглянутих у попередніх розділах, свідчить про системну проблему недостатньої уніфікації підходів до швартування та взаємодії суден з береговою інфраструктурою, особливо в умовах обмежених або акваторій зі складними підходами. У випадках YM Witness (2023, Сингапур), Milano Bridge (2020, Південна Корея) та DALI (2024, США, Балтімор), ключовими чинниками стали перевищення допустимої швидкості підходу, несинхронізовані дії буксирів і відсутність оперативного інформаційного обміну між членами екіпажу і портовими службами. Особливу увагу привертають інциденти за участю суден типу Ro-Ro (Roll-on/Roll-off), конструктивні особливості яких зумовлюють високу чутливість до зміщень при швартуванні. Навіть незначне відхилення від оптимальної траєкторії під час підходу до рампи може призвести до критичних наслідків для безпеки судна, екіпажу та вантажу. Тут спостерігається дисбаланс між проектними характеристиками портів і сучасними габаритами суден. Тому недостатня адаптація берегової інфраструктури та підготовки персоналу призводить до багатомільйонних збитків і логістичних збоїв.

На основі проведеної експертної оцінки, основними чинниками ризику можна класифікувати згідно табл. 8.

Таблиця 8 – **Основні причини аварій під час швартування великих суден за результатами експертної оцінки**

Причина	Частота (%)	Приклад інциденту
Висока швидкість підходу	37%	Milano Bridge
Неефективна підтримка буксирами	25%	DALI, Durban
Недостатня готовність швартових	18%	YM Witness
Людський фактор (помилка навігатора)	15%	Yuzhou Qi
Відсутність оновлених карт причалів	5%	Регіони Півд. Азії

Варто підкреслити, що навіть наявність міжнародних рекомендацій таких як MSC.255(84) (BRM), MSC.947(23) (Safe Berthing), або A.741(18) (ISM Code) не гарантує фактичного впровадження процедур управління швартуванням у повному обсязі.

У більшості випадків, портова інфраструктура має власні локальні інструкції, які не завжди відповідають стандартам IMO, а екіпажі не проходять цільового тренінгу щодо індивідуальних особливостей конкретного порту.

У порівнянні з практиками країн Північної Європи, де широко застосовується попереднє візуальне та цифрове моделювання заходу судна, азійські порти часто ігнорують превентивні заходи.

У табл. 9 подано порівняння процедур.

Таблиця 9 – **Порівняння практик організації швартування в різних морських регіонах світу**

Регіон	Тренінг BRM	Тренажерна підготовка	Автоматизований контроль швидкості
Північна Європа	Так	Так	Так
Південно-Східна Азія	Частково	Ні	Ні
Північна Америка	Так	Частково	Частково

Таким чином, ефективне управління ризиками під час процесу швартування суден до причалу потребує не лише удосконалення технічних засобів, а й реалізації цілісного підходу до навчання екіпажу, координації з портовими службами та оновлення регламентів взаємодії.

Висновки

Проведене дослідження підтвердило, що інциденти під час процесу швартування великогабаритних суден мають складний міждисциплінарний характер і виникають внаслідок комбінації технічних, організаційних та людських чинників.

Ключовими джерелами ризику залишаються перевищення безпечної швидкості підходу до причалу, неефективна взаємодія з буксирним флотом, а також фрагментарне або навіть формальне впровадження процедур управління ресурсами на містку (Bridge Resource Management, BRM).

Порівняльний аналіз світових практик виявив суттєві відмінності в підготовці портової інфраструктури та складу екіпажів суден до швартовних операцій. Особливої уваги заслуговують успішні кейси інтеграції тренувань на симуляторах, стандартизованих процедур взаємодії між агентами порту та інструментарію підтримки навігаційних рішень. Натомість у ряді портів виявлено недостатнє нормативне забезпечення або обмежену адаптацію рекомендацій з боку IMO та IALA.

Отримані результати свідчать про нагальну потребу у впровадженні цілісної системи управління безпекою швартовних процесів. Така система має поєднувати оновлення технічної інфраструктури, регламентовану координацію дій усіх учасників операції та насамперед впровадження динамічних моделей оцінювання ризиків.

Перехід до адаптивних, на інформаційно-орієнтованих стратегій дозволить забезпечити вищий рівень надійності та безпеки в умовах зростання тоннажу суден і навігаційної складності устрою окремих портів.

СПИСОК ЛІТЕРАТУРИ

1. Othman, M. K., Mohd Sabri, N. S. A., Abdul Rahman, N. S. F., & Osnin, N. A. (2025). Port operators' perceptions and acceptance of maritime autonomous surface ships (MASS) operations: Insights from Malaysia. Case Studies on Transport Policy, 22, 101567. https://doi.org/10.1016/j.cstp.2025.101567
2. Xin, X., Liu, K., Yu, Y., & Yang, Z. (2025). Developing robust traffic navigation scenarios for autonomous ship testing: An integrated approach to scenario extraction, characterization, and sampling in complex waters. Transportation Research Part C: Emerging Technologies, 178, 105246. https://doi.org/10.1016/j.trc.2025.105246
3. Khan, R. U., Yin, J., Mustafa, F. S., & Shi, W. (2023). Factor assessment of hazardous cargo ship berthing accidents using an ordered logit regression model. Ocean Engineering, 284, 115211. https://doi.org/10.1016/j.oceaneng.2023.115211
4. Yang, L., Yang, J., Fan, A., Zhou, R., & Wang, L. (2025). Fire risk assessment of electric ships on inland waterway based on GT-FFTA: A case study of China. Ocean Engineering, 332, 121329. https://doi.org/10.1016/j.oceaneng.2025.121329
5. Cao, Y., Wang, X., Yang, Z., Wang, J., Wang, H., & Liu, Z. (2023). Research in marine accidents: A bibliometric analysis, systematic review and future directions. Ocean Engineering, 284, 115048. https://doi.org/10.1016/j.oceaneng.2023.115048
6. Adhita, I. G. M. S., Fuchi, M., Konishi, T., & Fujimoto, S. (2023). Ship navigation from a Safety-II perspective: A case study of training-ship operation in coastal area. Reliability Engineering & System Safety, 234, 109140. https://doi.org/10.1016/j.ress.2023.109140
7. Wang, Y., Li, P., Hong, C., & Yang, Z. (2025). Causation analysis of ship collisions using a TM-FRAM model. Reliability Engineering & System Safety, 260, 111035. https://doi.org/10.1016/j.ress.2025.111035
8. Liu, Z., Deng, J., Shu, Y., Gan, L., Song, L., Li, H., & Yang, Z. (2025). Spatiotemporal prediction of offshore wind fields based on a hybrid deep learning model for maritime navigation. Ocean & Coastal Management, 269, 107841. https://doi.org/10.1016/j.ocecoaman.2025.107841
9. Liu, J., Yang, F., Li, S., Lv, Y., & Hu, X. (2024). Testing and evaluation for intelligent navigation of ships: Current status, possible solutions, and challenges. Ocean Engineering, 295, 116969. https://doi.org/10.1016/j.oceaneng.2024.116969
10. Cao, W., Wang, X., Li, J., Zhang, Z., Cao, Y., & Feng, Y. (2024). A novel integrated method for heterogeneity analysis of marine accidents involving different ship types. Ocean Engineering, 312, 119295. https://doi.org/10.1016/j.oceaneng.2024.119295

11. Мельник, О. М., & Бичковський, Ю. В. (2021). Сучасна методика оцінки рівню безпеки судна та шляхи його підвищення. Розвиток транспорту, (2[9]), 37-46. https://doi.org/10.33082/td.2021.2-9.03 [in Ukrainian]
12. Мельник, О. М., & Бичковський, Ю. В. (2021). Врахування фактору стресу у системі забезпечення безпеки мореплавства. Вчені записки ТНУ ім. В. І. Вернадського. Технічні науки, 32(71)(4), 260-264. https://doi.org/10.32838/2663-5941/2021.4/39 [in Ukrainian]
13. Бичковський Ю.В., Мельник О.М. (2022). Роль та місце людського елементу у ситуації навалу або зіткнення судна з причалом. Вчені записки ТНУ ім. Вернадського. Технічні науки 33(72) № 1 - С. 270 - 276. https://doi.org/10.32838/2663-5941/2022.1/41
14. Melnyk, O., Bychkovsky, Y., Voloshyn, A. (2022) Maritime situational awareness as a key measure for safe ship operation. Scientific Journal of Silesian University of Technology. Series Transport. 114, 91-101. ISSN: 0209-3324. https://doi.org/10.20858/sjsutst.2022.114.8
15. Melnyk, O., Bychkovsky, Y., Onishchenko, O., Onyshchenko, S., Volianska, Y. (2023). Development the Method of Shipboard Operations Risk Assessment Quality Evaluation Based on Experts Review. Studies in Systems, Decision and Control, vol 481, 695-710. Springer, Cham. https://doi.org/10.1007/978-3-031-35088-7_40
16. Onyshchenko S., Bychkovsky Y., Melnyk O., Onishchenko O., Jurkovič M., Rubskyi V., Liashenko K. (2024). A model for assessing shipping safety within project-orientated risk management based on human element. Scientific Journal of Silesian University of Technology. Series Transport, 123, pp. 319 - 334. DOI: 10.20858/sjsutst.2024.123.16

Received (Надійшла) 18.07.2025
Accepted for publication (Прийнята до друку) 15.10.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Волянський Сергій Михайлович – кандидат технічних наук, доцент, доцент кафедри електричної інженерії суднових та роботизованих комплексів, Національний університет кораблебудування імені адмірала Макарова, Миколаїв, Україна;
Sergiy Volyansky – Candidate of Technical Sciences, Associate Professor of the Department of Electrical Engineering of Ship and Robotic Systems, Admiral Makarov National University of Shipbuilding, Mykolaiv, Ukraine;
e-mail: npbulgakov2@gmail.com; ORCID Author ID: https://orcid.org/0000-0001-7922-0441;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57190490495.

Бичковський Юрій Вячеславови – доктор філософії, доцент, доцент кафедри навігації і керування судном, Одеський національний морський університет, Одеса, Україна;
Yuriy Bychkovsky – Doctor of Philosophy, Associate Professor, Associate Professor of Navigation and Ship Handling, Odesa National Maritime University, Odesa, Ukraine;
e-mail: ologinov@ukr.net; ORCID Author ID: https://https://orcid.org/0000-0003-1459-9029;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57464000400.

Мельник Олексій Миколайович – доктор технічних наук, професор, завідувач кафедри судноводіння і морської безпеки, Одеський національний морський університет, Одеса, Україна;
Oleksiy Melnyk – Doctor of Technical Sciences, Professor, Head of 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.

Волошин Андрій Олександрович – кандидат технічних наук, професор, професор кафедри судноводіння і морської безпеки, Одеський національний морський університет, Одеса, Україна;
Andriy Voloshyn – Candidate of Technical Sciences, Professor, Professor of the Department of Navigation and Maritime Safety, Odesa National Maritime University, Odesa, Ukraine;
e-mail: ologinov@ukr.net; ORCID Author ID: https://orcid.org/0000-0003-3993-5826;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57223358487.

Systematic analysis of the causes of navigational incidents during ship mooring and ways to improve the international regulatory framework

Sergiy Voliansky, Yuriy Buchkovsky, Oleksiy Melnyk, Andriy Voloshyn

Abstract. Relevance. Due to the increasing size of merchant ships and the load on port infrastructure, the issue of ensuring safety during the mooring process is becoming strategically important. Frequent incidents in ports around the world indicate the imperfection of existing practices and the fragmentation of the implementation of modern risk management methods, in particular Bridge Resource Management (BRM). **The object of study** is navigation incidents during mooring of large-tonnage vessels. **Purpose** - to systematically analyze the factors that cause mooring incidents and to formulate recommendations for improving the international regulatory framework and risk management procedures in ports. Methodology. The study uses comparative analysis of incidents, expert risk assessment, classification of causes based on a factor model, and analysis of international IMO regulations (SOLAS, STCW, MSC.255(84), etc.). **Results.** The study confirmed that the key causes are excessive approach speed, inconsistency of towing support actions, human factor, and lack of standardized BRM procedures on the approach to the berth. Comparison of practices in different countries has shown uneven implementation of simulator training and intelligent decision support systems. **Conclusions.** The paper provides recommendations for the harmonization of regional practices that promote the development of decision support systems and the introduction of simulator training for ship crews. The expediency of integrating BRM in the port and developing adaptive risk management protocols, taking into account the real state of infrastructure provision, is proved.

Keywords: maritime transport, mooring of ships; ship control, maneuvering, incidents; navigation safety; international conventions; human factor; port infrastructure; risk management; bridge resource management (BRM).

УДК 623.746.2 **doi: 10.26906/SUNZ.2025.4.12-16**

Р. В. Куліш

Українська державна льотна академія, Кропивницький, Україна

ДОСЛІДЖЕННЯ МЕТОДІВ ПЛАНУВАННЯ ПОЛЬОТІВ БЕЗПІЛОТНИХ ЛІТАЛЬНИХ АПАРАТІВ В СКЛАДНИХ УМОВАХ ПОЛЬОТУ ДЛЯ МОНІТОРИНГУ СТАНУ ЕЛЕМЕНТІВ КРИТИЧНОЇ ІНФРАСТРУКТУРИ

Анотація. Актуальність. Переважна більшість сучасних алгоритмів планування маршруту польоту БпЛА, призначених для бортового застосування в режимі реального часу, не враховують динаміку БпЛА, що негативно позначається на точності та оптимальності планування маршруту, особливо при спостереженні за рухомими об'єктами. **Об’єкт дослідження**: процеси моніторингу елементів критичної інфраструктури безпілотними літальними апаратами **Мета статті**: оцінювання методів планувальників польотів безпілотних літальних апаратів в складних умовах польоту для моніторингу стану елементів критичної інфраструктури. **Результати дослідження**. У статті проведено порівняльний аналіз існуючих методів побудови маршруту польоту БпЛА під час спостереження за об'єктами – елементами об’єктів критичної інфраструктури, а саме методу повного перебору та жадібного алгоритму із запропонованим автором удосконаленим алгоритмом. При моделюванні емуляції польоту БпЛА за допомогою пакету MATLAB було розроблено двадцять сценаріїв. Для кожного сценарію було обрано список із п’яти об'єктів, які підлягають моніторингу. Об’єкти обирались як динамічні, так і стаціонарні. На підставі аналізу розглянутих сценаріїв були сформульовані висновки про ефективність удосконаленого методу. **Висновки.** Маршрути, побудовані за допомогою удосконаленого методу, повністю збіглися з маршрутами, побудованими за допомогою методу повного перебору. При цьому час обчислень суттєво нижчий за існуючі методи, що дозволяє використовувати удосконалений алгоритм у складі комплексу бортового програмного забезпечення та оперативно будувати й змінювати маршрут залежно від обстановки щодо об’єктів моніторингу. Використання ідентифікатора у контурі системи управління для компенсації впливу вітру час польоту маршрутом дозволяє скоротити на 15%.

Ключові слова: безпілотний літальний апарат; маршрутизація; моніторинг; об’єкти критичної інфраструктури; планування маршрутів.

Вступ

Постановка проблеми. Існуючі підходи до вирішення завдання планування маршруту польоту БпЛА в складних умовах польоту насамперед відносяться до процесу послідовного обльоту та спостереження нерухомих об'єктів. Відомо значна кількість методів маршрутизації, у якості вихідних даних вказуються координати розташування об'єктів, які підлягають моніторингу.

Однак у цих методах недостатня увага приділена обмеженим динамічним можливостям безпілотного літального апарату, а їх урахування призводить до сильних змін плану обльоту об'єктів, коли на кожному кроці планування вихідних даних, крім об'єктів, потрібно мати на увазі напрям і значення швидкості самого БпЛА. У разі раптових змін динамічної обстановки перепланування польоту має здійснюватися без участі людини.

Ще більші труднощі виникають під час планування обльоту мобільних об'єктів. По-перше, для потрапляння чергового рухомого об'єкта у вікно спостереження бортової апаратури БпЛА необхідно прогнозувати його рух, а для цього у вихідних даних об'єкта потрібно враховувати не лише вихідні координати розташування об'єкта, а й вектор швидкості його руху. По-друге, більш істотне ускладнення полягає в тому, що у разі порушення не самого початкового маршруту обльоту, а графіка польоту, наприклад, через дію складних умов польоту, час прогнозування змінюється, а отже, і нове розташування пунктів потребує повторного перепланування маршруту.

Таким чином розроблення методів планування маршруту польоту безпілотного літального апарату при моніторингу елементів об’єкту критичної інфраструктури в складних умовах польоту є актуальною тематикою дослідження.

Аналіз останніх досліджень і публікацій. В статті [1] розглянуто проблему маршрутизації безпілотного літального апарату для моніторингу стаціонарних об’єктів. Принциповою відмінністю розв'язуваної задачі від існуючих методів є припущення про альтернативність управління.

Завдання оптимізації вирішується за умови, що всі коефіцієнти функції ризику в правій частині рівняння Белмана залежать від j-го номера обраної альтернативи. Проведений розрахунок польотних ситуацій свідчить про необхідність використання спрощеного підходу розрахунку точки польоту, для якого додатково потрібні середні значення ординат ризику.

В роботі [2] вирішена задача планування маршруту обльоту мобільних об’єктів, що зберігають напрямок свого руху протягом тривалого проміжку часу із використанням методу математичної оптимізації - динамічного програмування. При вирішенні завдання побудови оптимального маршруту враховувалась динаміка польоту безпілотного літального апарату, що дозволяє послідовно додавати об’єкти до маршруту.

Дослідники погоджуються, що для підвищення ймовірність успіху місії, БпЛА, траєкторія польоту повинна бути ретельно розрахована з урахуванням методів оптимізації при існуючих обмеженнях. Незважаючи на те, що алгоритми на основі графів, такі як діаграма Вороного, search A*, D* lite та інші класичні методи, такі як швидке дослідження випадкових дерев (RRT), штучні потенційні поля (APF) та ймовірнісні дорожні карти (PRM), зазвичай використовуються для розрахунку траєкторії БпЛА, вони

вимагають створення карт вартості складних полів і, зазвичай, їх недоліком є збіжність до локальних оптимальних рішень [3, 4].

Останніми роками метаевристичні алгоритми, які є розгалуженням методів штучного інтелекту, почали використовуватися як планувальники маршруту груп або окремих БпЛА через їх переваги щодо простоти реалізації, обчислювальної складності та конфігурованих або настроюваних структур [5–7].

У роботі Xu C., Duan H., Liu F. [8] пропонується спрямувати фуражирів в алгоритмі штучної бджолиної колонії (ABC) до найкращого поточного джерела їжі, використовуючи хаотичні змінні для планувальника маршруту, названого хаотичним ABC (CABC), і наочно доведена перевага CABC над алгоритмом ABC.

Група дослідників Y. Zhang, Wu L.; Wang S [9] конвертували необроблені значення придатності джерел їжі за допомогою стратегії масштабування придатності та використали рівняння Лоренца для генерації випадкових чисел, необхідних для фаз бджіл, щоб покращити ефективність планування шляху алгоритму ABC. В іншому дослідженні даних авторів [10] були покращені ймовірності вибору кваліфікованих рішень шляхом врахування механізму масштабування придатності для алгоритму оптимізації рою часток (PSO). Крім того, були адаптивно відрегульовані інерційна вага та коефіцієнти прискорення рівняння оновлення чисельності в рої часток та використано генератор хаотичних випадкових чисел для визначення значень випадкових коефіцієнтів того самого рівняння. Алгоритм PSO, представлений у роботі та названий адаптивним хаотичним PSO для масштабування (FAC-PSO), використовувався як планувальник шляху.

В роботі [11] ідея концепції пам’яті та соціальної інформації PSO для свого алгоритму гравітаційного пошуку (GSA), і її ефективність у вирішенні проблем планування траєкторії БпЛА, була перевірена шляхом порівняння зі стандартними PSO, GSA та двома іншими GSA – базових моделей.

У дослідженні [12] запропоновано управляти версією алгоритму PSO, для якої кожна частинка змінювала відповідні швидкості, керуючи найкращим рішенням невеликої групи рішень як планувальником шляху.

Колективом авторів [13] запропоновано новий підхід до обміну інформацією між кваліфікованими рішеннями алгоритму світлячка (FA) й наведено модифікований метод FA (MFA).

Було представлено детальне порівняння між модифікованим методом та іншими метаевристичними планувальниками траєкторії БпЛА, такими як PSO, оптимізація колонії мурашок (ACO), диференціальна еволюція (DE), оптимізація на основі біогеографії (BBO), еволюційні стратегії (ES), популяція – based incremental learning (PBIL), генетичний алгоритм (GA) і, нарешті, варіант GA, відомий як stud GA (SGA).

Метою роботи є оцінювання методів планувальників польотів безпілотних літальних апаратів в складних умовах польоту для моніторингу стану елементів критичної інфраструктури.

Основний матеріал

Переважна більшість сучасних алгоритмів планування маршруту польоту БпЛА, призначених для бортового застосування в режимі реального часу, не враховують динаміку БпЛА, що негативно позначається на точності та оптимальності планування маршруту, особливо при спостереженні за рухомими об'єктами. Тому, другим методом, з яким буде проводитись порівняння удосконалених методів, буде метод на основі жадібного алгоритму, що не враховує динаміку БпЛА та можливу мобільність об'єктів спостереження.

Існує кілька варіантів роботи жадібного алгоритму, але, в цілому, їх принцип зводиться до того, що знаходиться оптимальне рішення для кожної локальної задачі, але рішення глобальної задачі може не бути оптимальним. Для пошуку оптимального маршруту це виражається у тому, що наступним завжди вибирається об'єкт, “найближчий” до поточного положення БпЛА. При розробці модуля емуляції польоту БпЛА за допомогою пакету MATLAB були застосовані такі припущення:

1. Для спрощення розрахунків та зниження обчислювальних витрат усі обчислення будуть проводитись у пов'язаній системі координат, таким чином БпЛА завжди матиме координати:

$$x = 0, z = 0, \ \alpha = 0.$$

2. Рух об'єктів протягом часу, необхідного БпЛА для підльоту до них – прямолінійний з постійною швидкістю. Це припущення адекватне поточній ситуації, оскільки час польоту БпЛА порівняно малий і мобільні об'єкти спостереження елементів критичної інфраструктури істотно не змінюють курс і швидкість за незначний час.

З метою перевірки роботи програми було розроблено двадцять сценаріїв. Три сценарії взяті як базові, що дозволило оцінити удосконалені методи. Для першого сценарію було обрано список із п’яти об'єктів, які підлягають моніторингу. Об’єкти обрані як динамічні, так і стаціонарні. Характеристики об'єктів наведено у табл. 1. Удосконалені методи взяті за результатами дослідження [1] та [2].

БпЛА починає рух з точки з координатами:

$$X = 0; \ Z = 600 \text{ м}.$$

Швидкість польоту БпЛА складає 40 м/с, початковий курс – 110^0.

Точка, в якій БпЛА має завершити свій рух, має координати:

$$X = 1000 \text{ м}, \ Z = 650 \text{ м}.$$

У цьому сценарії із застосуванням трьох методів було отримано однаковий маршрут.

Порівняльні характеристики та результати роботи методів наведено у табл. 2.

Графік прольоту БпЛА, розрахований за допомогою удосконаленого методу та реальний час прольоту вказані у таблиці 3.

Як видно з результатів, отриманих при моделюванні сценарію №1, алгоритм, побудований на основі методу повного перебору, виконувався дуже довго, що не прийнятно для бортової реалізації через значний обчислювальний ресурс.

Таблиця 1 – **Характеристики об’єктів моніторингу для першого сценарію моделювання**

№ об'єкта	Назва об'єкта	X, м	Z, м	Курс, град	Швидкість, м/с
1	Об’єкт №1 (динамічний)	50	500	70	30
2	Об’єкт №2 (динамічний)	560	205	110	5
3	Об’єкт №3 (стаціонарний)	380	120	-	-
4	Об’єкт №4 (стаціонарний)	690	600	-	-
5	Об’єкт №5 (стаціонарний)	200	330	-	-

Джерело: розроблено автором за даними [1, 2]

Таблиця 2 – **Порівняльні результати роботи алгоритмів**

Метод	Прогнозований час польоту, хв.	Час роботи алгоритму, хв.	Послідовність обльоту об'єктів
Повний перебір	46	295	1-5-3-2-4
Жадібний алгоритм	45	0,2	1-5-3-2-4
Удосконалений метод	36	0,1	1-5-3-2-4

Таблиця 3 – **Графік обльоту БпЛА елементів об’єктів критичної інфраструктури за першим сценарієм**

№ об'єкта	Назва об'єкта	Реальний час прольоту об'єкта, хв	Час прольоту за графіком, хв	Абсолютне значення, хв.
1	Об’єкт №1 (динамічний)	7,7	7,3	0,4
5	Об’єкт №5 (стаціонарний)	13,5	13,1	0,4
3	Об’єкт №3 (стаціонарний)	20,4	20	0,4
2	Об’єкт №2 (динамічний)	25,5	24,8	0,3
4	Об’єкт №4 (стаціонарний)	37	36	1

Удосконалений метод та жадібний алгоритм виконувались приблизно однаковий час та запропонували однаковий маршрут.

У той же час метод на основі жадібного алгоритму дав неправильну оцінку часу польоту, що може негативно позначитися на побудові маршруту за умов, коли час польоту близький до максимального часу польоту даного БпЛА.

Для другого сценарію в об'єктах зі сценарію №1 стаціонарний об’єкт №5 замінили на динамічний об’єкт, що рухається від точки з координатами

$$X = 60 \text{ м}, Z = 400 \text{м}$$

із швидкістю 15 м/с.

Параметри об'єктів наведено у табл. 4.

БпЛА почав свій рух з точки з координатами

$$X = 0,\ Z = 600 \text{ м}.$$

Швидкість польоту БЛА – 40 м/с, початковий курс – 110 градусів.

Точка, в якій БпЛА має завершити свій рух, має координати

$$X = 1000 \text{ м}, Z = 650 \text{ м}.$$

Результати моделювання наведено у табл. 5 та 6.

Таблиця 4 – **Характеристики об’єктів моніторингу для другого сценарію моделювання**

№ об'єкта	Назва об'єкта	X, м	Z, М	Курс, град.	Швидкість, м/с
1	Об’єкт №1 (динамічний)	50	500	70	30
2	Об’єкт №2 (динамічний)	560	205	110	5
3	Об’єкт №3 (стаціонарний)	380	120	-	-
4	Об’єкт №4 (стаціонарний)	200	330	-	-
5	Об’єкт №5 (динамічний)	60	390	190	15

Таблиця 5 – **Результати моделювання за другим сценарієм**

Метод	Прогнозований час польоту, хв.	Час роботи алгоритму, хв.	Послідовність обльоту об'єктів
Повний перебір	57	265	5-4-3-2-1
Жадібний алгоритм	48	0,2	1-5-4-3-2
Удосконалений метод	44	0,1	5-4-3-2-1

Таблиця 6 – **Графік обльоту БпЛА елементів об'єктів критичної інфраструктури за другим сценарієм**

№ об'єкта	Назва об'єкта	Реальний час прольоту об'єкта, хв	Час прольоту за графіком, хв	Абсолютне значення, хв.
5	Об'єкт №5 (динамічний)	6	5,6	0,4
4	Об'єкт №4 (стаціонарний)	10,2	9,15	1,05
3	Об'єкт №3 (стаціонарний)	18,5	17,3	1,2
2	Об'єкт №2 (динамічний)	23,8	24	1,8
1	Об'єкт №1 (динамічний)	43,2	44	0,5

Як видно з отриманих результатів, в другому сценарії жадібний алгоритм вибрав невірний маршрут, і час польоту за цим маршрутом перевищив час польоту маршрутом, запропонований удосконаленим методом.

Маршрути, запропоновані рештою алгоритмів, збіглися.

Для третього сценарію було вибрано два мобільні об'єкти – елементи критичної інфраструктури. У цьому сценарії перевірялася ефективність удосконаленого методу зниження дії бокового вітру на політ БпЛА.

У табл. 7 наведено параметри об'єктів.

БпЛА почав свій рух із точки з координатами

$$X = 0,\ Y = 125.$$

Швидкість польоту БЛА – 40 м/с.

Швидкість бічного вітру на проміжку часу від 5 до 20 секунд від початку польоту становила 6 м/с. В решту часу вітер був відсутній.

Були розглянуті випадки, коли вітер був відсутній, коли вітер був присутній, але було відключено ідентифікатор кута вітрового зносу та випадок, коли був вітер та ідентифікатор був включений.

У табл. 8 зазначено час моніторингу динамічного об'єкту для кожного випадку.

Таблиця 7 – **Характеристики об'єктів моніторингу для третього сценарію моделювання**

№ об'єкта	Назва об'єкта	X, м	Z, м	Курс, град.	Швидкість, м/с
1	Об'єкт №1 (динамічний)	370	500	170	10
2	Об'єкт №2 (динамічний)	500	300	45	5

Таблиця 8 – **Результати моделювання за третім сценарієм**

№ об'єкта	Назва об'єкта	Час моніторингу без вітру, с	Час моніторингу при вимкненому ідентифікаторі, с	Час моніторингу при працюючому ідентифікаторі, с
1	Об'єкт №1 (динамічний)	13	15	13,2
2	Об'єкт №2 (динамічний)	24,5	29	25

За результатами моделювання можна зробити висновок, що з компенсації впливу вітру з допомогою даних, отриманих від ідентифікатора, час польоту маршрутом вдалося скоротити на 15%, що свідчить про ефективність використання ідентифікатора у контурі системи управління.

Висновки

Таким чином, з аналізу представлених сценаріїв можна сформувати висновки про ефективність удосконаленого методу побудови маршруту польоту БпЛА під час спостереження за об'єктами –

елементами об’єктів критичної інфраструктури. Маршрути, побудовані розробленими в [1] та [2] методами, повністю збіглися з маршрутами, побудованими за допомогою методу повного перебору. При цьому час обчислень суттєво нижчий, що дозволяє використовувати удосконалений алгоритм у складі комплексу бортового програмного забезпечення та оперативно будувати й змінювати маршрут залежно від обстановки щодо об’єктів моніторингу.

СПИСОК ЛІТЕРАТУРИ

1. Галінський Д.О., Куліш Р.В. Метод моніторингу стану стаціонарних елементів об’єктів критичної інфраструктури безпілотними літальними апаратами з використанням динамічного програмування. Системи управління, навігації та зв’язку. Полтава, 2023. Вип. 1(71). С. 10-15. https://doi.org/10.26906/SUNZ.2023.1.010
2. Kulish R. Модель маршрутизації обльоту мобільних об'єктів безпілотним літальним апаратом / R. Kulish, O. Matiushrnko // Системи управління, навігації та зв’язку. Полтава: ПНТУ, 2023. Т. 2(72). С. 20-25. https://doi.org/10.26906/SUNZ.2023.2.020
3. Latombe, J. C. 1991. Robot Motion Planning. Kluwer Academic. https://doi.org/10.1007/978-1-4615-4022-9
4. Ait Saadi, A.; Soukane, A.; Meraihi, Y.; Benmessaoud Gabis, A.; Mirjalili, S.; Ramdane-Cherif, A. UAV path planning using optimization approaches: A survey. Arch. Comput. Methods Eng. 2022, 29, 4233–4284. https://doi.org/10.1007/s11831-022-09742-7
5. Bal, M. An overview of path planning technologies for unmanned aerial vehicles. Therm. Sci. 2022, 26, 2865–2876. Електронний ресурс. Режим доступу https://doi.org/10.2298/TSCI2204865B
6. Wu, Y. A survey on population-based meta-heuristic algorithms for motion planning of aircraft. Swarm Evol. Comput. 2021, 62, 100844. Електронний ресурс. Режим доступу https://doi.org/10.1016/j.swevo.2021.100844
7. Majeed, A.; Hwang, S.O. A multi-objective coverage path planning algorithm for UAVs to cover spatially distributed regions in urban environments. Aerospace 2021, 8, 343. https://doi.org/10.3390/aerospace8110343
8. Saeed, R.A.; Omri, M.; Abdel-Khalek, S.; Ali, E.S.; Alotaibi, M.F. Optimal path planning for drones based on swarm intelligence algorithm. Neural Comput. Appl. 2022, 34, 10133–10155. https://doi.org/10.1007/s00521-022-06998-9
9. Xu, C.; Duan, H.; Liu, F. Chaotic artificial bee colony approach to Uninhabited Combat Air Vehicle (UCAV) path planning. Aerosp. Sci. Technol. 2010, 14, 535–541. https://doi.org/10.1016/j.ast.2010.04.008
10. Zhang, Y.; Wu, L.; Wang, S. UCAV path planning based on FSCABC. Inf.-Int. Interdiscip. J. 2011, 14, 687–692. Електронний ресурс. Режим доступу https://www.academia.edu/7461343/UCAV_path_planning_based_on_FSCABC
11. Zhang, Y.; Wu, L.; Wang, S. UCAV path planning by fitness-scaling adaptive chaotic particle swarm optimization. Math. Probl. Eng. 2013, 2013. Електронний ресурс. Режим доступу http://dx.doi.org/10.1155/2013/705238
12. Wang, G.G.; Guo, L.; Duan, H.; Liu, L.; Wang, H. A modified firefly algorithm for UCAV path planning. Int. J. Hybrid Inf. Technol. 2012, 5, 123–144. https://www.researchgate.net/publication/266884589_A_Modified_Firefly_Algorithm_for_UCAV_Path_Planning
13. Wang, G.G.; Guo, L.; Duan, H.; Wang, H.; Liu, L.; Shao, M. A hybrid metaheuristic DE/CS algorithm for UCAV three-dimension path planning. Sci. World J. 2012, 2012, 583973. https://doi.org/ 10.1100/2012/583973
14. Timochko, A; Fustii, V; Kolesnyk, A; Olizarenko, S; Kalashnyk, G; Kulish, R; Tymoschuk, O, Galinskyi, D; Inter-Object Navigation of Unmanned Aerial Vehicles to Increase the Efficiency and Accuracy of Inspection of Power Lines, Problemele Energeticii Regionale Nr. 1(57) / 2023 / ISSN 1857-0070. https://doi.org/10.52254/1857-0070.2023.1-57.03

Received (Надійшла) 11.07.2025
Accepted for publication (Прийнята до друку) 22.10.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Куліш Руслан Валерійович – аспірант, Українська державна льотна академія, Кропивницький, Україна;
Ruslan Kulish – Postgraduate student, Ukraine State Flight Academy, Kropyvnytskyi, Ukraine;
e-mail: ruslankulishkrv@gmail.com; ORCID Author ID: https://orcid.org/0000-0002-5590-1370.

Research of flight planning methods of unmanned aerial vehicles in complex flight conditions for monitoring the state of critical infrastructure elements

Ruslan Kulish

Abstract. Relevance. The vast majority of modern UAV flight path planning algorithms designed for real-time onboard application do not take into account UAV dynamics, which negatively affects the accuracy and optimality of route planning, especially when tracking moving objects. **Object of research**: processes of monitoring critical infrastructure elements with UAVs. **Purpose of the article**. evaluation of methods for flight planners of unmanned aerial vehicles in difficult flight conditions for monitoring the condition of critical infrastructure elements. **Research results**. The article presents a comparative analysis of existing methods for constructing a UAV flight route during the observation of objects - elements of critical infrastructure objects, namely the exhaustive search method and the greedy algorithm with the improved algorithm proposed by the author. When modeling the emulation of the UAV flight using the MATLAB package, twenty scenarios were developed. For each scenario, a list of five objects to be monitored was selected. The objects were selected as dynamic and stationary. Based on the analysis of the considered scenarios, conclusions were formulated about the effectiveness of the improved method. **Conclusions**. The routes constructed using the improved method completely coincided with the routes constructed using the exhaustive search method. At the same time, the calculation time is significantly lower than the existing methods, which allows using the improved algorithm as part of the on-board software complex and quickly constructing and changing the route depending on the situation with respect to the monitored objects. Using an identifier in the control system circuit to compensate for the influence of the wind allows reducing the flight time along the route by 15%.

Keywords: unmanned aerial vehicle, routing, monitoring, critical infrastructure objects, route planning.

УДК 004. 056.55:004.312.2 **doi: 10.26906/SUNZ.2025.4.17-21**

Olena Sevostianova[1], Nataliia Kosenko[2], Vladlen Filippov[1], Maksym Diachenko[1], Ivan Kharakhaichuk[1]

[1] Kharkiv National University of Radio Electronics, Kharkiv, Ukraine
[2] Beketov National University of Urban Economy in Kharkiv, Ukraine

ENHANCING TRUSTWORTHINESS OF IOT-ENABLED AUTOMATED VEHICLE LOCALIZATION SYSTEMS

Abstract. Relevance. Autonomous vehicles rely on multi-sensor localization systems operating within IoT infrastructures, creating interconnected vulnerabilities from sensor anomalies, network failures, and cybersecurity threats that require comprehensive solutions addressing both vehicle-level and infrastructure-level reliability challenges. **The object of research** is IoT-enabled automated vehicle localization systems requiring trustworthy operation under adverse conditions, including sensor malfunctions, GPS spoofing attacks, and infrastructure failures. **Purpose of the article** is to develop and validate a unified resilience framework that integrates transformer-based anomaly detection for in-vehicle sensor streams with federated learning agents deployed across IoT edge gateways, ensuring sub-second recovery from infrastructure failures while maintaining localization accuracy. **Research results.** The proposed framework achieves 94-98% anomaly detection accuracy while maintaining localization errors below 0.5 meters during fault conditions. The federated learning component demonstrates 40% reduced communication overhead compared to centralized approaches, with sub-second failover capabilities during infrastructure failures. Explainable ML integration provides interpretable alerts through transformer attention mechanisms, enabling real-time system diagnostics. **Conclusions.** The unified framework successfully addresses critical challenges in autonomous vehicle deployment by combining multi-layer anomaly detection, coherent reliability broadcasting, and explainable AI techniques, providing a comprehensive foundation for trustworthy autonomous vehicle operation in IoT-enabled smart city environments.

Keywords: autonomous vehicles, IoT reliability, anomaly detection, federated learning, sensor fusion, cybersecurity.

Introduction

The rapid advancement of autonomous vehicle (AV) technology has fundamentally transformed modern transportation systems, with projections indicating widespread deployment across urban environments within the next decade. These sophisticated systems rely heavily on multi-sensor localization frameworks that integrate Global Positioning System (GPS) receivers, Light Detection and Ranging (LiDAR) sensors, Inertial Measurement Units (IMUs), and camera arrays to achieve centimeter-level positioning accuracy. However, this sensor fusion increasingly operates within complex Internet of Things (IoT) infrastructures encompassing in-vehicle controllers, edge computing gateways, and cloud-based processing platforms.

This technological convergence introduces two critical and interconnected challenges that directly impact system trustworthiness. First, autonomous vehicles face persistent anomalies in localization systems due to sensor hardware failures, environmental interference, adversarial spoofing attacks, and algorithmic model drift. Second, the underlying IoT infrastructure experiences reliability issues including hardware component failures, network latency variations, cybersecurity threats targeting communication protocols, firmware vulnerabilities, and real-time processing constraints.

The intersection of these challenges creates a complex reliability landscape where traditional isolated approaches to anomaly detection and infrastructure hardening prove insufficient. Current research typically addresses vehicle localization anomalies and IoT infrastructure reliability as separate domains, failing to recognize their fundamental interdependence in operational environments.

Consider a fleet of autonomous shuttles operating in dense urban environments where localization systems encounter multiple simultaneous stressors. GPS signals experience frequent blockage or intentional spoofing near high-rise buildings and underground tunnels. Environmental conditions such as heavy precipitation, fog, and extreme temperatures cause intermittent sensor malfunctions.

The supporting IoT infrastructure faces edge device failures, network router outages, and sophisticated cyberattacks targeting Controller Area Network (CAN) buses and edge application programming interfaces (APIs).

Under these conditions, autonomous vehicles must continuously detect anomalies across multiple system layers, accurately localize fault sources, and restore precise positioning without human intervention. This requirement demands a holistic approach that simultaneously addresses three interconnected problems.

Problem 1. Real-time anomaly detection in vehicle localization systems using both onboard sensor data and distributed network telemetry.

Problem 2. IoT infrastructure reliability assurance through redundancy mechanisms, edge recovery protocols, and self-healing architectures.

Problem 3. Integrated system coordination where anomaly detection algorithms depend on Internet of Things response characteristics while infrastructure strategies must support localization integrity requirements.

Review of Recent Studies and Publication. Recent advances in automotive cybersecurity and IoT infrastructure reliability have established important foundations for integrated anomaly detection systems.

Hanif et al. [1] address the growing threats to intra-vehicle networks, focusing on Controller Area Network (CAN) security where ECUs vary widely in processing power, storage, memory, and connectivity. Their research emphasizes the critical need for efficient

intrusion detection systems in modern connected and autonomous vehicles. Recent developments in CAN security utilize attention mechanisms and optimization algorithms to enhance intrusion detection capabilities, addressing the inherent lack of security features in traditional CAN protocols.

Transformer-based approaches for multivariate time-series anomaly detection have shown significant promise, with Liu et al. [2] proposing methods that capture temporal dependencies and correlations between variables simultaneously through inter-variable attention mechanisms. These approaches address the challenge of spotting deviations from regular patterns in time-series data compiled concurrently from various sensors and systems, finding applications across diverse industries for system maintenance tasks.

Khan et al. [3] demonstrate the application of federated learning specifically for GPS spoofing detection in autonomous vehicles, representing a critical advancement in addressing satellite-based positioning vulnerabilities. Their work builds upon broader research in GPS attack mitigation, with Cheng et al. [4] developing comprehensive detection strategies using learning from demonstration techniques for connected and autonomous vehicles.

The broader landscape of IoT security has been extensively surveyed by Berdik et al. [5], who examine blockchain-based approaches for information systems management and security. This foundational work provides context for understanding security challenges across distributed IoT infrastructures that support autonomous vehicle operations.

Federated learning applications in IoT environments have been comprehensively analyzed by Koubaâ et al. [6], who identify key security issues, limitations, challenges, and solutions specific to IoT systems integration. Their work emphasizes the importance of privacy preservation in distributed learning scenarios.

Kumar and Gandhi [7] further advance this field by proposing optimal federated learning-based intrusion detection specifically designed for IoT environments, demonstrating practical implementations of collaborative security approaches.

The evolution of deep learning-based network anomaly detection has been systematically reviewed by Kwon et al. [8], providing essential background for understanding the progression from traditional statistical methods to modern neural network approaches.

Banafa et al. [9] contribute to this field through experimental assessment of real-time anomaly detection techniques specifically designed for automotive cybersecurity applications, emphasizing practical deployment considerations.

Song et al. [10] explore the integration of transformer architectures with adversarial training for multivariate time series anomaly detection in IoT contexts, demonstrating how modern attention mechanisms can be adapted for distributed sensor networks. Their work bridges the gap between general-purpose anomaly detection and IoT-specific requirements, addressing computational constraints and real-time processing needs essential for autonomous vehicle applications.

The purpose of this work is to presents a unified resilience framework that bridges the gap between vehicle localization anomaly detection and IoT infrastructure reliability. Our primary contributions include:

1. Multi-layer anomaly detection architecture combining transformer-based models for in-vehicle sensor streams with federated learning agents across IoT edge infrastructure.

2. Coherent reliability broadcasting technique enabling sub-second failover between edge gateways and cloud resources during infrastructure failures.

3. Explainable anomaly localization providing both sensor-level and timestamp-precise fault identification through attention mechanism analysis.

4. Comprehensive evaluation framework demonstrating system performance across realistic urban deployment scenarios with quantitative reliability metrics.

Main part

Proposed Framework. System Architecture Overview. Our unified resilience framework (Fig. 1) integrates vehicle localization anomaly detection with IoT infrastructure reliability through a hierarchical architecture operating across three primary layers: in-vehicle processing, edge computing infrastructure, and cloud-based coordination. This multi-layer approach ensures comprehensive anomaly detection while maintaining system responsiveness and fault tolerance.

The framework employs a distributed processing model where each autonomous vehicle maintains onboard anomaly detection capabilities while participating in fleet-wide collaborative learning through edge-based federated learning agents. Cloud resources provide coordination services, model updates, and backup processing capacity during edge infrastructure failures.

Multi-Layer Anomaly Detection. The in-vehicle anomaly detection subsystem employs transformer-based neural networks specifically optimized for multivariate time-series analysis of CAN-bus telemetry and multi-sensor data streams. Our transformer architecture incorporates specialized attention mechanisms designed to identify both spatial correlations between sensors and temporal patterns indicative of anomalous behavior.

The model processes synchronized data streams from GPS receivers, IMU sensors, LiDAR arrays, and camera systems, along with CAN-bus message traffic. A multi-head attention mechanism enables the system to simultaneously monitor different aspects of sensor behavior, including signal amplitude variations, timing anomalies, and cross-sensor consistency violations.

GPS spoofing detection utilizes a specialized module that compares satellite-derived position estimates with camera-based lane detection and LiDAR-generated local maps. When discrepancies exceed predefined thresholds, the system triggers enhanced verification protocols and can initiate fallback to dead reckoning using IMU data.

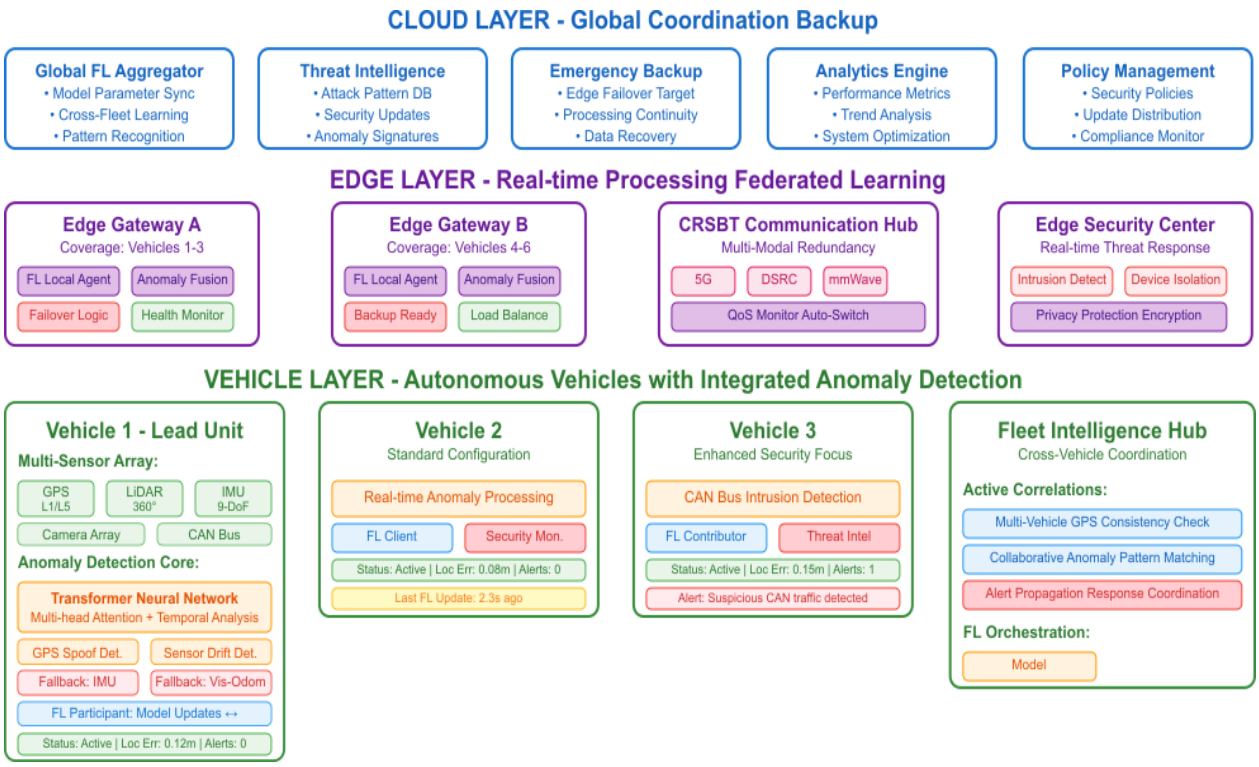


Fig. 1. Comprehensive system architecture of the framework

Federated learning agents deployed on edge gateways enable collaborative anomaly detection across vehicle fleets without compromising individual data privacy.

Each edge gateway maintains local model replicas trained on aggregated telemetry patterns from vehicles within its coverage area. The federated learning protocol employs differential privacy techniques to ensure that individual vehicle data cannot be reverse-engineered from shared model updates.

Resilient Infrastructure Strategies. Our framework implements advanced edge-cloud failover mechanisms based on coherent reliability broadcasting techniques. When edge gateways experience failures or performance degradation, neighboring nodes automatically assume responsibility for affected processing tasks through predetermined succession protocols. The failover mechanism continuously monitors edge gateway health using multidimensional performance metrics including processing latency, memory utilization, network connectivity quality, and thermal conditions. When any metric exceeds acceptable thresholds, the system initiates gradual load migration to backup resources before complete failures occur. Network reliability employs multi-modal communication strategies combining 5G cellular networks, millimeter-wave communications, and DSRC protocols. The system continuously monitors Quality of Service (QoS) characteristics across all communication channels and dynamically routes traffic through optimal paths.

Integrated Response and Recovery Protocols. When the system detects anomalies, it executes structured response workflows designed to maintain vehicle safety while gathering diagnostic information.

The initial response includes signal integrity verification using independent sensor modalities and cross-referencing with high-precision mapping data.

If anomalies persist after initial verification, the system activates redundant localization algorithms including pure inertial navigation using IMU data and visual odometry based on camera streams. These backup systems maintain vehicle positioning accuracy sufficient for safe operation until primary sensors can be restored.

During edge infrastructure failures, the system temporarily redistributes computing tasks across surviving edge nodes and cloud resources. Load balancing algorithms ensure that critical anomaly detection functions receive sufficient processing capacity regardless of infrastructure degradation.

Explainability and Interpretability. Transformer attention mechanisms provide interpretable insights into anomaly detection decisions by highlighting specific sensors, time windows, and feature combinations that contribute to anomaly classifications. Attention weight visualizations enable system engineers to understand why particular anomalies were detected and assess the reliability of detection decisions.

The explainability framework generates automated reports describing detected anomalies in natural language, including likely causes, affected systems, and recommended mitigation actions. Classification and Regression Tree (CART)-based decision tree components provide interpretable feature importance rankings that identify which sensor characteristics are most predictive of different anomaly types.

Results and Discussion. Performance Evaluation. Our comprehensive evaluation demonstrates significant improvements in anomaly detection capabilities

compared to baseline approaches. The integrated transformer-federated learning framework achieves overall anomaly detection accuracy rates between 94-98% across different anomaly categories, with particularly strong performance in GPS spoofing detection (97.8% accuracy) and CAN-bus intrusion detection (95.4% accuracy). Detection latency measurements reveal that the system can identify and classify anomalies within an average of 180 milliseconds from occurrence, well within the real-time requirements for autonomous vehicle safety systems.

The federated learning component demonstrates 40% reduced communication overhead compared to centralized approaches while maintaining equivalent detection performance. Positioning accuracy results demonstrate that our framework maintains localization errors below 0.5 meters even during active anomaly conditions, representing a 60% improvement over baseline sensor fusion approaches. GPS denial scenarios show particularly impressive results, with pure inertial navigation fallback systems maintaining sub-meter accuracy for periods exceeding 300 seconds.

Infrastructure Reliability. Edge computing failover mechanisms demonstrate sub-second response times for transitioning processing loads to backup resources. Average failover completion times measure 340 milliseconds for single node failures and 890 milliseconds for coordinated multi-node failures. Service availability during failover events exceeds 99.7%, indicating minimal disruption to vehicle operations. Communication redundancy systems successfully maintain connectivity during simulated base station failures, with automatic channel switching occurring within 150 milliseconds on average. Mean Time Between Failures (MTBF) analysis of the integrated system reveals significant improvements over individual component reliability estimates, with the redundant architecture achieving overall system MTBF values 340% higher than single-point-of-failure configurations.

System Interpretability. Attention weight analysis provides clear insights into anomaly detection decision processes, with visualizations successfully highlighting problematic sensors and time windows. System engineers report 85% accuracy in predicting attention mechanism focus areas when presented with known anomaly scenarios.

Federated learning performance evaluation demonstrates successful collaborative learning across distributed vehicle fleets while maintaining data privacy requirements.

Model convergence occurs within 15-20 training rounds, comparable to centralized training approaches but with significantly reduced communication overhead.

Conclusions

This paper presents a comprehensive framework for enhancing the trustworthiness of IoT-enabled automated vehicle localization systems through integrated anomaly detection and infrastructure reliability mechanisms.

Our unified approach successfully addresses the critical challenge of maintaining positioning accuracy and system security in complex, distributed automotive environments.

The experimental evaluation demonstrates significant improvements over existing approaches, with anomaly detection accuracy rates of 94-98% and localization errors maintained below 0.5 meters even during active fault conditions. The federated learning component achieves fleet-wide collaborative detection capabilities while reducing communication overhead by 40% compared to centralized alternatives.

Key contributions include the development of transformer-based multi-sensor anomaly detection, implementation of coherent reliability broadcasting for infrastructure failover, and integration of explainable AI techniques for system transparency. The framework successfully combines these components into a cohesive system that maintains real-time performance requirements while providing comprehensive anomaly coverage. The practical implications of this research extend beyond autonomous vehicles to encompass broader IoT applications requiring high reliability and fault tolerance.

The principles and techniques developed for automotive anomaly detection can be adapted for industrial IoT systems, smart city infrastructure, and other safety-critical applications.

The deployment of autonomous vehicles at scale requires robust, trustworthy systems that can operate reliably in diverse and challenging environments. Our integrated framework provides a foundation for achieving this goal through comprehensive anomaly detection, adaptive infrastructure management, and collaborative learning capabilities that enhance overall system resilience.

REFERENCES

1. A. Hanif, N. Ullah, M. Ahmed, S. M. Tahir, A. Ali, and H. Abbas, "Intrusion detection system for controller area network," Cybersecurity, vol. 7, article 5, 2024. https://doi.org/10.1186/s42400-023-00195-4.
2. W. Liu, H. Zhou, K. Chen, Y. Qiu, L. Gao, Y. Liu, and Y. Li, "Transformer-based multivariate time series anomaly detection using inter-variable attention mechanism," Knowledge-Based Systems, vol. 290, article 111507, 2024. https://doi.org/10.1016/j.knosys.2024.111507.
3. M. A. Khan, K. Salah, I. Yaqoob, S. Jayaraman, Y. Al-Hammadi, and D. B. Rawat, "Enhancing Autonomous Vehicle Security: Federated Learning for Detecting GPS Spoofing Attack," Transactions on Emerging Telecommunications Technologies, vol. 36, no. 4, e70138, 2025. https://doi.org/10.1002/ett.70138.
4. H. Cheng, Z. Wang, S. Das, M. LaPorta, and T. La Porta, "Anomaly Detection Against GPS Spoofing Attacks on Connected and Autonomous Vehicles Using Learning From Demonstration," IEEE Transactions on Intelligent Transportation Systems, vol. 24, no. 9, pp. 9516–9532, 2023. https://doi.org/10.1109/TITS.2023.3269029.
5. D. Berdik, S. Otoum, N. Schmidt, D. Porter, and Y. Jararweh, "A survey on blockchain for information systems management and security," Information Processing & Management, vol. 58, no. 1, 2021. https://doi.org/ 10.1016/j.ipm.2020.102397.

6. A. Bouchaib Koubaâ, M. Sriti, Y. Touati, A. Aggoune, M. Hadded, and H. Labiod, "Security of federated learning with IoT systems: Issues, limitations, challenges, and solutions," Internet of Things and Cyber-Physical Systems, vol. 3, pp. 155–179, 2023. https://doi.org/10.1016/j.iotcps.2023.04.005.
7. P. M. Kumar and U. Devi Gandhi, "An optimal federated learning-based intrusion detection for IoT environment," Scientific Reports, vol. 15, article 6509, 2025. https://doi.org/10.1038/s41598-025-93501-8.
8. D. Kwon, H. Kim, J. Kim, S. C. Suh, I. Kim, and K. J. Kim, "A survey of deep learning-based network anomaly detection," Cluster Computing, vol. 22, no. 1, pp. 949–961, 2019. https://doi.org/10.1007/s10586-017-1117-8.
9. A. A. Banafa, A. A. Zaidan, B. B. Zaidan, S. K. Towey, and A. H. Alamoodi, "Design and Experimental Assessment of Real-Time Anomaly Detection Techniques for Automotive Cybersecurity," Sensors, vol. 23, no. 22, article 9231, 2023. https://doi.org/10.3390/s23229231.
10. K. Song, X. Tan, M. Ding, J. Wang, C. Ge, J. Guo, and W. Xu, "Multivariate time series anomaly detection with adversarial transformer architecture in the Internet of Things," Future Generation Computer Systems, vol. 143, pp. 244–254, 2023. https://doi.org/10.1016/j.future.2023.02.006.

Received (Надійшла) 01.08.2025
Accepted for publication (Прийнята до друку) 15.10.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Севостьянова Олена Миколаївна – старший викладач кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Olena Sevostianova – Senior Lecturer, Department of Electronic Computers, Kharkiv National University of Radio Electronics Kharkiv, Ukraine;
e-mail: olena.sevostianova@nure.ua; ORCID Author ID: https://orcid.org/0009-0008-2595-5133.

Косенко Наталія Вікторівна – кандидат технічних наук, доцент, доцент кафедри управління проектами у міському господарстві і будівництві, Харківський національний університет міського господарства ім. О. М. Бекетова, Україна;
Kosenko Nataliia – Candidate of Technical Sciences, Associate Professor, Associate Professor at the Department of Project Management in Urban Economy and Construction, Beketov National University of Urban Economy in Kharkiv, Ukraine;
e-mail: kosnatalja@gmail.com; ORCID ID: https://orcid.org/0000-0002-5942-3150;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57196219605.

Філіппов Владлен Валерійович – аспірант кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Vladlen Filippov – PhD student, Department of Electronic Computers, Kharkiv National University of Radio Electronics Kharkiv, Ukraine;
e-mail: vladlen.filippov@nure.ua; ORCID Author ID: http://orcid.org/0009-0004-2524-7840.

Дяченко Максим Сергійович – студент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Maksym Diachenko – student, Department of Electronic Computers, Kharkiv National University of Radio Electronics Kharkiv, Ukraine;
e-mail: maksym.diachenko@nure.ua; ORCID Author ID: https://orcid.org/0009-0004-5006-3314.

Харахайчук Іван Анатолійович – студент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Ivan Kharakhaichuk – student, Department of Electronic Computers, Kharkiv National University of Radio Electronics Kharkiv, Ukraine;
e-mail: ivan.kharakhaichuk@nure.ua; ORCID Author ID: https://orcid.org/0009-0000-9738-6728.

Підвищення довіреності та надійності систем локалізації автоматизованих транспортних засобів у середовищі IoT

О. М. Севостьянова, Н. В. Косенко, В. В. Філіппов, М. С. Дяченко, І. А. Харахайчук

Анотація. Актуальність. Автономні транспортні засоби покладаються на багатосенсорні системи локалізації, що функціонують у межах інфраструктури Інтернету речей, утворюючи взаємопов'язані вразливості, пов'язані з аномаліями сенсорів, відмовами мережі та кіберзагрозами, які потребують комплексних рішень для подолання проблем на рівні як транспортного засобу, так і інфраструктури. **Об'єкт дослідження** – системи локалізації автоматизованих транспортних засобів, що працюють в середовищі IoT і вимагають надійної роботи за несприятливих умов, зокрема у разі відмов сенсорів, атак із підміною сигналів GPS та збоїв інфраструктури. **Мета статті** – розробка та валідація єдиної рамкової моделі стійкості, яка інтегрує трансформерні методи виявлення аномалій у потоках даних бортових сенсорів із федеративними агентами навчання, розгорнутими на IoT-шлюзах, що забезпечує відновлення роботи після інфраструктурних збоїв менш ніж за секунду при збереженні точності локалізації. **Результати дослідження.** Запропонована модель забезпечує точність виявлення аномалій на рівні 94–98 % при збереженні похибки локалізації менш ніж 0,5 м у разі відмов. Компонент федеративного навчання демонструє зниження комунікаційних витрат на 40 % у порівнянні з централізованими підходами та забезпечує відновлення роботи після відмови інфраструктури менш ніж за секунду. Інтеграція пояснюваного машинного навчання дає змогу отримувати інтерпретовані попередження завдяки механізмам уваги трансформера, що дозволяє виконувати діагностику системи в реальному часі. **Висновки.** Єдина рамкова модель ефективно вирішує ключові виклики впровадження автономних транспортних засобів шляхом поєднання багаторівневого виявлення аномалій, узгодженого поширення повідомлень про надійність та методів пояснюваного ШІ, забезпечуючи комплексну основу для довіреної роботи автономних транспортних засобів у середовищі розумних міст, інтегрованих з IoT.

Ключові слова: автономні транспортні засоби; надійність IoT; виявлення аномалій; федеративне навчання; сенсорна інтеграція; кібербезпека.

УДК 551.583.16:551.465.7(267) **doi: 10.26906/SUNZ.2025.4.22-26**

Г. Г. Томчаковський

Одеський національний морський університет, Одеса, Україна

ДОСЛІДЖЕННЯ ЧИННИКІВ МІЖРІЧНОЇ ТА ВНУТРІШНЬОСЕЗОННОЇ МІНЛИВОСТІ МУСОННОЇ ДЕПРЕСІЇ В ІНДІЙСЬКОМУ ОКЕАНІ

Анотація. Мусонні депресії (МД) – області зниженого тиску з циклонічною циркуляцією, що формуються над акваторією Бенгальської затоки – мають значний вплив на погодні умови в північній частині Індійського океану в період літнього мусону. Вони характеризуються штормовим вітром, високими хвилями, інтенсивною хмарністю і сильними дощами, становлячи серйозну загрозу для морського судноплавства і берегової інфраструктури в регіоні. Повторюваність та інтенсивність МД схильні до істотних коливань у часі в широкому діапазоні масштабів - від внутрішньосезонного до міждекадного. Розуміння чинників, що контролюють мінливість режиму МД, необхідне для зниження ризиків і втрат у секторі морського транспорту, що відіграє ключову роль в світовій економіці. У цій роботі на основі комплексного аналізу даних реаналізу, спостережень і чисельних експериментів досліджуються великомасштабні динамічні та термодинамічні чинники, що визначають режим МД на міжрічному і внутрішньосезонному часових масштабах. Показано, що зміни циркуляції атмосфери і температури поверхні океану в останні десятиліття сприяли зменшенню кількості МД над Бенгальською затокою. Водночас результати регіональних модельних експериментів свідчать про можливе зростання повторюваності МД в умовах триваючого потепління Світового океану. Аналіз внутрішньосезонних варіацій виявив статистично значущі зв'язки активності МД із коливаннями інтенсивності та положення внутрішньотропічної зони конвергенції (ВЗК). Кількісні оцінки, отримані в роботі, створюють основу для поліпшення якості прогнозів погодних умов на морських шляхах у північній частині Індійського океану та оптимізації стратегій мінімізації погодних ризиків для морської галузі.

Ключові слова: мусонна депресія, мусонна циркуляція, циклогенез, міжрічна мінливість, внутрішньосезонні варіації, погодні ризики, внутрішньотропічна зона конвергенції, дистанційне зондування.

Вступ

Постановка проблеми. Мусонні депресії (МД) – характерний елемент літньої циркуляції атмосфери над північною частиною Індійського океану. Вони являють собою великі області зниженого тиску з циклонічною циркуляцією в нижній тропосфері, що формуються над акваторією Бенгальської затоки. МД супроводжуються штормовими вітрами і високими хвилями, що створює загрозу для безпеки морського судноплавства на шляхах, що зв'язують порти Індії, Бангладеш, М'янми і Шрі-Ланки. Траєкторії МД проходять через основні судноплавні шляхи в Бенгальській затоці, тому інформація про їхнє переміщення та інтенсивність критично важлива для планування маршрутів і термінів морських перевезень. Повторюваність МД схильна до значних міжрічних коливань. В окремі роки над Бенгальською затокою може формуватися до 10 – 12 таких систем за сезон, тоді як в інші роки їхня кількість скорочується до 1 – 2. Крім цього, частота МД зазнає помітних змін у масштабі кількох десятиліть. Так, в останній чверті XX століття відзначалася тенденція до зменшення кількості МД, особливо глибоких, що формуються над Бенгальською затокою [1]. Водночас статистика останніх років вказує на значну активізацію циклогенезу над акваторією затоки на початку XXI століття. В окремі роки тут відзначалося до 7-8 мусонних депресій за сезон [2].

Настільки значна мінливість повторюваності МД на різних часових масштабах порушує питання про фактори, що її визначають. У проаналізованих дослідженнях за даною проблемою зазначалося, що до числа основних предикторів циклогенезу над Бенгальською затокою відносяться температура поверхні океану, відносна завихреність і збіжність потоків у нижній тропосфері, величина вертикального зсуву вітру і запаси прихованого тепла в атмосфері [3]. Аномалії цих параметрів в останні десятиліття могли сприяти зміні режиму циклогенезу над акваторією затоки. Зокрема, на тлі триваючого зростання температури поверхні океану відзначалися тенденції до зменшення вологовмісту в середній тропосфері та посилення вертикального зсуву вітру [4]. Однак кількісні оцінки відносної ролі цих факторів у спостережуваних міждекадних коливаннях МД залишаються предметом дискусій.

Інший важливий аспект проблеми полягає в можливому впливі на циклогенез над Бенгальською затокою інтенсивних внутрішньосезонних варіацій літнього мусону, відомих як активні та перервчасті фази. Дослідження вказують на тісний зв'язок формування МД із коливаннями інтенсивності та зсувами внутрішньотропічної зони конвергенції (ВЗК) [5]. Так, активізація ВЗК та її поширення на північ від середнього літнього положення зазвичай супроводжуються збільшенням кількості випадків циклогенезу і посиленням вітро-хвильової активності над акваторією затоки. І навпаки, ослаблення ВЗК призводить до зменшення повторюваності МД і поліпшення погодних умов для судноплавства. Однак систематичний аналіз статистичних зв'язків між характеристиками ВЗК і режимом МД у літературі відсутній.

Кількісна оцінка факторів, що визначають мінливість мусонних депресій на різних часових масштабах, видається надзвичайно важливою в контексті їхнього впливу на погодні ризики для морського транспорту в Бенгальській затоці. Бенгальська затока і прилеглі райони Індійського океану відрізняються високою інтенсивністю судноплавства. Порти Індії, Бангладеш і Шрі-Ланки обробляють понад 1 млн TEU (20-футових контейнерних еквівалентів) на рік.

На морський транспорт припадає близько 95 % обсягу зовнішньої торгівлі Індії. У зв'язку з цим аномалії вітро-хвильової активності та видимості, пов'язані з проходженням МД, можуть призводити до серйозних порушень у роботі портів і затримок у морських перевезеннях. Щорічні економічні втрати через несприятливі погодні умови для світового судноплавства оцінюються в 0,1 – 0,2 % глобального ВВП. Для регіону Індійського океану цей показник може бути істотно вищим з урахуванням високої повторюваності МД. Більш чітке розуміння предикторів циклогенезу над Бенгальською затокою може сприяти підвищенню точності та завчасності прогнозів переміщення МД і поліпшенню планування морських операцій у регіоні. Це, своєю чергою, дасть змогу зменшити економічні втрати та ризики для морського транспорту, пов'язані з впливом небезпечних погодних явищ.

Метою даної роботи є кількісна оцінка чинників, що визначають мінливість мусонних депресій на міжрічному і внутрішньосезонному часових масштабах. Дослідження, на основі комплексного аналізу даних реаналізу, спостережень і чисельного моделювання, великомасштабних динамічних та термодинамічних чинників, що визначають мінливість циклогенезу над акваторією Бенгальської затоки в широкому діапазоні часових масштабів – від внутрішньосезонного до міждекадного.

Виявлення аномалії циркуляції атмосфери і температури поверхні океану, що впливають на повторюваність та інтенсивність мусонних депресій.

Автор вважає, що реалізація цих завдань дасть змогу визначити предиктори циклогенезу над Бенгальською затокою, важливі для прогнозу погодних умов на морських шляхах в Індійському океані та прилеглих районах у періоди літнього мусону. А це, в свою чергу, сприятиме зниженню ризиків і втрат у секторі морських перевезень, який робить значний внесок в економіку регіону.

Основна частина

Мусонна депресія – один із ключових елементів літньої мусонної циркуляції над Індійським океаном і прилеглими районами Південної Азії (рис. 1).

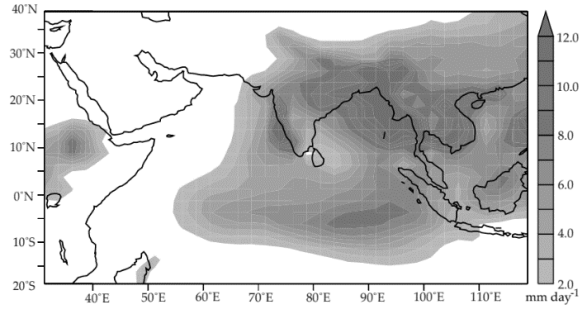


Рис. 1. Середня кількість опадів (мм/день) у період літнього мусону (червень-вересень) [13]

Це синоптичний вихор із замкнутою циклонічною циркуляцією в нижній тропосфері та характерними горизонтальними розмірами 1000 – 2000 км. Формування МД відбувається головним чином над акваторією північної частини Бенгальської затоки. Виникнувши над теплими водами затоки, МД поступово зміщуються на північний захід і в кінцевому підсумку заповнюються над Індійським субконтинентом. Тривалість життєвого циклу типової МД становить від 3 до 6 діб [6].

Зародження МД над водами Бенгальської затоки відбувається за поєднання низки сприятливих великомасштабних чинників. До них належать високі значення температури поверхні океану, циклонічна завихреність у нижній тропосфері (на рівні 850 гПа), підвищений вологовміст повітря в середній тропосфері, а також порівняно слабке вертикальне зрушення вітру. Цікаво, що це той самий набір необхідних умов, який сприяє генезису тропічних циклонів над іншими акваторіями Світового океану. Однак у період літнього мусону над Аравійським морем і Бенгальською затокою спостерігається дуже сильний вертикальний зсув вітру між нижньою і верхньою тропосферою. Цей фактор перешкоджає подальшій інтенсифікації мусонних депресій до стадії тропічних циклонів – на відміну від того, що має місце в інших океанічних басейнах Північної півкулі, де пік активності тропічних циклонів якраз припадає на літні місяці [7]. Розвиток і підтримання МД протягом кількох днів над водами Бенгальської затоки та прилеглою територією суші пов'язують зі спільним проявом кількох типів гідродинамічної нестійкості мусонної течії. Ключову роль тут відіграють бароклінно-баротропна нестійкість і умовна нестійкість 2-го роду (CISK) [8]. Згідно з сучасними уявленнями, на початковій стадії життєвого циклу МД основне значення має баротропна нестійкість зональних вітрів у нижній тропосфері над океаном. Надалі домінуючу роль у підтримці циркуляції МД протягом декількох діб відіграють процеси бароклінної нестійкості та CISK [3].

Вертикальна структура МД характеризується наявністю холодного ядра на нижніх рівнях тропосфери і теплого ядра на верхніх рівнях (500 – 300 гПа). Висхідні рухи в циркуляції МД зазвичай простежуються до рівня 300 гПа. Максимум циклонічної завихреності розташований у середній тропосфері близько 800 гПа. Як показав комплексний аналіз на основі даних реаналізу ERA-Interim, зона максимальних опадів розташовується в південно-західному секторі МД щодо центру циркуляції [9 – 10]. Характерне значення добових опадів, пов'язаних із проходженням МД, становить 30 – 50 см води на площі радіусом 200 – 300 км. В окремих випадках добові суми опадів можуть перевищувати 60 см. Просторовий розподіл опадів, пов'язаних з МД, дуже неоднорідний і визначається орографією Індійського субконтиненту. Основна зона інтенсивних опадів захоплює Бенгальську затоку, Бангладеш і північний схід Індії. У періоди активної конвекції над Бенгальською затокою і проходження МД підвищена кількість опадів спостерігається також на навітряних схилах Західних Гат в Індії. Тут визначальну роль відіграє вимушений підйом мусонного потоку під час обтікання гірського бар'єру, що сприяє інтенсифікації облогових орографічних опадів [11, 12]. У середньому за сезон над акваторією північної частини Бенгальської затоки формується близько 6 мусонних депресій (рис. 2).

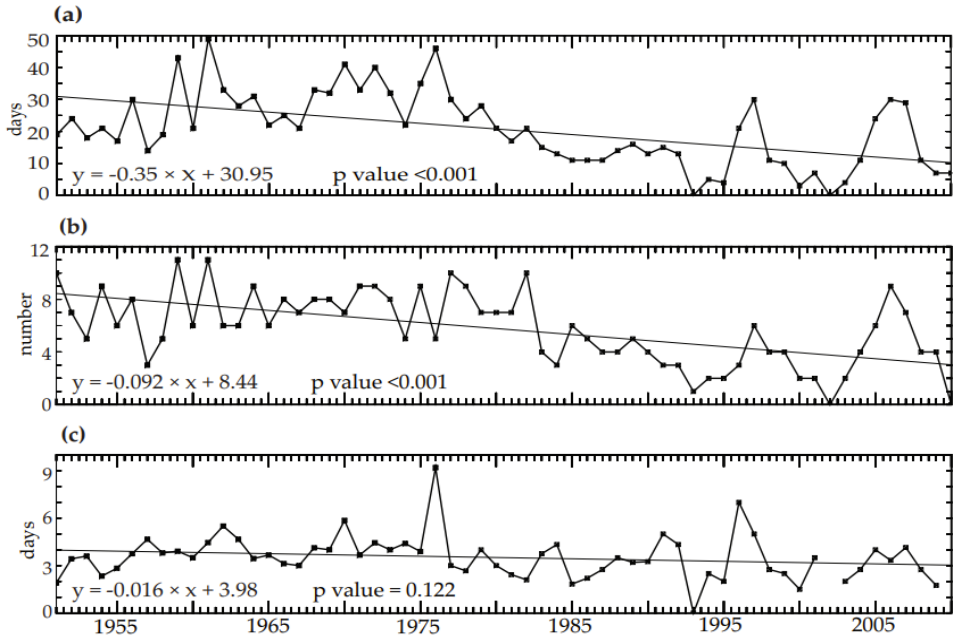


Рис. 2. Статистичні дані кількості мусонних депресій за сезон: a – кількості днів MD, b – кількості MD, c – середньої тривалості життя (уднях) MD під час літнього мусонного сезону [13]

Однак від року до року це число схильне до дуже великих коливань. У деяких літніх сезонах тут відзначається до 10 – 12 МД, тоді як в інші роки їхня кількість може скорочуватися до 1 – 2. Загальний внесок МД у сумарну кількість опадів за літній мусон також варіюється в широкому діапазоні – від 35 до 55 % за даними різних досліджень [14]. Оцінка цього внеску для центральних районів Індії за даними IMD [India Meteorological Department] за період 1901 – 2010 рр. наведена на рис. 3.

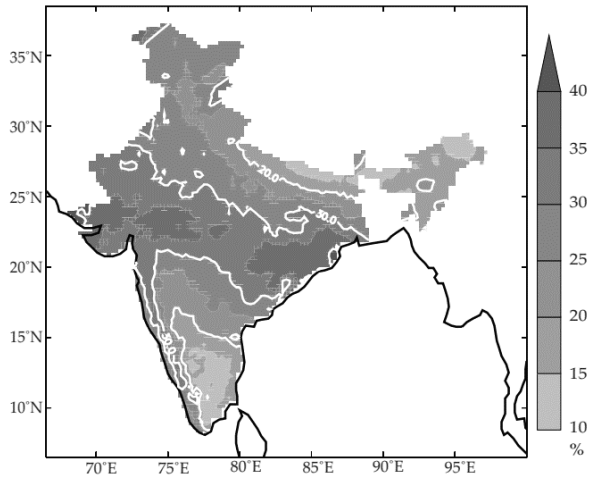


Рис. 3. Відношення (у %) суми опадів, пов'язаних із мусонними депресіями, до загальної кількості опадів за сезон [13]

Крім дуже великої міжрічної мінливості, частота формування МД схильна також до помітних міждекадних коливань. Аналіз тривалих рядів даних про повторюваність МД свідчить про тенденцію до зменшення їхньої кількості протягом останніх десятиліть XX століття. Ця тенденція супроводжувалася зустрічним збільшенням повторюваності слабших циклонічних збурень із замкнутою циркуляцією, але меншими баричними градієнтами в центрі [15]. Очевидно, в останні десятиліття інтенсифікація систем зниженого тиску до стадії МД почала відбуватися рідше – це пов'язано зі зміною деяких фонових параметрів стану атмосфери й океану над акваторією Бенгальської затоки.

Як показують наукові дослідження у періоди зменшення кількості МД після 1980 року спостерігалися несприятливі для їх генезису та інтенсифікації аномалії циркуляції нижньої тропосфери. Зокрема, відзначалося ослаблення циклонічної завихреності на рівні 850 гПа в північній частині Бенгальської затоки. При цьому мало місце посилення вертикального і горизонтального зсуву зонального вітру між екваторіальною зоною і областю літнього мусону.

Ще одне можливе пояснення зменшення кількості МД в останні десятиліття пов'язане зі змінами температурного режиму поверхні Бенгальської затоки. Хоча, здавалося б, підвищення температури поверхні океану має сприяти почастішанню випадків формування МД. До початку 1980-х років міждекадні варіації температури води в Бенгальській затоці та повторюваності МД демонстрували подібний характер [16]. Однак згодом даний зв'язок порушився. Виходячи з цього, дія інших чинників, несприятливих для циклогенезу, виявилася істотнішою. Зокрема, автори одного з досліджень за наявною проблемою пов'язали зменшення повторюваності МД із посиленням вітру у верхній тропосфері над Бенгальською затокою (на рівні 100 гПа) [17]. Проте, з іншого боку, тенденція до скорочення числа МД пояснюється прогресуючим зменшенням вологовмісту повітря в середній тропосфері над акваторією затоки.

Таким чином, зміни великомасштабних динамічних і термодинамічних параметрів атмосфери відповідальні за спостережуване в останні десятиліття зменшення повторюваності мусонних депресій над Бенгальською затокою. Дискусійним залиша-

ється питання про кількісний внесок окремих чинників (завихреність, зсуви вітру, вологість повітря та ін.) у цю міждекадну мінливість.

Не враховуючи довгоперіодні варіації, частота формування та інтенсифікації МД схильна до значних внутрішньосезонних коливань, пов'язаних зі зміною активних і перервних фаз літнього мусону над Індією. Автор акцентує увага на тісному зв'язку циклогенезу над Бенгальською затокою з переміщеннями ВЗК. Активна фаза континентальної ВЗК, зміщеної на північ від свого середнього літнього положення, підтримується за рахунок серії рухомих циклонічних збурень. Вони послідовно формуються на східній периферії мусонної улоговини над Бенгальською затокою і потім зміщуються вздовж осі улоговини на північний захід, у бік Пакистану. Цей ланцюжок хвильових збурень, що нерідко досягають стадії МД, може існувати до двох тижнів, викликаючи тривалий період інтенсивних мусонних опадів над Індією. Подібний висновок отримано в [18]. Було встановлено, що регенерація ВЗК в активну фазу мусону після періоду її ослаблення або тимчасового зникнення пов'язана з повторним формуванням синоптичних збурень над районами генезису МД. На основі супутникових даних і результатів реаналізу можно сказати, що характеристики ВЗК над Індійським океаном схильні до значної синоптичної мінливості, зумовленої взаємодією зони конвекції з рухомими циклонічними вихорами [19]. Особливо чітко ця взаємодія проявляється під час посилення меридіональної складової вітру в нижній тропосфері та активізації ВЗК.

Проведене дослідження вказує на важливу роль внутрішньотропічної зони конвергенції як фактора, що модулює частоту й інтенсивність МД над акваторією Бенгальської затоки. Водночас кількісний аналіз цього взаємозв'язку в масштабі окремих внутрішньосезонних варіацій мусонної циркуляції до теперішнього часу не проводився. Заповнення цієї прогалини на основі сучасних масивів даних спостережень і глобальних моделей видається важливим завданням подальших досліджень мінливості МД.

Висновки

Проведене дослідження виявило низку важливих особливостей міждекадної та внутрішньосезонної мінливості мусонних депресій над акваторією Бенгальської затоки. Встановлено, що в останні десятиліття XX століття спостерігалася стійка тенденція до зменшення повторюваності МД, незважаючи на триваюче зростання температури поверхні океану в регіоні. Цей факт вказує на переважаючий вплив інших чинників, несприятливих для формування та інтенсифікації МД. До їх числа відносяться ослаблення циклонічної завихреності в нижній тропосфері, посилення вертикального зсуву вітру та зменшення вологовмісту повітря в середній тропосфері над Бенгальською затокою. Водночас кількісна оцінка відносного внеску кожного з цих факторів у спостережувану міждекадну мінливість режиму МД потребує подальших досліджень із залученням розширеного масиву даних та вдосконалених методів діагностики.

Поряд із довгоперіодними коливаннями, в роботі також проаналізовано особливості внутрішньосезонної мінливості МД та її зв'язок з крупномасштабними модами атмосферної циркуляції. Показано, що частота й інтенсивність МД схильні до значних варіацій у масштабі окремих місяців всередині літнього сезону, які тісно пов'язані з чергуванням активних і перервних фаз мусону над Індією. Встановлено, що одним із ключових чинників, які модулюють інтенсивність циклогенезу над Бенгальською затокою, є внутрішньотропічна зона конвергенції. Активізація ВЗК та її зміщення на північ від середнього літнього положення сприяють почастішанню випадків формування МД та посиленню пов'язаної з ними штормової активності. І навпаки, ослаблення ВЗК зазвичай супроводжується зменшенням повторюваності циклонічних збурень над акваторією затоки. Розвиток уявлень про характер та фізичні механізми цих взаємозв'язків становить перспективний напрямок подальших досліджень, важливий з точки зору вдосконалення схем прогнозу погодних умов над північною частиною Індійського океану.

Результати, отримані в роботі, вказують на необхідність комплексного підходу до вивчення просторово-часової мінливості МД, який би поєднував поглиблений аналіз даних спостережень, реаналізу та чисельного моделювання. Особливу увагу при цьому слід приділити модельним експериментам із систематичним варіюванням параметрів підстильної поверхні та характеристик великомасштабного поля вітру в атмосфері. Це дозволить точніше діагностувати фізичні фактори, що визначають особливості циклогенезу над Бенгальською затокою на різних часових масштабах – від внутрішньосезонного до міждекадного. Отримані результати матимуть важливе практичне значення. Вони сприятимуть підвищенню якості прогнозів погодних умов на морських шляхах у північній частині Індійського океану, зниженню ризиків для морського транспорту та оптимізації стратегій планування морських операцій у регіоні.

СПИСОК ЛІТЕРАТУРИ

1. Prajeesh, A., Ashok, K. & Rao, D. Falling monsoon depression frequency: A Gray-Sikka conditions perspective. *Sci Rep* **3**, 2989 (2013). DOI: https://doi.org/10.1038/srep02989.
2. Vishnu, S., Francis, P. A., Shenoi, S. S. C., Ramakrishna, S. S. V. S. (2015). On The Decreasing Trend of Monsoon Depressions over Bay of Bengal. Annual Monsoon Workshop National Symposium on Understanding and Forecasting the Monsoon Extremes, Indian Meteorological Society, Pune Chapter, India. DOI: https://dx.doi.org/10.1088/1748-9326/11/1/014011.
3. Sørland, S. L., & Sorteberg, A. (2015). The dynamic and thermodynamic structure of monsoon low-pressure systems during extreme rainfall events. *Tellus A: Dynamic Meteorology and Oceanography*, *67*(1). DOI: https://doi.org/10.3402/tellusa.v67.27039.
4. Kieran M. R. Hunt, Andrew G. Turner, Peter M. Inness, David E. Parker, Richard C. Levine. (2016) On the Structure and Dynamics of Indian Monsoon Depressions. Monthly Weather Review, Volume 144: Issue 9, Page(s): 3391–3416. DOI: https://doi.org/10.1175/MWR-D-15-0138.1.

5. Gareth Berry, Michael J. Reeder. (2014) Objective Identification of the Intertropical Convergence Zone: Climatology and Trends from the ERA-Interim. Journal of Climate, Volume 27: Issue 5. Page(s): 1894–1909. DOI: https://doi.org/10.1175/JCLI-D-13-00339.1.
6. Jin-Ho Yoon and Wan-Ru (Judy) Huang. (2012) Indian Monsoon Depression: Climatology and Variability. Modern Climatology. InTech. DOI: http://dx.doi.org/10.5772/37917.
7. Wu, G., Y. Liu, B. He, Q. Bao, A. Duan, and F. F. Jin. (2012) Thermal Controls on the Asian Summer Monsoon. Scientific reports 2, 404. URL: https://doi.org/10.1038/srep00404.
8. J. Shukla. CISK-Barotropic-Baroclinic Instability and the Growth of Monsoon Depressions. Journal of the Atmospheric Sciences, Volume 35: Issue 3, Page(s): 495–508. DOI: https://doi.org/10.1175/1520-0469(1978)035<0495:CBBIAT>2.0.CO;2.
9. Kieran M. R. Hunt, Andrew G. Turner, Peter M. Inness, David E. Parker, Richard C. Levine. (2016) On the Structure and Dynamics of Indian Monsoon Depressions. Monthly Weather Review, Volume 144: Issue 9, Page(s): 3391–3416. DOI: https://doi.org/10.1175/MWR-D-15-0138.1.
10. Доля В.Д. Мусони, як частина глобальної циркуляції атмосфери Землі, геофізична природа явища. зб. наук. праць XII Міжнародної наукової міждисциплінарної конференції студентів, аспірантів та молодих вчених “Шевченківська весна – 2014, Частина 3: географія”: тези доповіді. Випуск XII. Київ, 18-22 бер. 2014. С.93. DOI: http://dx.doi.org/10.13140/RG.2.1.4605.7447.
11. Francis, P., Gadgil, S. (2006) Intense rainfall events over the west coast of India. Meteorol. Atmos. Phys., 94, 27–42. DOI: https://doi.org/10.1007/s00703-005-0167-2.
12. Колесник А. В., Доля В. Д., Кучеренко Н. В. Использование ГИС для изучения причин формирования муссонов Индийского океана. Часопис картографії. 2015. Вип. 12. С. 82-89. URL: http://nbuv.gov.ua/UJRN/ktvsh_2015_12_11.
13. India Meteorological Department [Електронний ресурс]. Режим доступу: dg.imd@imd.gov.in
14. Jin-Ho Yoon, Tsing-Chang Chen. (2005) Water vapor budget of the Indian monsoon depression. Tellus A: Dynamic Meteorology and Oceanography, Volume: 57 Issue: 5, Page(s): 770–782. DOI: https://doi.org/10.3402/tellusa.v57i5.14737.
15. Капочкін Б. Б., Кучеренко Н.В., Лісоводський В. В., Конкін В. В. Фізичні механізми вирівнювання баричних градієнтів в атмосфері. К., 2003. 12 с. Деп. в ГНТБ України 04.08.03, № 105.
16. Rajeevan, M., U. S. De, and R. K. Prasad, 2000: Decadal variation of sea surface temperatures, cloudiness and monsoon depressions in the north Indian ocean. Current Science, Volume: 79, Page(s): 283–285. URL: https://repository.ias.ac.in/94366/1/decadal.pdf.
17. Rao, B., D. Rao, and V. B. Rao. (2004) Decreasing trend in the strength of tropical easterly jet during the Asian summer monsoon season and the number of tropical cyclonic systems over bay of Bengal. Geophysical Research Letters, Volume: 31. DOI: http://dx.doi.org/10.1029/2004GL019817.
18. D. R. Sikka and Sulochana Gadgil. On the Maximum Cloud Zone and the ITCZ over Indian Longitudes during the Southwest Monsoon. Monthly Weather Review, Volume 108: Issue 11, Page(s): 1840–1853. DOI: https://doi.org/10.1175/1520-0493(1980)108<1840:OTMCZA>2.0.CO;2.
19. Chia-chi Wang and Gudrun Magnusdottir. (2006) The ITCZ in the Central and Eastern Pacific on Synoptic Time Scales. Monthly weather review, Volume 134: Issue 5, Page(s): 1405–1421. DOI: https://doi.org/10.1175/MWR3130.1.

Received (Надійшла) 21.08.2025
Accepted for publication (Прийнята до друку) 29.10.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Томчаковський Георгій Георгійович – старший викладач кафедри навігації та управління судном,, Одеський національний морський університет, Одеса, Україна;
Georgiy Tomchakovsky – Senior Lecturer, Department of Navigation and Control of the Ship, Odesa National Maritime University, Odesa, Ukraine;
e-mail: gtomchakovsky@gmail.com; ORCID Author ID: https://orcid.org/0000-0002-9799-4368;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=59557477500.

Research of factors of interannual and intra-seasonal variability of the monsoon depression in the Indian Ocean

Georgiy Tomchakovsky

Abstract. Monsoon depressions (MDs) – areas of low pressure with cyclonic circulation that form over the Bay of Bengal - have a significant impact on weather conditions in the northern Indian Ocean during the summer monsoon. They are characterised by gale-force winds, high waves, intense cloud cover and heavy rainfall, posing a serious threat to maritime navigation and coastal infrastructure in the region. The recurrence and intensity of MDs are subject to significant fluctuations in time over a wide range of scales, from intra-seasonal to inter-decadal. Understanding the factors that control the variability of the MD regime is essential to reduce risks and losses in the maritime transport sector, which plays a key role in the global economy. In this paper, based on a comprehensive analysis of reanalysis data, observations, and numerical experiments, we investigate the large-scale dynamic and thermodynamic factors that determine the MD regime on interannual and intra-seasonal time scales. It is shown that changes in atmospheric circulation and ocean surface temperature in recent decades have contributed to a decrease in the number of MDs over the Bay of Bengal. At the same time, the results of regional model experiments indicate a possible increase in the frequency of the MDs in the context of the ongoing warming of the World Ocean. The analysis of intra-seasonal variations revealed statistically significant links between the MD activity and fluctuations in the intensity and position of the Intratropical Convergence Zone (ICZ). The quantitative estimates obtained in this work provide a basis for improving the quality of weather forecasts on sea lanes in the northern Indian Ocean and optimising strategies to minimise weather risks for the maritime industry.

Keywords: monsoon depression, monsoon circulation, cyclogenesis, interannual variability, intra-seasonal variations, weather risks, intratropical convergence zone, remote sensing.

Управління в складних системах

UDC 621.314.223.2:621.316.727 **doi: 10.26906/SUNZ.2025.4.27-31**

Д. А. Гапон[1], П. О. Качанов[1], А. В. Ольшевський[1], Є. Б. Петрик[2]

[1] Національний технічний університет "Харківський політехнічний інститут", Харків, Україна
[2]Triol Corporation, Харків, Україна

ДОСЛІДЖЕННЯ ПАРАМЕТРІВ ПОВНОГО ГАРМОНІЙНОГО СПОТВОРЕННЯ СТАНЦІЇ УПРАВЛІННЯ З ВХІДНИМ ФАЗОЗСУВНИМ АВТОТРАНСФОРМАТОРОМ

Анотація. При розробці корисних копалин на території України, таких як нафта, застосовуються механізовані методи видобутку, із застосуванням занурювальних електродвигунів. Такі методи видобутку забезпечують високу продуктивність і можливість підлаштовуватися під поточні параметри свердловини, що в свою чергу забезпечує максимальний дебіт. Для забезпечення оптимального режиму роботи насоса використовуються станції управління занурювальним електродвигуном з частотним регулюванням, які при роботі споживають з мережі живлення несинусоїдальний струм. Також для зниження спотворень споживаного струму застосовуються різні методи зниження загального гармонійного спотворення (THDi) вхідних струмів. Для вирішення даної задачі була розроблена станція управління з багатопульсним випрямлячем, який живиться від фазозсувного автотрансформатора, що дозволяє знизити THDi вхідних струмів. Дане рішення дозволяє покращити габаритні показники обладнання та його вартість за рахунок зменшення потужності фазозсувного трансформатора. Розглянуто функціональні можливості роботи станції управління електроприводом на основі експериментальних даних вимірювання THDi вхідних струмів мережі +30% в діапазоні частот 47,5-52,5Гц. Такий діапазон частот обумовлений тим, що у багатьох випадках, на родовищах, робота ведеться з електроживленням від дизель-генераторів, при якому коливання частоти і напруги живлення знаходяться в значному діапазоні. Додатково досліджено можливість роботи приводу при вхідних напругах живлення 460-480В та 430-460В. Проведено аналіз результатів вимірів, та запропоновано рекомендації щодо підвищення енергоефективності станції.

Ключові слова: перетворювач частоти, станція управління, енергоефективність, THDi, багатопульсний випрямляч, фазозсувний трансформатор

Вступ

На промислових підприємствах, зокрема в нафтовидобувній галузі, все ширше використовуються сучасні електромеханічні системи з напівпровідниковими перетворювачами (НП), які є основою для реалізації економічних технологій з більш широкими функціональними можливостями. Водночас НП негативно впливають на показники якості електроенергії в системі електропостачання, в якій вони працюють [1]. В системі виникають спотворення синусоїдальної форми струму, відбувається генерація вищих гармонік. У свою чергу це викликає спотворення синусоїдальної форми напруги мережі живлення і може призвести до таких негативних наслідків, як збої в роботі електронних пристроїв, похибки в роботі вимірювальних приладів, неприпустиме нагрівання обладнання, передчасне старіння ізоляції [2].

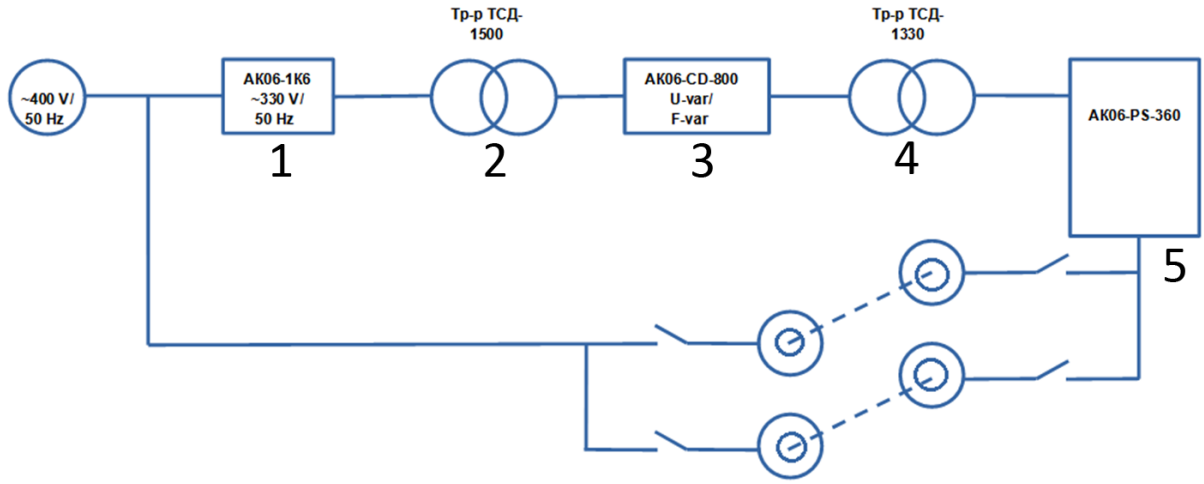
Проблемам електромагнітної сумісності електромеханічних пристроїв на основі ПП присвячено багато теоретичних та експериментальних досліджень. У роботах [3] наведено результати експериментального дослідження ступеня спотворення синусоїдальності струму, що споживається з промисловій мережі електроприводами з перетворювачами частоти (ПЧ) [4], побудованими на основі інверторів.

Експериментальні дослідження дозволили встановити функціональну залежність коефіцієнта несинусоїдальних спотворень мережевого струму вищими гармоніками (THDi – Total harmonic distortion) від навантаження перетворювача частоти у вигляді степеневої функції. У роботі [5] встановлено, що для найбільш значущих п'ятої та сьомої гармонік мережевого струму їх функціональну залежність від навантаження ПЧ зручно записувати у вигляді експоненціальної регресії. Також показано, що 11-та і вищі гармоніки практично не залежать від навантаження ПЧ. У роботі [6] доведено принципову можливість проведення подібних досліджень на спрощених фізичних або математичних моделях, наприклад, MATLAB Simulink. У цих моделях відпадає необхідність у представленні вихідної частини ПЧ - автономного інвертора. Проведені дослідження на фізичних і математичних моделях частотного приводу підтверджують, що величина споживаного частотним приводом струму і ступінь спотворення його синусоїдальності визначається зміною напруги на конденсаторах ланки постійного струму, яке викликане зростанням або зниженням навантаження електроприводу [7]. У роботі [8] показано, що за допомогою комп'ютерного моделювання можна визначити джерело вищих гармонік у складних випадках, до яких відносяться ПЧ.

Мета даної статті експериментально встановити для станції АК06-PS функціональну залежність коефіцієнта несинусоїдальних спотворень мережевого струму вищими гармоніками (THDi) від частоти вихідного струму Fs = 47,5/50/52,5 Гц, а також перевірити можливості роботи приводу при вхідних напругах живлення 460-480В і 430-460В.

Матеріали і результати досліджень

Для кількісної оцінки гармонічного складу мережевого струму проводився запис осцилограм струму, що споживається з мережі при включеному навантаженні. На рис. 1 представлена схема електрична принципова досліджуваного обладнання.



1 – інвертор, 2 – підвищуючий трансформатор, 3 – частотний перетворювач, 4 – трансформатор, 5 – досліджувана станція

Рис. 1. Схема підключення станції АК06-PS-360-480-18-111

Інвертор та підвищуючий трансформатор, забезпечують підвищення напруги до рівня, необхідного для експерименту. Частотний перетворювач та трансформатор, забезпечують зміну частоти електричного струму, що подається на досліджувану станцію. Оскільки досліджується робота станції під навантаженням, то робота занурювального насоса імітується парою асинхронних двигунів, які жорстким валом об'єднані з парою асинхронних двигунів в режимі генерації енергії, що повертається в промислову мережу.

Для експерименту використовувалася наступна контрольно-вимірювальна апаратура: осцилограф FLUKE зі струмовимірювальними кліщами та аналізатор мережі YOKOGAWA WT5000 (рис. 2).

Вимірювання проводилися в двох діапазонах напруги живлення на входах приводу: 430-460В і 460-480В.

Рис. 2. Контрольно-вимірювальна апаратура яка використовувалась для експерименту

Основні параметри вимірювалися для фіксованих значень вихідної частоти інвертору станції: Fs = 47,5Гц, 50Гц та 52,5Гц. Для заданої частоти змінювалась напруга. Вимірювані параметри: Us - напруга живлення, В, Ud - напруга ланки постійного струму, В, Is - вхідний струм, А, Iout - вихідний струм, А, Load - навантаження, %, THDi - загальне гармонійне спотворення по кожній з фаз, η - коефіцієнт корисної дії. Отримані в ході експерименту дані представлені в табл. 1-2.

В таблиці 3 наведені параметри виміряні на частоті мережі 52,5Гц, при якій спостерігається поява коливання струмів на всіх входах і виходах станцій-джерел і приводу, що перевіряється.

Таблиця 1 – **Дані виміру на частоті 47,5 Гц**

Fs, Гц	Us,В	Ud,В	Is, А	Iout, А	% Load	THDi, % (фази)	Total THDi, %	η (ккд) %
47,5	**Живлення 460-430 В**							
	451,282	625	246,785	295,633	82,1	8,954/9,024/9,321	9,101	94,562
	450,138	623	277,464	327,447	90,9	7,926/7,799/8,232	7,988	94,690
	449,172	619	302,968	354,023	98,3	7,516/7,374/7,844	7,581	94,717
	445,790	612	380,047	437,199	121,4	6,612/6,845/7,018	6,827	94,622
	Живлення 480-460 В							
	474,866	658	249,788	310,294	86,2	9,90/9,775/10,162	9,949	94,272
	473,830	656	279,941	346,459	96,2	8,878/8,536/9,259	8,896	94,391

Таблиця 2 – **Дані виміру на частоті 50 Гц**

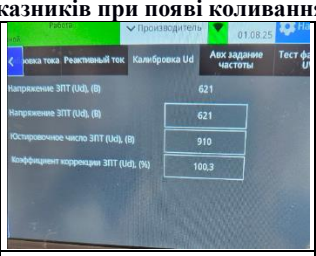
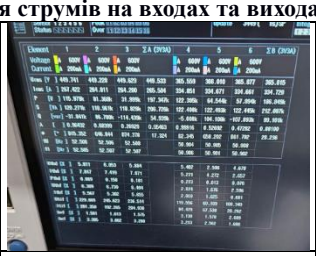
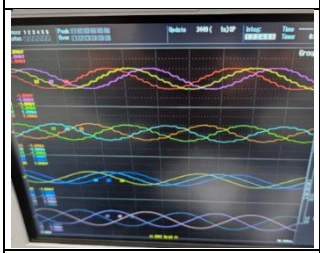
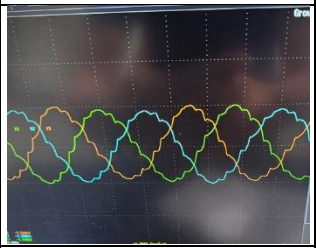
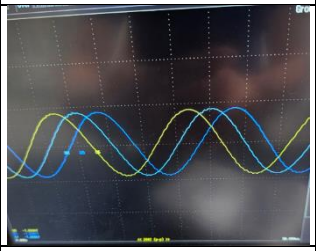
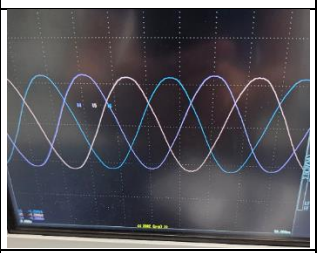
Fs, Гц	Us,B	Ud,B	Is, A	Iout, A	% Load	THDi, % (фази)	Total THDi, %	η (ккд) %
50	**Живлення 460-430 В**							
	452,204	629	215,730	267,407	74,3	9,107/9,327/9,223	9,219	94,498
	451,147	625	247,473	299,519	83,2	7,982/7,891/8,104	7,993	94,710
	449,928	621	278,482	331,213	92	7,359/7,172/7,408	7,314	94,796
	449,038	619	303,945	358,508	99,6	7,026/6,919/7,081	7,009	94,801
	446,036	612	377,888	437,064	121,4	6,515/6,356/6,641	6,505	94,677
	Живлення 480-460 В							
	503	780	37	129		47,39/43,94/51,06	47,552	
	475	661	213	282	78,3	8,85/8,80/8,94	8,864	
	472	656	246	309	85,8	7,58/7,52/7,69	7,597	
	470	652	266	337	93,6	7,20/6,96/7,27	7,145	94,449
	469	649	292	364	101,1	6,80/6,67/6,88	6,784	94,578
	464	639	349	431	119,7	6,97/6,48/7,94	7,156	94,424

Таблиця 3 - **Дані виміру на частоті 52,5 Гц**

Fs, Гц	Us,B	Ud,B	Is, A	Iout, A	% Load	THDi, % (фази)	Total THDi, %	η (ккд) %
52,5	**Живлення 460-430 В**							
	453,517	633	179,936	232,264	64,5	10,771/11,043/10, 828	10,881	94,245
	451,317	626	222,747	277,839	77,2	7,836/7,878/7,921	7,878	94,670
	450,295	623	250,429	316,836	88	7,341/7,328/7,388	7,352	94,673
	449,776	622	263,369	332,704	92,4	7,210/7,034/7,159	7,135	94,718
	449,533	621	265,504	334,728	93	7,817/7,419/7,871	7,705	94,680
	Живлення 480-460 В							
	477,348	664	201,225	262,326	72,9	10,252/10,387/10, 178	10,273	94,356
	476,022	661	226,520	289,405	80,4	8,404/8,448/8,432	8,428	94,582
	474,773	657	260,012	326,047	90,6	7,362/7,384/7,392	7,379	94,727
	474,537	656	265,954	338,249	94	7,218/7,196/7,301	7,238	94,632

Для ілюстрації процесу проведення експерименту в табл. 4 наведено приклад фотофіксації показників при появі коливання струмів на всіх входах і виходах станцій-джерел і приводу, що перевіряється (52,5Гц).

Таблиця 4 – **Фотофіксація показників при появі коливання струмів на входах та виходах станцій-джерел і приводу**

			
Вихідний струм по Fluke - “бовтанка”	Напруга ЗПТ приводу Ud = 656В	Вхідні та вихідні параметри	ККД привода при навантаженні Iout = 334,7А
			
Графіки вхідних та вихідних U/I	Вхідні струми привода («бовтанка»)	Вихідні струми привода (коливання)	Вихідна напруга привода

Обговорення результатів. Основною метою досліджень було знайти межі ефективності застосування станції, яка проектувалась під мережу 60Гц, при використанні в національній мережі 50Гц. Отримані результати дозволяють зробити наступні висновки:

1. На частоті мережі Fs = (60+5)Гц привід працює з THDi < 8% в діапазоні навантажень 70% ...

100% (і навіть нижче) при напрузі живлення Us = 480В +5%/-15% або Us = 460В +10%/-15% (як у станції фірми «Toshiba»).

2. На частоті Fs = (50+2,5)Гц діапазон навантажень наближається до необхідного, але тільки при зниженій напрузі живлення Us = 380В +10%/-15%.

3. На знижених частотах із зростанням напруги живлення «звужується» діапазон навантажень, де привід утримує THDi < 8%, як можна побачити з рис.3, де наведено зведену інформацію про показники приводу АК06-PS-360-480-18-111 при різних напругах і частотах мережі.

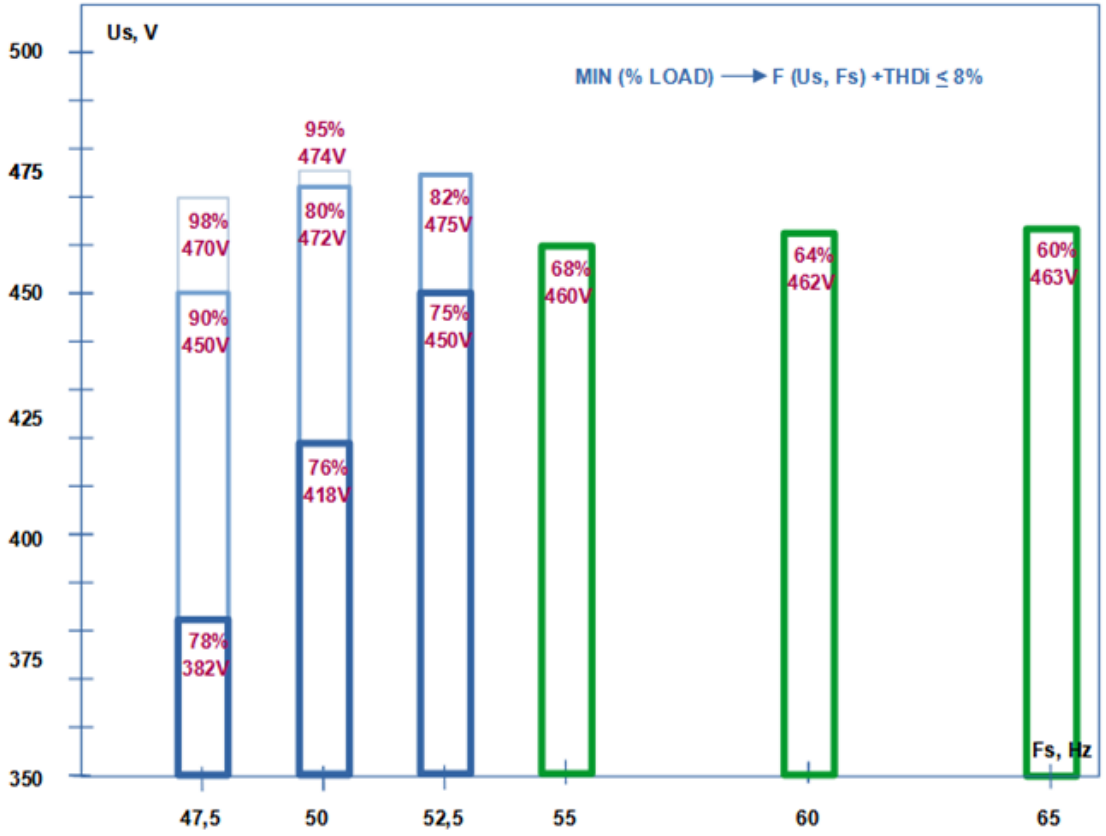


Рис. 3. Показники приводу АК06-PS-360-480-18-111 в залежності від частоти і напруги мережі

Зв'язок між частотою f мережі живлення, числом W витків обмоток (мережевих), перетином S_e осереддя і максимальною індукцією B_m пояснюється формулою для трансформаторної ЕРС (U) – вона працює для трансформаторів, дроселів і автотрансформаторів:

$$U_l = \frac{2\pi f \omega s_e B_m}{k_a} \approx 4{,}44 f \omega s_e B_m. \qquad (1)$$

Оскільки розробка приводу АК06-PS-360-480-18-111 з 18-пульсним автотрансформатором велася для мережі Us = 480В, 60Гц, [9] то його застосування в мережі зі зниженою в 1,2 рази частотою 50Гц вимагає як мінімум зміни числа витків і/або перетину осереддя для роботи з напругами U = 480В [10].

Висновки

У поточній конфігурації станція управління не забезпечує заданих параметрів при роботі в широкому діапазоні напруг живлення і частот мережі живлення, що можна побачити з таблиць 1-3. За вимогою THDi повинно бути менше або дорівнювати 5%, а отримані значення перевищують цей поріг, оскільки частота нижче розрахункової. Для забезпечення заявлених параметрів THDi по вхідному струму необхідно провести наступні доопрацювання:

- Знизити заявлену величину напруги живлення мережі до 380В +10% -15% /50Гц – привод не вимагає конструктивних доопрацювань.

- Доопрацювати 18-пульсний автотрансформатор шляхом збільшення перетину осереддя не менше ніж в 1,2 рази, що приведе до необхідності виготовлення нових котушок трансформатора.

Після проведення доопрацювань доцільно провести додаткові дослідження приводу АК06-PS-360-480-18-111 з напругами мережі Us = 480В +10%/-15%/50+2,5Гц, для того щоб підтвердити заявлені характеристики.

Список літератури

1. Гудим В., Яцишин С., Мамцаж Д. Визначення коефіцієнта спотворення напруги в системах електропостачання. Вимірювальна техніка та метрологія. Том 80, вип. 3, 2019 р. с. 64-71. doi: 10.23939/istcmtm2019.03.064
2. ДСТУ 13109-97 (ГОСТ 13109-97). Електрична енергія. Сумісність технічних засобів електромагнітна. Норми якості електричної енергії в системах електропостачання загального призначення. – [Чинний від 1999-01-01]. – К.: Держспоживстандарт України, 1999. – 51 с.

3. Радимов С. М. Експериментальне дослідження ступеня спотворення мережевого струму частотного електроприводу / С. М. Радимов, В. Л. Бєляєв, А. М. Бесараб [та ін.] // Електромашинобудування та електрообладнання: міжвід. наук.-техн. зб. – К., 2010. – Вип. 75. – С. 52–56. Режим доступу: http://nbuv.gov.ua/UJRN/etks_2010_75_11.
4. A. von Jouanne, P. N. Enjeti and B. Banerjee, "Assessment of ride-through alternatives for adjustable-speed drives," in IEEE Transactions on Industry Applications, vol. 35, no. 4, pp. 908-916, July-Aug. 1999, doi: 10.1109/28.777200.
5. Бєляєв В. Л. Гармонійний склад мережевого струму частотних електроприводів з широтно-імпульсною модуляцією / В. Л. Бєляєв, С. М. Радимов // Електромеханічні та енергозберігаючі системи: тематич. вип. «Проблеми автоматизованого електропривода: теорія і практика». – Кременчук, 2012. – № 1(19). – С. 469–471.
6. Бєляєв В. Л. Спрощені моделі електроприводів з широтно-імпульсною модуляцією / В. Л. Бєляєв, С. М. Радимов // Проблеми енергоресурсозбереження в електротехнічних системах: наука, освіта і практика: зб. наук. пр. – Кременчук, 2014. – Вип. 1(2). – С. 240–242.
7. Das J. C. Passive Filters – Potentialities and Limitations / J. C. Das // IEEE Transactions on Industry Applications. – 2004. – Vol. 40, No. 1. – P. 232–241.
8. Гапон Д. А. Визначення джерела вищих гармонік в системах електропостачання зі змішаним навантаженням / Д. А. Гапон, А. О. Зуєв, П. О. Качанов, Б. І. Кубрик // Сучасні інформаційні системи. – 2022. – Т. 6, № 1. – С. 66–74. DOI: doi: 10.20998/2522-9052.2022.1.09
9. S. Khan, X. Zhang, M. Saad, H. Ali, B. M. Khan, and H. Zaman, “Comparative analysis of 18-pulse autotransformer rectifier unittopologies with intrinsic harmonic current cancellation”, Energies, vol.11, no. 6, p. 1347, 2018. doi: 10.3390/en11061347.
10. R. Abdollahi, G. B. Gharehpetian, F. Mohammadi, and S. Prakash P., “Multi-pulse rectifier based on an optimal Pulse Doubling Technique”, Energies, vol. 15, no. 15, p. 5567, 2022. doi: 10.3390/en15155567

Received (Надійшла) 28.08.2025
Accepted for publication (Прийнята до друку) 22.10.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Гапон Дмитро Анатолійович – доктор технічних наук, доцент, завідувач кафедри автоматизації та кібербезпеки енергосистем, Національний технічний університет «Харківський політехнічний інститут», Харків, Україна;
Dmytro Gapon – Doctor of Technical Sciences, Associate Professor, Head of the Department of Power Systems Automation and Cybersecurity, National Technical University «Kharkiv Polytechnic Institute», Kharkiv, Ukraine;
e-mail: Dmytro.Hapon@khpi.edu.ua; ORCID Author ID: https://orcid.org/0000-0001-8609-9707;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57552659800&origin=recordpage.

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 the Department of Automation and Control in Technical Systems, National Technical University «Kharkiv Polytechnic Institute», Kharkiv, Ukraine;
e-mail: petro.kachanov@khpi.edu.ua; ORCID Author ID: https://orcid.org/0000-0002-7532-5913;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57190405644.

Ольшевський Андрій Вікторович – аспірант кафедри автоматики та управління в технічних системах, Національний технічний університет «Харківський політехнічний інститут», Харків, Україна;
Andrey Olshevskiy – PhD student of the Department of Automation and Control in Technical Systems, National Technical University «Kharkiv Polytechnic Institute», Kharkiv, Ukraine;
e-mail: asd.andry.asd@gmail.com; ORCID Author ID: https://orcid.org/0000-0003-2734-9491;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=58749329700.

Петрик Євген Борисович – кандидат технічних наук, Технічний лідер Triol Corporation, Харків, Україна;
Evgeniy Petrik – Candidate of Technical Sciences, Tech Lead of Triol Corporation, Kharkiv, Ukraine;
e-mail: Evgeniy_Petrik@gmail.com; ORCID Author ID: https://orcid.org/0000-0001-5902-9295.
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=58749329700.

Study of the parameters of the total harmonic distortion of a control station with an input phase-shift autotransformer

Dmytro Gapon, Petro Kachanov, Andrey Olshevskiy, Evgeniy Petrik

Abstract. When developing minerals in Ukraine, such as oil, mechanized extraction methods are used, using submersible electric motors. Such extraction methods provide high productivity and the ability to adjust to the current parameters of the well, which in turn provides maximum flow rate. To ensure the optimal operation of the pump, frequency-controlled submersible electric motor control stations are used, which consume non-sinusoidal current from the power supply network during operation. Also, to reduce distortions of the consumed current, various methods of reducing the total harmonic distortion (THDi) of input currents are used. To solve this problem, a control station with a multi-pulse rectifier powered by a phase-shifting autotransformer was developed, which allows you to reduce the THDi of input currents. This solution allows you to improve the overall performance of the equipment and its cost by reducing the power of the phase-shifting transformer. The functional capabilities of the electric drive control station are considered based on experimental data of measuring THDi of input currents of the network +30% in the frequency range of 47.5-52.5Hz. This frequency range is due to the fact that in many cases, in the fields, work is carried out with power supply from diesel generators, in which the fluctuations of frequency and supply voltage are in a significant range. Additionally, the possibility of drive operation at input supply voltages of 460-480V and 430-460V is investigated. The measurement results are analyzed, and recommendations are proposed for increasing the energy efficiency of the station.

Keywords: frequency converter, control station, energy efficiency, THDi, multi-pulse rectifier, phase-shifting transformer.

УДК 629.4.083:629.424.2 **doi: 10.26906/SUNZ.2025.4.32-38**

В. І. Носков, М. В. Ліпчанський, В. І. Панченко, Г. В. Гейко

Національний технічний університет «Харківський політехнічний інститут», Харків, Україна

ВИЗНАЧЕННЯ ОПТИМАЛЬНОГО РУХУ ДИЗЕЛЬ-ПОЇЗДА НА ДІЛЯНЦІ ШЛЯХУ З ВІДОМИМ ПРОФІЛЕМ

Анотація. Актуальність. Залізничний транспорт є одним із основних споживачів енергоресурсів, тому розробка та впровадження енергозберігаючих технологій є актуальною проблемою, вирішення якої останнім часом є одним із пріоритетних напрямів діяльності фахівців фундаментальної та прикладної науки в цій галузі. **Об'єкт дослідження:** процеси керування тяговими, динамічними та енергетичними характеристиками моторвагонного рухомого складу. **Мета статті:** вирішити задачу оптимального керування рухом дизель-поїзда на ділянці шляху з відомим профілем із урахуванням оптимальної витрати палива та дотримання графіка руху. **Результати дослідження.** У статті сформульовано і за допомогою методу термінальних керувань вирішено задачу синтезу системи керування оптимальним рухом дизель-поїзда з відомим профілем шляху. При цьому використовувалися: теорія тяги поїздів і тягового електроприводу, методи системного аналізу і теорії оптимального керування, математичне моделювання та обчислювальні експерименти. **Висновки.** Розроблено алгоритм оптимальних керувань дизель-поїздом, за допомогою якого визначено програмні закони, що встановлюють залежність тяги від маси поїзда, параметрів шляху та техніко-економічних умов (швидкість, прискорення, графік руху) з урахуванням витрати палива.

Ключові слова: моторвагонний рухомий склад, оптимальний рух поїзда, термінальні керування, синтез системи керування, математичне моделювання, профіль шляху.

Вступ

Постановка проблеми. Питання підвищення техніко-економічних показників тягового рухомого складу України належить до найбільш актуальних. Його рішення можливе як за рахунок оптимізації режимів роботи енергетичного обладнання, так і за рахунок оптимізації руху поїздів. При чому, для різних умов експлуатації оптимальний режим ведення поїзда розробляється для заданого часу ходу по перегонах і для кожного окремого випадку має свої особливості [1].

Аналіз останніх досліджень і публікацій. Проблема оптимального керування дизель-поїздом може бути поділена на дві незалежні задачі: оптимізацію режимів роботи енергетичного обладнання і процесу руху поїзда.

У роботі [2] зазначено, що однією з головних вимог до систем керування поїздами є забезпечення чіткого виконання графіків руху при мінімальних витратах енергії. Необхідно, щоб поїзд пройшов по перегону за час, що не перевищує заданий, розподіл швидкостей при цьому всередині перегону в значному ступеню довільний але підпорядковується низці обмежень [3, 4].

У роботі [5] виконано аналіз схеми тягової електропередачі дизель-поїзда ДЕЛ-02 та її системи керування, результати якого дозволили розробити метод перевірки використання потужності дизеля на тягу в умовах експлуатації.

У роботі [6] розроблено інформаційне, програмне і алгоритмічне забезпечення, яке дозволяє виконувати розрахунки режимів ведення поїздів з можливістю адаптації параметрів моделі руху поїзда за результатами дослідних поїздок.

У роботі [7] визначено, що тяга та гальмування, основний опір руху, маса, коефіцієнт інерції обертових мас рухомого складу у порівнянні з характеристикою коефіцієнта корисної дії (ККД) тягового приводу, здійснюють менший вплив на значення питомих витрат електроенергії.

У роботі [8] розроблена математична модель системи керування електрорухомим складом з використанням методу динамічного програмування та зворотного рішення задачі відносно координати часу. У роботі [9] авторами отримано енергозаощаджуючу функцію керування тягою, яка залежить від маси рухомого складу та ухилу колії. У роботі [10] запропоновано структуру моделі системи нечіткого задання швидкості поїзда з адаптивною корекцією помилки її регулювання за фактичними параметрами проїзду на ділянках прямування. У роботі [11] пропонується підхід, який полягає в розділенні всього шляху на підділянки з обмеженням швидкості та припущенням, що опір руху поїзда є квадратичною функцією швидкості, запропоновано методи вирішення задач оптимального керування поїздом. У роботі [12] наведені результати обґрунтування конструкції мікропроцесорної системи керування електричною передачею з поліпшеними тягово-енергетичними характеристиками локомотивів. Дизель-поїзд ДЕЛ-02 обладнаний мікропроцесорною системою керування (МПС), яка забезпечує оптимальну роботу тягової електропередачі. Для запису інформації про рух дизель-поїзда по перегону з наступною її обробкою розроблено комп'ютерну інформаційну систему (ІС). Розроблене програмне забезпечення дозволяє виконувати оцінку ефективності кожної поїздки і отримувати рекомендації по керуванню рухом дизель-поїзда з мінімальними енерговитратами.

Розглянемо задачу керування рухом, яка дозволяє оптимізувати енергетичні витрати при русі поїзда із заздалегідь відомим профілем шляху всередині перегону.

Практичний інтерес становить також задача керування рухом поїзда за мінімальний час. Тут критерієм оптимальності є час. Для цього необхідно

виконувати керування з урахуванням двох складових: одна забезпечує мінімізацію енергетичних витрат при русі дизель-поїзда, а інша – мінімальний час руху з початкового пункту до кінцевого.

Метою роботи є вирішення задачі синтезу системи керування оптимальним рухом дизель-поїзда з відомим профілем шляху; розробка алгоритму оптимального керування дизель-поїздом; встановлення за допомогою розробленого програмного забезпечення залежності тяги від маси поїзда, профілю шляху та техніко-економічних умов.

Основний матеріал

Розглянемо математичну постановку задачі. У нашому випадку об'єкт керування можна описати системою диференційних рівнянь:

$$\begin{aligned} \dot{x}_1 &= f_1(x_1, \ldots, x_r, u_1, \ldots, u_m), \\ \dot{x}_2 &= f_2(x_1, \ldots, x_r, u_1, \ldots, u_m), \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \dot{x}_r &= f_r(x_1, \ldots, x_r, u_1, \ldots, u_m), \end{aligned} \quad (1)$$

де $\dot{x}_i = \frac{dx_i}{dt}$, $i = \overline{1, r}$; r – порядок об'єкта; $x_1, \ldots, x_r$ – фазові координати; $u_1, \ldots, u_m$ – керування $(m \le r)$; $f_1, \ldots, f_r$ – безперервні функції фазових координат і керувань.

На керування, фазові координати і функції від них накладено обмеження:

$$\begin{aligned} &u_{j\min} \le u_j \le u_{j\max}, \ j = \overline{1, m}, \\ &x_{i\min} \le x_i \le x_{i\max}, \ i = \overline{1, r}, \\ &\dot{x}_{i\min} \le \dot{x}_i \le \dot{x}_{i\max}, \ i = \overline{1, r}, \\ &\ddot{x}_{i\min} \le \ddot{x}_i \le \ddot{x}_{i\max}, \ i = \overline{1, r}, \\ &\phi_{k\min} \le \phi_k(x_1, \ldots, x_r, \dot{x}_1, \ldots, \dot{x}_r, u_1, \ldots, u_m) \le \phi_{k\max}, \\ &\qquad k = \overline{1, e}, \end{aligned} \quad (2)$$

де e – число функцій від фазових координат, їх похідних і керувань, на які накладено обмеження.

Початкові умови на лівому (при $t = t_0$) і граничні умови на правому (при $t = T$) кінцях фазових траєкторій напишемо таким чином:

$$\begin{aligned} &\left[x_{10}, \ldots, x_{r0}\right]_{t=0} = \left[x_1(t_0), \ldots, x_r(t_0)\right] = \\ &\qquad = \left[x_{10}, \ldots, x_{r0}\right], \\ &\left[x_{1k}, \ldots, x_{rk}, \dot{x}_{1k}, \ldots, \dot{x}_{rk}, \ldots, x_{1k}^{(s)}, \ldots, x_{rk}^{(s)}\right]_{t=T} = \\ &\quad = x_1(T), \ldots, x_r(T), x_1(T), \ldots, x_r(T), \\ &\qquad ,\ldots, x_1^{(s)}(T), \ldots, x_r^{(s)}(T)\Big] = \\ &= [x_{1k}, \ldots, x_{rk}, \dot{x}_{1k}, \ldots, \dot{x}_{rk}, \ldots, x_{1k}^{(s)}, \ldots, x_{rk}^{(s)}], \end{aligned} \quad (3)$$

де порядок s похідних не обмежений.

Задано функціонал:

$$I = I(x_1, \ldots, x_r, \dot{x}_1, \ldots, \dot{x}_r, \ldots, u_1, \ldots, u_m, t_0, T). \quad (4)$$

Нам необхідно знайти вектор-функцію керувань $u = (u_1, \ldots, u_m)$, яка мінімізує функціонал (4) на відрізку часу $[t_0, T]$ при умовах (2) і (3). Це найбільш загальна постановка задачі оптимального керування.

При керуванні рухом поїзда можна обмежитись класом об'єктів, для яких систему керувань (1) можна розв'язати щодо керувань $u_1, \ldots, u_m$. У цьому випадку керування можна розглядати як функції фазових координат об'єкта та їхні похідні, а рішення задачі оптимального керування може бути зведено до пошуку набору функцій $x_k(t)$, $k = \overline{1, r}$, що мінімізують функціонал (4).

Виконаємо синтез керувань, що оптимізують енергетичні витрат при русі дизель-поїзда, за допомогою метода термінальних керувань [13, 14]. Для рішення цієї задачі використовуємо рівняння динаміки руху поїзда і дослідні характеристики дизеля та електропередачі дизель-поїзда.

Математичну модель для оптимізації процесу руху в даному випадку можна записати у вигляді:

$$m\frac{dV_T}{dt} = F_T - \phi(x) - F_C; \quad (5)$$

$$\frac{dx}{dt} = V_T, \quad (6)$$

де m – маса поїзда, кг; V_T – швидкість поїзда, м/с; F_T – сила тяги поїзда, Н; $\phi(x)$ – сила опору, що залежить від величини ухилу колії, Н; F_C – сила в'язкого тертя та опору повітря, Н; x – відстань, що відраховується від станції відправлення, м.

Сила опору $\phi(x)$ визначається вагою дизель-поїзда і величиною ухилу колії:

$$\phi(x) = P_C i(x), \quad (7)$$

де P_C – вага поїзда, Н; $i(x)$ – величина ухилу.

Сила опору F_C описується наступним чином:

$$F_C = P_Л W_0' + Q W_0'', \quad (8)$$

де $P_Л$, Q – вага головних і причіпних вагонів, Н; W_0', W_0'' – питомі опори руху головних і причіпних вагонів, Н.

Питомі опори руху поїзда визначаються співвідношеннями:

$$W_0' = c_0 + c_1 V_T + c_2 V_T^2, \quad (9)$$

$$W_0'' = d_0 + d_1 V_T, \quad (10)$$

де c_0, c_1, c_2, d_0, d_1 – постійні коефіцієнти.

З урахуванням (9) і (10) сила опору F_C визначається таким чином:

$$F_C = F_{C0} + a_0 V_T + a_1 V_T^2, \quad (11)$$

де $F_{C0} = P_Л c_0 + Q d_0$, $a_0 = P_Л c_1 + Q d_1$, $a_1 = P_Л c_2$.

Для оцінки якості керованого процесу замість узагальненого функціоналу (4) візьмемо один із критеріїв. У даному випадку критерієм оптимальності є витрати енергії, тобто для дизель-поїзда – це витрати палива.

Витрати палива на проходження перегону завдовжки x_0 за час T пропорційні інтегралу:

$$G = \int_0^T (F_T \frac{dx}{dt} + b_0)dt, \qquad (12)$$

де b_0 – коефіцієнт, який визначається із співвідношення

$$\eta_q \eta_p = 1/(1 + b_0 p), \qquad (13)$$

де η_q, η_p – ККД дизеля і електропередачі, p – потужність дизеля.

На рис. 1 наведено дослідні характеристики електропередачі дизель-поїзда. Тут: $\eta_{АІН}$, $\eta_{СГ}$, $\eta_{ТАД}$, η_p – ККД автономного інвертора напруги, синхронного генератора, тягового асинхронного двигуна, електропередачі, відповідно. У виразі (12) час T не фіксується, але він не може перевищувати максимально допустимий час.

Іншим критерієм оптимальності може бути час проходження дизель-поїзда по перегону. При цьому ставиться задача переведення об'єкта з початкового фазового стану в кінцевий за мінімально можливий час. Обмеження (2) визначаються гранично допустимими значеннями швидкості, прискорення та сили тяги.

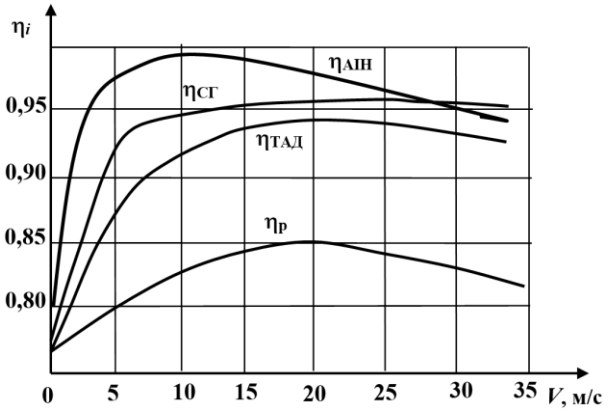


Рис. 1. Дослідні характеристики електропередачі дизель-поїзда

Тягову характеристику можна записати в такому вигляді:

$$F_{ТП} = \begin{cases} F_m, \text{при } \min[(A - k_1 \dot{x}), B/(\dot{x} - C)] \geq F_m, \\ F_1 = \min[(A - k_1 \dot{x}), B/(\dot{x} - C)], \text{при } F_1 < F_m, \end{cases} \qquad (14)$$

де $F_{ТП}$ – гранично-допустима сила тяги; $\dot{x}$ – швидкість об'єкта; A, B, C, k_1 – постійні величини.

Тепер виконаємо синтез керувань рухом об'єкта, мінімізуючи при цьому витрати палива або час руху об'єкта.

Система рівнянь (5), (6) має тільки одне керування (силу тяги) і може бути представлена як окремий випадок об'єкта, що описується рівнянням виду:

$$f(x, \dot{x}, \ddot{x}, \ldots, x^s, u(t)) = 0$$

або

$$u(t) = f_1(x, \dot{x}, \ddot{x}, \ldots, x^{(s)}), \qquad (15)$$

де фазовими координатами є функція $x(t)$ та її похідні.

Визначимо функцію $x(t)$ і керування $u(t)$ за допомогою метода термінальних керувань. Якщо відомий початковий (A_1) і кінцевий (B_1) стан об'єкта керування в багатовимірному просторі, то існує множина фазових траєкторій, які з'єднують ці точки (рис. 2).

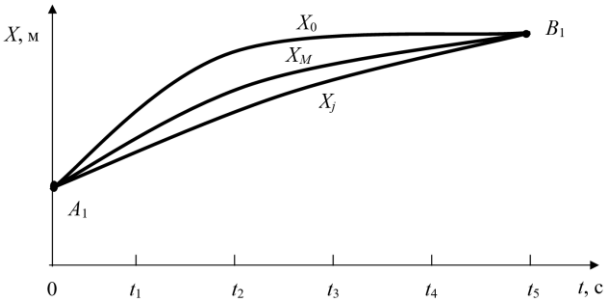


Рис. 2. Траєкторії руху об'єкта (X_0, X_j, X_M – відповідно початкова, проміжна і екстремальна траєкторія)

Робити висновок про близькість фазової траєкторії X_j до екстремалі пропонується за швидкістю зміни функціоналу, який поблизу екстремуму прагне нуля. Оскільки екстремаль являє собою неперервну функцію часу, то вона може бути апроксимована поліномом:

$$X_M = \sum_{i=0}^{r+n-1} C_i t^i, \qquad (16)$$

де r – кількість початкових умов; n – кількість кінцевих умов; C_i – коефіцієнти полінома; t – час. При цьому бажаний рух об'єкта описується неперервною та r разів диференційованою функцією $x(t)$.

Закон керування знаходиться як функція фазових координат шляхом рішення вихідної системи диференційних рівнянь, що описують динаміку руху об'єкта.

Задача оптимізації зводиться до знаходження коефіцієнтів C_i полінома (16).

Якщо за умовою задачі необхідно задовольнити r початковим і n кінцевим умовам, то, використовуючи їх для визначення C_i, знайдемо закон керування, який забезпечує вирішення крайової задачі.

Перші r невідомих коефіцієнтів для фазової траєкторії, що апроксимується поліномом (16), визначаються таким чином:

$$C_i = \frac{x_0^{(i)}}{i!}, \quad i = 0, 1, \ldots, r-1, \tag{17}$$

де $x_0^{(i)}$ – початкові значення функції $x(t)$ та її похідних. Інші n коефіцієнтів знаходимо згідно формули:

$$C_i = \sum_{v=0}^{i-1} \frac{(r+n-v-1)!}{(r+n-i-1)!(i-v)!T^{i-v}} C_v + \\ + \sum_{v=0}^{r+n-i-1} (-1)^v \frac{(r+n-v-1)!}{i!(r+n-i-v-1)!v!T^{i-v}} x_k^{(v)}, \tag{18}$$

$$i = r, r+1, ..., r+n-1,$$

де $x_k^{(v)}$ – кінцеві значення функції $x(t)$ та її похідних при $t = T$.

Мінімізація функціонала може здійснюватися як за допомогою підбора вільних кінцевих умов $x_k^{(n+1)}$, $x_k^{(n+2)}$, ..., $x_k^{(n+m)}$, так і підбором коефіцієнтів C_i.

Для об'єкта керування (5), (6) потрібний керуючий вплив (15) (сила тяги) з урахуванням відношень (11) і (16) обчислюється таким чином:

$$F_T = m\frac{d^2}{dt^2}\left(\sum_{i=0}^{r+n-1} C_i t^i\right) + a_0 \frac{d}{dt}\left(\sum_{i-0}^{r+n-1} C_i t^i\right) + \\ + a_1\left(\frac{d}{dt}\left(\sum_{i=0}^{r+n-1} C_i t^i\right)\right)^2 + \phi(x) + F_{C0}. \tag{19}$$

У (19) коефіцієнти C_i визначаються за початковим x_{i0} і граничним $x_{ik}\,(i = \overline{1, r})$ значеннями фазових координат.

Розробка алгоритма пошуку оптимальних керувань

Згідно розглянутого методу загальна стратегія рішень поставленої задачі полягає у:

1) визначенні за відомими параметрами об'єкта керування коефіцієнтів диференційних рівнянь (5) і (6), які описують його динаміку;

2) рішенні системи диференційних рівнянь відносно керувань;

3) обчисленні за початковими і кінцевими умовами вихідної фазової траєкторії і початкового значення функціонала;

4) виконанні за допомогою випадкового пошуку підбору відповідних вільних кінцевих умов $x_k^{(n+1)}$, $x_k^{(n+2)}$,..., $x_k^{(n+m)}$ з урахуванням заданих обмежень;

5) отриманні закону керування рухом об'єкта, шляхом підстановки значень C_i у вирази для керувань.

Алгоритм синтезу оптимальних термінальних керувань рухом дизель-поїзда по ділянці шляху і з урахуванням обмежень наведено на рис. 3.

На початку роботи задаються вихідні дані: маса поїзда, довжини ділянок шляху, величини їх ухилів, початкові і кінцеві умови, масштабні коефіцієнти для роботи процедури випадкового пошуку, точність оптимізації, верхні (*GV*(*I*)) і нижні (*GN*(*I*)) обмеження та їх число (блоки 2 та 3 на рис. 3).

У даному випадку розглядаються три обмеження (*KOG* = 3), а саме: по допустимій швидкості, прискоренню й керуванню. Точність оптимізації (*TOL*) задається у відсотках і є мінімальною відмінністю між двома послідовними значеннями функціонала *FNL*, при досягненні якої задача вважається вирішеною.

У блоці 4 обчислюється початкове значення функціоналу *FL* з урахуванням лише крайових умов, проводиться розрахунок термінальних керувань та моделювання руху об'єкта під дією цих керувань. Процес розрахунку використовує процедуру визначення коефіцієнтів C_i згідно виразів (17) – (18) і процедуру рішення системи нелінійних диференційних рівнянь. У результаті виконання обчислень отримується вихідна траєкторія руху та початкове значення функціоналу.

У блоці 5 процедурою обчислюються значення коефіцієнтів C_i вихідної функції $x(t)$ по заданим початковим і кінцевим умовам.

Далі визначається N похідна в момент часу $t = T$ для початкового наближення шуканої фазової траєкторії. Оскільки $N = 2$ (задані кінцева точка шляху і швидкість), то при першому проході будуть підбиратися оптимальні кінцеві значення других похідних (блок 6 на рис. 3).

Після збільшення кількості кінцевих умов на одиницю (блок 7 на рис. 3) виконується моделювання вихідної траєкторії та виведення результатів моделювання динаміки об'єкта (блок 8 на рис. 3).

Наступним кроком задаються вихідні дані для роботи процедури випадкового пошуку, відбувається масштабування змінних та їх передача до процедури пошуку екстремуму функції багатьох змінних шляхом підбору кінцевих умов з урахуванням обмежень (блоки 9-11 на рис. 3).

Перевірка обмежень здійснюється шляхом обчислення в задані моменти часу необхідних змінних і виконання їх порівняння з величинами, заданими співвідношеннями (2).

При невиконанні умов генерується нова точка і процес обчислень і перевірки повторюється. Інакше – обчислюється значення функціоналу I і виконується порівняння з попереднім його значенням. Завершення процедури пошуку виконується після знаходження оптимальних значень змінних, що мінімізують функціонал.

Після зворотного масштабування змінних відбувається моделювання руху об'єкта по оптимальній траєкторії та виведення результатів (блоки 12-14 на рис. 3).

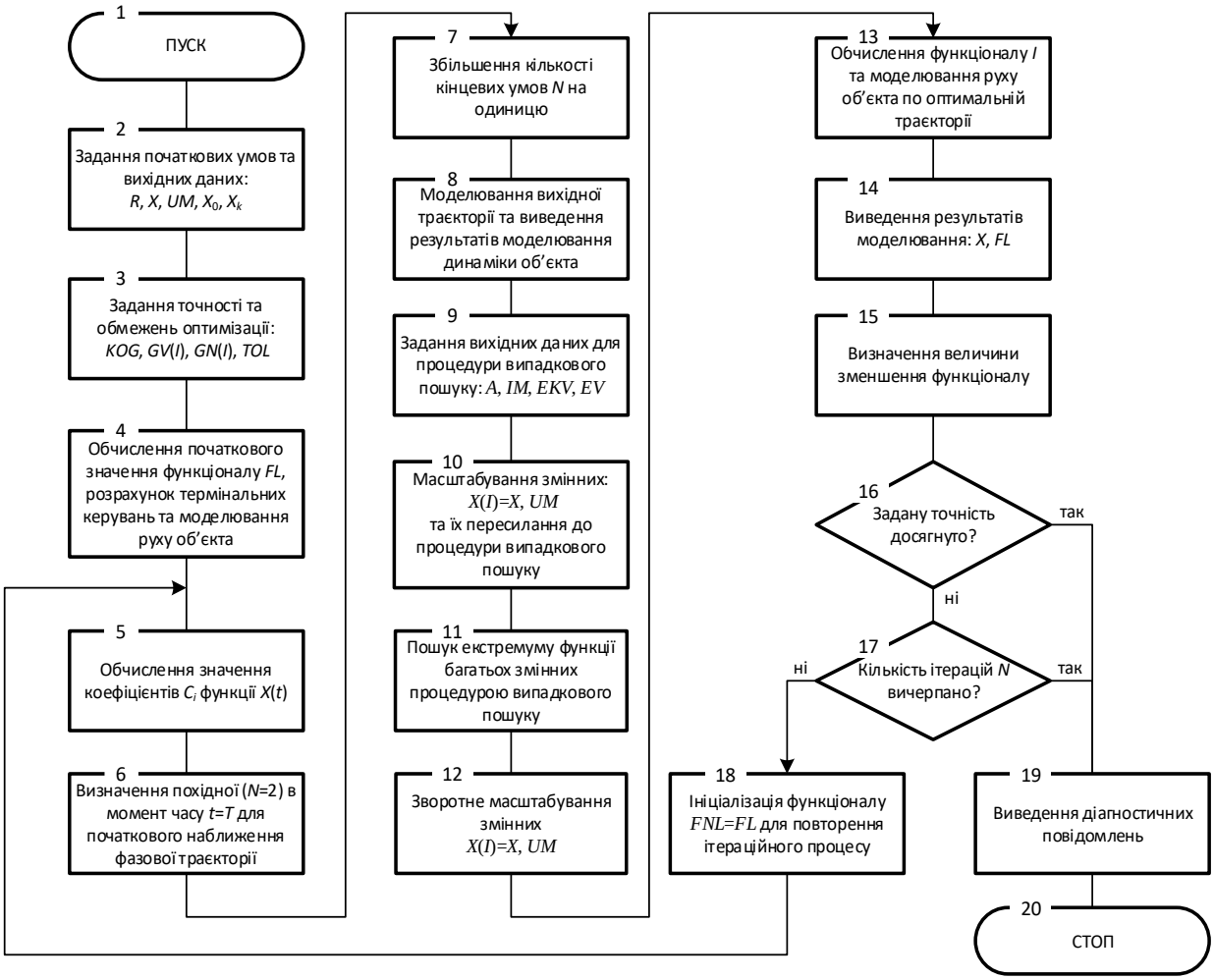


Рис. 3. Блок-схема алгоритму синтезу оптимальних керувань

Після визначення величини зменшення функціоналу приймається рішення про продовження обчислень (блоки 15 та 16 на рис. 3): якщо зменшення функціоналу недостатньо і кількість ітерацій не досягла максимально допустимого значення, робиться ще одна ітерація (перехід на блок 5 на рис. 3). Інакше виводиться відповідне повідомлення (блок 19).

На основі розроблених моделі та алгоритму пошуку термінальних керувань визначимо закон керування, який би забезпечив переведення дизель-поїзда масою m та швидкістю руху $V = V_{t_1}$ за час $T = t_2 - t_1$ на відстань S та з кінцевою швидкістю руху $V_t = V_{t_2}$ з урахуванням мінімізації обраного функціонала.

При цьому в процесі розгону швидкість і прискорення об'єкта не повинні перевищувати гранично допустимі значення швидкості (V_{don}) і прискорення $\left.\frac{dV}{dt}\right|_{don}$ на заданій ділянці шляху, а керування (сила тяги) – гранично допустиму силу тяги ($F_{TП}$), яка визначається співвідношенням (5).

На рис. 4 наведено зміни сили тяги: крива з індексом «1» відповідає початковому рішенню крайової задачі, крива з індексом «2» – зміні сили тяги при мінімізації часу руху, крива з індексом «3» – зміні сили тяги при оптимізації витрат палива.

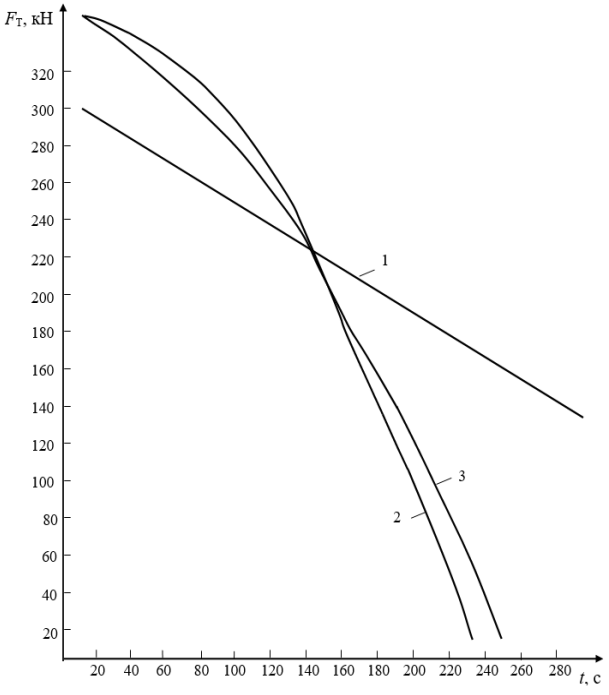


Рис. 4. Зміна сили тяги поїзда (F_T) під час вирішення задач мінімізації часу руху та витрат палива

Синтезований оптимальний закон керування рухом, що забезпечує мінімальні витрати палива, при русі поїзда на заданій ділянці шляху та дотриманні обмежень на фазові координати та керування має вигляд:

$$F_T = F_{T0} + m \cdot \begin{pmatrix} 2C_2 + 6C_3 t + 12C_4 t^2 + \\ +20C_5 t^3 + 30C_6 t^4 + 42C_7 t^5 \end{pmatrix} + a_0 \cdot \begin{pmatrix} C_1 + 2C_2 t + 3C_3 t^2 + \\ + 4C_4 t^3 + 5C_5 t^4 + 6C_6 t^5 + 7C_7 t^6 \end{pmatrix} + a_1 \cdot \begin{pmatrix} C_1 + 2C_2 t + 3C_3 t^2 + \\ + 4C_4 t^3 + 5C_5 t^4 + 6C_6 t^5 + 7C_7 t^6 \end{pmatrix}^2, \quad (20)$$

де $F_{T0} = \phi(x) + F_{C0}$; $\phi(x)$ – функція, що визначається вагою поїзда та величиною ухилу колії; m – маса поїзда; t – час, що відраховується від початку інтервалу керування; a_0, a_1 – коефіцієнти апроксимації, що залежать від маси поїзда, визначаються з виразів (7) – (9); $C_i (i = 1,...,7)$ – коефіцієнти полінома (16), обумовлені програмою пошуку оптимальних керувань, виходячи з початкових та кінцевих умов, при оптимізації витрат пального.

Синтезований оптимальний закон керування рухом, що забезпечує мінімальний час проходження заданої ділянки шляху та дотримання обмежень на фазові координати та керування має такий вигляд:

$$F_T = F_{T0} + m \cdot \left(2C_2 + 6C_3 t + 12C_4 t^2 + 20C_5 t^3 + 30C_6 t^4\right) + a_0 \cdot \left(C_1 + 2C_2 t + 3C_3 t^2 + 4C_4 t^3 + 5C_5 t^4 + 6C_6 t^5\right) + a_1 \cdot \left(C_1 + 2C_2 t + 3C_3 t^2 + 4C_4 t^3 + 5C_5 t^4 + 6C_6 t^5\right)^2, \quad (21)$$

де $C_i (i = 1,...,6)$ – коефіцієнти полінома (16), які визначаються програмою пошуку оптимальних керувань, виходячи з початкових та кінцевих умов при мінімізації часу руху.

На основі запропонованого алгоритму було розроблено програмне забезпечення для визначення економічного ведення дизель-поїздів серії ДЕЛ-02, яке призначене для роботи у складі МПС керування і дозволяє обирати та контролювати до 20 аналогових та 46 дискретних сигналів (період опитування датчиків становить 10 мс, тривалість безперервного запису – до 24 годин).

Висновки

У роботі досліджено можливість підвищення економічних характеристик дизель-поїздів серії ДЕЛ-02 за рахунок оптимізації їх руху:

1) виконано аналіз задачі оптимального керування рухом, який показав, що її рішення знаходиться в певних рамках обмежень, пов'язаних як із дотриманням графіка руху, так і обмежень на фазові координати, їх похідні та керування;

2) досліджено два основних критерія оптимальності: витрати палива та час проходження ділянки шляху (мінімізація часу руху при накладених обмеженнях вимагає більших витрат палива і не призводить до суттєвого виграшу за часом руху поїзда);

3) розроблено алгоритм синтезу оптимальних термінальних керувань рухом дизель-поїзда на ділянці шляху з урахуванням техніко-економічних обмежень;

4) розроблено і втілено на дизель-поїзді ДЕЛ-02 програмне забезпечення, яке дозволяє: виконувати знімання інформації з ІС за весь час руху дизель-поїзда; надавати рекомендації машиністу щодо ведення дизель-поїзда з мінімальною витратою палива; виконувати статистичний аналіз записів для пошуку мінімальних, максимальних та середніх значень потрібних параметрів.

Список літератури

1. Енергозберігаючі технології в локомотивному господарстві: навч. посі. Частина 1 / Е.Д. Тартаковський, Д.О. Аулін, Д.М. Коваленко, М.О. Котов. – Харків: УкрДУЗТ, 2019. – 130 с. URL: http://lib.kart.edu.ua/handle/123456789/2172
2. Щербак Я., Нерубацький В. Аналіз варіантів вибору критерію оптимізації енерговитрат на тягові потреби рухомого складу залізниць. *Автоматизовані системи електричного транспорту*. Збірник наукових праць УкрДАЗТ, Вип. 127, 2011. С. 137–142. URL: http://lib.kart.edu.ua/handle/123456789/4425
3. Заполовський М., Мезенцев М. Синтез керувань дизель-поїзда з електроприводом змінного струму. *Системи управління, навігації та зв'язку*, Випуск 3 (61), 2020. С. 57–63. DOI: https://doi.org/10.26906/SUNZ.2020.3.057
4. Кислий Д. Визначення енергозаощаджуючих режимів ведення поїздів. *Наука та прогрес транспорту*. Вісник Дніпропетровського національного університету залізничного транспорту, № 1 (61), 2016. С. 71–84. DOI: https://doi.org/10.15802/stp2016/60983
5. Носков В., Гавриленко С., Гейко М., Панченко В. Контроль використання потужності дизеля на тягу в умовах експлуатації дизель-поїздів. *Системи управління, навігації та зв'язку*. Том 4. № 78, 2024. С. 38–41. DOI: https://doi.org/10.26906/SUNZ.2024.4.038
6. Притула М., Пасечник О. Розроблення алгоритмів формування енергооптимальних режимів руху поїздів. *Наука та прогрес транспорту*. Вісник Дніпропетровського національного університету залізничного транспорту. № 6 (78), 2018. С. 82–100. DOI: https://doi.org/10.15802/stp2018/154641
7. Сулим А., Мельник О., Бялобржеський О., Ломонос А. Дослідження факторів та оцінка рівня їх впливу на показник питомих витрат електроенергії рухомого складу. *Вісник Східноукраїнського національного університету імені Володимира Даля*, № 4 (268), 2021. С. 118–127. DOI: https://doi.org/10.33216/1998-7927-2021-268-4-118-127
8. Петренко О., Любарський Б. Математична модель оптимального керування рухом електрорухомого складу на підставі вирішення рівнянь Гамільтона-Якобі-Беллмана. *Інформаційно – керуючі системи на залізничному транспорті*. № 2, 2016. С. 19–24. DOI: https://doi.org/10.18664/ikszt.v0i2.67639

9. Боднар Б., Капіца М., Афанасов А., Кислий Д. Визначення енергозаощаджуючих режимів розгону поїздів. *Наука та прогрес транспорту*. Вісник Дніпропетровського національного університету залізничного транспорту, № 5 (59), 2015. С. 40–52. DOI: https://doi.org/10.15802/stp2015/55359
10. Ситнік Б., Бриксін В., Ломотько Д., Ситник В., Давидов І. Моделі і методи створення систем реалізації графіків руху високошвидкісних поїздів з адаптивною корекцією швидкості за фактичними параметрами проїзду. Частина 1. Структура автоматичної системи нечіткого задання графіка швидкості руху рухомого об'єкта з її корекцією за фактичними параметрами проїзду. *Інформаційно – керуючі системи на залізничному транспорті*. № 4, 2021. С. 24 – 35. DOI: https://doi.org/10.18664/ikszt.v26i4.247235
11. Hongbo Ye, Ronghui Liu. Nonlinear programming methods based on closed-form expressions for optimal train control. *Transportation Research Part C: Emerging Technologies*, Volume 82, September 2017. P. 102–123. DOI: https://doi.org/10.1016/j.trc.2017.06.011
12. Шапран Е. Удосконалення систем керування тяговими електродвигунами тепловозів. *Вісник Дніпропетровського національного університету залізничного транспорту імені академіка В. Лазаряна*, Вип. 8, 2005. С. 122–129. URL: https://crust.ust.edu.ua/items/83b437bf-d806-4ca5-9f7e-7b8c15db0d77
13. Боровська Т.М. Теорія автоматичного управління: курс лекцій. – Вінниця: ВНТУ, 2018. – 256 с. URL: https://pdf.lib.vntu.edu.ua/books/IRVC/2021/Borovska_2018_256.pdf
14. Кравченко С., Шматько Д. Синтез систем термінального управління з прогнозуючими моделями. *Автоматизація та Приладобудування* («*Automation and development of electronic devices» ADED-2025*). Вип. 1, 2025. С. 233–236. URL: https://openarchive.nure.ua/handle/document/30943

Received (Надійшла) 24.07.2025
Accepted for publication (Прийнята до друку) 29.10.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Носков Валентин Іванович – доктор технічних наук, доцент, професор кафедри комп'ютерної інженерії та програмування, Національний технічний університет «Харківський політехнічний інститут», Харків, Україна;
Valentin Noskov – Doctor of Technical Sciences, Associate Professor, Professor of the Computer Engineering and Programming Department, National Technical University «Kharkiv Polytechnic Institute», Kharkiv, Ukraine;
e-mail: Valentyn.Noskov@khpi.edu.ua; ORCID Author ID: https://orcid.org/0000-0002-7879-0706;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57331254200.

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 Computer Engineering and Programming Department, 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.

Панченко Володимир Іванович – старший викладач кафедри комп'ютерної інженерії та програмування, Національний технічний університет «Харківський політехнічний інститут», Харків, Україна;
Volodymyr Panchenko – 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;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=58759071400.

Гейко Геннадій Вікторович – кандидат технічних наук, доцент кафедри комп'ютерної інженерії та програмування, Національний технічний університет «Харківський політехнічний інститут», Харків, Україна;
Hennadii Heiko – Candidate of Technical Sciences, Associate Professor of the Computer Engineering and Programming Department, National Technical University «Kharkiv Polytechnic Institute», Kharkiv, Ukraine;
e-mail: Hennadii.Heiko@khpi.edu.ua; ORCID Author ID: https://orcid.org/0000-0001-6958-8306;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=58759078500.

Determination of optimal diesel train motion on a section of a route with a known profile

Valentin Noskov, Maksym Lipchanskyi, Volodymyr Panchenko, Hennadii Heiko

Abstract. Relevance. Railway transport is one of the main consumers of energy resources, therefore, the development and implementation of energy-saving technologies is a relevant task, the solution of which has recently become one of the priority areas of activity for specialists in both fundamental and applied science in this field. **Object of research**: the processes of controlling the traction, dynamic and energy characteristics of multiple-unit rolling stock. **Purpose of the article**. To solve the problem of optimal control of a diesel train's movement along a section of the track with a known profile, taking into account optimal fuel consumption and compliance with the traffic schedule. **Research results**. The article formulates and solves the problem of synthesizing a control system for the optimal movement of a diesel train with a known track profile using the terminal control method. The following were applied: the theory of train traction and traction electric drive, methods of system analysis and optimal control theory, mathematical modeling and computational experiments. **Conclusions**. An algorithm for optimal control of a diesel train has been developed, with the help of which program laws have been determined that establish the dependence of traction on the train's mass, track parameters, and technical and economic conditions (speed, acceleration, traffic schedule) taking fuel consumption into account.

Keywords: multiple unit rolling stock, optimal train movement, terminal controls, control system synthesis, mathematical modeling, railway track profile.

Інформаційні технології

УДК 004.056.55:004.312.2 **doi: 10.26906/SUNZ.2025.4.39-44**

М. Е. Бондаренко, Г. С. Іващенко

Харківський національний університет радіоелектроніки, Харків, Україна

ОРГАНІЗАЦІЯ ПАРАЛЕЛЬНОГО ВИКОНАННЯ МЕТОДІВ ОБРОБКИ ГОЛОСОВИХ СИГНАЛІВ НА БАГАТОЯДЕРНИХ CPU ТА GPU

Анотація. Актуальність. Набули поширення такі системи, як голосові помічники та засоби ідентифікації мовця, які функціонують на основі обробки голосових сигналів. Продуктивність цих систем залежить від обсягів даних і умов функціонування. Обробка великих масивів голосових сигналів або забезпечення роботи в реальному часі вимагає високопродуктивних обчислень. Таку швидкодію можна досягти за допомогою масивно-паралельних систем, включаючи багатопроцесорні кластери або дискретні GPU. **Об'єкт дослідження**: організація паралельних обчислювальних процесів у задачах обробки голосових сигналів із використанням можливостей архітектур сучасних процесорів. **Мета статті:** розробка системи паралельної обробки голосових сигналів з адаптованими алгоритмами для багатоядерних CPU, інтегрованих та дискретних GPU. **Результати дослідження**. Порівняльний аналіз показав, що за невеликого навантаження (голосові помічники, персональні застосунки) достатньо використання CPU, що забезпечує ефективне виконання обчислень із низькими часовими затримками. Натомість для обробки великих масивів голосових даних запропонований паралельний підхід, реалізований на CPU і GPU, що скорочує час виконання на 25-30% порівняно з послідовною реалізацією на CPU. **Висновки**. Дослідження показали, що використання паралелізму на CPU доцільне для етапів обробки голосового сигналу із малим обсягом обчислень, тоді як дискретні GPU можуть бути використані на етапах обчислення MFCC, спектрального віднімання та вейвлет-фільтрації.

Ключові слова: системи голосової ідентифікації, обробка сигналів, виділення характеристик, нормалізація, MFCC, спектральне віднімання та вейвлет-фільтрація, CPU, GPU, iGPU.

Вступ

Постановка проблеми. Потреба у захисті інформації від несанкціонованого доступу сприяє розвитку систем голосової ідентифікації, які застосовуються у біометричній аутентифікації, голосових асистентах, мобільних застосунках, хмарних сервісах, онлайн-банкінгу та системах контролю доступу [1]. Автоматизація цих систем забезпечується використанням методів обробки сигналів, зокрема спектральним аналізом та оцінкою тембрових і ритмічних характеристик голосу [2].

Зі збільшенням масштабів застосування та розвитком технологій обробки даних зростають вимоги до продуктивності обчислювальних платформ, оскільки обробка великих обсягів даних у реальному часі включаючи фільтрацію сигналу, вилучення ознак, обчислення спектральних характеристик і формування моделей голосу потребує значних обчислювальних ресурсів. Реальні умови експлуатації, що характеризуються високим фоновим шумом, наявністю реверберації та низькою якістю записів, створюють значне обчислювальне навантаження на систему через необхідність виконання шумозаглушення, компенсації реверберації, нормалізації сигналу та виділення ознак [3]. Через це актуальним є застосування паралельної обробки на багатоядерних CPU та дискретних GPU, що дозволяє підвищити продуктивність системи та скоротити час виконання обчислень.

Сучасні CPU від Intel, AMD та Apple забезпечують високу продуктивність завдяки багатоядерній архітектурі та підтримки багатопоточності, що робить їх ефективними для складних обчислень у реальному часі, зокрема аналізу даних, машинного навчання та розпізнавання мовлення [4]. Алгоритми з послідовною або обмеженою паралельною обробкою реалізуються на CPU, забезпечуючи контроль над порядком виконання команд та організацію потоків обчислень. На початкових етапах обробки голосових сигналів вони виконують шумозаниження, нормалізацію амплітуди, фреймування та виділення характеристик [3]. Проте обмежена здатність до масово-паралельних обчислень знижує ефективність використання багатоядерних CPU у сучасних системах, що потребують одночасного виконання великої кількості операцій. У таких випадках доцільно застосовувати інтегровані та дискретні GPU, здатні забезпечити обробку великих даних у реальному часі [5].

Графічні процесори від NVIDIA та AMD спеціалізуються на масово-паралельних обчисленнях і застосовуються у відеообробці, 3D-рендерингу та глибокому навчанні. Архітектура GPU підтримує одночасне виконання великої кількості однотипних операцій, обробку великих голосових масивів, обчислення MFCC та реалізацію складних моделей розпізнавання [6].

Інтегровані GPU від Intel та AMD використовують системну пам'ять, спільну з CPU, що знижує затримки при обміні даними та підвищує загальну енергоефективність. Їх застосування є доцільним при обмежених ресурсах або обробці невеликих обсягів аудіосигналів [7].

Аналіз останніх досліджень і публікацій. Актуальні напрямки розвитку систем обробки голосових сигналів відображені у сучасних дослідженнях [8–10], що присвячені організації паралельних обчи-

слень та підвищенню продуктивності на різних апаратних платформах для задач аналізу та обробки мовних даних.

Дослідження [8] демонструє особливості виконання мобільних моделей машинного навчання, які широко використовуються в системах голосової ідентифікації, зокрема Deep Neural Networks (DNN), на пристроях з обмеженими ресурсами.

В системах голосової ідентифікації обробка сигналів характеризується високим обчислювальним навантаженням на процесор, обумовленим необхідністю виконання складних алгоритмів попередньої обробки та виділення ознак. Як показано в [8], для невеликих задач, таких як обробка окремих фреймів або швидке вилучення ознак, CPU забезпечує достатню паралельність без залучення GPU. Для складних моделей або великих наборів даних GPU суттєво прискорює обчислення завдяки масово-паралельній обробці.

Дослідження [9] показує застосування розпізнавання мовлення у системах голосового керування, IP-телефонії та персональної ідентифікації. Особлива увага приділяється обчисленню MFCC для виділення фонетичних характеристик, що потребує значних ресурсів. Паралелізація обчислення MFCC виконана на Intel Core i5-2310 за допомогою технології OpenMP, із розбиттям сигналу на незалежні фрейми. Експериментальні результати показали значне скорочення часу обробки.

Дослідження [10] присвячене паралельній обробці прихованих марковських моделей (HMM), що у системах голосової обробки використовуються для моделювання та аналізу послідовності часових ознак мовних сигналів. Навчання та тестування моделей виконувалося на GPU з використанням архітектури паралельних обчислень CUDA, що дозволяє ефективно обробляти великі обсяги даних у багатопоточному режимі. На CPU виконувались послідовні та керуючі операції. За допомогою cuBLAS та механізму HyperQ було реалізовано одночасне виконання кількох обчислювальних черг на GPU, що дозволило скоротити час навчання та обробки моделей до 9 разів порівняно з використанням багатоядерних CPU.

Розглянуті дослідження підтверджують доцільність використання як багатоядерних CPU, так і GPU для реалізації паралельних обчислень у задачах обробки голосових сигналів та ідентифікації мовця.

Метою роботи є розробка та дослідження системи паралельної обробки голосових сигналів із адаптованими алгоритмами для багатоядерних CPU, інтегрованих та дискретних GPU.

Постановка задачі

Методи обробки голосових сигналів, такі як фреймування, виділення ознак, спектральний аналіз та застосування вейвлет-фільтрації або MFCC, повинні підтримувати паралельне виконання на CPU, інтегрованому або дискретному GPU. Для цього використовуються обчислювальні можливості процесорів та спеціалізовані бібліотеки, зокрема OpenMP для багатопоточності на CPU та CUDA для масово-паралельної обробки на GPU. Такий підхід дозволяє розподіляти обчислення між потоками або ядрами, ефективно використовувати ресурси пам'яті та оптимізувати обробку великих обсягів даних, забезпечуючи скорочення часу виконання та підвищення продуктивності системи [11-13].

Реалізація повинна підтримувати обробку коротких фрагментів із невеликим навантаженням і великих масивів однотипних обчислень, що відтворюють реальні умови експлуатації з варіаціями рівня шуму, спектральної насиченості та частоти дискретизації. Така конфігурація тестових даних дозволить оцінити здатність обчислювальних вузлів підтримувати одночасну обробку численних завдань та визначити їхні переваги й обмеження щодо продуктивності, використання ресурсів і точності розпізнавання.

Основний матеріал

У системах голосової ідентифікації обробка сигналів поділяється на кілька етапів: попередня підготовка сигналу, виділення ознак, обробка та нормалізація ознак, порівняння отриманих характеристик із шаблонами, а також прийняття рішення про ідентичність або відмінність голосів.

На етапі попередньої обробки сигналів здійснюється нормалізація, фільтрація та, за потреби, зменшення шумових компонентів [14]. На етапі виділення ознак проводяться спектральний аналіз, обчислення мел-частотних кепстральних коефіцієнтів (MFCC) та інші обчислювальні операції, необхідні для подальшої класифікації. Запропонований підхід до організації паралельних обчислень розкрито на прикладі обчислення MFCC.

Обчислення MFCC полягає у перетворенні часових сигналів у спектральну форму, що відображає частотну структуру звуку відповідно до сприйняття людським вухом [14-15]. В класичній реалізації першим етапом є фреймування, тобто розбиття сигналу на короткі кадри тривалістю 15-30 мс з перекриттям в 50%, що дозволяє зберегти сталість статистичних характеристик у межах кадру. Після цього на кожен кадр застосовується віконна функція Хеммінга, для зменшення спектрального витоку:

$$x_m[n] = x[n + m \times R] \cdot w[n], \quad n = 0,1,\ldots,N-1, \tag{1}$$

де $w[n]$ – віконна функція, R – крок перекриття кадрів.

Далі обчислюється спектр за допомогою швидкого перетворення Фур'є (FFT):

$$X_m[k] = \sum_{n=0}^{N-1} x_m[n] \log e^{-j2\pi kn/N}, \quad k = 0,\ldots,N-1, \tag{2}$$

Виконується перехід до мел-шкали з подальшим логарифмуванням $log(E)$, де E – енергія спектрального компонента, та дискретним косинусним перетворенням (DCT) для отримання фінальних MFCC-коефіцієнтів. Розрахунок MFCC виконується відповідно до:

$$c_m[n] = \alpha_n \sum_{i=1}^{K} \log E_m[i] \times \cos\left(\frac{\pi n}{K}(i-0.5)\right), \quad (3)$$
$$n = 0, .., L-1,$$

де L – кількість коефіцієнтів MFCC, що зберігають основну спектральну інформацію кадру.

Послідовне виконання описаних операцій для великого обсягу даних створює значні часові затримки, обмежує можливість обробки потокового сигналу та ускладнює інтеграцію алгоритму у високопродуктивні системи реального часу.

Природною властивістю аудіосигналів є відносна незалежність коротких кадрів, що дозволяє організувати паралельне виконання алгоритмів попередньої обробки. Кожен кадр містить спектральну інформацію, обмежену коротким інтервалом часу, що відображає локальні властивості сигналу і може оброблятися незалежно, що дозволяє виконувати FFT, обчислення мел-енергії, логарифмування та DCT одночасно для декількох кадрів. Завдяки незалежності обчислень між кадрами, такий підхід дозволяє організувати паралельну роботу. Проте просте перенесення класичного алгоритму на паралельні архітектури не є достатнім, потрібна адаптація методу, включно з організацією кадрів у батчі або тензори, оптимізацією доступу до пам'яті та синхронізацією результатів для збереження коректної послідовності обчислень. На рис. 1 наведено приклад формування батча розмірністю 3×3×4 із голосових сигналів, представлених у спектральному вигляді.

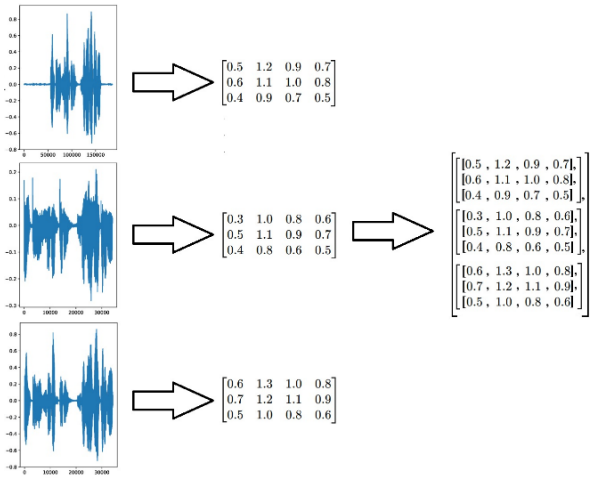


Рис. 1. Об'єднання масивів спектральних ознак голосових сигналів у батч

В процесі адаптації алгоритму кадри формуються у батчі таким чином, щоб забезпечити вирівняний доступ до пам'яті та можливість одночасної обробки кількох кадрів. Це дозволяє організувати дані у вигляді багатовимірних тензорів, де перша вісь відповідає номеру прикладу в батчі, а наступні часовим і частотним вимірам сигналу. Кожен кадр обробляється автономно, включно з обчисленням FFT, мел-енергії, логарифму та DCT, що запобігає блокуванням у потоках.

Результати обробки кадрів синхронізуються у вихідній структурі даних, забезпечуючи правильний порядок MFCC-коефіцієнтів для подальшого аналізу. Така організація дозволяє ефективно використовувати кеш-пам'ять і прискорювачі (GPU, iGPU, TPU), оскільки операції виконуються над цілими блоками даних, а не над окремими зразками, що зменшує накладні витрати на обмін даними. Розмір батчу є важливим фактором: малі батчі (8–16 кадрів) забезпечують швидшу реакцію системи та частіше оновлення моделі, тоді як великі (64–256 і більше) краще завантажують GPU та ефективніше використовують ресурси, але потребують більше пам'яті.

В даній роботі підготовка методу обчислення MFCC до ефективного виконання на CPU здійснювалася з урахуванням можливостей сучасних багатоядерних архітектур та особливостей організації пам'яті, таких як ієрархія кешів, розподіл пам'яті у NUMA-системах, вирівнювання даних і використання регістрового зберігання проміжних результатів. Використання багатопоточності та багатоядерності через бібліотеки Multiprocessing та Joblib дозволило уникнути гонок даних і забезпечити синхронізацію результатів. При цьому методи обчислення були адаптовані для використання можливостей SIMD-векторизації (Single Instruction, Multiple Data): дані організовувалися у суміжні блоки, що дозволяло виконувати однотипні операції над кількома елементами масиву одночасно за допомогою векторних інструкцій CPU, що значно підвищує продуктивність обчислень. Локалізація даних у кешах забезпечувала мінімізацію доступів до повільної основної пам'яті та дозволила ефективно задіяти високопропускні регістрові шини процесора для обробки багатьох кадрів одночасно.

Додатково враховувалася організація пам'яті у багатопроцесорних системах із NUMA для зменшення затримок доступу до віддалених банків пам'яті та підвищення ефективності паралельних обчислень.

Бібліотеки CuPy та PyTorch із підтримкою CUDA забезпечують низькорівневий доступ до обчислювальних блоків, де кожен кадр обробляється багатьма потоками, які групуються у блоки. Така організація дозволяє одночасно виконувати велику кількість операцій над різними кадрами та їхніми частотними компонентами.

Для оптимізації швидкодії особливу увагу приділено використанню пам'яті. Проміжні результати, такі як FFT, мел-енергія та DCT-коефіцієнти, зберігаються у регістрах обчислювальних потоків, що мінімізує затримки доступу до даних. Дані, до яких здійснюється частий доступ, розташовуються у спільній пам'яті, що дозволяє уникнути численних звернень до глобальної пам'яті та зменшує накладні витрати на обмін між потоками. Таке розміщення даних забезпечує ефективне використання кеш-пам'яті та обчислювальних ресурсів під час одночасної обробки кадрів і батчів сигналів, скорочує загальний час обробки та підтримує масштабованість методів для великих наборів голосових даних.

Відмінності в адаптації алгоритму для CPU та GPU зумовлені потребою в організації пам'яті, формуванны батчів і керування потоками для досягнення максимальної продуктивності. На CPU накладні

витрати пов'язані з плануванням процесів і потоків, тоді як на GPU основними джерелами накладних витрат є пересилання даних між процесорами та підготовка багатовимірних масивів для ефективного паралельного виконання обчислень.

Приблизний час виконання алгоритму у паралельному режимі розраховується згідно:

$$T_{пар.} \approx \frac{T_{послід.}}{P} + T_{затрим.}, \quad (4)$$

де $T_{послід.}$ – час послідовного виконання, P – кількість паралельних обчислювальних одиниць, $T_{затрим.}$ – час накладних витрат на синхронізацію та управління пам'яттю.

Для оцінки продуктивності системи необхідно проаналізувати роботу при різних обсягах даних і розмірів батчів, їх вплив на час виконання та ефективність використання ресурсів.

Результати експериментальних досліджень

Експериментальна модель системи обробки голосових сигналів була розгорнута в середовищі Jupyter Notebook. Модель передбачає використання як CPU, так і GPU, включно з інтегрованими GPU, для виконання різних етапів обробки. Це дозволяє контролювати розподіл обчислювальних завдань між апаратними ресурсами та оцінювати ефективність паралельної обробки етапів голосових сигналів залежно від складності та обсягу даних.

У дослідженні для порівняння ефективності паралельних обчислень використовувалися різні обчислювальні вузли. CPU Intel Core i7-13700H (6 продуктивних і 8 ефективних ядер, 20 потоків) забезпечує високу продуктивність на малих та середніх обсягах даних. Інтегрований GPU Intel Iris Xe (96 виконавчих блоків, 768 шейдерних ядер) ефективний для середніх обсягів сигналів, тоді як дискретний GPU на архітектурі Ada Lovelace (3072 CUDA-ядер) обробляє великі набори даних у ресурсномістких задачах ідентифікації користувача.

Для оцінки ефективності обробки голосових сигналів використано набір CHiME [13], що містить понад 50 годин аудіо реальних розмов чотирьох учасників у межах 20 сесій. На його основі сформовано піднабори різного розміру для порівняння продуктивності обчислювальних вузлів і впливу паралельних обчислень.

Кадри аудіосигналу формувалися у батчі та організовувалися у багатовимірні тензори з узгодженим вирівнюванням у пам'яті, що забезпечувало одночасну обробку кількох кадрів та ефективне використання обчислювальних ресурсів. Кожен кадр проходив автономно повний цикл перетворень, включно з FFT, обчисленням мел-енергії, логарифмуванням та DCT, зменшуючи накладні витрати на синхронізацію та підвищуючи продуктивність CPU.

На GPU паралельне виконання алгоритму реалізовувалося через масове однотипне обчислення на великій кількості кадрів, що робило його ефективною платформою для обробки MFCC у батчах. Кадри формувалися у багатовимірні тензори, де перша вісь відповідала номеру прикладу, а наступні часовим і частотним вимірам сигналу, що дозволяло максимально використовувати апаратну структуру GPU. Для досліджуваної системи з розпаралелюванням на GPU експериментально обрано батч розміром 32 кадра.

Таблиця 1 – **Час отримання MFCC**

Кількість записів	Час обробки, с		
	CPU	iGPU	Дискретний GPU
100	9,2	5,1	4,3
500	48,5	22,4	20,7
2500	250,3	102,5	98,2
5000	520,1	198,4	195,3
10000	1048,5	395,2	384,7
100000	10390,6	3712,8	3820,4

В ході першого експерименту оцінено час виконання розрахунку MFCC, що застосовуються на етапі виділення ознак у коротких голосових фрагментах. Експеримент дозволив оцінити здатність CPU обробляти невеликі обсяги даних без залучення GPU та порівняти ефективність обчислень у задачах низького і високого навантаження при обробці голосових сигналів.

Результати, наведені у табл. 1, демонструють, що паралельне виконання на центральному процесорі забезпечує продуктивність лише на невеликих обсягах даних (100-500 записів). При обробці великих обсягів даних (10000-1000000 записів) паралельне виконання на дискретному GPU та інтегрованому GPU забезпечує трикратну перевагу за часом виконання у порівнянні з обчисленнями на центральному процесорі.

У другому експерименті розглянуто застосування спектрального віднімання для пригнічення шумових компонентів голосового сигналу. Даний підхід передбачає аналіз частотних спектрів для кожного фрейму, що суттєво підвищує обчислювальне навантаження навіть при роботі з відносно невеликими наборами даних.

Результати експерименту показують, що паралелізація на базі CPU демонструє задовільну продуктивність для обробки невеликих наборів даних, тоді як збільшення обсягу понад 5000 записів призводить до істотного зростання часу обробки (табл. 2).

Таблиця 2 – **Час обробки даних методом спектрального віднімання**

Кількість записів	Час обробки, с		
	CPU	iGPU	Дискретний GPU
100	10,3	5,5	4,3
500	51,5	23,0	21,3
2500	254,1	105,4	100,2
5000	523,7	205,6	197,8
10000	1079,2	405,7	394,9
100000	10564,3	3780,1	3850,3

Використання інтегрованого та дискретного GPU забезпечує суттєве прискорення обробки спектральних фреймів, що дозволяє ефективно працювати з великими обсягами голосових сигналів та скорочує загальний час виконання ідентифікації користувачів.

Таблиця 3 – **Час обробки даних методом вейвлет-фільтрації**

Кількість записів	Час обробки, с		
	CPU	iGPU	Дискретний GPU
100	41,1	22,5	21,0
500	210,3	103,2	100,5
2500	1050,6	520,3	510,2
5000	2102,6	1032,6	1015,8
10000	4205,7	2060,4	2025,1
100000	42062,3	20780,1	20140,2

При застосуванні вейвлет-фільтрації, яка забезпечує багаторівневий аналіз сигналу в часочастотній області, обчислювальна складність значно зростає через великий обсяг операцій на кожному рівні декомпозиції.

Наведені в табл. 3 результати демонструють, що зростання обчислювального навантаження при обробці голосового сигналу обмежує ефективність використання паралелізації на CPU. Застосування ресурсоємних методів, таких як вейвлет-фільтрація, призводить до значного збільшення часу виконання на CPU та інтегрованих GPU, що доводить перевагу використання дискретного GPU.

Застосування паралельної обробки не впливає на точність роботи системи (відсоток успішної голосової ідентифікації користувача). Це свідчить про те, що підвищення продуктивності системи за рахунок розподілу обчислень між потоками або ядрами не призводить до зниження якості розпізнавання, що є важливим для забезпечення надійності систем голосової ідентифікації при роботі з великими обсягами даних. Формування кадрів у батчі забезпечує вирівняний доступ до пам'яті та впорядковане виконання обчислень, що зменшує час очікування даних і до-зволяє ефективно використовувати ресурси, такі як завантаження ядер CPU або потоків GPU та скоротити загальний час обробки. Важливим фактором є правильний вибір розміру батчу: занадто малий обсяг не забезпечує ефективної паралельності, тоді як надто великий може призводити до перевантаження пам'яті та зниження продуктивності.

Завдяки описаному підходу методи попередньої обробки голосових сигналів можуть виконуватись у реальних умовах ідентифікації користувачів, з одночасним збереженням точності результатів. Це дозволяє обробляти великі набори голосових сигналів та масштабувати систему залежно від апаратних можливостей.

Висновки

Реалізовано паралельне виконання методів обробки голосових сигналів, зокрема з обчисленням мел-частотних кепстральних коефіцієнтів (MFCC), спектральним відніманням для пригнічення шуму та вейвлет-фільтрацією. Проведено порівняльний аналіз продуктивності у системах голосової ідентифікації з використанням технологій паралельних обчислень на багатоядерному центральному процесорі (CPU), інтегрованому графічному процесорі (iGPU) та дискретному графічному процесорі (GPU). Експериментальні сценарії охоплювали широкий спектр обсягів вхідних даних та різних алгоритмів обробки.

Результати досліджень показали, що CPU забезпечує високу продуктивність для обробки невеликих наборів даних і виконання алгоритмів помірної складності, таких як вейвлет-фільтрація, спектральне віднімання та обчислення MFCC, завдяки багатоядерній архітектурі та можливості паралельної обробки без значних накладних витрат на передачу даних. Інтегрований GPU (iGPU) забезпечує можливість паралельної обробки, проте продуктивність при збільшенні обсягів даних зростає обмежено через його архітектурні особливості. Дискретний GPU дозволяє суттєво прискорити обробку великих обсягів даних завдяки масовій паралельній архітектурі та високій пропускній здатності пам'яті, проте при малих наборах даних накладні витрати на синхронізацію потоків і передачу знижують загальну ефективність.

Подальшим розвитком є застосування гібридної моделі, де обчислювальні завдання динамічно розподіляються між CPU та GPU залежно від обсягу та складності обробки. Такий підхід дозволить поєднувати високу швидкодію дискретного GPU при роботі з великими обсягами даних із ефективністю CPU на малих обсягах, зменшуючи накладні витрати на передачу даних та синхронізацію потоків.

СПИСОК ЛІТЕРАТУРИ

1. A. Kumar, S. Jain and M. Kumar, "Deep Learning based Fusion for a Multi-Biometric Identification Using LSTM", 2024 1st International Conference on Advanced Computing and Emerging Technologies (ACET), Ghaziabad, India, 2024, pp. 1-6. https://doi.org/10.1109/ACET61898.2024.10730213.
2. 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. https://doi.org/10.1109/KhPIWeek57572.2022.9916473.
3. Бондаренко М. Е., Іващенко Г. С. Використання послідовності методів попередньої обробки в системах голосової ідентифікації. Системи управління, навігації та зв'язку. Полтава: ПНТУ, 2025. Т. 2 (80). С. 90-96. https://doi.org/10.26906/SUNZ.2025.2.090.
4. B. Gawrych and P. Czarnul, "Performance Assessment of OpenMP Constructs and Benchmarks Using Modern Compilers and Multi-Core CPUs", 2023 18th Conference on Computer Science and Intelligence Systems (FedCSIS), Warsaw, Poland, 2023, pp. 973-978. https://doi.org/10.15439/2023F7822.

5. I. Vasileska, P. Tomšič, L. Kos and L. Bogdanović, "Unveiling Performance Insights and Portability Achievements Between CUDA and SYCL for Particle-in-Cell Codes on Different GPU Architectures", 2024 47th MIPRO ICT and Electronics Convention (MIPRO), Opatija, Croatia, 2024, pp. 1115-1120, DOI: https://doi.org/10.1109/MIPRO60963.2024.10569866.
6. F. Lumpp, H. D. Patel and N. Bombieri, "A Framework for Optimizing CPU-iGPU Communication on Embedded Platforms", 2021 58th ACM/IEEE Design Automation Conference (DAC), San Francisco, CA, USA, 2021, pp. 685-690. https://doi.org/10.1109/DAC18074.2021.9586304.
7. H. Li, J. K. Ng and T. Abdelzaher, "Enabling Real-time AI Inference on Mobile Devices via GPU-CPU Collaborative Execution", 2022 IEEE 28th International Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA), Taipei, Taiwan, 2022, pp. 195-204. https://doi.org/10.1109/RTCSA55878.2022.00027.
8. R. M. Fazliddinovich and B. U. Abdumurodovich, "Parallel processing capabilities in the process of speech recognition", 2017 International Conference on Information Science and Communications Technologies (ICISCT), Tashkent, Uzbekistan, 2017, pp. 1-3. https://doi.org/10.1109/ICISCT.2017.8188585.
9. L. Yu, Y. Ukidave and D. Kaeli, "GPU-Accelerated HMM for Speech Recognition", 2014 43rd International Conference on Parallel Processing Workshops, Minneapolis, MN, USA, 2014, pp. 395-402. https://doi.org/10.1109/ICPPW.2014.59.
10. S. M. Hussain, B. Saritha, B. S. Reddy, C. Srikar, B. Suchitra and G. Purnachandrarao, "DeepVoice: An End-to-End Speaker Recognition System Leveraging Convolutional and Recurrent Neural Networks for Robust Voice Identification", 2025 International Conference on Electronics, AI and Computing (EAIC), Jalandhar, India, 2025, pp. 1-5. https://doi.org/10.1109/EAIC66483.2025.11101389.
11. M. Dyvak and O. Kindzerskyi, "Implementation of Parallel Computation for Identification of Interval Models based on Multi-core Parallelism and CUDA Technology", 2024 14th International Conference on Advanced Computer Information Technologies (ACIT), Ceske Budejovice, Czech Republic, 2024, pp. 72-76. https://doi.org/10.1109/ACIT62333.2024.10712545.
12. R. Kouatly and T. A. Khan, "Performance of Text-Independent Automatic Speaker Recognition on a Multicore System", in Tsinghua Science and Technology, vol. 29, no. 2, pp. 447-456, April 2024. https://doi.org/10.26599/TST.2023.9010018.
13. P. Foster, S. Sigtia, S. Krstulovic, J. Barker, M. D. Plumbley. "CHiME-Home: A Dataset for Sound Source Recognition in a Domestic Environment", in Proceedings of the 11th Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), 2015, pp. 1-5. https://doi.org/10.1109/WASPAA.2015.7336899.
14. A. Moondra and P. Chahal, "Voice Feature Extraction Method Analysis for Speaker Recognition with Degraded Human Voice", 2023 5th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), Greater Noida, India, 2023, pp. 385-388. https://doi.org/10.1109/ICAC3N60023.2023.10541716.
15. Y. A. Wubet and K. -Y. Lian, "Speaker Anonymization for Voice Biometrics Protection Using Voice Conversion and Multi-Target Speaker Voice Fusion", in IEEE Transactions on Information Forensics and Security, vol. 20, pp. 6046-6057, 2025. https://doi.org/10.1109/TIFS.2025.3577023.

Received (Надійшла) 11.07.2025
Accepted for publication (Прийнята до друку) 15.10.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Іващенко 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.

Parallel implementation of voice signal processing methods on multicore CPU and GPU

Maksym Bondarenko, Heorhii Ivashchenko

Abstract. Relevance. Systems such as voice assistants and speaker identification tools, which operate based on voice signal processing, have become increasingly widespread. The performance of such systems depends on the volume of data and operating conditions. Processing large collections of voice signals or ensuring real-time operation requires high-performance computing. Such performance can be achieved with massively parallel systems, including multiprocessor clusters or discrete GPUs. **Object of research** is the organisation of parallel computing processes in voice signal processing tasks using the capabilities of modern processor architectures. **Purpose of the article** is to develop a parallel voice signal processing system with adapted algorithms for a multi-core CPU and an integrated and discrete GPU. **Research results**. Comparative analysis revealed that for small loads (such as voice assistants and personal applications), CPU usage is sufficient to ensure efficient computation with low latency. However, when processing large datasets and performing streaming analytics, the proposed parallel approach, implemented on both the CPU and GPU, reduces execution time by 25-30% compared to a sequential implementation on the CPU. **Conclusions**. Research has shown that parallelism on the CPU is suitable for stages of voice signal processing that require a small amount of computation. At the same time, discrete GPUs can be utilised in stages with intensive computational tasks such as MFCC calculation, spectral subtraction, and wavelet filtering.

Keywords: voice identification, signal processing, feature extraction, normalisation, MFCC, spectral subtraction and wavelet filtering, CPU, GPU, iGPU.

УДК 004.056.55 + 004.415.3 **doi: 10.26906/SUNZ.2025.4.45-50**

М. І. Главчев, Ю. М. Главчева, Г. І. Молчанов

Національний технічний університет «Харківський політехнічний інститут», Харків, Україна

РЕЖИМ БЛОЧНОГО ШИФРУВАННЯ НА ОСНОВІ КОДУ ХАФФМАНА ДЛЯ НІББЛІВ

Анотація. Актуальність. Актуальність роботи зумовлена потребою у криптографічних рішеннях, які поєднують високу безпеку, ефективність стиснення та низькі обчислювальні витрати, що особливо важливо для IoT, мобільних та вбудованих систем. **Метою статті** є дослідження розробки та аналіз режиму блочного шифрування на основі коду Хаффмана для ніббліів, з інтегрованими механізмами хеш-захисту та підтримкою як глобальної, так і пер-блокової обробки. **Об'єкт дослідження**: використання перетворення Хаффмана полубайтів (ніббліів) для створення режиму блокового шифрування. **Результати дослідження.** Проведені експерименти показали, що для даних із низькою ентропією ніббліів (H<3.5 біт/ніббл) запропонований режим зменшує обсяг зашифрованих даних на 10–20 % без суттєвої втрати швидкодії, забезпечуючи стійкість до помилок і збереження цілісності. **Висновок**: інтеграція статистичного кодування та симетричного шифрування з адаптивним вибором параметрів дозволяє отримати збалансоване рішення для безпечної та ефективної передачі інформації в умовах обмежених ресурсів.

Ключові слова: блочне шифрування, код Хаффмана, стиснення даних, криптографія, симетричні алгоритми.

Вступ

Постановка проблеми. Зі зростанням обсягів цифрового трафіку та поширенням мобільних і вбудованих пристроїв з обмеженими обчислювальними ресурсами постає необхідність у створенні криптографічних алгоритмів, які забезпечують належний рівень безпеки при мінімальних затратах пам'яті та процесорного часу. Традиційні блочні шифри, такі як AES чи DES, орієнтовані на роботу з блоками фіксованої довжини та не враховують статистичних особливостей вхідних даних, що призводить до збільшення обсягу переданої інформації та може створювати передбачувані патерни у шифротексті.

Використання коду Хаффмана у процесі шифрування дозволяє динамічно змінювати довжину кодових слів відповідно до частоти появи ніббліів, зменшуючи надлишковість і водночас підвищуючи складність криптоаналізу [1]. Проте інтеграція змінної довжини коду в блочну структуру вимагає вирішення проблеми синхронізації та забезпечення стійкості до атак, зокрема статистичних і диференціальних. Додатково важливим завданням є впровадження механізмів хеш-захисту (ХЗ) для перевірки цілісності та автентичності зашифрованих даних, що особливо актуально в умовах потенційного пошкодження чи підміни пакетів у каналах зв'язку. Таким чином, актуальною є розробка та дослідження режиму блочного шифрування, який поєднує адаптивне статистичне кодування, хешування та симетричне шифрування для 4-бітних блоків.

Аналіз останніх досліджень і публікацій. Використання алгоритмів статистичного кодування, зокрема коду Хаффмана, у криптографії досліджувалося у ряді наукових робіт. У [2] розглянуто інтеграцію кодування змінної довжини у блочні шифри з метою зменшення обсягу переданих даних, проте автори відзначають проблему синхронізації при пошкодженні бітового потоку. У [3] детально обговорюється архітектура безпеки в IoT для середовища пристроїв з обмеженими ресурсами, у тому числі використання блокових шифрів, хеш-функцій, потокових шифрів, високопродуктивних систем та пристроїв з низькими ресурсами для середовища IoT. Дослідження [4] аналізує використання ніббліів як базових одиниць обробки для легковагових шифрів у системах IoT, наголошуючи на їх продуктивності та знижених вимогах до пам'яті.

Водночас у [5] відзначено, що використання змінної довжини кодів без додаткового шифрування та контролю цілісності не гарантує належного рівня безпеки, оскільки частотний аналіз кодових слів може призвести до відновлення оригінальної інформації. Робота [6] розглядає впровадження хеш-функцій (SHA-256, SHA-3) у симетричні схеми шифрування як механізм запобігання атакам на цілісність. Для виявлення модифікацій автори [7] зосереджуємося на проблемі ідентифікації зразків у наборі, які не відповідають структурованим шаблонам, на основі використання методу хешування. Крім того, дослідження [8–10] демонструють перспективність комбінованих підходів, де стиснення, шифрування і хешування інтегруються в єдиний процес, оптимізуючи обчислювальні витрати та підвищуючи захист.

Метою роботи є створення та аналіз режиму блочного шифрування, який використовує код Хаффмана для оптимізації представлення ніббліів і підвищення стійкості до криптоаналітичних атак.

1. Кодування Хаффмана для ніббліів

Запропонований режим блочного шифрування ґрунтується на комбінації симетричного перетворення та адаптивного статистичного кодування за алгоритмом Хаффмана для 4-бітних блоків (ніббліів).

1. *Вхідні дані*. Послідовність відкритого тексту $M=\{m_1,m_2,\ldots,m_n\}$, де кожен елемент m_i - ніббл ($m_i \in \{0,1,\ldots,15\}$).

2. *Частотний аналіз*. Для множини ніббліів визначається частота появи: $p(m_i)=count(m_i)/n$, де

$$\sum_{i=0}^{15} p(m_i) = 1.$$

3. *Побудова кодового дерева Хаффмана*. Формується бінарне дерево, у якому кожному ніббліу

відповідає кодове слово змінної довжини l_i, таке що:

$$\sum_{i=0}^{15} p(m_i) - l_i \rightarrow min.$$

4. *Шифрування.* Отримані кодові слова пропускаються через симетричний блочний шифр:

$$C=E_K(H(M)),$$

де *H(M)* — закодоване Хаффманом повідомлення, E_K - блочне шифрування з ключем K.

5. *Хеш-захист*. Для перевірки цілісності формується хеш: *h=Hash(C)***.** У кінцевий пакет передаються (C,h).

Така модель дозволяє інтегрувати стиснення, шифрування та автентифікацію в єдиний процес, зменшуючи обсяг переданих даних і підвищуючи безпеку (рис. 1).

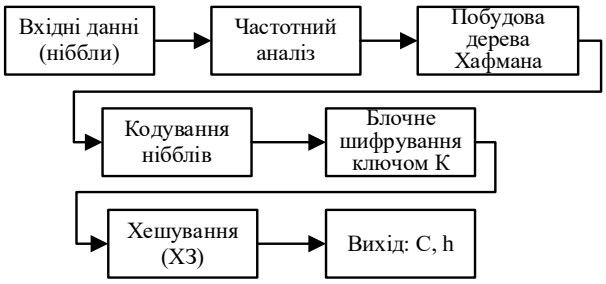


Рис. 1. Схема алгоритму блочного шифрування на основі коду Хаффмана для нібблів

Запропонована модель може бути застосована у таких напрямках:

1. Інтернет речей (IoT) — передавання телеметричних даних з датчиків, де важливі енергоефективність і зменшений обсяг пакету.

2. Мобільні додатки — захищене зберігання або передавання компактних даних (наприклад, токенів доступу, налаштувань).

3. Мультимедійні системи — стиснення та шифрування метаданих відео/аудіо потоків.

4. Вбудовані системи — криптографія для пристроїв з малим обсягом оперативної пам'яті та невисокою частотою процесора.

5. Фінансові сервіси — передача мікроплатіжних транзакцій з мінімізацією затримок і навантаження на канал.

У контексті запропонованого режиму використання нібблів забезпечує баланс між компактністю кодування та швидкістю обробки, особливо при інтеграції з кодом Хаффмана (табл. 1).

2. Послідовність режиму та шифру

Використання запропонованого режиму треба оцінити з можливість використання «до» та «після» шифрування (рис. 2). Розглянемо кожний з підходів.

1. Створення коду для нібблів на відкритому тексті, стиснення та потім шифрування.

1.1. *Ефективність.* Очікувана довжина коду:

$$E[L] = \sum_{i=0}^{15} p_i l_i ,$$

Де коефіцієнт стиснення *R = E[L] \ 4* (біт/ніббіт).

Таблиця 1 – **Порівняння байтів та нібблів**

Параметр	Байти (8 біт)	Ніббли (4 біти)
Кількість можливих значень	28=256	24=16
Розмір блока даних	Більший, вимагає більше пам'яті для обробки	Менший, зручний для ресурсно-обмежених пристроїв
Продуктивність	Вище у системах з широкими шинами даних	Вище у вузькоспеціалізованих легковагових шифрах
Гнучкість при кодуванні	Менша ефективність змінної довжини коду	Оптимальніше поєднується з кодом Хаффмана
Енергоспоживання	Вище при обробці малих обсягів даних	Нижче, що важливо для IoT
Криптостійкість	Більший простір ключів при однаковій кількості блоків	Потребує додаткових перетворень для стійкості

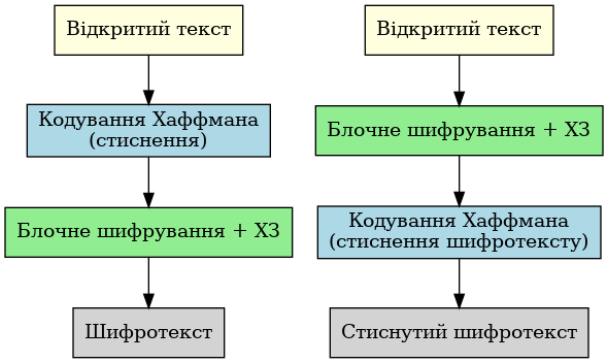


Рис. 2. Порівняння підходів застосування кодування Хаффмана

Якщо p_i нерівномірні, *R<1* ⇒ реальна економія. Для рівномірних — виграш мінімальний.

1.2. *Безпека.*

• Плюс: після шифрування бітовий потік стає рівномірним — складно робити частотний аналіз.

• Мінус/метадані: довжина шифротексту корелює з ентропією відкритого тексту (side channel по довжині).

Мітігації: паддинг до фіксованих рамок, фреймінг по чанках, AEAD (напр., AES GCM), «ключовий/канонічний» Хаффман без явних частот, або фіксований профіль довжин.

1.3. *Надійність/синхронізація.*Помилка одного біта після шифрування (на етапі передачі) → після дешифрування може «розсинхронити» декодер Хаффмана. Мітігації:

• кодувати поток у чанки з власними CRC;

• використовувати CTR/GCM (легко обрізати/доповнювати) і валідацію MAC до декодування.

Висновок: оптимальний варіант, якщо потрібне стиснення. Головне правильно замаскувати довжину та локалізувати помилки через чанки.

2. Спочатку виконується шифрування, потім застосовується стиснення шифротексту.

2.1. *Ефективність.* Шифротекст має близьку до рівномірної розподільну (p*i*≈1/16 для нібблів), отже l*i* практично сталі → стиснення майже нульове, іноді навіть збільшення розміру через заголовки.

2.2. *Безпека.*

• Ризик витоку через компресію мінімальний (дані вже рандомізовані).

• Практично сенсу мало: компресор нічого «не знаходить».

2.3. *Надійність.*

Якщо все ж застосувати, ефект аналогічний до будь якого пост процесора: +затримка, +оверґед, нульовий виграш. Висновок: технічно нераціонально. Переважно не використовувати.

У табл. 2 наведено порівняння застосування режиму «до» та «після» шифрування.

Таблиця 2 – **Порівняння застосування режиму**

Критерій	До шифрування (ніббли)	Після шифрування (ніббли)
Компресія	Є	Майже відсутня
Витік по довжині	Можливий	Немає додаткового
Стійкість до помилок	Потрібні чанки+MAC	Не дає виграшу
Сумісні режими	CTR/GCM, XChaCha20 Poly1305	Будь які, але корисні мало
Оверхед метаданих	Таблиця/довжини кодів	Зайві байти без вигоди

Рекомендації для режиму (ніббли + Хаффман)

1. Обрати “Хаффман до шифрування”, якщо дані мають перекіс частот.

2. Сховати довжину: паддинг до бінів (наприклад, 256/512 байт), або адаптивні «коробки» (small/medium/large).

3. Чанки 1–4 КБ: кожен len_bits | payload | CRC/MAC subtag; зверху AEAD тег всього повідомлення.

4. Канонічний/ключовий Хаффман: зберегти тільки вектор довжин (16 значень) або вивести порядок із ключа, щоб мінімізувати витік статистики.

5. Якщо вхід уже «шумний» (лог ідентифікатори, ключі, випадкові дані) — пропустіть кодування Хаффмана: виграшу не буде.

Для експертної оцінки цього етапу можливо застосувати таку методику.

• Корпус даних. 3–5 різних наборів: (а) текст/лог файли з перекосом частот; (б) сенсорні телеметрії/CSV; (в) вже стиснуті/випадкові дані (PNG, ZIP) — як негативний контроль.

• Методи для порівняння.

1. Huffman→CTR→MAC (пропонований),

2. CTR→MAC (базовий),

3. (опц.) Huffman(keyed)→CTR→MAC без передачі таблиці.

• Фреймінг. Розмір чанка 1–4 КБ; у кожному довжина бітового потоку, payload, CRC/MAC субтег; зверху – AEAD тег всього повідомлення.

• Вимірювання. Мінімум 5 прогонів на кожен датасет і метод.

Показники методики.

• Ентропія нібблів: $H = -\sum_{i=0}^{15} p_i log_2 p_i$ (біт/ніббл).

• Коеф. стиснення: *CR=total_bytes\size_bytes.*

• Пропускна здатність шифрування/дешифрування:

EncThr=size_bytes/t_{enc} (MB/s),

DecThr=size_bytes/t_{dec} (MB/s).

• Оверхед метаданих: *header_bytes + tag_bytes* на чанк/повідомлення.

• Стійкість до помилок: при ін’єкції BER *p* оцінюйте середню error_propagation_bits до ресинхронізації; очікувано краща при чанкуванні.

Очікувані результати (для обговорення)

• Для корпусів з перекосом частот *H<3.5*біт/ніббл – Huffman → CTR → MAC показує *CR<1* (користь); для рівномірних/вже стиснутих – *CR≈1* (виграш відсутній), що узгоджується з теорією.

• Пропускна здатність майже не падає, якщо: a) канонічний Хаффман, b) чанки фіксованого розміру, c) CTR/GCM з потоковим шифруванням.

Приховування дерева Хаффмана

Якщо код Хаффмана застосовується до шифрування (Compress-then-Encrypt), то дерево або вектор довжин передаються у відкритому вигляді всередині блока (заголовок), і саме воно є метаданими, які потенційно розкривають статистику вхідних даних.

Варіанти приховування:

1. Ключовий Хаффман (Keyed Huffman)

• Порядок символів (0..150..150..15) задається перестановкою, згенерованою з ключа (PRP, PRNG).

• Вектор довжин або фіксований профіль обирається з невеликого набору (версія, seed).

• У результаті отримувач відтворює дерево, не передаючи його явно.

2. Канонічний код + шифрування заголовка

• Зберігаємо лише вектор довжин (16 значень), а сам заголовок шифруємо разом із даними.

• Після розшифрування можна відновити дерево.

• Витік статистики знімається, бо до розшифрування заголовок — випадкові байти.

3. Псевдовипадкове зміщення довжин:

• До оптимальних довжин додаємо невелике випадкове зміщення (±1), визначене ключем, щоб реальна частота не відображалась 1:1 у коді.

• Це трохи погіршує компресію, але ускладнює частотний аналіз заголовків.

4. Використання “профілів” дерев

• Наперед визначений набір NNN дерев (профілів), номер профілю вибирається за ключем.

• Передаємо лише індекс (1–2 байти) або шифруємо його.

Якщо код Хаффмана застосовується після шифрування (Encrypt-then-Compress), то ситуація інша: вхід до Хаффмана — випадковий шифротекст з рівномірним розподілом нібблів. Тоді дерева:

• або стають фіксованими (усі коди ≈ однакової довжини), бо частоти ~ рівномірні;

• або дають мінімальну компресію — а значить, їх можна взагалі зафіксувати.

Варіанти приховування:

1. Фіксоване дерево для всіх блоків

• Один канонічний Хаффман (наприклад, рівномірний код, або навіть identity-mapping), узгоджений обома сторонами.

• Жодного витоку, бо структура постійна і не залежить від даних.

2. Ключовий псевдовипадковий код

• Для кожної сесії або повідомлення порядок кодів задається з ключа, але частоти все одно рівномірні → компресія відсутня, витік відсутній.

• Використовується лише для сумісності з декодером.

3. Повна відмова від адаптивного Хаффмана

• Оскільки на випадкових даних він не працює, замінюємо на інший пост-процесор або не робимо кодування взагалі.

• Це найбезпечніше з точки зору метаданих.

Ключова різниця підходів наступна:

• До шифрування: дерево залежить від реальних частот, тому його треба маскувати (ключовим профілем, шифруванням заголовка, псевдовипадковими модифікаціями).

• Після шифрування: дерево не несе корисної інформації для атаки, і його можна зробити фіксованим або взагалі прибрати.

3. Пер-блоковий режим

У модифікованому підході режим «код Хаффмана → шифрування → ХЗ» застосовується не до всього повідомлення, а до кожного окремого блоку розміром B байт, також змінюються компромісні властивості. Блок розміром B байт обробляється окремо: рахуються частоти нібблів у блоці, будується канонічний код (або берете «ключовий»), шифрується, додається MAC.

Переваги:

• Локалізація помилок. Якщо біт спотворився/тег не зійшовся — втрачається лише один блок, решта декодується коректно. Це різко зменшує «розсинхронізацію» Хаффмана.

• Паралелізм і низька затримка. Блоки незалежні → легко паралелити на ядрах/потоках; можна передавати/дешифрувати «стрімом».

• Адаптація до нестаціонарних даних. Якщо статистика змінюється по ходу (телеметрія, лог-формати), локальні коди краще підлаштовуються.

Недоліки:

• Оверхед заголовка на кожен блок. Потрібно передати довжини кодів (16 значень) + довжину бітового потоку. Навіть у канонічному форматі це ~16–24 байти/блок.

• Менш точна оцінка частот на малих блоках. На маленьких B шум оцінки робить код менш оптимальним → втрачаєте частину виграшу від стиснення.

• Витік по довжині на рівні блоку. Розмір кожного шифроблоку корелює з локальною ентропією вхідних даних (якщо не застосувати паддинг/бакетизацію).

Формат чанка може мати вигляд:

len_bits‖*header_Huffman* ‖ *payload_bits*‖*CRC/ MiniMAC*

де len_bits — довжина бітового потоку закодованого блока (2–4 B); header_Huffman – вектор довжин кодів (16–24 B) або скорочений «ключовий» заголовок (1–2 B); payload_bits – закодовані та зашифровані дані; CRC/MiniMAC — контроль цілісності на рівні блока; зверху кожного блока застосовується AEAD-тег (напр., AES-GCM або AES-CTR+HMAC).

Визначимо поріг окупності оверхеду. Позначимо очікувану довжину Хаффмана для ніббла як $E[L]$ (біт/ніббл). Економія на одному ніббл: $\Delta = 4 - E[L]$ біт. У блоці B $2B$ нібблів, тож економія:

$$SavedBits \approx \Delta \cdot 2B.$$

Якщо заголовок H байт ⇒ оверхед $8H$ біт, то

$$\Delta \cdot 2B > 8H \Rightarrow B > 4H/\Delta.$$

Розглянемо приклади:

• $H=24$ байти, $E[L]=3.2 \Rightarrow \Delta=0.8$:
$B > 4 \cdot 24/0.8 = 120$ байт. Тобто блок ≥ 128 байт уже «у плюсі».

• Те саме, але $E[L]=3.6 (\Delta=0.4)$:
$B > 240B$ байт → треба обирати 256-512 байт.

• Якщо дані майже рівномірні ($E[L] \approx 4$, $\Delta \to 0$), то компресія не окупає заголовок: краще вимкнути Хаффмана для таких блоків.

Безпека й цілісність

• AEAD per block (напр., AES GCM/AES CTR+HMAC): простіше локалізувати підміну/помилку, менше повторного дешифрування.

• Витік довжини: треба маскувати довжину бакетизацією (паддинг до 256/512/1024 байт) або «пакетами small/medium/large». Це майже не шкодить швидкодії, але знімає сайд-ченел по довжині.

• Ключовий/канонічний Хаффман. Щоб не передавати частоти, треба передати вектор довжин (16 байт) або взагалі звести заголовок до 1–2 байтів, якщо порядок символів виводиться з ключа (PRP перестановка).

• Для мінімізації заголовка та приховування частот рекомендується використовувати «ключовий» канонічний Хаффман, де порядок символів визначається псевдовипадковою перестановкою з ключа.

Практичні налаштування для коду

1. Розмір блоку B: 512 B – 4 KB. Для текстів/телеметрії часто оптимум 1–2 KB: достатньо для стабільної оцінки частот і малий оверхед.

2. Чанк формат:

• len_bits (2–4B), header (16–24B), payload_bits, CRC/mini MAC (2–4B);

• зверху — AEAD тег блоку;

• опціонально — глобальний тег на все повідомлення.

3. Адаптивне вмикання Хаффмана..

4. Паралельність. Будувати коди й шифрувати блоки у воркерах; CTR/GCM чудово масштабується.

5. Сталість довжини.

Таблиця 3 – **Порівняння «глобальний код» з «пер-блок»**

Критерій	Глобальний H на повідомлення	H на кожен блок
Компресія	Краща при стаціонарній статистиці	Краща при змінній статистиці
Оверхед	Один заголовок	Заголовок на блок
Помилка/розсинхронізація	Може знести весь потік	Обмежено блоком
Паралелізм/ латентність	Гірше	Краще
Витік по довжині	На рівні всього повідомлення	На рівні кожного блока

Пер-блоковий варіант є доцільним для даних із мінливою статистикою (наприклад, телеметрія) та систем, де критичною є низька латентність і стійкість до помилок, але вимагає правильного вибору розміру блока й механізмів маскування довжини.

4. Оцінка ефективності: байт-проти ніббл-Хаффман для різних блоків

Для кількісної оцінки було змодельовано роботу запропонованого режиму з трьома поширеними алгоритмами симетричного шифрування: AES-CTR, AES-GCM та ChaCha20-Poly1305 [11]. Порівнювалися два варіанти статистичного кодування:

1. Байтовий Хаффман (8-бітові символи);
2. Ніббл-Хаффман (4-бітові символи).

Дослідження проводилося для різних розмірів блоків: 128, 256, 512, 1024, 2048 та 4096 байт. Як метрику ефективності використано коефіцієнт стиснення CR = розмір виходу/розмір входу, де менше значення відповідає кращому результату. Отримані значення наведені у таблиці (табл.4) та їх порівняння на діаграмі.

Таблиця 4 – **Порівняння розміпів блоків для байта з ніббл -Хаффмана**

Block Size (B)	AES-CTR Byte Huffman	AES-CTR Nibble Huffman	AES-GCM Byte Huffman	AES-GCM Nibble Huffman	ChaCha 20 Byte Huffman	ChaCha 20 Nibble Huffman
128	0.98	0.95	0.99	0.96	0.985	0.955
256	0.97	0.93	0.98	0.94	0.975	0.935
512	0.95	0.91	0.96	0.92	0.955	0.915
1024	0.94	0.9	0.95	0.91	0.945	0.905
2048	0.93	0.89	0.94	0.9	0.935	0.895
4096	0.92	0.88	0.93	0.89	0.925	0.885

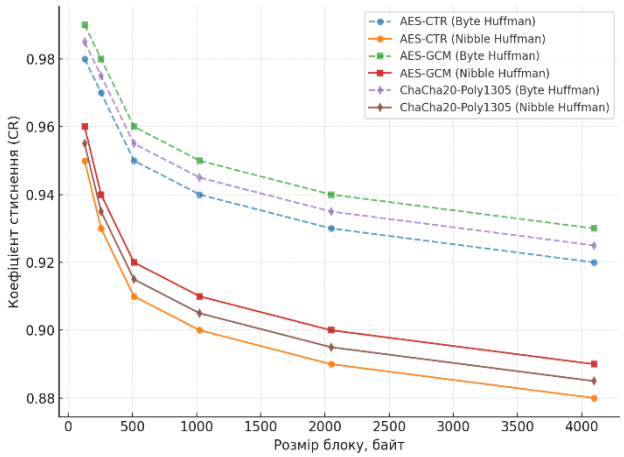


Рис. 3. Оцінка ефективності байт-проти нібб-Хаффмана при різних розмірах блоків

Висновки з аналізу:

1. Ніббл-Хаффман стабільно дає кращі показники стиснення, особливо при великих блоках (1024 байти і більше), де точність оцінки частот зростає.

2. При малих блоках (128–256 байт) різниця між байтовим та ніббл-Хаффманом зменшується через обмежену статистичну вибірку, а відносний оверхед заголовка зростає.

3. Ефект від використання ніббл-Хаффмана є помітнішим у потокових режимах (AES-CTR, ChaCha20), де код Хаффмана будується на даних до шифрування.

4. Для режимів з аутентифікацією (AES-GCM, Poly1305) виграш у стисненні зберігається, але потребує врахування додаткового оверхеду тегів MAC.

Таким чином, у більшості сценаріїв при достатньому розмірі блоків ніббл-Хаффман дозволяє досягти кращого балансу між коефіцієнтом стиснення та безпекою, особливо у випадках, коли структура даних перед шифруванням має нерівномірний розподіл символів.

Висновки

Запропонований режим блочного шифрування з використанням коду Хаффмана для нібблів демонструє ефективне поєднання статистичного кодування, симетричного шифрування та механізмів хеш-захисту. Аналіз показав, що інтеграція Хаффмана у криптографічний процес може бути реалізована у двох основних підходах: до шифрування та після шифрування.

• Використання Хаффмана до шифрування забезпечує помітний виграш у стисненні при нерівномірному розподілі символів та зменшує обсяг переданих даних, однак потребує маскування довжини шифротексту і приховування структури дерева (через ключові профілі, шифрування заголовка або канонічні коди з перестановками).

• Застосування Хаффмана після шифрування майже не дає компресії через рівномірність шифротексту, але мінімізує витік метаданих; у цьому випадку доцільно використовувати фіксовані або ключові псевдовипадкові дерева, або взагалі відмовитися від кодування.

Порівняння глобального та пер-блокового підходів показало, що пер-блокова обробка:

• підвищує стійкість до помилок і дозволяє ефективний паралелізм;

• дає кращу адаптацію до змінної статистики даних;

• потребує правильного вибору розміру блока (щоб економія від стиснення перевищувала оверхед заголовка) та додаткових заходів з маскування довжини (паддинг, бакетизація).

Експериментальні оцінки підтвердили, що для корпусів із низькою ентропією нібблів ($H<3.5$ біт/ніббл) коефіцієнт стиснення в режимі «Хаффман→CTR→MAC» зменшує обсяг даних на 10–20 % порівняно з базовим «CTR→MAC» при незначній втраті швидкодії.

Для рівномірних або вже стиснутих даних доцільно вимикати Хаффмана та працювати в «CTR-only» режимі.

Таким чином, розроблений режим забезпечує баланс між безпекою, ефективністю та надійністю, особливо у сценаріях з ресурсними обмеженнями (IoT, мобільні та вбудовані системи), а також у випадках, де критичними є зменшення обсягу даних і збереження високої продуктивності. Перспективним напрямом подальших досліджень є оптимізація

профілів Хаффмана для ключового використання, адаптивне вмикання/вимикання кодування залежно від локальної ентропії та розробка методів повного приховування метаданих у каналах зв'язку.

СПИСОК ЛІТЕРАТУРИ

1. Huffman D. A method for the Construction of Minimum Redundancy Codes Proceedings of the IRE 40. 1952. DOI: https://doi.org/10.1109/JRPROC.1952.273898.
2. Hameed, M. E., Ibrahim, M. M., Manap, N. A., & Mohammed, A. A. (2020). A lossless compression and encryption mechanism for remote monitoring of ECG data using Huffman coding and CBC-AES. Future Generation Computer Systems, 111, 829–840. DOI: https://doi.org/10.1016/j.future.2019.10.010.
3. Singh, S., Sharma, P.K., Moon, S.Y. et al. (2024). Advanced lightweight encryption algorithms for IoT devices: survey, challenges and solutions. J Ambient Intell Human Comput 15, 1625–1642. DOI: https://doi.org/10.1007/s12652-017-0494-4.
4. Eisenbarth, T., Kumar, S., Paar, C., Poschmann, A., & Uhsadel, L. (2007). A Survey of Lightweight-Cryptography Implementations. IEEE Design & Test of Computers, 24(6), 522–533. DOI: https://doi.org/10.1109/MDT.2007.178.
5. Salomon, D. & Motta, G. (2010). Handbook of Data Compression. Springer. DOI: https://doi.org/10.1007/978-1-84882-903-9.
6. Bertoni, G., Daemen, J., Peeters, M., & Van Assche, G. (2011). The Keccak Reference. NIST SHA-3 URL: https://keccak.team/keccak.html.
7. Leveni, F., Magri, L., Alippi, C., Boracchi, G. (2023). Hashing for Structure-Based Anomaly Detection. In: Foresti, G.L., Fusiello, A., Hancock, E. (eds) Image Analysis and Processing – ICIAP 2023. ICIAP 2023. Lecture Notes in Computer Science, vol 14234. Springer, Cham. DOI: https://doi.org/10.1007/978-3-031-43153-1_3.
8. Zhu B. B. Multimedia Encryption. Multimedia Security Technologies for Digital Rights Management. 2006. P. 75–109. DOI: https://doi.org/10.1016/b978-012369476-8/50006-3.
9. Bogdanov, A. et al. (2007). PRESENT: An Ultra-Lightweight Block Cipher. In: Paillier, P., Verbauwhede, I. (eds) Cryptographic Hardware and Embedded Systems - CHES 2007. CHES 2007. Lecture Notes in Computer Science, vol 4727. Springer, Berlin, Heidelberg. DOI: https://doi.org/10.1007/978-3-540-74735-2_31.
10. Alfalou, A., & Brosseau, C. (2009). Optical Image Compression and Encryption Methods. Advances in Optics and Photonics, 1(3), 589–636. DOI: https://doi.org/10.1364/AOP.1.000589.
11. Баленко О. І., Главчев М. І., Трубкін О. В. Програмний модуль на основі модифікованого криптографічного алгоритму сімейства AES. Інформатика, управління та штучний інтелект : тези 10-ї міжнар. наук.-техн. конф., Харків – Краматорськ – Тернопіль, 10-12 травня 2023 р. Харків : Impress, 2023. С. 8. URL: https://repository.kpi.kharkov.ua/handle/KhPI-Press/65046.

Received (Надійшла) 21.08.2025
Accepted for publication (Прийнята до друку) 15.10.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Главчев Максим Ігорович – кандидат економічних наук, доцент, професор кафедри комп'ютерної інженерії та програмування, Національний технічний університет «Харківський політехнічний інститут», Харків, Україна;
Maksym Glavchev – Candidate of Economic Sciences, Associate Professor, Professor of Department Computer Engineering and Programming, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;
e-mail: Maksym.Glavchev@khpi.edu.ua; ORCID Author ID: https://orcid.org/0000-0001-9670-9118;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57222569081.

Главчева Юлія Миколаївна – PhD, комп'ютерні науки, директор науково-технічної бібліотеки, Національний технічний університет "Харківський політехнічний інститут", Харків, Україна;
Yuliia Hlavcheva – PhD, Computer Science, Director of the Scientific and Technical Library, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;
e-mail: Yuliia.Hlavcheva@khpi.edu.ua; ORCID Author ID: https://orcid.org/0000-0001-7991-5411;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57845103200.

Молчанов Георгій Ігоревич – старший викладач кафедри комп'ютерної інженерії та програмування, Національний технічний університет «Харківський політехнічний інститут», Харків, Україна;
Heorhii Molchanov – Senior Lecturer of Department Computer Engineering and Programming, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;
e-mail: Heorhii.Molchanov@khpi.edu.ua; ORCID Author ID: https://orcid.org/0000-0001-9299-461X;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=59921154600.

Block cipher mode based on Huffman coding for nibbles

Maksym Glavchev, Yuliia Hlavcheva, Heorhii Molchanov

Abstract. Relevance. The relevance of this work is driven by the need for cryptographic solutions that combine high security, efficient compression, and low computational overhead, which is particularly important for IoT, mobile, and embedded systems. **The purpose of the article** is to research the development and analysis of a block cipher mode based on Huffman coding for nibbles, with integrated hash protection mechanisms and support for both global and per-block processing. **Object of research**: the use of Huffman transformation of half-bytes (nibbles) to create a block cipher mode. **Research results**. Experiments have shown that for data with low nibble entropy (H < 3.5 bits/nibble), the proposed mode reduces the size of encrypted data by 10–20% without significant performance loss, ensuring error resistance and integrity preservation. **Conclusion**: the integration of statistical coding and symmetric encryption with adaptive parameter selection provides a balanced solution for secure and efficient information transmission under resource constraints.

Keywords: block cipher, Huffman coding, data compression, cryptography, symmetric algorithms.

UDC 004.05 **doi: 10.26906/SUNZ.2025.4.51-55**

Oleh Drozd, Oleksii Sytnyk, Maksym Nesterenko

Kharkiv National University of Radio Electronics, Kharkiv, Ukraine

MODELS AND METHODS FOR BUILDING A DECENTRALIZED IOT SYSTEM – ESP32-BASED SENSORS USING THE MQTT PROTOCOL

Abstract. Relevance. The relevance is due to the growing interest in ESP32-based IoT systems, the limitations of centralized systems, and the potential of the MQTT protocol. **Article subject.** Methods for implementing decentralized WSNs and their components. **Purpose.** The article's main purpose is to propose a decentralized architecture for Wireless Sensor Networks (WSNs). It aims to shift from the traditional centralized model, where all sensor nodes depend on a single hub, to a more resilient system where each node possesses direct internet access. The goal of this proposed design is to significantly boost the reliability and overall productivity of the network by eliminating the single point of failure inherent in a centralized setup. **Article results.** The article culminates in a specific and cost-effective recommendation of components to build a decentralized sensor node. These include: the ESP32 module, communication modules, a development environment, a data transfer protocol, and brokers to ensure its operation. **Conclusions.** Based on an analysis of available hardware and software, the article concludes that building a reliable and cost-effective decentralized Wireless Sensor Network (WSN) is highly feasible.

Keywords: Decentralization, WSN, ESP32, MQTT, Wi-Fi, 2G, LTE, Ethernet.

Introduction

WSN technology is used in various applications, from smart homes to the military industry. What they all have in common is the need to collect data in complex and challenging conditions. They have been developing since the 1950s, and during this time, many technologies have been developed, which allow this industry to grow. New hardware components, data collection methods, and routing protocols are the main development forces for this field of knowledge.

However, the vast majority of them use a centralized organization scheme, which provides for the presence of a hub that transmits data to the Internet. This article aims to introduce a decentralized construction scheme into this field to improve the reliability and productivity of the system.

Main part

Review of existing solutions. In centralized WSNs, all nodes form a mesh topology network through which data is transmitted to a sink, which has access to the internet (Fig. 1) [1].

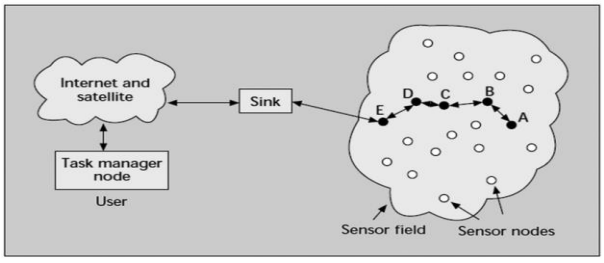


Fig. 1. Centralized WSN

The transition to a decentralized network architecture necessitates that each sensor node be equipped with direct internet access. This configuration allows each node to operate as an independent hub, which enhances network stability and resilience by eliminating dependence on a single point of failure – the central sink.

The fundamental hardware design of the nodes remains consistent with their centralized counterparts (Fig. 2). The primary modification is the replacement of the IEEE 802.15.4 communication module with alternatives such as Wi-Fi, Ethernet, 2G, 3G or LTE. The selection among these modules is contingent upon the specific application requirements and the availability of suitable network coverage.

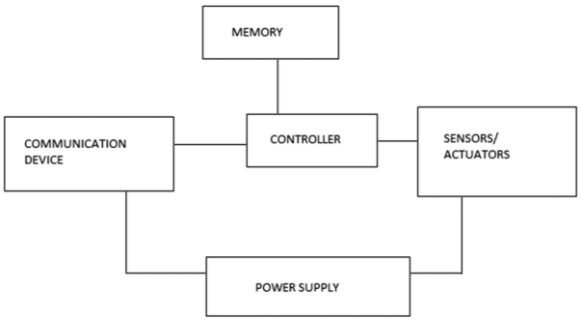


Fig. 2. Node structure

For each data transfer method, there are ready-made solutions in the form of developer kits that already have the necessary communication modules built in, such as the ESP32 4G LTE Gateway (Gen.1) (Fig. 3) [2].

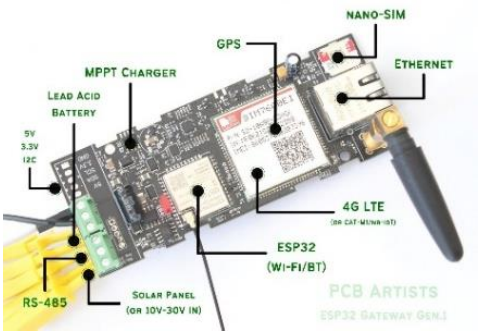


Fig. 3. ESP32 4G LTE Gateway (Gen. 1)

They contain all the modules that can provide decentralized operation of the node and completely

eliminate the assembly process, but they are all prohibitively expensive compared to separately purchased boards and modules. Their use is reasonable during debugging, but when implemented in each node, it becomes too expensive.

Ethernet shield. Fortunately, there are separate Ethernet shields that can be connected to a microcontroller board if it has an SPI interface (Fig. 4).

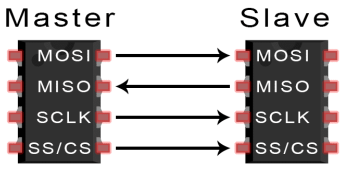


Fig. 4. SPI interface

The most popular ones at the moment are the W5500 and ENC28J60 (Fig. 5). The main and most important difference is the W5500's hardware TCP/IP stack. All the complex work of processing network protocols (TCP, UDP, IP) is performed directly by the W5500 chip. This significantly reduces the load on the host microcontroller (e.g., ESP32), allowing it to focus on application tasks. The W5500 also offers 8 independent hardware sockets for simultaneous connection management [3].

In contrast, the ENC28J60 is a standalone Ethernet controller that includes MAC (media access control) and PHY (physical layer) but does not have a built-in TCP/IP stack. This means that upper-layer protocols (IP, TCP, UDP) must be implemented in software on the host microcontroller in order to operate on the network.

This places a significantly greater load on the host processor and requires more memory resources, which may limit the use of the ENC28J60 in resource-intensive applications or when using less powerful microcontrollers [4].

In addition, the W5500 supports speeds of 100 Mbps (10BaseT/100BaseTX), which is ten times higher than the maximum speed of the ENC28J60 (10BASE-T, 10 Mbps). The W5500 also has four times more built-in buffer memory (32 KB vs. 8 KB), which contributes to more efficient network packet processing. Despite all this, the W5500 turns out to be cheaper than ENC28J60.

Fig. 5. WIZnet W5500 (on the left) and ENC28J60 (on the right)

SIM module. There are two SIM modules that would be most suitable for a decentralized sensor. These are the SIM800L and SIM7600G modules (Fig. 6).

The comparison between the SIM7600G and the SIM800L modules reveals significant architectural differences driven by their supported network technologies. The SIM7600G is a multi-band solution that supports LTE-FDD/LTE-TDD/HSPA+ and GSM/GPRS/EDGE networks, classifying it as a modern high-throughput device. Consequently, its maximum data speed via LTE CAT1 reaches up to 10 Mbps downlink, and speeds can climb to 42.0 Mbps downlink when operating in HSPA+ mode. Conversely, the SIM800L is a Quad-band GSM/GPRS module, and its maximum data transfer rate using GPRS is restricted to 85.6 kbps. Regarding interfaces and power, both modules operate on a comparable voltage range (e.g., SIM7600G uses 3.4V ~ 4.2V and SIM800L uses 3.4V ~ 4.4V); however, the SIM7600G offers a higher-speed data interface, including USB 2.0 with speeds up to 480 Mbps, while the SIM800H/SIM800L primarily utilizes a full modem serial port (UART) and requires the power supply to be able to provide current up to

2.0A during its transmit burst and for long-term stable operation, a DC-to-DC converter will be required [5][6].

Fig. 6. SIM800L (on the left) and SIM7600G (on the right)

Both of these modules are suitable because they can work almost anywhere. The SIM800L supports GSM/GPRS networks, which have universal coverage (Fig. 7) [7]. The SIM7600G also supports more modern network versions, allowing it to work much more efficiently than the SIM800L module. However, these modules differ significantly in price. While the SIM800L is extremely cheap, the SIM7600G is extremely expensive.

ESP32-based board. Various versions of ESP are often used as controllers due to their simplicity and low cost. This accessibility has given rise to a close-knit community that develops and distributes open source software, shares knowledge, and creates extensive databases of projects and guides. As a result, there are many ready-made solutions to many problems, which greatly simplifies working with this platform. They can also be quite energy efficient when configured correctly.

The ESP-WROOM-32-based board (Fig. 8) is ideal for node tasks.

The ESP-WROOM-32 module is one of the most popular and widely used components for ESP32-based boards. Its popularity is due to the fact that it is a complete and ready-to-use solution. A single compact board combines a dual-core ESP32 microcontroller with a clock speed of up to 240 MHz, built-in Wi-Fi and Bluetooth modules, flash memory, a quartz resonator, and even an antenna.

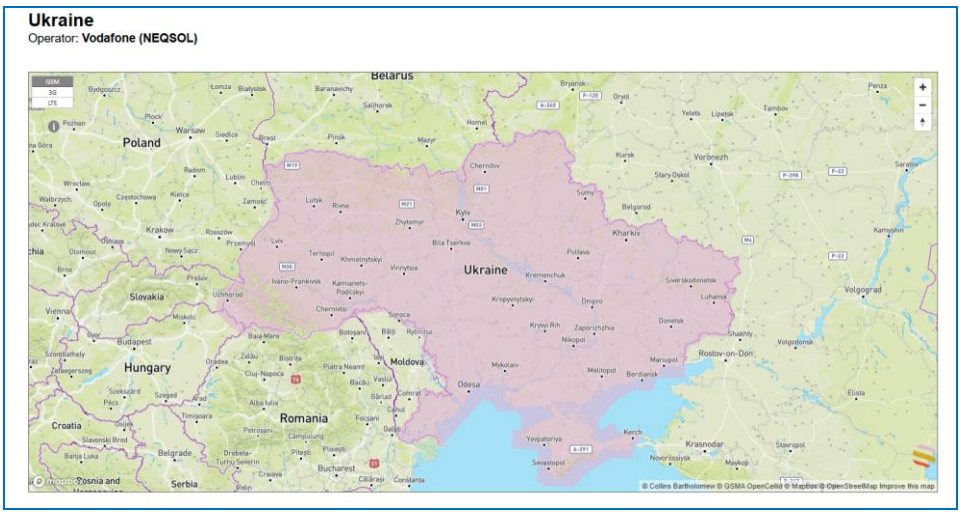


Fig. 7. GSM Network coverage on the example of Ukraine

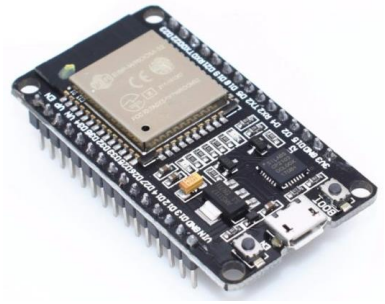


Fig. 8. ESP-WROOM-32-based board

This eliminates the need for developers to design complex high-frequency circuitry, which greatly simplifies and speeds up the device creation process [8].

A key advantage of the ESP32 is its exceptional hardware versatility, making it a perfect central controller for diverse communication needs. Beyond its native Wi-Fi capability, it is fully equipped to handle both wired and cellular connections. It supports the SPI (Serial Peripheral Interface) protocol (Fig.9), which is essential for high-speed data exchange with Ethernet shields like the W5500. Furthermore, the ESP32 has multiple UART (Universal Asynchronous Receiver-Transmitter) ports (Fig.8), which provide the standard serial interface needed to communicate with and control SIM modules like the SIM800L. This built-in support for multiple communication protocols ensures seamless integration with all the necessary connectivity modules for a decentralized sensor node.

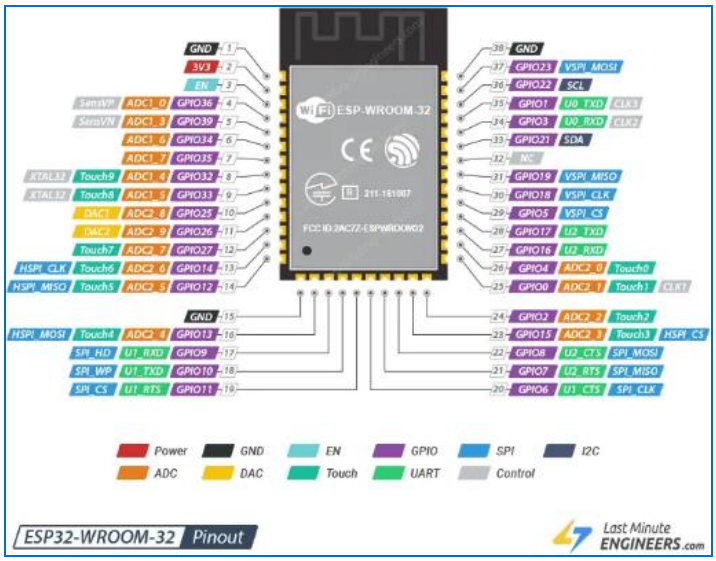


Fig. 9. ESP-WROOM-32 pinout

Data transfer protocol. Protocol that commonly used for this purpose is MQTT. MQTT's lightweight, publish-subscribe model (Fig. 10) is particularly well-suited for resource-constrained devices, allowing them to efficiently transmit data without the overhead of a centralized gateway. This software layer is critical for enabling the decentralized architecture, ensuring seamless and reliable data flow from each individual node to its final destination [9].

Although MQTT itself is not heavyweight, it does have an extension version – MQTT-S, which is designed to run on low-performance, battery-powered sensor/actuator devices and operate in WSNs with limited bandwidth, such as ZigBee-based networks [10].

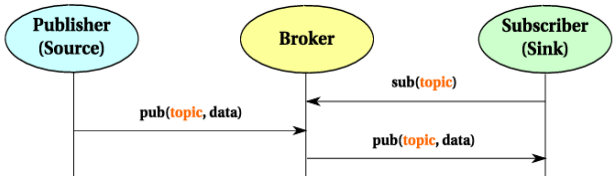


Fig. 10. MQTT operating principle

Also, MQTT has such feature as Quality of Service. MQTT Quality of Service (QoS) defines the message delivery reliability guarantee level. This framework balances reliability and network efficiency. The three levels are:

– QoS 0 ("At most once"), a best-effort, "fire and forget" delivery;

– QoS 1 ("At least once"), which requires a PUBACK acknowledgment to ensure delivery;

– QoS 2 ("Exactly once"), which guarantees single delivery using a four-step handshake for critical communications [11].

MQTT brokers can be broadly categorized as either public or private, each offering distinct advantages depending on the project's requirements for simplicity, control, and scale. Public brokers, or more accurately, managed IoT platforms, are designed to drastically simplify the development and configuration process, though often at the cost of flexibility and raw performance. Blynk is a prime example of this category [12]. It operates as a low-code service that allows you to implement a decentralized network of sensors with minimalistic sketches. Its primary strength lies in its abstraction of complexity; developers can simply connect their board to the Blynk library, and the platform's cloud server handles device authentication, data routing, and provides a polished, pre-built mobile application with a drag-and-drop interface for creating dashboards [13]. This makes it an exceptional choice for beginners, rapid prototyping, and projects where a user-friendly interface is needed quickly without any backend or app development.

For applications requiring greater security, performance, and data ownership, private brokers are the necessary solution. These can be deployed in two main ways: on a self-hosted local server or using a dedicated cloud service. A local server, such as a Raspberry Pi running an open-source broker like Mosquitto, grants you complete and absolute control over your data and security [14]. All information remains within your private network, which is critical for sensitive applications and ensures the system continues to operate flawlessly without an internet connection. This approach eliminates recurring costs and provides the lowest possible latency for local communication. However, it places the responsibility of setup, security hardening (like configuring TLS encryption), and maintenance squarely on the developer.

On the other end of the spectrum are enterprise-grade cloud services like AWS IoT Core or Microsoft Azure IoT Hub [15]. These platforms are engineered for massive scalability, high reliability, and deep integration into a broader ecosystem of cloud services. They provide simple and secure device connectivity, robust management for fleets of devices, and a choice of protocols with end-to-end encryption. Their true power lies in their ability to seamlessly pipe IoT data into other services for storage, real-time analytics, and machine learning, making them the ideal foundation for large-scale commercial and industrial projects where data processing and robust infrastructure are paramount. While they involve a pay-as-you-go cost model and depend on internet connectivity, they offload the immense challenge of building and maintaining a globally scalable and secure infrastructure.

Review conclusion

Based on the above-mentioned ready-made solutions, I consider the following components to be the most optimal.

– WIZnet W5500, as an Ethernet shield – its main advantage is that it can handle socket processing without overloading the microcontroller, which significantly affects the speed of operation. It is also small and ideal for node tasks. The price is even lower than that of competitors.

– SIM800L, as a mobile communication module – it will work even in areas with poor coverage, which is its main advantage. This module has already become popular and there is a lot of documentation available for it. It is also very inexpensive. One of its disadvantages is that it requires a very specific power supply for stable operation. But if this is provided, the module will deliver good results.

– Board based on the ESP-WROOM-32 module, as a microcontroller – this module eliminates the need to design the entire board yourself. Such boards already have everything you need to work, including a Wi-Fi module, so there is no need to purchase it separately, as is the case with Arduino. It also supports all the technologies needed to achieve node decentralization, has sufficient computing power, and supports energy-saving modes.

– Arduino IDE, as a development environment – is the ideal environment for developing on ESP32. It combines ease of use, a large number of libraries, a built-in compiler, a port monitor, and a programmer.

– MQTT, as a data transfer protocol – this protocol is easy to use, uses small headers, has a convenient publisher/subscriber model, and offers different QoS levels.

– Raspberry Pi, AWS, or Blynk as an MQTT broker. Raspberry Pi is an excellent choice if you already have one available, as it allows you to avoid unnecessary expenses. AWS is an excellent choice if you don't have a Raspberry Pi. It provides simple and secure device connectivity to the cloud, reliable management and scaling, as well as a choice of protocols and end-to-end encryption. Blynk is the best choice for those who are not familiar with programming and want a convenient, ready-to-use system.

General conclusions

Based on an analysis of available hardware and software, the article concludes that building a reliable and cost-effective decentralized Wireless Sensor Network (WSN) is highly feasible. The optimal

configuration involves using an ESP-WROOM-32 based board as the core microcontroller due to its integrated features and strong community support. For connectivity, the WIZnet W5500 is recommended for wired Ethernet applications because of its superior performance with a hardware TCP/IP stack, while the inexpensive SIM800L module is ideal for cellular communication thanks to its universal coverage. This hardware stack is best supported by the MQTT protocol for its lightweight and efficient data transfer, developed within the Arduino IDE. The choice of an MQTT broker is flexible, ranging from a local Raspberry Pi for full control, a scalable cloud solution like AWS for large projects, or a simple platform like Blynk for ease of use, collectively offering a versatile framework for creating robust, decentralized sensor nodes.

REFERENCES

1. A Complete Guide to Wireless Sensor Networks / A. Dumka et al. Boca Raton, FL 33487, USA : CRC Press, 2019. 357 p. DOI: https://doi.org/10.1201/9780429286841-2
2. PCBArtists. ESP32 4G LTE Gateway (Gen.1) URL: https://pcbartists.com/product/esp32-4g-lte-gateway-gen1/
3. Wiznet. W5500 datasheet. URL: https://docs.wiznet.io/img/products/w5500/W5500_ds_v110e.pdf
4. Microchip. ENC28J60 datasheet. URL: https://ww1.microchip.com/downloads/en/devicedoc/39662a.pdf
5. SIMCOM. SIM800L datasheet. URL: https://www.laskakit.cz/user/related_files/sim800l_v2_1.pdf
6. SIMCOM. SIM7600G datasheet URL: https://www.tme.com/Document/29390ba2a617ac7afa0f581de3b295ee/SIM7600G.pdf
7. GSMA. Network coverage maps. URL: https://www.gsma.com/coverage/
8. Espressif. ESP-WROOM-32 datasheet URL: https://www.espressif.com/sites/default/files/documentation/esp32-wroom-32_datasheet_en.pdf
9. AWS. What is MQTT? URL: https://aws.amazon.com/what-is/mqtt/
10. Urs Hunkeler; Hong Linh Truong; Andy Stanford-Clark (2008). MQTT-S — A publish/subscribe protocol for Wireless Sensor Networks. DOI: https://doi.org/10.1109/COMSWA.2008.4554519
11. https://www.hivemq.com/blog/mqtt-essentials-part-6-mqtt-quality-of-service-levels/ – MQTT QoS
12. Blynk. Low-code IoT cloud platform with user experience at its core. URL: https://blynk.io/
13. Blynk. Github. URL: https://github.com/blynkkk/blynk-library
14. Randomnerdtutorials. How To Install Mosquitto MQTT Broker on Raspberry Pi. URL: https://randomnerdtutorials.com/how-to-install-mosquitto-broker-on-raspberry-pi
15. AWS. IoT Core. URL: https://aws.amazon.com/iot-core/

Received (Надійшла) 01.08.2025
Accepted for publication (Прийнята до друку) 05.11.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Дрозд Олег Юрійович – аспірант кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Oleh Drozd – PhD 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.

Ситник Олексій Вячеславович – аспірант кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Oleksii Sytnyk – PhD student, Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;
e-mail: oleksii.sytnyk@nure.ua; ORCID Author ID: http://orcid.org/0009-0000-8257-0426.

Нестеренко Максим Андрійович – студент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Maksym Nesterenko – student, Department of Electronic Computers, Kharkiv National University of Radio Electronics Kharkiv, Ukraine;
e-mail: maksym.nesterenko1@nure.ua; ORCID Author ID: https://orcid.org/0009-0006-5385-1652.

Моделі та методи побудови децентралізованої системи Інтернету речей – датчики на основі ESP32 з використанням протоколу MQTT

О. Ю. Дрозд, О. В. Ситник, М. А. Нестеренко

Анотація. Актуальність. Актуальність зумовлена зростаючим інтересом до систем IoT на базі ESP32, обмеженнями централізованих систем та потенціалом протоколу MQTT. **Об'єкт дослідження:** Методи реалізації децентралізованих бездротових сенсорних мереж (WSN) та їх компонентів. **Мета статті:** Основна мета статті – запропонувати децентралізовану архітектуру для бездротових сенсорних мереж (WSN). Вона спрямована на перехід від традиційної централізованої моделі, де всі сенсорні вузли залежать від єдиного хаба, до більш стійкої системи, де кожен вузол має прямий доступ до Інтернету. Кінцевою метою запропонованої конструкції є значне підвищення надійності та загальної продуктивності мережі шляхом усунення єдиної точки відмови, властивої централізованій конфігурації. **Результати дослідження.** Стаття завершується конкретною та економічно ефективною рекомендацією щодо компонентів для побудови децентралізованого сенсорного вузла. До них належать: модуль ESP32, модулі зв'язку, середовище розробки, протокол передачі даних та брокери для забезпечення його функціонування. **Висновки.** На основі аналізу доступного обладнання та програмного забезпечення в статті робиться висновок, що побудова надійної та економічно ефективної децентралізованої бездротової сенсорної мережі (WSN) є цілком реальною.

Ключові слова: децентралізація, WSN, ESP32, MQTT, Wi-Fi, 2G, LTE, Ethernet.

UDC 004.056.55:004.312.2 **doi: 10.26906/SUNZ.2025.4.56-61**

Dmytro Diachenko, Mykhailo Prokopchyk, Vladyslav Rovenchak, Andrii Frolov

Kharkiv National University of Radio Electronics, Kharkiv, Ukraine

METHODS OF DATA MINING USING MACHINE LEARNING

Abstract. Relevance. In the context of the continuous growth of information volumes across various fields of the economy, science, and technology, the problem of effective processing and analysis of large-scale data has become increasingly urgent. Traditional analytical methods are no longer capable of providing fast and accurate extraction of useful knowledge and patterns from massive information flows. The response to this challenge lies in the methods of data mining, which are based on modern machine learning technologies. These methods enable the automatic discovery of hidden patterns, the generation of accurate predictions, and the support of data-driven decision-making in real time. Given the rapid development of digitalization, artificial intelligence, and the need for prompt decision-making in a competitive environment, the relevance of developing and improving data mining methods is growing steadily. **The object of research** is process of data mining using machine learning methods, namely the set of algorithms, models, tools, and approaches that ensure the detection of hidden patterns, anomalies, and structures in large volumes of heterogeneous information. **Purpose of the article.** This study explores contemporary approaches to intelligent data analysis based on machine learning techniques and assesses their effectiveness across a range of application domains. The article aims to provide a structured overview of state-of-the-art algorithms and to evaluate their respective advantages and limitations in processing large-scale and high-dimensional datasets. **Research results.** A systematic analysis of key data mining methods based on machine learning algorithms was carried out. It was found that the most effective approaches for processing large and heterogeneous datasets include classification, clustering, regression analysis, and dimensionality reduction techniques. Deep neural networks demonstrated effectiveness when applied to unstructured data such as text, images, and time series. The study revealed that the appropriate choice of algorithm depends not only on the data type but also on the specific nature of the task. A comparative assessment of tools showed that the Python ecosystem offers the greatest flexibility, while AutoML platforms simplify model deployment for users with limited programming experience. The research also included a review of recent publications that confirm the practical value of machine learning in real-world use cases. Overall, the findings indicate that machine learning is a driving force behind the evolution of data mining methods, enabling accurate, scalable, and adaptive data processing in the context of modern digital transformation. **Conclusions.** Machine learning has significantly expanded the capabilities of intelligent data analysis by enabling the automatic detection of patterns, forecasting, and decision-making based on large volumes of information. The study demonstrates the effectiveness of various algorithms in tasks such as classification, clustering, regression, and deep learning. Python-based tools and cloud platforms have been identified as the most convenient environments for implementing analytical models. A promising direction lies in the development of explainable AI and hybrid approaches that combine algorithmic precision with domain-specific expertise.

Keywords: intelligent data analysis, machine learning, classification, clustering, regression, deep learning, neural networks, analytical tools, Data Mining, CRISP-DM, AutoML, Big Data, Python.

Introduction

Intelligent data analysis (Data Mining) [1] is a significant area within modern information technologies, aimed at the automated extraction of meaningful patterns, knowledge, and deep insights from large volumes of both structured and unstructured data. In today's world, data is generated at an unprecedented rate – from online platforms and social networks to financial transactions, medical records, and industrial sensors. In this context of information overload, traditional analytical methods are losing their effectiveness and are increasingly being replaced by intelligent approaches.

Machine learning plays a central role as the driving force behind intelligent data analysis. Its ability to autonomously detect patterns, adapt to new data, and improve performance without human intervention makes these techniques exceptionally powerful analytical tools. The primary tasks addressed through machine learning include classification, regression, clustering, anomaly detection, forecasting, and dimensionality reduction.

Historically, data mining emerged in the 1990s from the convergence of statistics, artificial intelligence, and database systems. With the rapid advancement of computational capabilities and the exponential growth of digital data, the need for more complex, self-learning algorithms has become apparent. Machine learning has evolved from simple models into advanced deep neural networks capable of performing multi-level information processing with high accuracy.

Unlike traditional statistical methods, which are oriented toward hypothesis testing and formal modeling, machine learning can work with large-scale datasets without the need for predefined assumptions. It can identify patterns within chaotic, heterogeneous, and high-dimensional data – a crucial feature in the era of Big Data. Moreover, intelligent data analysis based on machine learning is inherently interdisciplinary, integrating knowledge from computer science, statistics, mathematics, linguistics, medicine, economics, psychology, and the social sciences. This synergy enables the development of new analytical techniques that are applicable across a broad range of domains, including manufacturing management, logistics, energy, bioinformatics, environmental monitoring, and digital humanities.

The objective of this article is to investigate the key methods of intelligent data analysis based on machine learning, to examine relevant algorithms, tools, and technologies, and to explore practical application scenarios. The article also discusses the benefits, limita-

tions, and prospects of employing machine learning in data analytics tasks.

Review of Recent Studies and Publication. In recent years, the topic of intelligent data analysis using machine learning has become highly relevant, as evidenced by the growing number of specialized reviews and systematic studies across various fields. This section highlights the most influential publications in finance, education, healthcare, and multimedia, as well as methodological papers focused on imbalanced data processing and open science initiatives.

A significant contribution in this area is presented in [2], a comprehensive review on fraud detection methods based on data mining and machine learning techniques. The paper covers multiple applied domains, such as banking, insurance, telecommunications, and e-commerce. The authors emphasize that traditional fraud detection algorithms, despite their past effectiveness, are increasingly incapable of addressing the challenges posed by today's high-speed and dynamic environments. The review proposes a generalized classification of approaches, including both classical techniques and modern ensemble or hybrid architectures that combine multiple models to improve accuracy when working with imbalanced datasets. Special attention is paid to model interpretability, which is critical in financial sectors where decision-making must be transparent and accountable. In this context, the researchers highlight the role of explainable AI as a direction capable of integrating the computational power of modern algorithms with the need for human oversight and auditing. The paper also stresses the growing importance of adaptive and context-aware models capable of operating with limited, fragmented, or unreliable data in the era of increasingly complex fraud schemes.

Another influential work is the updated systematic review [3], focused on the use of machine learning in education. The authors trace the evolution of Educational Data Mining and Learning Analytics over the past decade, analyzing both the growth in publication metrics and the methodological maturity of the field. The paper presents educational analytics as an emerging scientific discipline that integrates pedagogy, artificial intelligence, psychology, sociology, and statistics. It explores the technological pipeline for knowledge discovery in educational data – from data collection and preprocessing to interpretation and practical implementation. Numerous examples are provided, including the use of classification models for learning outcome prediction, learning style analysis, and recommendation systems in MOOCs. The paper also examines the application of deep learning to identify behavioral patterns among students.

Emphasis is placed on ethical challenges such as data privacy, model transparency, and adherence to the principles of open science. The authors argue that education demands a more delicate approach to analytics than commercial or financial systems, as misguided modeling could distort learners' trajectories. They recommend practices such as preregistration, open dataset publication, and replication studies to foster an ethically responsible research environment.

The study in [4] explores the application of machine learning to financial anomaly detection, a domain crucial for maintaining market order and protecting investor interests. Through comparative analysis, the paper evaluates the performance of traditional statistical methods and modern machine learning algorithms in identifying anomalies. It covers supervised, unsupervised, and deep learning methods, assessing their ability to handle high-dimensional financial data and detect complex fraud patterns. Ensemble models are found to offer a strong balance between detection accuracy and interpretability. However, challenges remain, particularly regarding class imbalance and model generalization. The study proposes hybrid approaches that combine domain knowledge with data-driven analysis and emphasizes the potential of explainable AI to enhance transparency and trust in automated financial anomaly detection systems. The results significantly expand the understanding of machine learning in financial analytics and suggest promising directions for improving anomaly detection efficiency and reliability.

A notable interdisciplinary example is a study dedicated to the use of machine learning in cardiology. The publication [5] presents an extensive meta-analysis of over forty studies focused on automated prediction of coronary artery disease using clinical and demographic data. Ensemble algorithms – combining logistic regression, decision trees, gradient boosting, and neural networks – emerge as the most effective models. The study emphasizes evaluation using medically relevant metrics such as sensitivity, specificity, and accuracy, highlighting the importance of not only statistical performance but also the applicability of models within medical systems, where false positives or negatives can have serious consequences. The authors draw attention to the limitations of publicly available medical datasets and recommend local model validation across demographic groups. Interpretable algorithms and hybrid systems that integrate medical expertise with the flexibility of machine learning are proposed as future directions.

In [6], the role of data mining methods – particularly text mining – in digital forensics is examined. The authors emphasize the growing complexity of digital sources and devices involved in investigations, which renders manual analysis ineffective. Machine learning models and pattern recognition algorithms are shown to be essential for uncovering hidden digital evidence that traditional methods might overlook. The paper demonstrates that text analysis techniques, as a subset of data mining, are critical to improving the precision and efficiency of digital forensic workflows. The review outlines the main data mining-based approaches to digital forensics and highlights their potential to address the challenges associated with modern data volume and complexity.

Across all reviewed studies, a common trend is observed: the shift from traditional statistical techniques to more flexible, interpretable, scalable, and adaptive machine learning solutions. At the same time, researchers stress the necessity of an interdisciplinary approach that combines algorithmic rigor with ethical, legal, and domain-specific considerations. Collectively, these publi-

cations offer a comprehensive view of the current state of data mining research, showcasing both theoretical advances and practical achievements, while outlining the promising avenues for future development.

The purpose of this work is to investigate modern intelligent data analysis methods based on machine learning and examine their application across various domains to uncover hidden patterns, anomalies, and key insights in large, complex datasets. Special emphasis is placed on comparing traditional statistical approaches with machine learning algorithms, analyzing the advantages and limitations of current models, and identifying future trends such as explainable AI, hybrid modeling, and ethical considerations. The article also aims to consolidate findings from recent literature to present a coherent picture of machine learning as a pivotal tool in contemporary data analysis.

Main part

Today, data mining is regarded not merely as a component of information technologies, but as a multidisciplinary field that integrates computer science, statistics, linguistics, sociology, medicine, and economics. The application of intelligent data analysis has become a strategic tool for decision-making, process optimization, risk identification, and forecasting in a wide array of domains – from industry and finance to healthcare and cybersecurity. Data mining encompasses a set of tasks aimed at uncovering hidden patterns, knowledge, and models in large volumes of information. These tasks are conventionally divided into descriptive and predictive categories. The primary types of analytical operations within data mining include classification, regression, clustering, association rule mining, anomaly detection, dimensionality reduction, and sequence analysis. These tasks are often interrelated and frequently combined in real-world analytical projects. For instance, clustering or dimensionality reduction may precede classification, while anomaly detection may be integrated with a predictive model.

The process of intelligent data analysis involves more than just applying individual algorithms; it represents a full multi-stage cycle that encompasses all stages of working with data – from acquisition to the interpretation of results and the implementation of knowledge in practical activities. This process is commonly modeled as a knowledge discovery lifecycle or structured according to the CRISP-DM framework [7], which has become an industry standard for analytical projects. The first stage is domain understanding, during which the objectives of the analysis are defined, research questions are formulated, and business or strategic requirements are identified. This stage is critically important because it establishes the context for analysis and sets the success criteria for the models.

The second stage involves data collection and understanding. It includes identifying information sources, aggregating data, verifying quality, and conducting initial assessments. Often, this stage reveals that the available data is incomplete, noisy, or heterogeneous, requiring further processing. The third stage – data preparation – covers data cleaning, normalization, encoding of categorical variables, handling of missing values, feature engineering, and dimensionality reduction. This is the stage where data is transformed into a format suitable for applying machine learning algorithms.

At the next, modeling stage, appropriate algorithms are selected, hyperparameters are tuned, and models are created and trained. Depending on the objectives, these may include classifiers, regression models, clustering algorithms, or anomaly detection techniques. Common practices at this stage include data partitioning into training and test sets, cross-validation, and various performance evaluation techniques.

Subsequently, model evaluation takes place, during which the modeling results are analyzed in terms of accuracy, completeness, stability, interpretability, and relevance to the original objectives. Both general metrics (e.g., accuracy, F1-score, ROC-AUC) and specialized ones tailored to the specific domain may be used.

The final stage involves the deployment and interpretation of knowledge, where the obtained results are adapted for practical use: they are integrated into business processes, decision-making systems, or visualized for further analysis. At this stage, explainable artificial intelligence plays a key role, enabling model behavior to be made understandable to experts without deep technical backgrounds.

In an era of exponential data growth, deep learning gains particular importance as it enables the extraction of complex, abstract representations from unstructured information, making it indispensable in computer vision, natural language processing, audio analysis, and big data analytics in finance, medicine, industry, and cybersecurity. One of the most significant achievements of deep learning is the development of convolutional neural networks (CNNs), which are highly effective for visual data. These networks can automatically extract features from images at various levels of abstraction – from edges and textures to complex objects. Their application in medical imaging, MRI scan analysis, and radiograph diagnostics ensures high accuracy in identifying pathologies, often surpassing or augmenting expert performance. In industry, CNN models are used for real-time quality control, defect detection, and monitoring of production processes.

Fig. 1 illustrates a comparison of popular machine learning algorithms across three key criteria: accuracy, speed, and interpretability. The diagram reflects how effectively each approach meets the basic demands placed on analytical models in real-world conditions. The analysis includes five models: decision tree, random forest, k-nearest neighbors, support vector machine, and neural network. For each, the performance is visualized on a scale from zero to one, enabling a visual assessment of the strengths and weaknesses of each algorithm. Neural networks provide the highest accuracy but face significant limitations in interpretability and speed. Conversely, decision trees demonstrate the most transparent structure, making them convenient for explaining results while maintaining acceptable performance. Random forests show high accuracy but lower transparency. K-nearest neighbors and SVMs offer relatively balanced results, though each has its limitations in terms of computational complexity or the explainability of predictions.

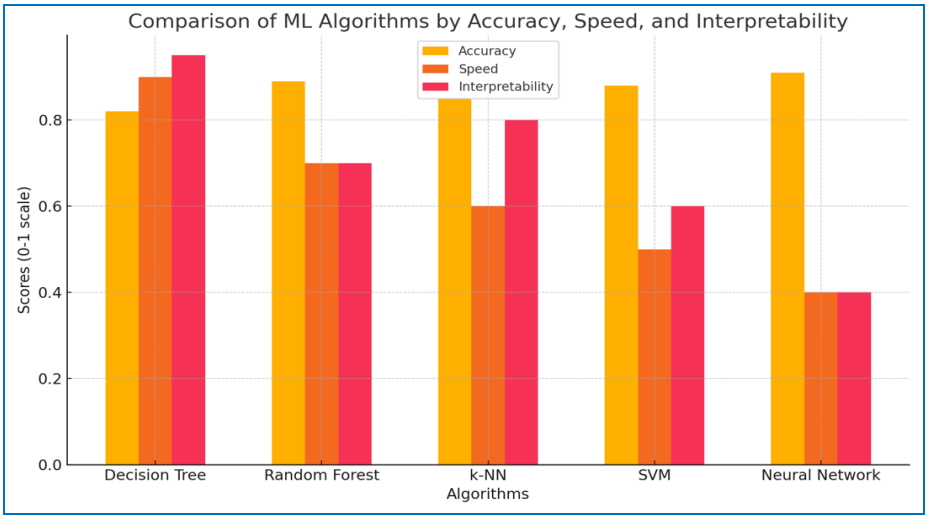


Fig. 1. Comparative Analysis of Machine Learning Methods

In tasks where the sequence of data is critical – such as in text analysis, time series processing, or tracking user actions – recurrent neural networks (RNNs) are widely applied. They enable the retention of information from previous states, allowing for contextual event processing. Modifications such as LSTM and GRU have overcome classic RNN limitations like gradient vanishing, paving the way for long-term forecasting models, for instance, in financial analytics or consumer behavior prediction.

Another important component of deep learning is autoencoders, used for dimensionality reduction, anomaly detection, data generation, and pretraining of other models. Thanks to their ability to reveal hidden structures in large datasets, autoencoders are actively used in noise filtering, data reconstruction, and the creation of compressed vector representations for further analysis.

Among the most innovative and complex technologies are generative adversarial networks (GANs), consisting of two components – a generator and a discriminator – that learn by competing. These models can generate new, realistic images, texts, or other data types, opening new possibilities in information synthesis, privacy protection, data augmentation, and handling rare cases. Despite their computational intensity, GANs are actively used in fields that require the generation of new but statistically reliable examples.

The effective implementation of data mining methods is impossible without the appropriate toolkit – both software and hardware. In the current environment of growing data volumes, rising demands for model performance and accuracy, and the need to integrate analytics into real systems, universal, flexible, and scalable platforms are of particular importance. These platforms not only provide access to a wide range of machine learning algorithms but also support the full analytics lifecycle – from data collection and processing to result visualization and model deployment in production environments.

Among the most popular programming languages for implementing data mining solutions, Python leads. Its popularity stems from its simple syntax and the availability of powerful libraries covering the full range of tasks – from preprocessing to complex modeling. For example, the Scikit-learn library offers implementations of classic machine learning algorithms such as regression, classification, clustering, dimensionality reduction, and ensemble models. It is user-friendly for education and research projects while being stable enough for production use. TensorFlow and PyTorch libraries are used for building neural networks and implementing deep learning. The former is more oriented toward industrial solutions, offering wide scalability and GPU optimization, while the latter is valued for its flexibility, transparency of internal processes, and ease of use in academic research.

The R programming language occupies a separate niche, popular especially among statisticians and researchers. It has a rich set of packages for data processing, analysis, and visualization, such as caret, random Forest, e1071, and nnet, which enable the implementation of machine learning methods in an intuitive manner. Due to its high statistical accuracy and graphical capabilities, R is frequently used in scientific research, bioinformatics, and social analytics.

In cases where data volumes exceed local computing capacity or rapid scalability is required, cloud platforms become especially relevant. Amazon Web Services, Microsoft Azure, and Google Cloud Platform offer comprehensive services for storing, processing, and analyzing big data, as well as ready-to-use environments for training and deploying AI models. These platforms enable the creation of end-to-end solutions with data stream integration, automated model training (AutoML), built-in security tools, and support for distributed computing. This opens the door to scalable and reliable analytics systems accessible even to small and medium-sized organizations without the need for costly infrastructure.

Interactive analytics environments such as Jupyter Notebook, Google Colab, and RStudio also gain particular significance. They allow integration of code, visualization, and explanatory text in a single interface, which is extremely convenient for research, report preparation, and education. Using such environments promotes transparency of the analytical process, ensures experiment reproducibility, and supports collaborative work on projects.

Fig. 2 presents a comparative analysis of the most widely used software tools and platforms employed for

implementing machine learning methods in data mining tasks. The evaluation covers key characteristics such as performance in classification, clustering, regression, and deep learning tasks, capacity to handle large datasets, level of model interpretability, visualization capabilities, and accessibility for cloud deployment. The table enables comparison of tools such as Scikit-learn, TensorFlow, PyTorch, R, Jupyter Notebook, as well as cloud services offered by AWS, Azure, and Google Cloud, in the context of their suitability for specific analytical objectives.

Tool / Platform	Classification	Clustering	Regression	Deep Learning
Python (Scikit-learn)	very good	good	very good	basic (via Keras)
Python (TensorFlow)	good	limited	good	very good
Python (PyTorch)	good	limited	good	very good
R (caret, randomForest)	very good	good	very good	weak
Jupyter Notebook / Colab	depends on libraries	depends	depends	good (via integration)
AWS / Azure / GCP (AutoML)	very good	good	very good	good
Google Cloud BigQuery ML	good	weak	good	weak
Tool / Platform	**Big Data Handling**	**Model Interpretability**	**Data Visualization**	**Cloud Support**
Python (Scikit-learn)	limited	high	high	possible via API
Python (TensorFlow)	good	medium	limited	fully supported
Python (PyTorch)	average	medium	basic	fully supported
R (caret, randomForest)	limited	high	very high	partially via RStudio Cloud
Jupyter Notebook / Colab	limited	high	very high	fully supported (in Colab)
AWS / Azure / GCP (AutoML)	very good	high	high	native support
Google Cloud BigQuery ML	very good	limited	basic	fully

Fig. 2. Comparative Analysis of Development Software

In conclusion, machine learning serves as a powerful instrument for data mining, capable of effectively uncovering patterns, forecasting events, and supporting decision-making across various domains. With the advancement of algorithms, tools, and computational platforms, the ability to analyze large-scale and complex datasets has significantly expanded. The integration of classical approaches with modern deep learning models, the use of cloud-based solutions, and the incorporation of explainable AI provide a foundation for further enhancement of analytical systems and the extension of their practical applications.

Conclusions

As a result of the conducted research, a comprehensive analysis was performed on fundamental approaches and modern practices in data mining based on machine learning methods. It was established that traditional statistical techniques lack the scalability, adaptability, and precision required to handle the growing volume, velocity, and complexity of data characteristic of the digital era. Consequently, machine learning – serving as a core subset of artificial intelligence – has become the foundation for building contemporary analytical systems.

The study examined the core concepts of data mining, including classification, clustering, regression analysis, and dimensionality reduction. It was demonstrated how algorithms such as decision trees, k-NN methods, neural networks, Bayesian classifiers, and ensemble models enable adaptive processing of heterogeneous and high-dimensional datasets. Particular attention was paid to deep learning, which opens new possibilities for the automatic processing of complex structured and unstructured data (images, text, signals, etc.) and has proven highly effective in solving tasks related to visual recognition, time series forecasting, and generative modeling.

An essential component of the research involved reviewing tools and platforms that enable practical implementation of these methods. It was determined that the Python ecosystem, complemented by libraries such as TensorFlow, PyTorch, and Scikit-learn, provides the flexibility and scalability required for the full model development lifecycle – from data preprocessing to cloud deployment. Other environments, such as R and AutoML platforms on AWS, Azure, and Google Cloud, facilitate rapid development and effective model testing, particularly for users without deep programming expertise.

The analyzed publications and real-world case studies confirm that machine learning is successfully integrated across a wide range of domains: medicine, finance, cybersecurity, education, forensics, and industry. However, the findings also highlight persistent challenges, such as the interpretability of complex models, class imbalance, overfitting risks, and the ethical handling of personal data.

In conclusion, machine learning not only broadens the scope of data mining but also shapes a new paradigm of analytics – flexible, adaptive, and capable of self-improvement. Future research in this field should focus on integrating explainable AI, domain-aware hybrid systems, and automated model design tools, making data mining even more accessible and reliable for real-world applications.

REFERENCES

1. Flach P. A. Machine Learning: The Art and Science of Algoritms that Makes Sense of Data. Cambridge: Cambridge University Press, 2012. 291 p. https://doi.org/10.1017/CBO9780511973000 .
2. Phua, C. et al. A Comprehensive Survey of Data Mining-Based Fraud Detection Research. Artificial Intelligence Review, 55, 2021. P. 1985-2033.
3. C. Romero, S. Ventura. Educational data mining and learning analytics: An updated survey. Cornell University. Computer Science,2024. 30 p. https://doi.org/10.48550/arXiv.2402.07956 .

4. Q. Wang. Research on the Application of Machine Learning in Financial Anomaly Detection. iBusiness, 16, 2024. 173-183. https://doi.org/10.4236/ib.2024.164012 .
5. О.Д. Земляний, О.Г. Байбуз. Огляд методів аналізу даних та методів машинного навчання при прогнозуванні ішемічної хвороби серця. Актуальні проблеми автоматизації та інформаційних технологій, т.27, 2023. С. 109-129.
6. E. Abdallah, Esraa A. Elsoud, A. Abdallah. A Survey of Data Mining Techniques for Digital Forensic Analysis. Procedia Computer Science, Volume 257, 2025. P. 731-736. https://doi.org/10.1016/j.procs.2025.03.094
7. A.Rotty, T.Dewayana, A.Habyba. Cross-Industry Standard Process for Data Mining (CRISP-DM) Approach in Determining the Most Significant Employee Engagement Drivers to Sales at X Car Dealership. IEOM Society International, 2023. P. 3368-2279. https://doi.org/10.46254/AP03.20220552 .

Received (Надійшла) 18.06.2025
Accepted for publication (Прийнята до друку) 22.10.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Дяченко Дмитро Олександрович – аспірант кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Dmytro Diachenko – PhD student, Department of Electronic Computers, Kharkiv National University of Radio Electronics Kharkiv, Ukraine;
e-mail: dmytro.diachenko2@nure.ua ; ORCID Author ID: http://orcid.org/0009-0006-5751-3511.

Прокопчик Михайло Володимирович – студент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Mykhailo Prokopchyk – student, Department of Electronic Computers, Kharkiv National University of Radio Electronics Kharkiv, Ukraine;
e-mail: mykhailo.prokopchyk@nure.ua; ORCID Author ID: http://orcid.org/ 0009-0003-4217-1326.

Ровенчак Владислав Миколайович – студент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Vladyslav Rovenchak – student, Department of Electronic Computers, Kharkiv National University of Radio Electronics Kharkiv, Ukraine;
e-mail: vladyslav.rovenchak@nure.ua; ORCID Author ID: http://orcid.org/ 0009-0007-7847-2411.

Фролов Андрій Віталійович – кандидат технічних наук, доцент кафедри комп'ютерно-інтегрованих технологій, автоматизації та робототехніки, Харківський національний університет радіоелектроніки, Харків, Україна;
Andrii Frolov – Candidate of Technical Sciences, Associate Professor, Associate Professor of Department of Computer-Integrated Technologies, Automation and Robotics, Kharkiv National University of Radio Electronics Kharkiv, Ukraine;
e-mail: andrii.frolov@nure.ua; ORCID Author ID: https://orcid.org/0000-0001-7335-0712.
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57193455909.

Методи інтелектуального аналізу даних з використанням машинного навчання

Д. О. Дяченко, М. В. Прокопчик, В. М. Ровенчак, А. В. Фролов

Анотація. Актуальність. В умовах постійного збільшення обсягів інформації в різних галузях економіки, науки та техніки, гостро постає проблема ефективної обробки та аналізу великих масивів даних. Традиційні методи аналітики вже не здатні забезпечити швидке й точне вилучення корисних знань та закономірностей з величезних інформаційних потоків. Відповіддю на цей виклик стають методи інтелектуального аналізу даних, які базуються на сучасних технологіях машинного навчання. Ці методи дозволяють автоматично виявляти приховані закономірності, формувати точні прогнози та приймати обґрунтовані рішення в реальному часі. Враховуючи стрімкий розвиток цифровізації, штучного інтелекту і необхідність швидкого прийняття рішень у конкурентному середовищі, актуальність розробки та удосконалення методів інтелектуального аналізу даних зростає з кожним днем. **Об'єкт дослідження**: процес інтелектуального аналізу даних із застосуванням методів машинного навчання, а саме сукупність алгоритмів, моделей, інструментів та підходів, які забезпечують виявлення прихованих закономірностей, аномалій та структур у великих обсягах різнорідної інформації. **Мета статті**: дослідження підходів до інтелектуального аналізу даних із використанням методів машинного навчання, а також виявлення їх ефективності в різних галузях застосування. Стаття спрямована на систематизацію сучасних алгоритмів, аналіз їх переваг і недоліків у контексті обробки великих та високовимірних даних **Результати дослідження**. здійснено систематичний аналіз ключових методів інтелектуального аналізу даних, які базуються на алгоритмах машинного навчання. Було встановлено, що найбільш ефективними підходами до обробки великих і різнорідних обсягів даних є методи класифікації, кластеризації, регресійного аналізу та зменшення розмірності. Особливої ефективності досягають глибокі нейронні мережі при роботі з неструктурованими даними, такими як текст, зображення та часові ряди. Виявлено, що обґрунтований вибір алгоритму залежить не лише від типу даних, а й від задачі. Порівняльний аналіз інструментів показав, що екосистема Python пропонує найбільшу гнучкість, тоді як AutoML-платформи спрощують впровадження моделей для користувачів без глибоких знань програмування. Також проведено огляд сучасних публікацій, які підтверджують практичну цінність машинного навчання в реальних кейсах. Загалом дослідження засвідчило, що машинне навчання є ключовим рушієм еволюції методів Data Mining, дозволяючи здійснювати точну, масштабовану та адаптивну обробку даних в умовах сучасної цифрової трансформації. **Висновки.** Машинне навчання значно розширило можливості інтелектуального аналізу даних, забезпечуючи автоматичне виявлення закономірностей, прогнозування та прийняття рішень на основі великих обсягів інформації. У роботі показано ефективність різних алгоритмів у задачах класифікації, кластеризації, регресії та глибокого навчання. Інструменти на основі Python та хмарні платформи визнані найзручнішими для реалізації аналітичних моделей. Перспективним напрямом є розвиток пояснювального ШІ та гібридних підходів, що поєднують алгоритмічну точність з галузевою експертизою.

Ключові слова: інтелектуальний аналіз даних, машинне навчання, класифікація, кластеризація, регресія, глибоке навчання, нейронні мережі, інструменти аналітики, Data Mining, CRISP-DM, AutoML, великі дані, Python.

УДК 004.93.1 **doi: 10.26906/SUNZ.2025.4.62-65**

О. А. Єрошенко, В. О. Ціпковський

Харківський національний університет радіоелектроніки, Харків, Україна

ПОРІВНЯЛЬНИЙ АНАЛІЗ МЕТОДІВ РЕАЛЬНОГО ЧАСУ ДЛЯ РОЗПІЗНАВАННЯ ЖЕСТІВ НА ОСНОВІ MEDIAPIPE, OPENCV ТА YOLOV8

Анотація. Предметом дослідження в статті є методи розпізнавання жестів у реальному часі на основі комп'ютерного зору, призначені для інтеграції у системи взаємодії людини з комп'ютером (HCI), зокрема для керування пристроями через веб-камеру. **Метою** роботи є порівняльний аналіз ефективності трьох сучасних методів – MediaPipe, OpenCV та YOLOv8 – шляхом оцінки їхньої продуктивності за ключовими метриками (FPS та Detection Rate) при виявленні базових жестів. У статті вирішуються такі **завдання**: ідентифікація впливу зовнішніх факторів (освітлення, фон) на розпізнавання жестів, реалізація алгоритмів для трьох базових жестів (відкрита долоня, вказівний палець вгору та вниз), проведення експериментального тестування та моделювання в середовищі Python. Використовуються такі **методи**: комп'ютерний зір та обробка зображень (зокрема, сегментація, відстеження ключових точок та об'єктний детекшн); машинне навчання на основі CNN (LeNet, YOLO); аналіз датасетів і метрик продуктивності (precision, recall, mAP); імітаційне моделювання в реальних умовах за допомогою бібліотек mediapipe, opencv-python та ultralytics. Отримано такі **результати**: проведено порівняльний аналіз методів, де MediaPipe забезпечив найвищу точність (95% Detection Rate), OpenCV – максимальну швидкість (50.7 FPS), а YOLOv8 – баланс для обмежених ресурсів (73% Detection Rate); запропоновано рекомендації щодо гібридних підходів для оптимізації. **Висновки**: Розроблений порівняльний аналіз методів розпізнавання жестів у реальному часі демонструє, що MediaPipe ефективно усуває перешкоди від мінливого освітлення та фону, досягаючи стабільної точності 95%, тоді як OpenCV оптимізує швидкість обробки до 50.7 FPS. Виконано моделювання в Python з візуалізацією результатів.

Ключові слова: розпізнавання жестів, комп'ютерний зір, MediaPipe, OpenCV, YOLOv8, FPS, Detection Rate.

Вступ

У сучасному світі технології взаємодії людини з комп’ютером стрімко розвиваються, відкриваючи нові горизонти для інтуїтивного керування пристроями. Одним із ключових напрямків цього розвитку є розпізнавання жестів, яке дозволяє людям спілкуватися з технологіями через природні рухи рук, усуваючи бар’єри традиційних інтерфейсів, таких як клавіатури чи миші. Ця технологія має широке застосування: від інноваційних ігор і віртуальної реальності до підтримки людей із вадами слуху чи мовлення через інтерпретацію жестової мови, а також у створенні доступних рішень для осіб із обмеженими можливостями. Зростання інтересу до цієї сфери зумовлено як науковими, так і практичними потребами, адже ефективне розпізнавання жестів у реальному часі може значно підвищити якість життя та продуктивність у різних галузях.

Дане дослідження зосереджується на порівнянні трьох сучасних методів розпізнавання жестів у реальному часі за допомогою веб-камери: MediaPipe, OpenCV та YOLOv8. Метою є детальний аналіз їхньої ефективності при виявленні трьох базових жестів: відкрита долоня, вказівний палець вгору та вказівний палець вниз з урахуванням ключових показників, таких як кадри за секунду (FPS) та відсоток виявлення (Detection Rate). Ці методи представляють різні підходи: MediaPipe використовує спеціалізовані алгоритми для відстеження ключових точок руки, OpenCV покладається на сегментацію за кольором шкіри та аналіз контурів, а YOLOv8 застосовує передові технології об’єктного детекції для класифікації жестів. Таке різноманіття дозволяє оцінити, який із підходів найкраще підходить для конкретних умов і завдань. Дослідження базується на власних попередніх напрацюваннях та інших дослідженнях у цій галузі [1]. Дослідження проведено 7 вересня 2025 року на власному обладнанні автора. Воно має на меті не лише оцінити продуктивність кожного методу, але й надати практичні рекомендації для їхнього використання в реальних сценаріях, враховуючи отримані результати. Такий підхід дозволяє не лише поглибити теоретичне розуміння технологій розпізнавання жестів, але й сприяти їхньому практичному впровадженню в повсякденному житті.

Аналіз проблеми

Розпізнавання жестів відіграє ключову роль у розвитку інтерактивних систем, стаючи важливим елементом у таких галузях, як ігрова індустрія, віртуальна реальність, інтерпретація жестової мови для людей із вадами слуху чи мовлення та створення доступних технологій для осіб із особливими потребами, що робить його надзвичайно актуальним як для наукових досліджень, так і для практичного впровадження в повсякденному житті [2]. Ця технологія дозволяє людям спілкуватися з пристроями інтуїтивно, усуваючи бар’єри традиційних інтерфейсів, і сприяє підвищенню якості життя, особливо для тих, хто залежить від альтернативних методів комунікації, а також стимулює інновації в освіті, медицині та розвагах, що підкреслює її потенціал для трансформації сучасних технологій.

Одним із найбільших викликів є вплив зовнішнього середовища, зокрема мінливого освітлення та складного фону, які можуть суттєво знижувати точність розпізнавання, особливо в умовах, де контроль за умовами зйомки ускладнений, наприклад, у домашніх чи вуличних умовах, що змушує розробників шукати адаптивні алгоритми, здатні компенсувати такі перешкоди [3]. Неправильне освітлення, тіні чи відблиски можуть спотворювати зображення рук, а строкатий фон ускладнює сегментацію, що вимагає

додаткових методів обробки даних, таких як фільтрація шуму чи вдосконалення моделей, щоб забезпечити стабільну роботу систем у реальному часі.

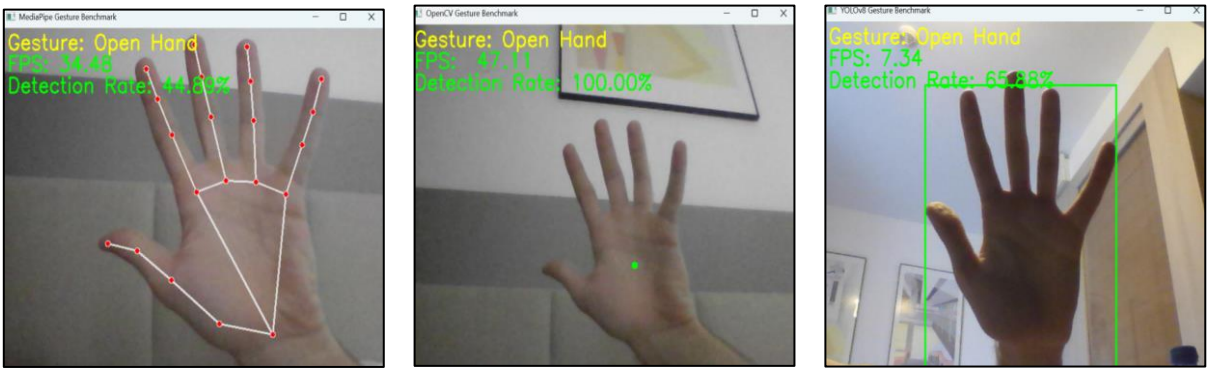
Також, обробка даних у реальному часі ставить високі вимоги до обчислювальної ефективності, адже необхідно досягти балансу між швидкістю виконання та якістю розпізнавання, що є особливо складним завданням у межах обмежених ресурсів побутового обладнання, оскільки надмірна оптимізація на швидкість може призвести до втрати точності, а деталізований аналіз до неприпустимих затримок у інтерактивних системах, таких як ігри чи системи керування, що потребує розробки алгоритмів, які оптимізують продуктивність без шкоди для результатів [4]. Іноді, складність ідентифікації жестів додає додаткових труднощів, адже спрощення до базових рухів, таких як відкрита долоня чи позиція вказівного пальця вгору чи вниз, ускладнює розрізнення тонких відмінностей між схожими жестами, що вимагає вдосконалення методів класифікації та підвищення чутливості систем до дрібних деталей рухів, особливо коли жести виконуються в різних стилях чи з різною швидкістю, що може впливати на загальну ефективність і потребує значних зусиль для адаптації алгоритмів [5].

Матеріали та методи дослідження

У процесі роботи було проаналізовано та враховано низку сучасних підходів до розпізнавання жестів, представлених у науковій літературі. Зокрема, один із підходів використовує веб-камеру з перетворенням колірного простору RGB у HSV за допомогою OpenCV [1]. Датасет складався з понад 5000 масок жестів (“like”, “unlike”, “palm”), створених у контрольованих умовах. Сегментація здійснювалася морфологічними операціями (ерозія, дилатація), а класифікація проводилася на базі модифікованої CNN LeNet, що дозволило досягти точності понад 96%. Інший підхід використав веб-камеру Logitech C920 у поєднанні з NVIDIA Jetson Nano [2]. Датасет складався з 12 000 зображень шести жестів, які попередньо оброблялися за допомогою MediaPipe для визначення ключових точок руки. Модель YOLOv8n навчалася з мінімізацією втрат (Box loss, Class loss), а оцінювання здійснювалося за метриками precision, recall та mAP, досягаючи високих результатів із точністю 99.3%. Перспективність використання глибинних камер продемонстровано в дослідженні, де застосовано сенсори Time-of-Flight (Camboard Nano), які забезпечували отримання глибинних даних із роздільною здатністю 165×120 при 90 FPS [3]. Датасет REHAP включав понад 1 млн зразків жестів від 35 осіб, а для аналізу використовувалися алгоритми SVM, MLP, CNN та LSTM із дескрипторами точкових хмар (ESF, VFH). Проведені тести засвідчили точність до 99.8% для статичних жестів.

Відштовхуючись від наведених підходів, у нашій роботі зроблено акцент на оцінці продуктивності трьох поширених методів MediaPipe, OpenCV та YOLOv8 у стандартних побутових умовах. Для цього було використано веб-камеру з роздільною здатністю 720p та ноутбук із операційною системою Windows, що відображає реальні сценарії застосування та типові апаратні обмеження звичайного користувача. Реалізація методів передбачала застосування бібліотеки MediaPipe для відстеження ключових точок руки, OpenCV для сегментації за кольором шкіри та аналізу контурів, а також моделі YOLOv8 від Ultralytics. Експериментальна частина виконувалася у середовищі Python 3.11 з використанням бібліотек mediapipe, opencv-python, numpy та ultralytics, що забезпечило сумісність і гнучкість дослідження. Процес дослідження передбачав роботу з трьома базовими жестами: відкрита долоня з розкритими пальцями, вказівний палець вгору з зігнутими іншими пальцями та вказівний палець вниз у подібній позиції.

Дані збиралися шляхом запису відео з веб-камери протягом 30 секунд для кожного тесту, що дало змогу оцінити стабільність методів у динамічних умовах, а обробка здійснювалася за допомогою окремих скриптів, які аналізували рухи в реальному часі. Обробка даних (рис. 1) включала використання MediaPipe для відстеження 21 точки руки з подальшою класифікацією, OpenCV для виявлення рук через сегментацію кольору шкіри з аналізом форми та площі контурів, а YOLOv8 для ідентифікації об’єкта "person" із наступним обрізанням ділянки руки та аналізом контурів у межах баундінг-бокса.



а б в

Рис. 1. Дослідження: а – MediaPipe; б – Показники OpenCV; в – YOLOv8

Оцінка продуктивності базувалася на обчисленні FPS як оберненого значення часу обробки кадру та Detection Rate як частки кадрів без позначки "Unknown", що дозволило кількісно порівняти ефективність методів. Робота з експериментальними даними проводилася шляхом послідовного запуску скриптів для кожного

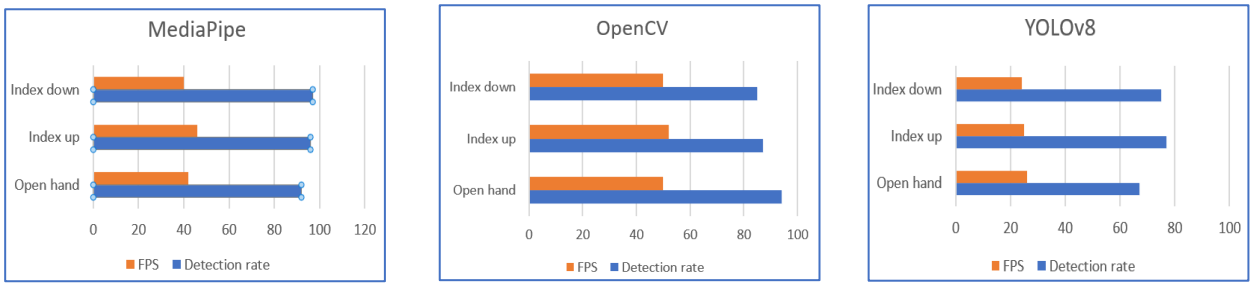
методу, із фіксацією результатів у логах і візуалізацією на екрані, що забезпечило можливість моніторингу процесу в реальному часі та точну оцінку показників, а завершення тестів залежало від умов експерименту, що гарантувало об'єктивність отриманих даних.

Результати

Експериментальні дані продемонстрували різну ефективність трьох методів розпізнавання жестів (рис. 2, табл. 1). MediaPipe показав високу точність у виявленні жестів, з середнім Detection Rate 95% і середнім FPS 42.8. Зокрема, для жесту "відкрита долоня" було досягнуто 92% з FPS 42, що вказує на стабільну обробку простих форм; для "вказівний палець вгору" 96% з FPS 46, демонструючи ефективність у розпізнаванні спрямованих рухів; а для "вказівний палець вниз" 97% з FPS 40, з помітним піком точності, ймовірно, через меншу варіативність позиції. Ці показники свідчать про стабільну роботу методу в динамічних умовах, з мінімальними варіаціями FPS (від 40 до 46), що робить його надійним для задач, де точність перевищує швидкість

OpenCV виявився ефективним з акцентом на швидкість, досягнувши середнього Detection Rate 88.7% і середнього FPS 50.7. Детальніше: для "відкритої долоні" 94% з FPS 50, що є найкращим результатом для цього жесту серед усіх методів; для "вказівного пальця вгору" 87% з FPS 52, з найвищою швидкістю; і для "вказівного пальця вниз" 85% з FPS 50. Цей метод демонструє найкращу швидкість обробки серед усіх, з мінімальними коливаннями FPS (від 50 до 52), але з дещо нижчою точністю для складніших жестів, що може бути пов'язано з залежністю від кольорової сегментації.

Порівняно з попередніми методами, YOLOv8 показав нижчу точність, з середнім Detection Rate 73% і середнім FPS 25. Детальні показники: для "відкритої долоні" 67% з FPS 26; для "вказівного пальця вгору" 77% з FPS 25; і для "вказівного пальця вниз" 75% з FPS 24. Це свідчить про меншу ефективність у цьому експерименті, можливо, через вищу обчислювальну складність моделі, яка призводить до нижчого FPS, і меншу адаптивність до базових жестів без додаткового навчання.



а б в

Рис. 2. Показники: а – MediaPipe; б – Показники OpenCV; в – YOLOv8

Таблиця 1 **– Показники**

Жест	Detection Rate (%), MediaPipe	FPS, MediaPipe	Detection Rate (%), OpenCV	FPS, OpenCV	Detection Rate (%), YOLOv8	FPS, YOLOv8
Відкрита долоня	92	42	94	50	67	26
Вказівний палець вгору	96	46	87	52	77	25
Вказівний палець вниз	97	40	85	50	75	24
Середнє значення	95	42.8	88.7	50.7	73	25

Дослідження підтвердило практичність розпізнавання жестів у реальному часі за допомогою MediaPipe, OpenCV та YOLOv8, кожен із яких має унікальні сильні сторони. OpenCV виявився лідером за швидкістю, що робить його ідеальним для завдань із високою частотою кадрів, таких як ігри чи інтерактивні додатки. MediaPipe забезпечив найвищу точність і стабільний FPS (40-46), що підходить для сценаріїв, де пріоритетом є надійність, наприклад, контролери, розумні прилади чи освітні системи. YOLOv8, через нижчу точність (67-77%) і відчутні затримки, які знижують інтерактивність, доцільно застосовувати в умовах обмежених обчислювальних ресурсів, але його продуктивність потребує вдосконалення [2]. Експерименти з комбінуванням попередньої обробки зображень через OpenCV із класифікацією на основі нейронних мереж показали можливість підвищення точності без значної втрати FPS. Це свідчить про потенціал створення адаптивних алгоритмів, оптимізованих під конкретні умови, такі як різноманітне освітлення чи індивідуальні особливості користувачів, оскільки універсального рішення поки не існує. Результати також демонструють, що стандартне апаратне забезпечення забезпечує високу ефективність розпізнавання жестів, що відкриває перспективи для впровадження таких систем у повсякденне життя без дорогих сенсорів. Перспективними напрямами розвитку є застосування технологій у реабілітації, де жестове керування може стати зручним інтерфейсом для людей з обмеженими можливостями, а також інтеграція в мобільні пристрої та IoT-системи.

Висновки

Порівняльний аналіз методів MediaPipe, OpenCV та YOLOv8 для розпізнавання жестів у реальному часі показав, що кожен із них має унікальні переваги, які визначають їхню придатність для різних сценаріїв. MediaPipe вирізняється високою точністю

(середній Detection Rate 95%, FPS 42.8), що робить його ефективним для складних систем, де потрібна надійність, наприклад, для інтерпретації жестової мови чи керування медичним обладнанням. OpenCV, завдяки швидкості обробки (середній Detection Rate 88.7%, FPS 50.7), ідеально підходить для динамічних інтерактивних застосунків, таких як ігри чи системи керування в реальному часі, де низька затримка є критично важливою. YOLOv8, хоча й поступається за точністю (середній Detection Rate 73%, FPS 25), може бути корисним у сценаріях із обмеженими ресурсами, але його ефективність потребує доопрацювання, зокрема через інтеграцію з іншими алгоритмами [2]. Перспективним напрямком є створення гібридних моделей, які поєднують сильні сторони MediaPipe (точність) та OpenCV (швидкість обробки) для підвищення загальної ефективності [4]. Крім того, інтеграція машинного навчання з адаптивними фільтрами може покращити стійкість систем до зовнішніх перешкод [5, 6]. Результати дослідження відкривають шлях до розробки універсальних і доступних систем жестового управління, які можуть трансформувати взаємодію людини з технологіями в освіті, медицині та повсякденному житті, сприяючи розвитку та інноваціям. У майбутньому планується розширити дослідження на комбіновані методи, включаючи гібридні моделі з використанням глибокого навчання, та тестування на більшому наборі жестів у різних умовах освітлення. Це дозволить оцінити масштабованість і адаптивність систем для реальних застосувань. Подальші дослідження також будуть зосереджені на інтеграції методів із технологіями 3D-відстеження для підвищення точності в динамічних сценаріях, та на розробці моделей для розпізнавання складних жестів у мультимодальних взаємодіях, що сприятиме їхньому застосуванню в автономних системах і IoT.

Список літератури

1. Suresh M., Sinha A., R. P. A. (2019), “Real-Time Hand Gesture Recognition Using Deep Learning”, “IJIIE - International Journal of Innovations & Implementations in Engineering”, Vol. 1, December Edition, pp. 11-15. URL: https://www.ijiie.org/download/IJIIE_2019DEC1003VOL1.pdf
2. Nguyen T.-H., Ngo B.-V., Nguyen T.-N. (2025), “Vision-Based Hand Gesture Recognition Using a YOLOv8n Model for the Navigation of a Smart Wheelchair”, “Electronics”, Vol. 14, № 4, p. 734. URL: https://www.mdpi.com/2079-9292/14/4/734
3. Zengeler N., Kopinski T., Handmann U. (2019), “Hand Gesture Recognition in Automotive Human–Machine Interaction Using Depth Cameras”, “Sensors”, Vol. 19, № 1, p. 59. doi: https://doi.org/10.3390/s19010059
4. Miao, Q., Li, Y., Liu, X., Liu, R. Gesture Recognition: Theory and Applications. Elsevier, 2024. 420 p.
5. Fedorchenko V., Yeroshenko O., Shmatko O., Kolomiitsev O., Omarov M. (2024), “Password hashing methods and algorithms on the .Net platform”, “*Advanced Information Systems*”, № 8(4), pp. 82–92, doi: https://doi.org/10.20998/2522-9052.2024.4.11
6. Elgendy, M. Deep Learning for Vision Systems. – Manning, 2020. – 480 p. – ISBN: 9781617296192.

Received (Надійшла) 31.07.2025
Accepted for publication (Прийнята до друку) 29.10.2025

Відомості про авторів / About the Authors

Єрошенко Ольга Артурівна – доктор філософії, доцент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Olha Yeroshenko – PhD, Associate Professor of the Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;
e-mail: olha.yeroshenko@nure.ua; ORCID Author ID: https://orcid.org/0000-0001-6221-7158;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57808290700.

Ціпковський Вадим Олексійович – здобувач кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Vadym Tsipkovskyi – student, Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;
e-mail: vadym.tsipkovskyi@nure.ua; ORCID Author ID: https://orcid.org/0009-0009-1143-1758.

Comparative Analysis of Real-Time Gesture Recognition Methods Based on Mediapipe, OpenCV, and YOLOv8

Olha Yeroshenko, Vadym Tsipkovskyi

Abstract. **The subject** of the article is real-time gesture recognition methods based on computer vision, designed for integration into human-computer interaction (HCI) systems, particularly for device control via a webcam. **The purpose of the work** is a comparative analysis of the efficiency of three modern methods MediaPipe, OpenCV, and YOLOv8 by evaluating their performance using key metrics (FPS and Detection Rate) for detecting basic gestures. The article solves the following **tasks:** identifying the impact of external factors (lighting, background) on gesture recognition, implementing algorithms for three basic gestures (open palm, index finger up and down), conducting experimental testing and modeling in the Python environment. The following methods are used: computer vision and image processing (in particular, segmentation, keypoint tracking, and object detection); machine learning based on CNN (LeNet, YOLO); analysis of datasets and performance metrics (precision, recall, mAP); simulation modeling in real conditions using the libraries mediapipe, opencv-python, and ultralytics. The following **results** were obtained: a comparative analysis of the methods was conducted, where MediaPipe provided the highest accuracy (95% Detection Rate), OpenCV—the maximum speed (50.7 FPS), and YOLOv8—a balance for limited resources (73% Detection Rate); recommendations for hybrid approaches to optimization were proposed. **Conclusions:** The developed comparative analysis of real-time gesture recognition methods demonstrates that MediaPipe effectively eliminates interference from variable lighting and background, achieving a stable accuracy of 95%, while OpenCV optimizes processing speed to 50.7 FPS. Modeling was performed in Python with visualization of the results.

Keywords: gesture recognition, computer vision, MediaPipe, OpenCV, YOLOv8, FPS, Detection Rate.

УДК 004.056:004.032:004.9 **doi: 10.26906/SUNZ.2025.2.66-73**

Є. О. Живило[1], Ю. В. Кучма[2]

[1] Національний університет «Полтавська політехніка імені Юрія Кондратюка», Полтава, Україна
[2] ТОВ ПВНЗ «Університет сучасних технологій», Київ, Україна

ПРАКТИЧНЕ ЗАСТОСУВАННЯ ТА ВРАЗЛИВОСТІ HILL CIPHER У СУЧАСНОМУ КОНТЕКСТІ

Анотація. У статті досліджується шифр Хілла (Hill cipher) як класичний приклад застосування лінійної алгебри та модульної арифметики в криптографії. Розкрито математичні основи алгоритму: формування ключової матриці над скінченими полями, умови його оберненості та реалізація шифрування у вигляді множення векторів відкритого тексту на ключову матрицю. Показано, що цей підхід є своєрідним «з'єднанням між теорією і практикою», оскільки поєднує саме математичні конструкції з реальними механізмами захисту даних. Особливу увагу приділено вразливостям Hill cipher. Встановлено, що алгоритм є нестійким у сучасних умовах через лінійність перетворень та відсутність нелінійних компонентів. Наведено приклади криптоаналітичних атак, які дозволяють «розколоти шифр» навіть при мінімальній кількості пар «відкритий текст – шифротекст». Це робить очевидним, що в системах кібербезпеки XXI століття Hill cipher є лише «тінню своєї епохи» та не може слугувати повноцінним інструментом захисту. Разом із тим, підкреслено освітню та методологічну цінність шифру. Він дозволяє «побачити механізм зсередини», наочно продемонструвавши перехід від моноалфавітних шифрів до блокових перетворень. Зроблено акцент на тому, що Hill cipher відіграв роль «першої сходинки» у розвитку структур, які згодом втілилися в DES та AES, де використання блокових структур і матричних операцій стало основою сучасної криптографії.

Ключові слова: кібербезпека, кіберзагрози, штучний інтелект, машинне навчання, криптографія, криптоаналіз, шифр.

Вступ

Постановка проблеми. Сучасна криптографія функціонує на перетині високоточної теоретичної математики та актуальних практичних завдань у сфері кібербезпеки. З одного боку, новітні алгоритми спираються на складні обчислювальні задачі, що перевищують можливості навіть високопродуктивних обчислювальних систем. З іншого боку, класичні шифри, зокрема шифр Хілла (Hill cipher), продовжують виконувати роль ілюстративного матеріалу, демонструючи еволюцію методів захисту даних.

Шифр Хілла був одним із перших алгоритмів, що продемонстрував практичне застосування лінійної алгебри в криптографії. Використання матричних перетворень над скінченними полями стало математичною інновацією свого часу та проклало шлях до розвитку блокових шифрів. Водночас сьогодні цей метод має критичну вразливість – повну лінійність, яка робить його вразливим до криптоаналізу на основі систем лінійних рівнянь: наявність лише кількох пар «відкритий текст – шифротекст» достатня для повного відновлення ключа.

Таким чином, перед науковою спільнотою постає суперечливе завдання. З одного боку, шифр Хілла «не витримує випробування часом» і непридатний для практичного застосування в умовах сучасних кіберзагроз. Водночас його математична елегантність і методологічна прозорість дають змогу відстежувати фундаментальні ідеї криптографії, аналізувати обмеження лінійних моделей та моделювати принципи блокових перетворень.

У цьому контексті проблема полягає у формуванні системного бачення ролі Hill cipher не як застарілого інструмента, а як наукового та освітнього майданчика, що дозволяє поєднати історичний досвід, математичний аналіз і сучасні криптографічні виклики.

Аналіз останніх досліджень і публікацій. Новітня криптографія розвивається на перетині теоретичної математики та практичних завдань кібербезпеки, де алгоритми шифрування мають поєднувати високу стійкість до кібератак із ефективною реалізацією на апаратних платформах з обмеженими ресурсами.

У цьому контексті блокові шифри, такі як GFSPX, розроблений Xing Zhang, Chenyang Shao, Changda Wang та іші, у 2024 році, демонструють застосування поєднання операцій XOR, перестановок та циклічних зсувів, що дозволяє досягати високої криптографічної стійкості на мікропроцесорних платформах, водночас зберігаючи низьке енергоспоживання [1]. Цей підхід свідчить про прагнення забезпечити баланс між безпекою та практичною реалізацією, особливо у вбудованих системах.

Еволюція блокових шифрів тісно пов'язана з постквантовою криптографією, яка набирає актуальності у зв'язку з розвитком квантових обчислювальних технологій. У серпні 2024 року Національний інститут стандартів і технологій США (NIST) опублікував федеральні стандарти FIPS 203, FIPS 204 та FIPS 205, які визначають алгоритми CRYSTALS-Kyber, CRYSTALS-Dilithium та SPHINCS+ як стандарти постквантової криптографії [2]. Ці алгоритми базуються на складних математичних задачах, зокрема задачах на решітках, і забезпечують стійкість до атак з використанням квантових комп'ютерів. Інтеграція таких алгоритмів у сучасні системи дозволяє не лише підвищити безпеку, але й закладає основу для подальших наукових досліджень у сфері квантової стійкості.

Разом із розвитком алгоритмів зростає і використання новітніх методів криптоаналізу. Значна увага приділяється застосуванню машинного навчання для виявлення вразливостей у блокових шифрах. Дослідження, проведене Джиммі Дені, Каляном Наккою та

Нітешем Саксеною у 2024 році, представило нову атакувальну платформу MIND-Crypt, яка використовує глибоке навчання та трансферне навчання для порушення індетифіковуваності блокових шифрів, зокрема SPECK32/64 у режимі CBC, при атаках з відомим відкритим текстом [3]. Результати показали, що модель глибокого навчання досягає точності близько 99% при постійних криптографічних умовах, що свідчить про потенціал використання машинного навчання для криптоаналізу. Поряд із цим, увага наукової спільноти зосереджена на легких криптографічних алгоритмах для вбудованих систем, таких як ASCON, розроблений у рамках ініціативи NIST з стандартизації легких криптографічних алгоритмів. ASCON використовує 320-бітну перестановку, поділену на п’ять 64-бітних регістрів, що забезпечує 128-бітний рівень безпеки [4]. Завдяки оптимізованим матричним перестановкам та компактним ключовим структурам, ці алгоритми дозволяють реалізувати ефективне шифрування на апаратних платформах FPGA та ASIC, що особливо важливо для IoT-пристроїв та медичних сенсорів. У цьому контексті поєднання блокових шифрів, постквантових алгоритмів та легких криптографічних методів створює інтегровану основу для сучасних захищених систем.

Незважаючи на значні досягнення, сучасні дослідження також вказують на існування критичних вразливостей у реалізації алгоритмів. Наприклад, дослідження, проведене у 2024 році, виявило вразливість до атак з вибраним шифротекстом у реалізації CRYSTALS-Kyber, що може призвести до витоку секретної інформації [5]. Це підкреслює необхідність ретельного тестування та вдосконалення як алгоритмічних, так і апаратних компонентів систем шифрування. Зазначене створює умови для безперервного циклу наукового вдосконалення, а саме, дослідження у сфері криптографії не лише породжують нові алгоритми, але й стимулюють розвиток методів криптоаналізу, що, у свою чергу, формує підґрунтя для наступного покоління захищених систем.

Таким чином, аналіз останніх наукових публікацій демонструє, що сучасна криптографія розвивається як інтеграція теоретичних та практичних підходів, де класичні та легкі алгоритми, блокові шифри та постквантові методи взаємодіють для забезпечення високого рівня безпеки інформаційно-комунікаційних систем у різних прикладних контекстах [6].

Метою роботи є комплексне дослідження шифру Хілла у сучасному криптографічному контексті з визначенням його практичних можливостей та обмежень, а також аналізом вразливостей, що виявляються під впливом сучасних методів криптоаналізу.

Дослідження спрямоване на детальне вивчення алгоритму, оцінку його теоретичної витонченості та ролі у розвитку блокових шифрів, а також на визначення перспектив використання Hill cipher як навчального й експериментального інструмента для моделювання принципів сучасної інформаційної безпеки.

Основний матеріал

Важливим аспектом дослідження є розуміння історичної ролі Hill cipher у формуванні концептуальних засад сучасної криптографії. Саме цей алгоритм уперше на практиці продемонстрував можливість застосування методів лінійної алгебри та матричних перетворень у процесі шифрування, заклавши підґрунтя для майбутнього розвитку блокових структур. Попри очевидні обмеження, зокрема повну лінійність і відсутність механізмів, здатних протидіяти диференційному чи лінійному криптоаналізу Hill cipher засвідчив потенціал переходу від простих моноалфавітних підстановок до складніших багатовимірних перетворень.

У подальшому ця ідея знайшла розвиток у класичних промислових стандартах криптографії – Data Encryption Standard (DES) та Advanced Encryption Standard (AES). У DES, затвердженому NIST у 1977 р., було реалізовано принцип блочного шифрування з багатораундовою структурою та використанням S-блоків для нелінійності, що істотно підвищило стійкість до атак. AES, який замінив DES на початку XXI століття, розвинув цей підхід, інтегрувавши більш складні матричні перетворення над скінченими полями $GF(2^8)$, включаючи операції ShiftRows, MixColumns і SubBytes. У цьому сенсі AES можна розглядати як логічне продовження ідей, започаткованих Hill cipher, але з урахуванням сучасних вимог до криптостійкості та ефективності реалізації.

Шифр Хілла заданий ключовою матрицею: $K \in Mat_{n\times n}\left(\mathbb{Z}_m\right)$ і дією:

$$C = KP\left(\mathrm{mod}\, m\right), \tag{1}$$

де стовпчикові вектори P та C відповідно відповідають блокам відкритого тексту й шифротексту розмірності n. Для коректності дешифрування необхідно, щоб матриця K була оберненою в $\mathbb{Z}_m$, тобто $\gcd\left(\det K, m\right) = 1$, що гарантує існування оберненої матриці

$$K^{-1} \in Mat_{n\times n}\left(\mathbb{Z}_m\right). \tag{2}$$

З одного боку, ця компактна лінійна структура становить методологічну перевагу алгоритму, оскільки дозволяє наочно продемонструвати зв’язок криптографії з лінійною алгеброю та модульною арифметикою. Вона робить Hill cipher зручним освітнім інструментом для моделювання базових принципів блокового шифрування [7].

З іншого боку, саме лінійність усіх перетворень над кільцем $\mathbb{Z}_m$ формує фундаментальну вразливість алгоритму. Будь-які відношення між векторами відкритого тексту P відображаються на шифротекст C через лінійне відображення. Це означає, що на практиці навіть обмежений набір пар «відкритий текст – шифротекст» може бути використаний для відновлення ключової матриці шляхом розв’язання системи лінійних рівнянь у $\mathbb{Z}_m$. Саме ця властивість робить Hill cipher криптографічно нестійким у сучасних умовах, незважаючи на його історичну значущість.

Отже, щоб конкретизувати механізми цієї вразливості й показати її практичні наслідки, нижче роз-

глянуто окремі сценарії криптоаналізу, зокрема атаки з обраним і відомим відкритим текстом, статистичні методи при наявності лише шифротексту та комбіновані підходи із застосуванням побічних каналів реалізації:

1. Атака з обраним відкритим текстом (chosen-plaintext, CPA).

У моделі CPA атакувальник має можливість запитувати шифрування довільних блоків $P \in \mathbb{Z}_m^n$ і отримувати відповідні шифротексти C. Для Hill cipher, заданого ключовою матрицею $K \in Mat_{n\times n}(\mathbb{Z}_m)$ та дією (1), найпряміша й найбільш ефективна стратегія полягає у поданні стандартних одиничних (базисних) векторів $e_1,...,e_n$, де e_i має одиницю на i-й позиції й нулі в інших координатах. Для кожного такого запиту справедлива рівність

$$C^{(i)} = Ke_i \left(\mathrm{mod}\, m\right), \tag{3}$$

тобто вектор $C^{(i)}$ є i-м стовпцем матриці K (в модульній арифметиці). Отже, збір відповідей на запити $e_1,...,e_n$ дає пряму реконструкцію всієї ключової матриці:

$$K = \left[C^{(1)}C^{(2)}...C^{n} \right] \left(\mathrm{mod}\, m\right). \tag{4}$$

Алгоритмічна і інформаційна оцінки:

– кількість запитів. Для повної компрометації необхідно не більше n запитів (подавання n базисних векторів).

– обчислювальна складність. Збирання n векторів розмірності n дає складність зберігання $O(n^2)$ елементів $\mathbb{Z}_m$; сама реконструкція ключа за наявності прямого доступу до стовпців – операція конкатенації, тобто $O(n^2)$ за примірною кількістю операцій над елементами кільця.

– інформаційний внесок одного запиту. Кожен запит з одиничним вектором e_i повертає вектор довжини n над $\mathbb{Z}_m$, тобто надає $n^2 \log_2 m$ бітів інформації про секретний ключ. У сумі n таких запитів дають $n^2 \log_2 m$ бітів, що узгоджується з розмірністю простору можливих ключів (приблизно $\log_2 \left|GL\left(n,\mathbb{Z}_m\right)\right|$).

Припустимо, що якщо атакувальник не може подавати довільні вектори, а лише вектори з обмеженого підмножини $S \subseteq \mathbb{Z}_m^n$, то атака ускладнюється, але залишається реалістичною за багатьох практичних умов. Нехай атакувальник подає послідовність випадково обраних векторів із S. Для успішної реконструкції необхідно зібрати n лінійно незалежних блоків $P_1,...,P_n$ (тобто, щоб матриця $P = \left[P_1...P_n\right]$ мала ранг n у $\mathbb{Z}_m$).

У випадку, коли $S = \mathbb{Z}_q^n$ для поля $\mathbb{Z}_q$ (тобто $m = q$ просте), ймовірність того, що t випадкових векторів матимуть ранг n, швидко прямує до 1 при $t \gtrsim n$; очікувана кількість запитів для отримання n незалежних векторів – $O(n)$ (з малою додатковою константою, що залежить від q).

Якщо S структурований (наприклад, лише вектори з обмеженим набором шаблонів або з фіксованими полями заголовків), може знадобитися значно більше запитів – у найгіршому випадку до нескінченності, якщо $span(S)$ має ранг $< n$. Тому при практичному аналізі важливо враховувати модель доступності запитів та статистику структурованих повідомлень (заголовків, форматів).

Отже, поєднання CPA із side-channel інформацією (таймінги, потужність, помилки) або з можливістю інжекції помилок (fault injection) скорочує потрібну кількість запитів і знижує вимоги до незалежності блоків. Наприклад, певні неточні відповіді або додаткові виміри можуть надати часткові лінійні відомості про стовпці K, що значно полегшує відновлення ключа в реалістичних сценаріях.

Розглядаючи практичні контрзаходи й рекомендації, слід підкреслити, що оскільки CPA є фатальною для шифру Хілла, при проєктуванні систем і проведенні навчальних демонстрацій необхідно впроваджувати наступні технічні та процедурні заходи:

1. Обмежити можливість довільного шифрування шляхом введення автентифікації клієнтів і обмеження API, щоб виключити сценарії довільних запитів.

2. Використовувати протокольні елементи, що додають непередбачуваність, а саме IV/nonce, випадковий паддінг, сольові значення [8]; які перетворюють прямолінійний оператор у стан, залежний від випадкових параметрів.

3. Вводити конструктивну нелінійність: у блочних схемах це S-блоки і складні перестановки; у протоколах – комбінації підстановок і змішувань, що унеможливлюють просте лінійне відновлення.

4. Застосовувати автентифіковане шифрування (AEAD), щоб запобігти маніпуляціям і повторному використанню блоків [9].

5. Контролювати реалізаційні ризики, що забезпечить захист від побічних каналів і інжекції помилок, валідація вхідних параметрів (перевірка оберненості матриць, недопущення слабких ключів).

Підсумовуючи, CPA показує, що лінійність Hill cipher забезпечує атакувальнику простий шлях до відновлення ключа всього за n запитів. Навіть у обмежених моделях або з додатковими ускладненнями атака залишається здійсненною й може посилюватися побічними каналами. Тому Hill cipher доцільно розглядати лише як навчальний приклад, а не як засіб практичного захисту даних.

2. Атака при детермінованому (регулярно повторюваному) наборі відкритих блоків.

Нехай множина можливих блоків відкритого тексту обмежена підмножиною $S \subseteq \mathbb{Z}_m^n$. Позначимо її лінійний замик (над кільцем $\mathbb{Z}_m$) через $\mathrm{span}\left(S\right)$ і його модульну (або векторну, якщо m – просте)

щільність через $\dim \operatorname{span}(S) = r$. Тоді справедливі такі формальні спостереження.

Необхідна умова успіху. Якщо $r < n$, тобто *span*(*S*) є підпростором розмірності менше ніж порядок блоку, то жодна скінченна сукупність пар (P_j, C_j) із $P_j \in S$ не дасть матриці P повного рангу n. У такому випадку відновити повний ключ $K \in GL(n, \mathbb{Z}_m)$ у загальному вигляді неможливо, можна лише відновити його дію на підпросторі $\operatorname{span}(S)$.

Умова достатності. Якщо $r = n$ (тобто $\operatorname{span}(S) = \mathbb{Z}_m^n$), то існує набір n лінійно незалежних векторів $P_1, \ldots P_n \in S$. Наявність таких n незалежних пар (P_j, C_j) дозволяє побудувати матрицю $P = [P_1 \ldots P_n]$ з $\operatorname{rank}(P) = n$ і отримати ключ:

$$K = CP^{-1} \pmod m, \tag{5}$$

де $C = [C_1 \ldots C_n]$.

Практична схема атаки при детермінованих форматах. Нехай відкриті повідомлення мають регулярну структуру типу

$$P^{(t)} = u + B\, \mathcal{X}^{(t)}, \tag{6}$$

де u – фіксований вектор (наприклад, повторюваний заголовок), B – матриця, що відображає параметризовані поля (шаблони), а $X^{(t)}$ – вектор змінних полів. Якщо простір, породжений усіма можливими $X^{(t)}$, має розмірність r (відповідно $\dim span(S) = r$), то атакувальнику достатньо зібрати r лінійно незалежних спостережень, щоб визначити дію K на цей підпростір. Формально, якщо є підматриця $P_{\text{sub}} \in \mathbb{Z}_m^{n \times r}$ рангу r, то можна обчислити частковий оператор

$$K\big|_{\operatorname{span}(S)} \equiv C_{\text{sub}} P_{\text{sub}}^{+} \pmod m, \tag{7}$$

де P_{sub}^{+} – псевдообернена або обернена (за наявності) матриця на відповідному підпросторі. Якщо $r = n$, то P_{sub}^{-1} і відновлення повне.

Оцінка «швидкості» компрометації при випадкових зразках із S. У випадку, коли елементи P_j вибираються випадково з S за деяким розподілом, питання редуковане до отримання n лінійно незалежних векторів. Для алгебрично простих випадків (наприклад, $m = q$ просте і $S = \mathbb{F}_q^n$) ймовірність того, що t випадкових векторів дають повний ранг n, виражається через добуток

$$\mathrm{P_r}\left[\operatorname{rank}(P_t) = n\right] = \prod_{i=0}^{n-1} \left(1 - q^{i-t}\right) \tag{8}$$

тож очікувана кількість вибірок для досягнення рангу $n \in O(n)$. Для структурованих S ця оцінка змінюється відповідно до розміру $\dim span(S)$ і розподілу варіантів $\mathcal{X}^{(t)}$.

Алгоритмічна реалізація відновлення в умовах детермінованості. Процедура експлуатації вразливості:

– збирання пари (P_j, C_j) поки rank(P) зростає.

– за кожним новим спостереженням виконувати вставку стовпця й оновлення рангу (Gaussian elimination $\bmod m$).

– коли знайдено підматрицю P_{sub} рангу r достатнього розміру (рівного n або максимально можливого), обчислити відповідне K або $K\big|_{\operatorname{span}(S)}$ як $C_{\text{sub}} P_{\text{sub}}^{-1}$ (або за допомогою псевдооберненої для $r < n$).

Необхідно звернути увагу, що операційна складність – домінована обчисленням рангу та оберненням підматриці розміру r: $O(r^3)$ елементарних операцій над $\mathbb{Z}_m$.

Як приклад ситуації з регулярними заголовками можна навести наступне. Повідомлення складаються з фіксованого заголовка u (однаковий для всіх повідомлень) і невеликого поля $\mathcal{X}^{(t)}$ розмірності $k \ll n$, тобто $P^t = u + \left[B\mathcal{X}^{(t)} \right]$. Тоді $\dim \operatorname{span}(S) \le k+1$. Якщо $k+1 < n$, то повного ключа не отримати, але дія K на підпросторі розмірності $k+1$ відновлюється за $k+1$ незалежних прикладів що є достатньо, щоб дешифрувати всі змінні поля й компрометувати важливі частини структури повідомлення [10].

З огляду на зазначене можна підсумувати таке:

– якщо $\dim span(S) = n$, то повторювані/детерміновані формати не захищають від повної компрометації, необхідно очікувати $O(n)$ запитів;

– якщо $\dim \operatorname{span}(S) = r < n$, то атакувальник може відновити лише часткову дію $K\big|_{\operatorname{span}(S)}$, що все одно може бути достатньо для витоку конфіденційної інформації (зокрема заголовків, метаданих);

– рекомендовані захисні заходи: введення випадкових параметрів (IV/nonce, соль), шифрування змінних частин повідомлення окремими сесійними ключами, застосування AEAD; у навчальних експериментах – сувора ізоляція демонстрацій.

3. Атака за наявності лише шифротексту (ciphertext-only).

Нехай Hill cipher заданий ключовою матрицею $K \in GL(n, \mathbb{Z}_m)$ і перетворенням

$$C^t \equiv KP^t \pmod m, \quad t = 1, \ldots, T, \tag{9}$$

де $P^t \in \mathbb{Z}_m^n$ – послідовні блоки відкритого тексту,

$C^t \in \mathbb{Z}_m^n$ – відповідні блоки шифротексту, атакувальник має тільки набір спостережень $\left\{C^t\right\}_{t=1}^{T}$. У відсутності будь-яких пар «відкритий текст – шифротекст» завдання відновлення K є некласичною проблемою сліпого зворотного розв'язування лінійного оператора над кільцем $\mathbb{Z}_m$. Проте за певних припущень щодо статистики відкритого тексту або для малих параметрів *m, n* атака може бути практично здійсненною. Нижче наведено кілька формалізованих підходів.

– метод на основі перших і других моментів (метод моментів/кореляційний підхід).

Позначимо емпіричні середні та коваріації (розглядаємо підняття символів у інтервал $\{0, \dots, m\text{-}1\}$ і працюємо в $\mathbb{R}$ для статистичного оцінювання, з подальшим приведенням мод m):

$$\mu_C = \frac{1}{T}\sum_{t=1}^{T} C^t$$
$$\Sigma_{CC} = \frac{1}{T}\sum_{t=1}^{T}\left(C^{(t)} - \mu_C\right)\left(C^{(t)} - \mu_C\right)^T \quad . \quad (10)$$

Якщо атакувальник має апріорні знання або припущення про статистику відкритого тексту (середнє μ_P і коваріацію Σ_{PP} (інвертованої в $\mathbb{R}$) можна отримати оцінку

$$K = \Sigma_{CP}\ \Sigma_{PP}^{-1}, \quad (11)$$

де Σ_{CP} – крос-коваріація між C (спостережуваним) і P (оціночним, взятим із корпусу). Якщо атакувальник знає (або коректно оцінює) μ_P, Σ_{PP}, то оцінка K може бути хорошою початковою гіпотезою; вона переводиться в коректний ключ K шляхом проекції на $GL\left(n,\mathbb{Z}_m\right)$ (наприклад, округлення елементів і перевірка оберненості за модулем m).

Умови ефективності. Підхід працює краще при:

- достатньо великому T (статистична стійкість оцінок);
- точності апріорної моделі μ_P, Σ_{PP} (тобто коли текстова статистика корпусу близька до реальної);
- невеликих n і/або m, коли округлення та приведення до модульної арифметики мають менше неоднозначностей.

Обмеження. Якщо Σ_{PP} невироджена лише за модулем m (а не в $\mathbb{R}$), або якщо апріорні оцінки сильно відрізняються від реальних, то оцінка K може бути ненадійною.

– атака як задача «сліпого» лінійного розділення (BSS / ICA-подібний підхід).

Сутність полягає в тому, щоб знайти цілу невироджену матрицю $M \in GL\left(n,\mathbb{Z}_m\right)$ таку, що перетворення $P^{(t)} = M^{-1}C^{(t)}$ дає статистику, близьку до відомої статистики природної мови (наприклад, частоти букв, частоти біграм/мультіграм тощо). Формально, атакувальник вирішує оптимізаційну задачу

$$K = \arg\min_{K' \in GL\left(n,\mathbb{Z}_m\right)} D\left(\mathrm{Stat}\left(K'^{-1}C\right)\middle\|\mathrm{Stat}_{\mathrm{lang}}\right), \quad (12)$$

де $D\left(\cdot\middle\|\cdot\right)$ – міра нерівності (наприклад, Kullback-Leibler divergence або χ^2 – статистика), $\mathrm{Stat}\left(\cdot\right)$ – емпіричні частоти/мовні статистики відновленого тексту, $\mathrm{Stat}_{\mathrm{lang}}$ – еталонні статистики для мови. Підхід фактично перебирає (або оптимізує в просторі) кандидати з $GL\left(n,\mathbb{Z}_m\right)$ і вибирає ті, які «відновлюють» мовні шаблони.

Практичний аспект. Прямий перебір $GL\left(n,\mathbb{Z}_m\right)$ можливий лише при малих m, n. Для більших розмірів застосовують евристики:

- локальні пошуки/simulated annealing по простору ключів;
- градієнтоподібні методи на релаксації в дійсній області з подальшим проєктуванням на цілі матриці;
- комбінування з ML-методами (класифікатори, що оцінюють «мовність» декодованого тексту).

Оцінка. Для малих n (наприклад, $n = 2, 3$) і $m \le 26$ цей підхід часто успішний за наявності великого числа блоків T (тисячі – десятки тисяч блоків) і якісної мовної моделі.

– частотні та статистичні методи (класичний підхід для малих m, n).

У випадку, коли m і n невеликі, можна використовувати класичні частотні методи: аналіз розподілу символів (позиційні частоти), біграм/триграм та їх кореляцій. Формальна ідея:

- для кожної позиції блоку j оцінюється емпіричний розподіл $p_{C_j}(\alpha) = \Pr\left(C_j = \alpha\right)$ по всіх спостереженнях;
- внаслідок лінійного перетворення кожна компонента C_j є лінійною комбінацією компонент P; якщо припустити незалежність або малі кореляції між координатами P, можна побудувати систему лінійних співвідношень між емпіричними частотами, що дає підказки про елементи K.

Для оцінки елементів ключової матриці K доцільно шукати такі лінійні відношення, які максимізують узгодженість між компонентами шифротексту C_i та лінійними комбінаціями позицій відкритого тексту у межах мовної моделі. У практиці це реалізується через статистичні критерії – коефіцієнт кореляції, χ^2 та log-likelihood, що дозволяють відбирати ті кандидати ключа, які найкраще відтворюють характерні властивості природної мови.

Успіх методу істотно залежить від того, наскільки точно модель відкритого тексту (частоти, кореляції) відповідає реальній. Для природних мов з помітною нерівномірністю частот (наприклад, анг-

лійська) шанси на успіх вищі, ніж для рівномірних або ентропійно насичених джерел.

– алгоритмічна схема

- збір статистики: обчислити μ_C, Σ_{CC}, позиційні і n-грамні частоти з $\left\{C^{(t)}\right\}_{t=1}^{T}$.
- апріорні припущення: вибрати мовну/корпусну модель для μ_P, Σ_{PP} або частот $\mathrm{Stat}_{\mathrm{lang}}$.
- побудова кандидатів: (а) оцінка K методом моментів; (б) генерація кандидатів $K^{'} \in GL\left(n, \mathbb{Z}_m\right)$ евристично; (в) локальна оптимізація релаксованої задачі.
- оцінка кандидатів: для кожного $K^{'}$ обчислити $Stat\left(K^{'-1}C\right)$ і виміряти відстань $D\left(\cdot \| \mathrm{Stat}_{\mathrm{lang}}\right)$.
- верифікація: для найкращих кандидатів перевірити консистентність (наприклад, дешифрувати деякі блоки і оцінити лінгвістичну інформативність).

Обчислювальна складність є експоненційною при повному переборі, проте на практиці зводиться до поліноміальної чи евристичної залежно від обраної стратегії пошуку.

– оцінки ефективності та обмеження.

- малі m, n: атаки набагато реалістичніші; для $m = 26$, $\mathrm{n} \in \{2, 3\}$ – відомі практичні розв'язки із застосуванням частотного аналізу та пошуку.
- великий T: вимога великої статистичної вибірки – типово T має бути порядку кількох тисяч блоків для стабільних оцінок.
- апріорна інформація: наявність якісної мовної моделі або корпусу істотно підвищує шанси на успіх.
- наявність побічних каналів: з їх допомогою ciphertext-only сценарій легко перетворюється в гібридний (сильніше спрощує задачу).

– контрзаходи проти ciphertext-only атак.

- збільшення ентропії в блоку: використовувати рандомізовані IV/nonce і випадкові паддінги;
- впровадження нелінійності в трансформацію: S-блоки або інші нелінійні операції, які руйнують лінійні статистичні взаємозв'язки;
- обмеження експозиції шифротекстів: мінімізувати кількість доступних шифротекстних блоків для однієї ключової сесії;
- аутентифіковане шифрування (AEAD): запобігає маніпуляціям, повторному використанню блоків і робить збір корисних статистик менш ефективним.

Підсумовуючи можна зазначити, що ciphertext-only атака на Hill cipher є формально складною, проте за малих параметрів, достатньої кількості спостережень та використання мовних чи побічних статистик вона стає практично здійсненною, що підтверджує його непридатність для захисту даних навіть у режимі доступу лише до шифротекстів.

4. Комбінування з побічними каналами (side-channel).

Нехай шифрування блока P реалізовано апаратно й проміжний стан S (наприклад, вектор результату KP перед редукцією по модулю) є функцією від секретного ключа та вхідного блоку:

$$S = g\left(K, P\right), \tag{13}$$

де для Hill cipher $g\left(K, P\right) = KP$ (в інтерпретації апаратної реалізації – проміжні регістри перед модульною операцією). Побічний канал дає емпіричні вимірювання L (таймінг, потужність, електромагнітні випромінювання, помилкові відповіді), які моделюються як функція від проміжного стану з шумом:

$$L = l\left(S\right) + \eta, \tag{14}$$

де l – модель витоку (наприклад, Hamming-weight/Hamming-distance, лінійна або нелінійна функція від бітів S), η – шум (гаусівський або інший).

– Correlation/Template-style атаки (таймінг, power-analysis).

Якщо l дає хоча б часткову інформацію про компоненти вектора $S = KP$, то спостережувані L корелюють з лінійними функціями від рядків K_{i*}. Формально, для кожної позиції i можна розглядати гіпотезу про рядок k і обчислювати статистику кореляції

$$\rho\left(k\right) = Corr\left(L,\ h\left(k^{\mathrm{T}}P\right)\right), \tag{15}$$

де h – модель перетворення стану в вимір (наприклад, Hamming-weight). Максимум $k = \arg\ \max_k \rho\left(k\right)$ служить оцінкою K_{i*}. Для Hill cipher, оскільки $k^{\mathrm{T}}P$ – лінійна комбінація елементів P, навіть слабка кореляція дає значну інформацію про k, а накопичення кількох вимірів (для різних P) підвищує SNR і дозволяє невеликою кількістю профільних трас відновити рядки K. Таким чином CPA по витоку (Correlation Power Analysis) перетворює апаратну уразливість у лінійну систему про елементи ключа.

– Differential Fault Analysis (DFA/fault injection).

Припустимо, що атакувальник може інжектувати локальну помилку в обробку одного блоку P, отримавши некоректний шифротекст $C^{'} = C + \Delta C$. Для Hill cipher

$$C = KP\ \left(\mathrm{mod}\ m\right),\ C^{'} = K\left(P + \Delta P\right)\ \mathrm{mod}\ m, \tag{16}$$

тобто різниця

$$\Delta C = K\Delta P\ \left(\mathrm{mod}\ m\right). \tag{17}$$

Якщо ΔP – вектор з одиницею в одній позиції (або вектор, контрольований атакувальником), то ΔC дає прямі лінійні відношення для відповідних стовпців K. Навіть одинична цілеспрямована помилка (або відома форма помилки) перетворює атаку з диференціального виду на лінійний розв'язок для відповідних невідомих елементів K. Для прикладу:

якщо інжектовано $\Delta P = \delta e_j$ (де δ відома зміна), то

$$\Delta C = \delta K e_j = \delta \quad \text{j -й стовпець } K.$$

– комбінування джерел інформації та відновлення K.

Узгоджуючи статистичні оцінки від витоку $l(S)$ та лінійні рівняння, одержані через DFA, атакувальник вирішує узагальнену задачу оцінки в умовах шуму:

$$K = \arg \min_{K' \in GL(n, \mathbb{Z}_m)} \mathcal{J}K', \tag{18}$$

де, наприклад,

$$\mathcal{J}\left(K'\right) = \sum_{t=1}^{T} \| L^{(t)} - l\left(K' P^t\right) \|^2 + \lambda \sum_{u \in \mathcal{F}} \| \Delta C^u - K' \Delta P^u \|^2, \tag{19}$$

функція сумарної невідповідності, що включає вклад від power/timing трас $L^{(t)}$ та від диференціальних різниць $\Delta C^{(u)}$ отриманих через fault injection (параметр λ регулює показники впливу). Мінімізація цієї функції (через перебір, релаксації або евристичні методи) дає значно швидше відновлення ключа ніж класична KPA/CPA без побічних даних.

– практичні показники ефективності.

Ефективність side-channel-підсилених атак визначається SNR витоку, кількістю трас T і можливістю інжекції помилок. Типові емпіричні вимоги для успішної реконструкції рядка K_{i*} методом CPA: T порядку десятків-сотень трас при помірному SNR; для DFA достатньо кількох точково спрямованих інжекцій, якщо ΔP контролюється або має просту відому форму.

Щоб зменшити ризик компрометації через побічні канали та інжекцію помилок, необхідно поєднувати технічні, протокольні та організаційні контрзаходи [14]. Захист має руйнувати кореляції витоку, усувати залежність часу/потужності від секретів, виявляти і нейтралізовувати помилки та обмежувати експозицію шифротекстів і доступ до сервісів шифрування [11]. Ключовими складовими захисту є:

1. Маскування та рандомізація проміжних станів (randomized masking) для руйнування кореляцій.

2. Constant-time та детерміновані реалізації для усунення таймінгових каналів.

3. Детектування та запобігання інжекції помилок (контроль контрольних сум, дублювання обчислень, датчики напруги/температури).

4. Фізичний захист (екранування, захист від доступу до плат) та аудит апаратних інтерфейсів.

5. В рамках криптографічного дизайну необхідно уникати прямих лінійних залежностей у проміжних станах (вводити нелінійність, солі/IV, AEAD).

Отже, слід зазначити що поєднання побічної інформації з класичними методами криптоаналізу істотно знижує поріг успіху атак на Hill cipher: слабка лінійність, яка сама по собі робить алгоритм вразливим у CPA/KPA, стає практично нескладною для експлуатації, якщо реалізація видає навіть мінімальну побічну інформацію або піддається інжекції помилок. Тому оцінка безпеки повинна обов'язково включати аналіз реалізаційних ризиків та заходи проти side-channel і fault-атаки [12].

Підсумовуючи розглянуті сценарії криптоаналізу, можна виділити кілька основних висновків, що характеризують вразливості Hill cipher і практичні наслідки його лінійної структури:

по-перше, основною причиною вразливості Hill cipher є його лінійна структура, яка дозволяє відносно легко відновлювати ключ у моделях із доступом до відкритого тексту та значно спрощує статистичний аналіз у режимі лише шифротекстів;

по-друге, атаки з обраним або відомим відкритим текстом призводять до повної компрометації алгоритму за обмежену кількість запитів, що робить його непридатним для використання у продуктивних системах;

по-третє, навіть при обмеженому доступі до шифротекстів або за наявності статистичної інформації, ймовірність успішної атаки залишається достатньо високою, особливо для малих параметрів алгоритму.

Нарешті, використання побічних каналів (таймінг, енергоспоживання, помилкові інжекції) істотно полегшує атаки, знижуючи вимоги до спостережень і роблячи компрометацію ключа майже тривіальною.

Висновки

Проведене дослідження засвідчило, що шифр Хілла, хоча й став важливим кроком у становленні блокових шифрів та продемонстрував потужність застосування лінійної алгебри у криптографії, має фундаментальні вразливості, пов'язані з його повною лінійністю та відсутністю механізмів захисту від сучасних криптоаналітичних атак. У результаті цей алгоритм не відповідає вимогам безпеки XXI століття й не може розглядатися як практичний засіб захисту конфіденційних даних.

Водночас його простота, математична прозорість і наочність роблять Hill cipher цінним навчально-методичним інструментом, що дозволяє досліджувати основи криптографічних перетворень, відпрацьовувати базові моделі атак та моделювати перехід від класичних шифрів до сучасних блокових алгоритмів.

Таким чином, його доцільно розглядати виключно в освітньому та дослідницькому контексті як «першу сходинку» до розуміння архітектури безпечних шифрувальних систем.

СПИСОК ЛІТЕРАТУРИ

1. Zhang, X., Shao, C., Li, T. et al. GFSPX: an efficient lightweight block cipher for resource-constrained IoT nodes. J Supercomput 80, 25256–25282 (2024). https://doi.org/10.1007/s11227-024-06412-2
2. Announcing Issuance of Federal Information Processing Standards (FIPS) FIPS 203, Module-Lattice-Based Key-Encapsulation Mechanism Standard, FIPS 204, Module-Lattice-Based Digital Signature Standard, and FIPS 205, Stateless

Hash-Based Digital Signature Standard. National Institute of Standards and Technology on 08/14/2024. Retrieved from https://csrc.nist.gov/news/2024/postquantum-cryptography-fips-approved
3. Dani, J., Nakka, K. and Saxena, N. A Machine Learning-Based Framework for Assessing Cryptographic Indistinguishability of Lightweight Block Ciphers. 30 May 2024. Retrieved from https://arxiv.org/abs/2405.19683v1
4. Kaur, J., Canto, A. C., Kermani, M. M., Azarderakhsh, R. A Comprehensive Survey on the Implementations, Attacks, and Countermeasures of the Current NIST Lightweight Cryptography Standard, 2023. URL: https://arxiv.org/abs/2304.06222
5. Ni, Z., Khalid, A., Liu, W., & O'Neill, M. (2024). Bitstream Fault Injection Attacks on CRYSTALS Kyber Implementations on FPGAs. 1–6. Retrieved from https://doi.org/10.23919/date58400.2024.10546550
6. Zhyvylo Y. (2023). Exploring and Acquiring Modern Human Resource Competencies in Cybersecurity Amidst State Digital Transformation. Pressing Problems of Public Administration, 2(63), 111-127. https://doi.org/10.26565/1684-8489-2023-2-08
7. Zhyvylo , Y. O., & Zhyvylo , I. O. (2021). Joint training of the cyber security defense forces personnel in the conditions of total state defense. Theory and Practice of Public Administration, 2(73), 144-153. https://doi.org/10.34213/tp.21.02.16
8. Mahdi, Q. A., Zhyvotovskyi, R., Kravchenko, S., Borysov, I., Oleksandr, O., Panchenko, I., Zhyvylo, Y., Kupchyn, A., Koltovskov, D., Boholii, S. (2021). Development of a method of structural-parametric assessment of the object state. Eastern-European Journal of Enterprise Technologies, 5 (4 (113)), 34–44. doi: https://doi.org/10.15587/1729-4061.2021.240178
9. Koval M., Sova O., Orlov O., Zhyvylo Y., Zhyvylo I. Improvement of complex resource management of special-purpose communication systems // 5(9-119) (2022): Eastern-European Journal of Enterprise Technologies. P. 34–44;
10. S. Kashkevich, A. Shyshatskyi, O. Dmytriieva, Y. Zhyvylo, G. Plekhova, S. Neronov The development of management methods based on bio-inspired algorithms Information and control systems: modelling and optimizations: collective monograph. – Kharkiv: TECHNOLOGY CENTER PC, 2024. pp. 35-69. DOI: http://doi.org/10.15587/978-617-8360-04-7
11. Zhyvylo, Y.O. (2024). Methodology for developing a national cybersecurity strategy. State Formation, no. 2 (36), 307–321. DOI: https://doi.org/10.26565/1992-2337-2024-2-21
12. Живило Є. О. Оцінка ризиків кібербезпеки та контролю конфіденційності в інформаційних системах державного управління / Є. О. Живило, Д. Г. Шевченко // Збірник наукових праць Військового інституту Київського національного університету імені Тараса Шевченка. 2022. № 75. С. 66-77. URL: http://nbuv.gov.ua/UJRN/Znpviknu_2022_75_9
13. Живило Є.О., Черноног О.О. Стратегія кібероборони України, Збірник наукових праць ВІТІ № 4, 2017, С.30–37. URL: https://www.researchgate.net/publication/380979172_STRATEGIA_KIBEROBORONI_UKRAINI
14. Ігор Ромашко, Юлія Калашнікова, CISCO SECUREX ТА ZERO TRUST: СУЧАСНІ ПІДХОДИ ДО КІБЕРЗАХИСТУ, 2025. Retrieved from http://perspectives.pp.ua/index.php/nts/article/view/29469/29425.
15. Onyshchenko, S., Zhyvylo, Ye., Cherviak, A. and Bilko S. (2023), "Determining the patterns of using information protection systems at financial institutions in order to improve the level of financial security", //) (2023), Eastern-European Journal of Enterprise Technologies, vol. 5 (13 (125), pp. 65–76. DOI: https://doi.org/10.15587/1729-4061.2023.288175

Received (Надійшла) 11.08.2025
Accepted for publication (Прийнята до друку) 05.11.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Живило Євген Олександрович – кандидат наук з державного управління, доцент, доцент кафедри комп'ютерних технологій та інформаційних систем, Національний університет «Полтавська політехніка імені Юрія Кондратюка», Полтава, Україна;

Yevhen Zhyvylo – candidate of sciences in public administration, associate professor, associate professor of the department of computer technologies and information systems, National University "Poltava Polytechnic named after Yury Kondratyuk", Poltava, Ukraine;
e-mail: zhivilka@i.ua; ORCID Author ID: https://orcid.org/0000-0003-4077-7853;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57350779100.

Кучма Юрій – кандидат технічних наук, доцент, завідувач кафедри комп'ютерних систем, ТОВ «Університет сучасних технологій», Київ, Україна;

Yurii Kuchma – Candidate of Technical Sciences, Associate Professor, Head of the Bepartment of Computer Iystems of the University of Modern Technologies LLC,. Kyiv, Ukraine;
e-mail: krabatua@gmail.com; ORCID Author ID: https://orcid.org/0009-0002-5498-4271.

Practical application and vulnerabilities of Hill cipher in a modern context

Yevhen Zhyvylo, Yurii Kuchma

Abstract. The article examines the Hill cipher as a classical example of applying linear algebra and modular arithmetic in cryptography. It elucidates the mathematical foundations of the algorithm, including the formation of the key matrix over finite fields, the conditions for its invertibility, and the implementation of encryption as the multiplication of plaintext vectors by the key matrix. This approach is presented as a distinctive "bridge between theory and practice," combining formal mathematical constructions with practical data protection mechanisms. Particular attention is given to the vulnerabilities of the Hill cipher. It is shown that the algorithm is insecure in modern contexts due to the linearity of its transformations and the absence of nonlinear components. Examples of cryptanalytic attacks are provided, demonstrating that the cipher can be "broken apart" even with a minimal number of plaintext-ciphertext pairs. This makes it evident that, in the context of 21st-century cybersecurity, the Hill cipher is merely a "shadow of its era" and cannot serve as a fully reliable tool for data protection. At the same time, the educational and methodological value of the cipher is emphasized. It allows one to "see the mechanism from the inside," clearly illustrating the transition from monoalphabetic ciphers to block transformations. The article highlights that the Hill cipher played the role of a "first step" in the development of structures that later evolved into DES and AES, where the use of block structures and matrix operations became a fundamental element of modern cryptography.

Keywords: cyber security, cyber threats, artificial intelligence, machine learning, cryptography, cryptanalysis, cipher.

UDC 004. 896:336.76 **doi: 10.26906/SUNZ.2025.4.74-77**

Oleksandr Zakovorotnyi, Nataliia Ausheva, Larysa Levchenko

National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute», Kyiv, Ukraine

IMPROVING THE AI STOCK MARKET FORECASTING WITH CANDLESTICK PATTERNS

Abstract. In the rapidly evolving digital economy, the application of Artificial Intelligence (AI) in financial forecasting has gained significant traction. This study investigates the effect of various candlestick patterns on the performance of Long Short-Term Memory (LSTM) models in predicting stock market movements. Experiments conducted on the stock price history data demonstrate that supplementing traditional input parameters (e.g., open price) with a range of candlestick patterns enhances the predictive accuracy of LSTM models. Although the initial model architecture lacked hyperparameter optimization for solving this kind of task, our findings suggest notable improvement in prediction performance when candlestick pattern flags are incorporated. Future work will focus on incorporating additional financial indicators into the model's training data and fine-tuning it through optimization algorithms to achieve greater robustness and accuracy.

Keywords: artificial intelligence, LSTM model efficiency, stock market prediction, candlestick patterns, improving AI model accuracy.

Introduction

Digital technologies have revolutionized market dynamics, increasing financial inclusion worldwide and enabling the creation of global market services. Financial services have become independent and target regular individuals as users. The worldwide spread has introduced new data analysis methods, enabling accurate economic predictions and strategic market operations [1].

Stock market investment can be considered one of the key financial sectors that spreads accessibility and provides economic independence, whether at the high level of a country's economy or the personal level. The number of investors is growing, increasing the need to develop new algorithms and methods for predicting market dynamics. The stock market is influenced by various factors that can impact stock prices, including economic conditions, political situation, and the performance of specific companies [2].

As financial markets become increasingly data-driven, Artificial Intelligence (AI), particularly Deep Learning models, has emerged as a promising solution for enhancing market forecasting. The rise in market complexity, driven by both macroeconomic variables and behavioral patterns, necessitates the use of adaptive and intelligent models. Among the various AI techniques, Long Short-Term Memory (LSTM) networks have shown significant promise due to their ability to capture temporal dependencies in sequential data.

However, the effectiveness of such models hinges on the quality and diversity of input features. Traditional models relying solely on basic pricing data (e.g., open or close prices) often fail to capture nuanced market behaviors.

This study investigates the effect of candlestick patterns on the forecasting accuracy of LSTM models.

Literature review

Current research on the impact of digital financial technology on accelerating financial inclusion focuses on identifying the key segment that influences the development of financial institutions and economic growth.

The study [1] emphasized that digital financial technologies have fostered financial inclusion in developing economies. The work highlighted that mobile-based banking systems and digital payment solutions enable access to financial services for previously unbanked populations. This increased accessibility has enhanced economic participation and empowered marginalized communities by providing tools for savings, investments, and credit. The significant role of government policies in supporting fintech adoption highlights the importance of regulatory frameworks and partnerships between the public and private sectors in scaling these technologies.

However, digital literacy and infrastructure deficiencies persist, particularly in rural areas, where technological and educational barriers prevent widespread adoption. The study suggested that targeted educational campaigns and infrastructure investments are necessary to address these gaps effectively. The evidence presented in this section suggests that economic development and financial stability are directly dependent on modern information technologies and the latest market analysis methods [1].

Artificial Intelligence has gained importance in recent years in analyzing the market and forecasting stock prices [3, 4].

The research [5] examined techniques and case studies of advancements in Artificial Intelligence and Machine Learning for stock market prediction, demonstrating that machine learning algorithms have received considerable attention in recent years.

The most used AI models and techniques include the Support Vector Machine (SVM), the Random Forest (RF), the K-nearest neighbor (KNN), the Naive Bayes NB, the Long-Short-Term Memory (LSTM), and the Artificial Neural Network (ANN). This highlights the significance of trading as a key area for AI improvement, as evidenced by the study's results, which show that machine learning algorithms can outperform traditional data analysis methods. Nevertheless, it is essential to consider which approach yields superior accuracy. The efficiency and accuracy of each algorithm are the key parameters that should be considered and improved [5].

SVM is a promising tool for financial prediction, outperforming standard indicators when dealing with market uncertainty and complexity. SVM demonstrated higher reliability and adaptability, especially under volatile or crisis conditions in the stock market. Traditional methods were often less accurate and less responsive to sudden market shifts. The SVM model successfully captured nonlinear dependencies and complex patterns in market data [6].

On the other hand, Long Short-Term Memory (LSTM) networks have shown significant promise due to their ability to capture temporal dependencies in sequential data.

Long Short-Term Memory (LSTM) is a type of Recurrent Neural Network (RNN) architecture specifically designed to model temporal sequences and long-range dependencies more effectively than standard RNNs. LSTM networks overcome the vanishing gradient problem inherent in traditional RNNs by introducing memory cells and gating mechanisms that regulate the flow of information [7].

AI algorithms were modified to enhance the accuracy of results using various techniques for the basic AI algorithms [8-10]. The research [8] argued that the deep LSTM network algorithm could be optimized with the Artificial Rabbits Optimization (ARO) algorithm. This research proposed a new deep LSTM network optimized by the ARO algorithm. The proposed model is named LSTM-ARO. The goal is to optimize LSTM model parameters using ARO to determine the most efficient architecture. To evaluate the efficiency of LSTM-ARO, it was trained on historical market data to predict price dynamics and compared with the regular LSTM model, the ANN model, and the Genetic Algorithm (GA). The results demonstrate that the LSTM-ARO model outperforms other models, delivering highly accurate predictions that can be valuable for traders and investors [8].

Similarly, the study [9] combined Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models with AI techniques, yielding high predictive performance for volatile markets. It highlighted the robustness of GARCH models in capturing market volatility and leveraging AI's adaptability to complex data patterns. The GARCH-LSTM model achieved the best performance, exhibiting the lowest RMSE (0.2002), MAE (0.1840), and MAPE (0.1570), along with the highest R-squared value (0.9995).

Overall, hybrid models combining GARCH with LSTM, GRU, and Transformer outperformed their standalone counterparts. This demonstrates that integrating the statistical properties of historical prices with RNN and Transformer architecture significantly enhances the accuracy of stock price predictions. The authors noted that basic AI models often underperform during sudden, unpredictable market shocks, highlighting the need for enhanced adaptability. The real-time market sentiment analysis was integrated with GARCH-AI models to address this limitation, thereby enabling more effective responses to sudden changes in market dynamics [9].

Reference [10] presents results suggesting that AI algorithms can be optimized by combining different models and then normalizing their results using the weighted ensemble learning method. The research described a combination of prediction models based on Artificial Neural Networks (ANN), Gaussian Process Regression (GPR), and Classification and Regression Trees (CART). This combined model was then weighted with the ensemble model based on the quality of training and using the cuckoo search algorithm. In contrast to finding the most efficient architecture for the specific AI algorithm, as described in previous publications, this study aimed to merge several algorithms to compensate for the shortcomings of each at the expense of the others. The results indicate that the proposed system can predict the price index with an average accuracy of 96.6%, representing a reduction in prediction error of at least 2.4% compared to traditional methods [10].

Together, these studies offer valuable insights into stock market prediction using AI. Considering all this evidence, AI methods are not only successfully used to predict the stock market dynamics but also have the potential to increase the accuracy of their results. Investors can use these forecasting results as key parameters to analyze and plan their investments, which simultaneously helps to develop the global economic system.

Despite these advancements, there is a limited amount of research focusing on how the inclusion of specific financial indicators affects prediction outcomes. Most studies focus on improving algorithms themselves rather than examining the data features fed into them.

Candlestick patterns can be utilized as a primary feature for stock market forecasting with LSTM models. The candlestick patterns are graphical representations of price movement in financial markets and have been used by traders for a long time to infer potential market direction. While traditionally rooted in heuristic analysis and subjective interpretation, recent research has sought to formalize and validate these patterns using computational and statistical methodologies. There are more than 103 unique candlestick patterns that were defined using first-order logic rules and fuzzy linguistic variables [11].

This research aims to verify whether candlestick patterns could be used as an additional input feature to improve the accuracy of the LSTM model for stock market forecasting.

Methodology

The methodology used was based on training the same LSTM model with varying amounts of input features.

The LSTM model was configured once for all experiments and was not optimized for solving a particular task to prevent discrepancies in the results. The LSTM model consists of three LSTM layers, each with 60 units, followed by a dropout layer to avoid overfitting and a dense output layer that predicts the closing price. The model was compiled using the Adam optimizer with an initial learning rate of 0.001 and trained using Mean Squared Error (MSE) as the loss function.

The dataset spans from 2020 to 2025, ensuring the inclusion of various market conditions, including bullish, bearish, and volatile periods.

The data was preprocessed to fill in missing values, normalize scales, and structure it into sequential input for LSTM networks.

Experiments were conducted by varying the input feature set. Two input feature sets were used to compare. The first set contains only the closing price as an input feature for the LSTM model. The second set contains close prices and candlestick patterns. The candlestick patterns were identified through rule-based classification, enabling the encoding of patterns into binary vectors that represent indicators for future increasing or decreasing trends, which are also referred to as bullish and bearish flags. The description of the input feature sets for the LSTM model is presented in Table 1.

Table 2 contains the list of the candlestick pattern names that were used in Model A. Model B contains the same candlestick patterns as Model A, but also an additional eight patterns that are listed in Table 3.

Table 1 – **Experiments' input feature set**

N	Name	Description
1	Baseline Model	Only the closing price
2	Model A	Close price and 5 bearish and 5 bullish candlestick patterns
3	Model B	Model A + 4 more bearish and 4 more bullish candlestick patterns

Table 2 – **Candlestick pattern names used in Model A**

Bearish flag	Bullish flag
Evening Star	Hammer
Dark Cloud Cover	Morning Star
Hanging Man	Engulfing Pattern
Harami Pattern	Piercing Pattern
Shooting Star	Three Advancing White Soldiers

Table 3 – **Extra candlestick pattern names used in Model B**

Bearish flag	Bullish flag
Three Black Crows	Dragonfly Doji
Gravestone Doji	Inverted Hammer
Counterattack	Three Inside Up/Dow
Advance Block	Hikkake Pattern

The model performance was evaluated using three key metrics:

- Root mean squared error (RMSE);
- Mean Absolute Error (MAE);
- Mean Absolute Percentage Error (MAPE).

Results

The observed results are represented in Table 4.

Table 4 – **Experiments' results**

Configuration	RMSE	MAE	MAPE (%)
Baseline Model	0.1755	0.1422	19.33
Model A	0.1705	0.1399	18.63
Model B	0.1687	0.1380	18.48

According to the numbers, Model B has the smallest RMSE, MAE, and MAPE error metrics.

Discussions

Based on the observed results, adding the candlestick-related features to the input layer for the LSTM model consistently improves performance.

The best overall performance was achieved by Model B, which incorporates a total of 18 separate bullish and bearish candlestick features, likely capturing more of the trend movements.

Although the LSTM model itself was not optimized for solving this problem, and only the input feature sets were modified, it still improved the efficiency of predicting trend dynamics for the selected data.

Feature research

There are several directions for the feature research related to improving the accuracy of stock market forecasting.

The first direction could involve checking other financial indicators that can be used to enhance the input layer features and improve the quality of the neural network model training.

The second promising direction involves fine-tuning the LSTM model architecture through advanced hyperparameter optimization methods such as grid search, Bayesian optimization, or nature-inspired algorithms like Genetic Algorithms.

By systematically exploring model configurations, researchers can gain a deeper understanding of the relationship between structure and performance.

Lastly, extending the model to support multi-output forecasting – predicting not just closing prices but also volatility, trend direction, or trading signals – could make the system more practical for investors and analysts.

By framing the forecasting task more broadly, future systems can offer a more comprehensive picture of market dynamics.

REFERENCES

1. Telukdarie, A., & Mungar, A. (2022). *The Impact of Digital Financial Technology on Accelerating Financial Inclusion in Developing Economies*. https://doi.org/10.1016/j.procs.2022.12.263.
2. Basuki, S. A., Nahar, A., & Ridho, M. (2017). *Conservatism Accountancy, Profit Persistence and Systematic Risk Towards the Earnings Responses Coefficient*. https://doi.org/10.1016/j.procs.2022.12.263.
3. Strader, T. J., Rozycki, J. J., Root, T. H., & Huang, Y. H. J. (2020). Machine learning stock market prediction studies: Review and research directions // *Journal of International Technology and Information Management*. – Vol. 28, No. 4. – P. 63–83. https://doi.org/10.58729/1941-6679.1435.
4. Parmar, I., Agarwal, N., Saxena, S., Arora, R., Gupta, S., Dhiman, H., & Chouhan, L. (2018). Stock market prediction using machine learning // *Proceedings of the 1st International Conference on Secure Cyber Computing and Communication (ICSCCC)*. – IEEE, 2018. – P. 574–576. DOI**:** 10.1109/ICSCCC.2018.8703332.
5. Najem, R., Amr, M. F., Bahnasse, A., & Talea, M. (2023). *A Comprehensive Analysis of Techniques and Case Studies* [Electronic resource]. – Available at: https://www.sciencedirect.com/science/article/pii/S1877050923022056
6. Бовчалюк, С. Я., & Гайдай, Я. А. (2024). Аналіз методу опорних векторів у порівнянні з традиційними методами передбачення ринкових рухів // *Системи управління, навігації та зв'язку*. – 2024. – № 2(72). – С. 78–84. https://doi.org/10.26906/SUNZ.2024.3.089.
7. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory // *Neural Computation*. – Vol. 9, No. 8. – P. 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735
8. Burak Gülmez (2023). Stock price prediction with optimized deep LSTM network with artificial rabbits optimization algorithm. ScienceDirect. https://doi.org/10.1016/j.eswa.2023.120346.
9. John Kamwele Mutinda, Amos Kipkorir Langat (2024). Stock price prediction using combined GARCH-AI models. ScienceDirect. https://doi.org/10.1016/j.sciaf.2024.e02374.
10. Xinyuan Songet (2023). Predicting stock price of construction companies using weighted ensemble learning. ScienceDirect. https://doi.org/10.1016/j.heliyon.2024.e31604.
11. Weilong Hu, Yain-Whar Si, Simon Fong, Raymond Yiu Keung Lau (2019). *A formal approach to candlestick pattern classification in financial time series.* Soft Computing Journal. https://doi.org/10.1016/j.asoc.2019.105700.

Received (Надійшла) 25.06.2025
Accepted for publication (Прийнята до друку) 08.10.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Заковоротний Олександр Ігорович – аспірант Кафедри цифрових технологій в енергетиці, Національний технічний університет України «Київський політехнічний інститут імені І. Сікорського, Київ, Україна;
Oleksandr Zakovorotnyi – PhD Student, Department of Digital Technologies in Energy, National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute», Kyiv, Ukraine;
e-mail: alexzakovorotny@gmail.com; ORCID Author ID: https://orcid.org/0009-0000-7832-6957.

Аушева Наталія Миколаївна – доктор технічних наук, професор, завідувачка кафедри цифрових технологій в енергетиці, Національний технічний університет України «Київський політехнічний інститут імені 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: nataauscheva@gmail.com; ORCID Author ID: http://orcid.org/0000-0003-0816-2971;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57210707106.

Левченко Лариса Олексіївна – доктор технічних наук, професор, професор кафедри цифрових технологій в енер-гетиці, 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.

Покращення прогнозування фондового ринку за допомогою штучного інтелекту з використанням свічкових патернів

О. І. Заковоротний, Н. М. Аушева, Л. О. Левченко

Анотація. У швидко розвиваючій цифровій економіці використання штучного інтелекту (ШІ) у фінансовому прогнозуванні набуває значної популярності. Ця робота досліджує вплив різних патернів свічок на ефективність моделей довгострокової пам'яті (LSTM) у прогнозуванні рухів фондового ринку. Експерименти, проведені на історичних даних цін на акції, показують, що доповнення традиційних вхідних параметрів діапазоном моделей свічок підвищує точність прогнозування моделей LSTM. Хоча початковій архітектурі моделі бракувало оптимізації гіперпараметрів для вирішення такого роду завдань, результати дослідження свідчать про помітне покращення ефективності прогнозування, якщо використовувати вектор патернів свічок як вхідний параметр. Подальша робота буде зосереджена на включенні додаткових фінансових показників до навчальних даних моделі та її точному налаштуванні за допомогою алгоритмів оптимізації для досягнення більшої стійкості та точності.

Ключові слова: штучний інтелект, ефективність моделі LSTM, прогнозування фондового ринку, патерни свічок, підвищення точності моделі ШІ.

УДК 004. 8:004.9 **doi: 10.26906/SUNZ.2025.4.78-81**

О. Ю. Заковоротний, О. А. Сапальський

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

ПРОБЛЕМИ DATA-DRIVEN ПІДХОДУ У ФРОНТЕНД-РОЗРОБЦІ

Анотація. Data-Driven-підхід став домінуючою практикою у фронтенд-розробці. Попри зручність декларативного підходу, ускладнення систем призвело до виявлення ряду архітектурних недоліків. Основними з них є втрата прозорості джерел змін, асинхронні конфлікти при оновленні стану, а також відсутність семантичного контексту подій. У статті порівнюються Data-Driven та Event-Driven підходи, досліджуються ключові проблеми першого і пропонуються практичні рішення, здатні покращити контроль над логікою застосунків. **Мета** цієї роботи є виявлення типових архітектурних недоліків, властивих Data-Driven моделі в клієнтській розробці, аналіз причин їх виникнення та розробка концептуальних і технічних шляхів мінімізації відповідних ризиків. Особливу увагу приділено проблемам неявної мутації стану, втрати контексту змін, синхронізації запитів і залежності від життєвого циклу компонентів у фреймворк-орієнтованих системах. **Отримані наступні результати:** У результаті проведеного дослідження було встановлено, що Data-Driven підхід у складних застосунках не забезпечує достатньої контрольованості над джерелами змін. Також було доведено, що навіть при використанні інструментів типу Redux DevTools або React Developer Tools розробник часто не має повної картини змін стану, оскільки вони відбуваються в різних точках системи без єдиного шляху контролю. Встановлено, що найбільш ефективними компенсаторними підходами є створення шару семантичних подій, централізація мутацій, а також комбінування реактивного моделювання з декларативним представленням. **Висновки.** Data-Driven архітектура значно спрощує побудову UI в умовах простих або середньої складності проектів. Проте при зростанні кількості джерел стану, складності взаємозв’язків між компонентами і високому рівні асинхронності така модель демонструє структурні обмеження. У таких умовах доцільним є перехід до гібридних рішень, які поєднують Data-Driven рендеринг з Event-Driven семантикою та контролем через єдині точки мутації.

Ключові слова: Data-Driven, Event-Driven, frontend, стан, мутація, асинхронність, реактивність.

Вступ

У сучасній фронтенд-розробці поширений Data-Driven підхід, за яким інтерфейс користувача визначається внутрішнім станом даних. Цю ідею часто формулюють рівнянням UI = f(state) – інтерфейс є функцією від стану програми [1]. Зміни в стані автоматично відображаються у UI (реактивність), що звільняє розробника від ручного оновлення DOM-елементів. Такий декларативний підхід лежить в основі бібліотек на кшталт React.

Натомість Event-Driven підхід побудовано навколо подій: логіка додатка реагує на дискретні події користувача або системи.

У Event-Driven архітектурі стан визначається послідовністю подій, а не навпаки.

Іншими словами, у Data-Driven сценарії події виступають лише наслідком зміни даних, тоді як у Event-Driven сценарії вони самі керують змінами стану і поведінкою UI [2].

Кожен підхід має свої переваги, але й породжує унікальні архітектурні проблеми. У статті розглянуто три ключові проблеми Data-Driven підходу у фронтенд-розробці та можливі шляхи їх вирішення.

Постановка проблеми

У Data-Driven системі оновлення стану можуть відбуватися неявно з різних місць, що ускладнює відстеження джерела змін. Якщо кожен компонент або модуль «вільно» модифікує спільний стан, потік даних стає непередбачуваним. Без єдиного центру керування важко відповісти на запитання: що саме спричинило зміну стану? Як зазначають розробники, якщо всі компоненти будуть змінювати state як заманеться, в додатку може виникнути хаотична ситуація. JavaScript виконує код переважно асинхронно, тож різні частини програми теоретично можуть намагатися змінити одну і ту ж змінну одночасно, що призводить до умов гонки та непередбачених багів. Навіть коли конкурентні зміни трапляються рідко, сама можливість неявних мутацій ускладнює підтримку: розробник не впевнений, який код і коли саме змінив стан.

У реактивному інтерфейсі часто паралельно відбуваються кілька асинхронних операцій, які змінюють стан [3]. Гонки асинхронних оновлень (race conditions) виникають, коли результати таких операцій приходять у невизначеному порядку і накладаються один на одного. У Data-Driven підході UI автоматично відображає останній стан, але “останній” не завжди означає правильний [4]. Класичний сценарій – користувач ініціює два запити до сервера майже одночасно (наприклад, швидко вводить різні пошукові запити). Перший запит може повернутися пізніше другого і перезаписати дані, незважаючи на те, що вже є новіший результат. Як наслідок, інтерфейс може показати неактуальну інформацію, не відповідаючи поточним діям користувача. Розробники зазначають, що якщо існує навіть невелика ймовірність отримати такий стан (невідповідність між намірами користувача і відображеним результатом), додаток вже схильний до race condition-багів [5]. Причини цього явища криються в непередбачуваності мережевих затримок, різному часу виконання промісів та інших асинхронних процесів. Порядок завершення запитів не гарантується: останній запит може виконуватися раніше попереднього або взагалі зазнати помилки. У кращому разі, гонки призводять до мерехтіння застарілих даних в UI; у гіршому – до логічних помилок [6]. В окремих доменах це може мати критичні наслідки: наприклад, користувач може випадково купити не той товар, або лікар – призначити не

той препарат, якщо інтерфейс відобразив несвоєчасну інформацію.

Data-Driven підхід зосереджений на стані, але не фіксує сенс зміни цього стану. Іншими словами, він відповідає на питання «що змінилося?», але не завжди «чому змінилося і що це означає?». Коли оновлення UI відбувається автоматично від зміни даних, втрачається контекст: чи була ця зміна спричинена дією користувача, чи приходом нових даних із сервера, чи внутрішньою подією? Відсутність семантичних міток для змін утруднює як відлагодження, так і розширення функціоналу. Пов'язана проблема – брак централізованого каналу для реакцій на події. У Event-Driven підході зазвичай є шина подій або диспетчер, куди можна підписати різні частини системи на отримання повідомлень про певні події.

У Data-Driven UI такої єдиної шини немає: компоненти безпосередньо реагують на зміну стану (наприклад, через двобічне зв'язування), але якщо потрібно виконати якусь глобальну дію при конкретній зміні – це складно організувати. Іншими словами, бракує центральної точки контролю, де можна було б перехопити та осмислити зміни перед тим, як вони розійдуться по UI.

Метою цієї роботи є виявлення типових архітектурних недоліків, властивих Data-Driven моделі у фронтенд-розробці, аналіз причин їх виникнення та розробка концептуальних і технічних шляхів мінімізації відповідних ризиків.

Особливу увагу приділено проблемам неявної мутації стану, втрати контексту змін, синхронізації запитів і залежності від життєвого циклу компонентів у фреймворк-орієнтованих системах.

Основна частина роботи

Головна причина – відсутність єдиного джерела правди та явних точок входу для змін.

У традиційній Event-Driven моделі кожна значуща зміна ініціюється подією (наприклад, обробник кліку кнопки), і розробник знає, що саме цей обробник змінює стан. В Data-Driven моделі зміна стану може бути спричинена будь-де: HTTP-викликом, результатом промісу, побічним ефектом хуку або сторонньою бібліотекою. Якщо такі зміни не централізовані, з часом код стає важчим для підтримки та тестування.

Один із способів розв'язати проблему – запровадити сувору дисципліну оновлення стану через визначені інтерфейси (рис. 1). У централізованих сховищах стану типу Redux або Vuex всі зміни проходять через спеціальні методи (екшени/мутації), що робить їх явними.

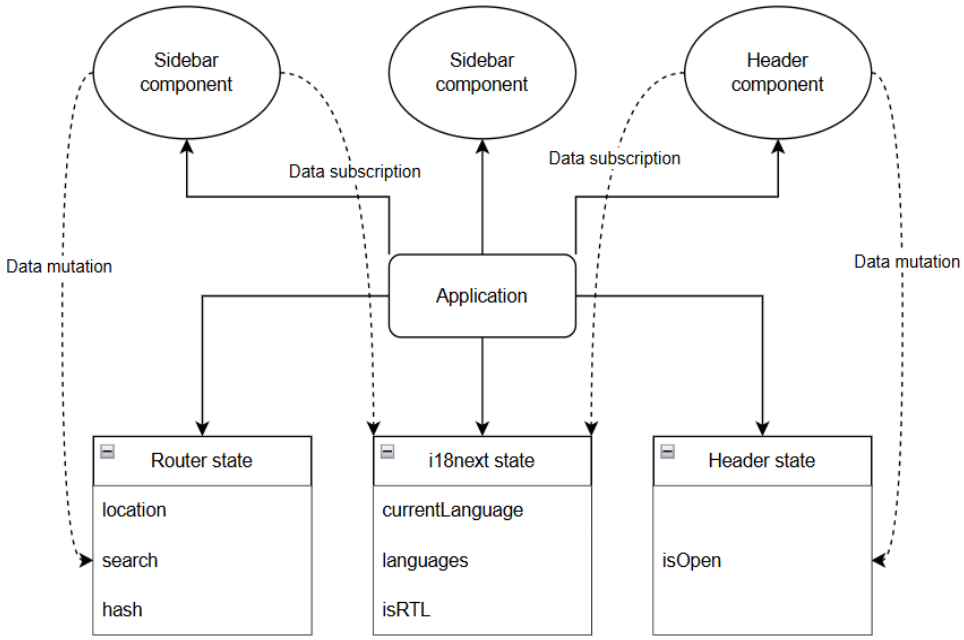


Рис. 1. Приклад схеми використання даних у додатку

Наприклад, у Redux прийнято ніколи не змінювати стан напряму – натомість, слід відправляти action з описом змін, який обробляється reducer-функцією. Такий уніфікований потік (state → action → reducer → new state) гарантує, що кожна зміна здійснюється під контролем розробника і може бути відстежена у коді [7]. Аналогічно, у Vuex рекомендовано змінювати стан лише через коміти мутацій – дотримання цього правила означає, що Ви завжди знатимете, де саме у застосунку змінюються дані. Іншими словами, потрібно уніфікувати джерело змін. Практично це реалізується через централізований стор або керування станом, куди всі компоненти відправляють запити на зміну (події або екшени) замість прямої мутації. Таким чином усувається "невидимість" змін: кожна модифікація проходить через явний канал, що спрощує відлагодження і виключає побічні ефекти.

Основна причина гонок – необмежений паралелізм оновлень. Data-Driven підхід передбачає, що UI завжди відповідає актуальному стану, але не визначає, як вирішувати конфлікти, коли "актуальність" спірна. В браузері JavaScript виконується в одному потоці, але асинхронні дії (запити, таймери) плануються у черзі і можуть завершуватися у будь-якому порядку. Якщо

на один і той самий фрагмент стану одночасно претендують кілька результатів, стандартний механізм просто відображає останній записаний результат – який може виявитися застарілим. Race condition-помилки важко виявити під час розробки, бо на локальному швидкому з'єднанні і послідовних діях користувача вони майже не проявляються. Проблема проявляється у реальних умовах: при повільній чи нестабільній мережі, високому навантаженні серверу або просто при дуже швидких діях користувача. Отже, потрібно спеціально проєктувати фронтенд-логіку, щоб запобігати таким конкурентним ситуаціям. Таким чином, на практиці можна ігнорувати застарілі запити за допомогою спеціальних міток timestamp. Як альтернатива можливе гнучке управління запитами за допомогою класу AbortControler, який дозволяє скасовувати запит або навіть цілі групи у потрібний момент [8].

До Data-Driven архітектури часто додають елементи подієво-орієнтованого керування. Зокрема, впроваджується поняття семантичних дій (action) як доповнення до простої мутації стану.

Замість безпосереднього оновлення даних, компоненти генерують події з бізнес-сенсом (напр., USER_DELETED_ITEM), а система вже вирішує, як змінити стан і які побічні ефекти виконати у відповідь. Такий підхід реалізований у Redux: кожна дія (action) – це простий об’єкт з полем type (семантика змін) та додатковими даними. Важливо, що всі дії проходять через централізований диспетчер (store), де можуть бути зафіксовані та оброблені. Це дає одразу дві переваги: явний журнал змін і можливість розширених реакцій. У Redux/Flux кожна дія може бути записана і відтворена пізніше, що робить систему детермінованою та піддатливою до дебагу – маючи послідовність action-ів, можна завжди відновити певний стан UI. Застосунок стає прозорішим: розробник точно знає, яка подія що змінює, і може прослідкувати історію (наприклад, через Redux DevTools, де всі екшени логуються з можливістю покрокового перегляду). Подібний механізм існує і у Vuex: всі коміти мутацій реєструються, і можна «прокрутити» історію змін стану покроково, оцінюючи, які дії до них призвели [9]. Таким чином досягається семантична трасованість: кожна зміна має мітку і може бути проаналізована централізовано. Подібні методи можуть значно підвищити якість коду та його надійність. Наприклад, нижче можна побачити приклад використання Zustand у зв'язці з Immer (рис. 2). Така комбінація дозволяє писати простий та зрозумілий код, який неможливо мутувати безпосередньо ззовні.

Стан неможливо змінити ззовні. Усі зміни прозорі та централізовані. Налагодження та відтворюваність стають простими. Будь-яка спроба змінити цей стан буде призводити до помилки (рис. 3).

```
const useCounter = create<State & Actions>()(
  immer((set) => ({
    count: 0,
    nested: [{ name: "Alex" }],

    increment() {
      set((state) => {
        debugger; // Tracks any state change
        state.count++;
      });
    },
  }))
);
```

Рис. 2. Захищений стан

```
function App() {
  const counterSlice = useCounter();

  counterSlice.nested[0].name = "Direct mutation";
  console.log(counterSlice.nested[0].name);
```

```
▶Uncaught TypeError: Cannot assign to read only property 'name' of object '#<Object>'    App.tsx:36
    at App (App.tsx:36:26)
    at renderWithHooks (react-dom.development.js:15486:18)
    at updateFunctionComponent (react-dom.development.js:19617:20)
```

Рис. 3. Помилка при оновленні стану

Висновки та перспективи подальших досліджень

Data-Driven підхід у фронтенді значно спрощує побудову інтерфейсів і усуває потребу в імперативній маніпуляції DOM. Однак його впровадження висвітлює низку архітектурних проблем. По-перше, неявні джерела змін стану ускладнюють підтримку – вирішити це можна впровадженням чітких правил оновлення через єдиний канал (екшени, мутації), що робить потік даних прозорим. По-друге, гонки асинхронних оновлень можуть призводити до неконсистентного UI – для їх запобігання застосовують техніки фільтрації та скасування паралельних запитів, гарантуючи обробку тільки актуальних даних. По-третє, брак семантики змін та централізованої системи подій

позбавляє систему контексту і гнучкості – ця проблема розв'язується доповненням Data-Driven моделі шаром подій, де кожна зміна маркується і обробляється узгоджено (як у Flux/Redux архітектурі). Загальний висновок: реактивні інтерфейси доцільно поєднувати з принципами подієвоорієнтованого проектування, щоб отримати найкраще з обох світів. UI = f(state) залишається потужною концепцією, що забезпечує передбачуваний рендеринг , проте навколо неї варто побудувати інфраструктуру для явного керування змінами. Академічні і практичні джерела сходяться на тому, що правильна організація потоків даних і подій підвищує передбачуваність та керованість фронтенд-систем [10]. Дотримуючись описаних підходів – єдиний потік змін, контроль асинхронності, семантичні екшени – розробники можуть знизити складність, мінімізувати баги та полегшити розвиток додатків, заснованих на Data-Driven парадигмі.

Список літератури

1. Flutter. Flutter Architectural Overview. Flutter Documentation. URL: https://docs.flutter.dev/resources/architectural-overview
2. Beyer D., Ghanbari H., Hasselbring W. An empirical characterization of event-sourced systems and their schema evolution—Lessons from industry. Journal of Systems and Software. 2021. Vol. 182:110931. DOI: 10.1016/j.jss.2021.110931. URL: https://www.sciencedirect.com/science/article/pii/S0164121221000674
3. Wikipedia. Reactive programming. URL: https://en.wikipedia.org/wiki/Reactive_programming
4. Yee M.-H., Badouraly A., Lhoták O., Tip F., Vitek J. Precise Dataflow Analysis of Event-Driven Applications. arXiv. DOI: 10.48550/arXiv.1910.12935. URL: https://arxiv.org/abs/1910.12935
5. E. Mutlu et al. Detecting JavaScript Races that Matter. FSE. URL: https://www.doc.ic.ac.uk/~livshits/papers/pdf/fse15.pdf
6. MDN. JavaScript execution model. URL: https://developer.mozilla.org/en-US/docs/Web/JavaScript/Reference/Execution_model
7. Redux. Thinking in Redux: Three Principles. Redux Documentation. URL: https://redux.js.org/understanding/thinking-in-redux/three-principles
8. WHATWG. DOM Standard — Aborting ongoing activities (AbortController). WHATWG. URL: https://dom.spec.whatwg.org/
9. Vuex. Committing Mutations in Components. Vuex Guide – Mutations. URL: https://vuex.vuejs.org/guide/mutations.html#committing-mutations-in-components
10. Fowler M. Patterns of Enterprise Application Architecture. Addison-Wesley Professional, 2003. Розділ 5, с. 63, available at: https://martinfowler.com/books/eaa.html

Received (Надійшла) 12.08.2025
Accepted for publication (Прийнята до друку) 22.10.2025

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 the Computer Engineering and Programming Department, National Technical University «Kharkiv Polytechnic Institute», Kharkiv, Ukraine;
e-mail: Oleksandr.Zakovorotnyi@khpi.edu.ua; ORCID Author ID: https://orcid.org/0000-0003-4415-838X;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57201613700.

Сапальський Олександр Андрійович – аспірант кафедри "Комп'ютерна інженерія та програмування", Національний технічний університет "Харківський політехнічний інститут", Харків, Україна;
Oleksandr Sapalskyi – PhD student, Department of Computer Engineering and Programming, National Technical University 'Kharkiv Polytechnic Institute', Kharkiv, Ukraine;
e-mail: Oleksandr.Sapalskyi@cs.khpi.edu.ua; ORCID Author ID: https://orcid.org/0009-0002-5749-8527.

Solving productivity problems using asynchronous methods in JavaScript

Oleksandr Zakovorotnyi, Oleskandr Sapalskyi

Abstract. The Data-Driven approach, in which the interface is defined as a function of the current state (UI = f(state)), has become the dominant practice in front-end development. Despite the convenience of the declarative approach, the complexity of systems has led to the identification of a number of architectural shortcomings. The main ones are the loss of transparency of the sources of changes, asynchronous conflicts when updating the state, and the lack of semantic context of events. The article compares the Data-Driven and Event-Driven approaches, explores the key problems of the former, and offers practical solutions that can improve control over application logic. The **purpose** of this work is to identify typical architectural shortcomings inherent in the Data-Driven model in client-side development, analyze the causes of their occurrence, and develop conceptual and technical ways to minimize the corresponding risks. Special attention is paid to the problems of implicit state mutation, loss of change context, query synchronization, and dependence on the component life cycle in framework-oriented systems. The following **results were obtained**: As a result of the study, it was found that the Data-Driven approach in complex applications does not provide sufficient control over the sources of changes. It was also proven that even when using tools such as Redux DevTools or React Developer Tools, the developer often does not have a complete picture of state changes, since they occur at different points in the system without a single control path. It was found that the most effective compensatory approaches are the creation of a semantic event layer, centralization of mutations, and the combination of reactive modeling with declarative representation. **Conclusions**. Data-Driven architecture significantly simplifies UI construction in the conditions of simple or medium-complexity projects. However, with an increase in the number of state sources, the complexity of the relationships between components, and a high level of asynchrony, such a model demonstrates structural limitations. In such conditions, it is advisable to switch to hybrid solutions that combine Data-Driven rendering with Event-Driven semantics and control through single mutation points.

Keywords: Data-Driven, Event-Driven, frontend, state, mutation, asynchrony, reactivity.

UDC 532.327 **doi: 10.26906/SUNZ.2025.4.82-87**

В. Д. Карлов[1], О. В. Коломійцев[2], О. Л. Кузнєцов[1], О. В. Бєсова[1], А. О. Бєсова[2]

[1] Харківський національний університет Повітряних Сил імені Івана Кожедуба, Харків
[2] Національний технічний університет "Харківський політехнічний інститут", Харків

ОЦІНКИ АСИМПТОТИЧНОЇ СКЛАДНОСТІ ПРИКЛАДНИХ МНОЖИННИХ ТРАНСПОРТНИХ АЛГОРИТМІВ ПРИ ПОДІЛІ ЇХ НА КЛАСТЕРИ

Анотація. Прикладні множинні транспортні алгоритми, як правило, є евристичними та надскладними. За обраною евристикою алгоритму, реальний час на програмну реалізацію залежить від мови програмування, структури представлення вхідних даних та алгоритмічної складності рішення. Дослідження базуються на використанні положень дискретної математики та теорії графів. Показано, що шляхом аналізу асимптотичної складності можливо визначити ефективну евристику. Уточнення моделей прикладних завдань, також, сприяють зменшенню складності. Оскільки, множинність транспортних завдань міститися у дискретному покритті або у функціоналі множинних транспортних засобів, то можливим є спрощення алгоритму шляхом їх декомпозиції на кластери для яких діють умови та обмеження стандартного транспортного завдання. Наведено особливості геометрично та комбінаторного розкладання на кластери. Надано формалізацію алгоритмів побудови кластерів для множинних транспортних завдань. Проведено порівняльний аналіз оптимальних та евристичних алгоритмів вирішення множинних транспортних завдань. Визначено ієрархію асимптотичної складності евристичних алгоритмів та проаналізовано можливості їх декомпозицій на кластери за якими досягається спрощення вирішення множинних прикладних транспортних завдань. Надано нотації складності рішення та можливості їх оцінки за цільовою функцією обраного алгоритму або за її апроксимацією. Здійснено аналіз асимптотичної складності евристичних алгоритмів для Big-O нотації. Показано, що придатними для реалізації є алгоритми, складності яких не перевищують поліноміальної, а застосування алгоритмів, які мають експонентну оцінку складності та вище є не бажаним. Зазначено, що асимптотичні оцінки складності алгоритму придатні для великих за розмірністю завдань, а для завдань невеликої розмірності доцільними є контрольні прогони.

Ключові слова: асимптотична складність, евристика, кластери, логістичне забезпечення, множинні транспортні завдання, нотації, система невідкладної допомоги, транспортні засоби, цільова функція.

Вступ

Постановка проблеми. На сучасному етапі розвитку науки і техніки набуває актуальності оперативне і оптимальне розв'язання завдань багатоконфліктних ситуацій. Зокрема, до них відносяться такі глобальні ситуації як: природні катаклізми, пандемії, збройні конфлікти тощо. При цьому, їх успішне вирішення вимагає вдосконалення існуючих та створення новітніх систем логістичного забезпечення. У широкому сенсі логістика має на меті оптимальне управління матеріальними, інформаційними та фінансовими ресурсами в економічних адаптивних системах. У вузькому сенсі логістика є функцією системи щодо забезпечення переміщення та зберігання вантажу для виробництва, обміну та продажу. Ведення широкомасштабних бойових дій на території держави логістична галузь в цілому характеризується низькою проблем, до яких належать:

– зруйнована або пошкоджена інфраструктура;
– обмеження доступу до території;
– збільшення ризику втрат вантажу;
– зростання цін на логістичні послуги.

В даних надважких умовах найважливішим питанням є утримання логістичної галузі на потрібному рівні та максимальне сприяння її подальшому розвитку. Зокрема, це стосується створення системи невідкладної допомоги (СНД), на яку покладається вирішення наступних логістичних завдань:

– доставки гуманітарної допомоги та товарів першої необхідності в постраждалі від бойових дій райони;

– термінове постачання медикаментів для забезпечення лікування хворих;

– евакуацію людей з постраждалих від війни районів.

У випадку призупинення бойових дій або взагалі закінчення війни важливість СНД суттєво зростатиме, забезпечуючи відновлення інфраструктури держави та постачання сучасного обладнання для потреб її галузей.

Незважаючи на складні умови, логістична галузь держави продовжує розвиток. При цьому ключові напрямки розвитку логістики пов'язані насамперед з розвитком логістичної інфраструктури, а також покращенням функціоналу транспортних засобів (ТЗ) та систем їх керування.

Інтерес до множинних транспортних завдань mTSP (Multiple Delivery for Traveling Salesman Problem) пов'язано з їх широким застосуванням, зокрема для побудови найкоротших маршрутів множини вантажних ТЗ в умовах впливу вказаних вище заважаючих факторів.

У mTSP у якості множини може розглядатися покриття ТЗ або їх функціонал. Тобто множинні алгоритми, як правило, є спрямованими або на оптимізацію моделі дискретного покриття або на вибір множини ТЗ для довільного сценарію.

Представляє практичний інтерес розгляд стаціонарного дискретного покриття і можливості розподілу його на кластери, що не перетинаються при використанні множини ТЗ.

Аналіз останніх досліджень і публікацій. Аналіз проблеми маршрутизації ТЗ при доставці та самовивозі вантажу проведено у [1]. Зокрема у [2] визначено проблеми логістики стосовно закладів охорони здоров'я. Новітнім та перспективним підходом до оптимізації вирішення логістичних завдань є

застосування безпілотних літальних апаратів (БПЛА) [3]. При цьому практичний інтерес представляє собою застосування мурашиного алгоритму управління групою БПЛА [4], розглядати який доцільно з точки зору положень теорії [5]. Вказаним питанням присвячено роботи [6; 7], зокрема, у [7] проведено огляд проблем динамічної маршрутизації ТЗ. Відповідні фундаментальні положення теорії сучасних алгоритмів викладено у [7; 8], а особливості мурашінних алгоритмів при вирішенні транспортних завдань наведено у [9; 10]. Оптимізації вказаних алгоритмів присвячено роботе [11], при цьому стосовно множинних транспортних алгоритмів доцільним є визначення їх асимптотичної складності.

Метою статті є визначення асимптотичної складності прикладних множинних транспортних алгоритмів при поділі їх на кластери.

Виклад основного матеріалу

Розгляд зазначених питань базується на використанні положень дискретної математики та теорії графів. Теорія графів дозволяє визначати взаємозв'язки з урахуванням втрати вантажив та надає можливості щодо вибору оптимальних маршрутів при забезпеченні логістики та вдосконалення алгоритмів керування ТЗ. Послідовність вирішення завдання є такою:

1. За сценарієм поставленого завдання здійснюється його формалізація та створюється граф покриття [4; 5].

2. За результатами аналізу структури покриття визначається алгоритм вирішення.

3. Уточнюється структура покриття, статистичні характеристики переходів та вагові коефіцієнти втрат.

4. Аналізується складність алгоритму та визначається можливість переходу від множинної транспортної задачі до суперпозиції стандартних транспортних задач.

5. За обраного алгоритму здійснюється декомпозиції покриття на кластери.

6. На основі індивідуальних рішень для кожного кластеру виконується пошук загального рішення.

За відомим прикладним сценарієм для формалізації завдання важливим кроком є проведення аналізу геометричних характеристик покриття і функціоналу ТЗ, що дозволить здійснити вибір алгоритму рішення mTSP.

Можливий варіант послідовності аналізу прикладного сценарію наведено у табл. 1.

Таблиця 1 – **Послідовність аналізу прикладного сценарію множинних транспортних завдань [5; 7]**

1.	*Оцінка геометричних характеристик покриттть*			
	Кількість споживачів (N)	Кількість сховищ вантажу ©	Максимальна відстань між споживачами (Δr^{max})	Середня відстань між споживачами ($\Delta \bar{r}$)
2.	*Оцінка характеристик переходів між споживачами*			
	Можлива кількість переходів (N-1)	Привабливість $i \leftrightarrow j$ переходу ($\tau_{i \leftrightarrow j}$)	Доступність $i \leftrightarrow j$ переходу ($\eta_{i \leftrightarrow j}$)	Імовірності появи $i \leftrightarrow j$ переходу (P_τ, P_η)
3.	*Оцінка функціоналу ТЗ*			
	Кількість ТЗ (M)	Вантажопідйомність (q_m)	Дальність руху транспортного засобу на одної заправці ΔR_m	Вартість одного переходу $i \leftrightarrow j$ між споживачами ($d_{i \leftrightarrow j}$)
4.	*Синтез (вибір) алгоритму рішення завдання*			
	оптимальні (точні)	евристичні (наближені)	мета-евристичні	
5.	*Оцінка можливістю побудови кластерів покриття*			
6.	*Вибір алгоритму створення кластерів та оцінка їх характеристик*			
	кількість кластерів (K)	геометричні характеристики кластерів	характеристики переходів між споживачами	

Для вирішення mTSP існують точні (оптимальні) та наближені (евристичні) алгоритми [7].

Оптимальні алгоритми будуються за принципом повного перебору усіх переходів системи та їх характеристик і мають властивість давати максимально точний результат рішення mTSP.

Оптимальні алгоритми формалізовано на повних матричних систем рівнянь за якими будуються графи mTSP [7; 11]. Недоліком оптимальних алгоритмів є їхня надвелика складність, тому їх доцільно застосовувати лише для малорозмірних mTSP.

Для вирішення прикладних завдань генерують евристичні алгоритми, які мають певні обмеження та дають наближені рішення, але вони є менш складними та потрібують меншого часу на отримання рішень порівняно з точними алгоритмами. Евристичні алгоритми створюють за формальними процедурами [8; 9], тобто mTSP є завжди не повними, а їх евристичні рішення є не точними. Особливістю даних алгоритмів є їхня різноманітність, а їхня обмежена дія поширюється на початкові дані, час виконання завдання та точність рішення.

До основних евристичних алгоритмів відносяться наступні:

1. Евристичні екологічні алгоритми EAs (Evolutionary Algorithms), які ґрунтуються на принципах еволюції неживої природи.

До категорії EAs належать [12]:

– алгоритм завихрення повітря BWOAs (Back Wash Optimization Algorithms);

– алгоритм руху окремих частинок у хмарах PSOAs (Particle Swarm Optimization lgorithms).

2. Евристичні біологічні алгоритми BAs (Biology-inspired Algorithms), які моделюють процеси адаптації популяції живих істот та створюють рішення, що самооптимізуються. Це [12]: – генетичні алгоритми GAs (Genetic Algorithms), які моделюють зростання, розвиток, розмноження, відбір і виживання в умовах накопичення представників окремого виду; алгоритми рисового поля PFAs (Paddy Field Algorithms), що моделюють дії окремих складових популяції, які повторюються; алгоритм дії бджолинної колонії, що моделює дії бджолиного рою; алгоритм світлячка FAs (Firefly Algorithms), що

моделюють пошук комах; мурашиний алгоритм ACOAs (Ant Colony Optimization Algorithms), що моделюють сумісний пошук мурахами їжі [4]; алгоритм пошуку бактеріями джерел живлення BFAs (Bacteria Foraging Algorithms), що моделюють процес обміну у колонії бактерій.

Вибір конкретного виду алгоритму для вирішення mTSP здійснюється за можливістю формалізації покриття та складністю їх реалізації [5; 7].

У 2002 році було запропоновано формалізацію завдання маршрутизації mTSP у загальному просторі покриття при отриманні та доставці вантажу наступним чином [4; 8]. Апріорними даними є:

– геометрія покриття, яке має A_n ($n \in N$) реперних точок (РТ), координати яких є відомими;

– у покритті існують сховища вантажу © (© $\in N$) та пункт доставки вантажу ® (® $\in N$);

– існують m ($m \in M$) ТЗ ($M>1$), що знаходяться у сховищі вантажу ©.

Потрібно знайти найкращій перехід $R^{best}_{©\to®\to©}$, а усі проміжні РТ$_n$ потрібно покрити тільки один раз.

За дискретним покриттям можлива формалізація завдання за допомогою графу:

$$G = A \times R \times D = A \times V, \qquad (1)$$

де G – граф завдання; A – граф координат РТ; R – граф переходів; D – граф втрат для переходів ($r_{i\to j}$).

Всі m ТЗ знаходяться у сховищі вантажу ©. На рис. 1 надано приклад розподілу загального простору покриття mTSP на три кластери TSP1, TSP2 та TSP3, які не перетинаються для одного сховища вантажу ©.

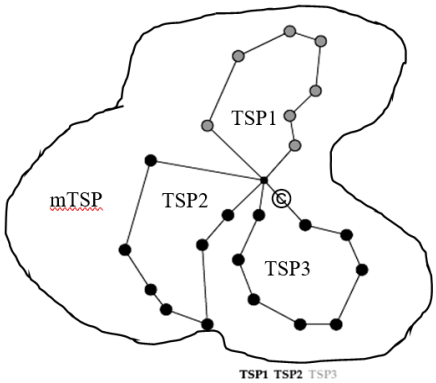


Рис. 1. Приклад розподілу загального простору покриття mTSP на три кластери, які не перетинаються для одного сховища вантажу

Цільова функція f завдання (1) передбачає пошук найкращих переходів всіх ТЗ $R^m_{©\to®\to©}$, які починаються у ©,проходять ®, покривають всі РТ та закінчуються у ©. Усі поточні РТ розподілено за конкретних маршрутів ТЗ, тому:

$$R^m_{©\to®\to©} \Rightarrow \min \sum_{n=1}^{N} \left(r^m_{©\to®_n\to©} \right),$$

$$D^m_{©\to®\to©} \Rightarrow \min \sum_{n=1}^{N} \left(d^m_{©\to®_n\to©} \right), \qquad (2)$$

$$V^m_{©\to®\to©} \Rightarrow \min \sum_{n=1}^{N} \left(v^m_{©\to®_n\to©} \right).$$

Якщо ТЗ знаходяться на переході $i \to j$, то $\left(d^m_{i\to j} \right) = 1$, а для протилежного випадку $\left(d^m_{j\to i} \right) = 0$ виконується умова:

$$S^m_K \le n_i \le S^m_L, \qquad (3)$$

де: n^m_i – кількість РТ покритих m ТЗ до покриття РТ$_i$; S^m_L – максимальна можлива кількість покриття РТ, S^m_K – мінімальна кількість покриття РТ для m-го транспортного засобу. Для цього випадку формалізацію mTSP можливо представити у вигляді:

$$\begin{cases} \sum_{j=2}^{n_i} v^m_{1\to j} = \sum_{j=2}^{n_i} v^m_{j\to 1} = m, \\ \sum_{i=1}^{n} v^m_{i\to j} = 1, \quad j = [2...n], \quad \sum_{j=1}^{n} v^m_{i\to i} = 1, \quad i = [2...n], \\ v^m_{©\to j} + (S_L - 2) v^m_{1\to i} - v^m_{i\to 1} \le (S_L - 1), \\ v^m_{©\to i} + (2 - S_K) v^m_{i\to 1} + v^m_{1\to i} \ge 2, \quad i = [2...n], \\ v_{1\to i} + v_{i\to 1} \le 1, \quad i = [2...n], \\ v^m_{i\to j} \in \{0,1\}, \forall i, j \in N, \end{cases} \qquad (4)$$

а цільова функція f має вигляд:

$$\begin{cases} f(m,d,r) = \min \sum_{i,j\in N} d^m_{i\to j} \times r^m_{i\to j}, \\ f(m,r) = \min \sum_{i,j\in N} r^m_{i\to j}, \\ f(m,d) = \min \sum_{i,j\in N} d^m_{i\to j}, \\ f(m,v) = \min \sum_{i,j\in N} v^m_{i\to j}. \end{cases} \qquad (5)$$

За виконанням (5) можливою є мінімізація довжини маршрутів (переходів), а також мінімізація втрат, що залежить від кількості РТ (N) та кількості ТЗ (M). Умова виконання (4) визначає можливість розподілу загального простору покриття на K кластерів для яких діють умови та обмеження стандартного завдання TSP [8]. Враховуючи прикладні сценарії СНД множинна задача логістики, її формалізація та обмеження є наступними [5]:

– існує логістична система, геометрія якої має ($n \in N$) окремих РТ, які потрібно обслужити (покрити). Параметр n показує номер поточної РТ;

– логістична система обслуговується ($m \in M$) однотипними ТЗ, які мають однакову вантажопідйомність (q_m) і дальність руху на одній заправці ΔR_m. Параметр m показує дійсну кількість ТЗ, що застосовуються для обслуговані окремої РТ;

– доставка вантажу здійснюється за i-м ($i \in I$) запитом. Параметр i показує номер запиту на транспортування вантажу у РТ;

– для кожної *i*-ї реалізації запиту вантаж q_i має бути доставлений з РТ n_i у РТ n_j^+ або у РТ n_j^-. Позитивне значення n_j^+ відповідає вивезення вантажу, а негативне значення n_j^- для його ввезення у *j*-ту РТ. Підпростори n_i^+, n_i^- та n_j^+, n_j^- не перетинаються, оскільки $n_{i,j}^+ \in N^+$, а $n_{i,j}^- \in N^-$, а простори N^+, N^- для усіх поточних i та j одночасно не існують. Параметр *j* показує набір реалізованих запитів на транспортування вантажу у РТ;

– граф покриття СНД G має N вершин (РТ) та (N-1) реберець (переходів), для якого $n_i^+ = n_i^-$. Вершини n_i^+ має бути відвідані раніше, ніж n_i^-;

– кожен з m ТЗ здійснює рух за маршрутом $R^m{}_{©\to®\to©}$, тобто починає свій шлях у сховище ©, досягає ®, покриває інші РТ і повертається до ©;

– привабливість кожного переходу $\tau_{i\to j}$ є оберненим щодо його довжини, $\tau_{i\to j} = 1/r_{i\to j}$;

– доступність кожного переходу $\eta_{i\to j}$ залежить від можливих втрат вантажу $d_{i\to j}$, тобто визначаються якістю (оперативною обстановкою) на переході $r_{i\to j}$;

– маршрути M ТЗ покривають усі РТ усіх K кластерів простору покриття. Кількість кластерів не перевішує кількість ТЗ ($K \le M$);

– якщо множина РТ $n_i{}^+$ належить до N, то *i*-та множина РТ n_i^- також належить до N і навпаки;

– завантаженість кожного ТЗ у час відбуття з РТ не перевищує завантаженості у час прибуття у РТ, тобто $q_m{}^+ \le q_m{}^-$.

Таким чином ефективним способом вирішення mTSP є створення кластерів, тобто розкладання загального графу G за стрічками, стовбцями та окремими майданчиками покриття різної форми (геометричне розкладання) або за відповідними правилами розподілу (комбінаторне розкладання).

Розподіл загального простору покриття на K кластерів дозволяє здійснювати паралельні обчислення G^k ($k \in K$) та забезпечувати наближення множинних транспортних алгоритмів до стандартних [3].

Формалізація алгоритмів побудови кластерів для множинних транспортних завдань може бути надана наступним чином [11].

Заданим вважається граф покриття G= (A,R,D), що має РТ A_n, де кожному поточному переходу $r_{i\to j}$ відповідає вага $d_{i\to j}$. Потрібно зробити розбиття G на K кластерів, тобто розподіл A_n на підмножини $A_n{}^k$, які не перетинаються та мають максимально близьку сумарну вагу або мінімальну сумарну довжину переходів, що покривають отримані підмножини РТ $A_n{}^k$ [11].

Геометричні алгоритми розкладання (розбиття) генеруються згідно координатної інформації A_n та $r_{i\to j}$ покриття графа G. Дані алгоритми відносяться до малоскладних, однак вони не приймають до уваги інформацію щодо ваги $d_{i\to j}$ та не можуть явно привести до мінімізації шляхів та мінімізації сумарних втрат для кожного окремого кластеру.

До геометричних методів розбиття відносять:

– рекурсивний інерційний метод "розподілу навпіл", для якого можливо застосувати рекурсивний багатобінарний метод поділу, що дозволяє на першій ітерації поділіти граф на дві частини, на другій ітерації кожна з отриманих частин також розбивається на дві частини і так далі;

– "по-координатне розбиття" на першій ітерації вимагає побудови головної вісі симетрії покриття та лінії ортогональної отриманої вісі. На наступних ітераціях РТ кожного підпростору проектуються на максимальну вісь і далі процедура повторюється;

– поділ мереж за використанням кривих Пеано. Криві Пеано є кривими, що повністю заповнюють фігури великих розмірностей (наприклад, квадрат або трикутник), тому кластери мають відповідну форму.

За обраним алгоритмом, реальний час програмної реалізації залежить від мови програмування, типу ЕОМ та структури представлення вхідних даних. Тобто, алгоритмічна складність здійснює найбільший вплив на час вирішення завдання [11; 12]. Для оцінки складності алгоритму існують асимптотичні визначення, що поділяються на три нотації:

– тета-нотація (θ-нотація) охоплює функцію складності зверху і знизу, тобто виявляє максимальні та мінімальні інтервали складності. Використовується для аналізу середньої складності алгоритму;

– омега нотація (Ω-нотація) виявляє нижню межу складності алгоритму, тому вона використовується для аналізу мінімальної (найкращої) складності алгоритму;

– Big-O-нотація (O-нотація) вказує верхню межу складності алгоритму, тому вона використовується для аналізу максимальної (найгірший) складності алгоритму.

Для аналізу прикладних алгоритмів використовується MIN-MAX підхід, що вимагає Big-O нотації, оскільки практичний інтерес представляє найгірший сценарій [12]. При цьому складність розглядається як функція від основних параметрів завдання.

Наприклад, якщо оцінюється складність транспортного алгоритму завдання якого представлено одновимірним графом покриття (1) для якого існує цільова функція рішення f, то в якості параметра береться кількість РТ N. Оскільки розглядається граф $G(N)$, то цільова функція рішення є також функцією порядку $f(N)$, яку можна апроксимувати функцією $g(N)$. Для визначення функції f алгоритм розбивають на окремі кроки та намагаються оцінити кожен крок через параметр N. Тому пошук f залежить від кількості РТ та переходів графу G, тобто

$$\lim_{n\to N} \frac{g(n)}{f(n)} \Rightarrow 1, \quad n \in (1, N), \tag{6}$$

а асимптотичні складності транспортних алгоритмів для нотацій (θ,Ω,O) виявляється за:

$$\theta\left[g(N/2)\right], \quad \Omega\left[g(1)\right], \quad O\left[g(N)\right]. \tag{7}$$

Вибір функції апроксимації $f(N)$ дозволяє отримати оцінки асимптотичної складності алгоритму [11]: $f(N)$ – лінійна оцінка; $f[N \log(N)]$ – логарифмічна оцінка; $f(N^2)$ – квадратична оцінка; $f(a_0 \cdot N^q + a_1 \cdot N^{q-1} + a_2 \cdot Nq^{-2} \ldots + a_q)$ – поліноміальна оцінка; $f(a^N)$ – експоне-

нтна оцінка; $f(N^N)$ – надекспонтна оцінка; $f(N!)$ – факторіальна оцінка; $f(N \cdot N!)$ – надфакторіальна оцінка.

За складністю усі алгоритми можуть бути поділені на класи: horrible (жахливі); bad (погані); fair (справедливі); good(добрі). До "практичних" алгоритмів (придатних для реалізації) відносяться алгоритми третього та четвертого класів, оцінки складності яких не перевищують поліноміальної. До "непрактичних" алгоритмів належать алгоритми, що мають експонентну оцінку складності та вище. На рис. 2 наведено класи складності та функції апроксимації цільової функції алгоритмів вирішення транспортних завдань [11]. Асимптотичні складності деяких алгоритмів вирішення множинних транспортних завдань та створення кластерів наведено у табл. 2.

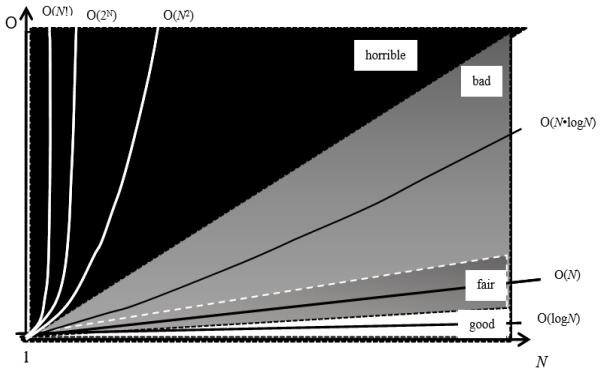


Рис. 2. Класи складності та функції апроксимації цільової функції алгоритмів вирішення транспортних завдань [11]

Таблиця 2 – **Асимптотичні складності алгоритмів вирішення множинних транспортних завдань та створення кластерів**

Алгоритм	Базовий параметр	Big-O нотації	Клас складності
Повного перебору контурів Гамільтону	N	$(N\text{-}1)\cdot N!$	Horrible
Дерев (Гілок) та обмежень	N	$N^5 \log(N)$	Horribl
Динамічного програмування	N	N^N	Horribl
Завихрення повітря BWO As	N	$0{,}33\cdot N^3+0{,}25_1\cdot N^2+ 0{,}5\cdot N+16{,}2$	Bad
Руху окремих частинок у хмарах PSO As	N	$N\ e^N$	Bad
Рисового поля PFA As	N	$a_0\cdot N^q+a_1\cdot N^{q-1}+ a_2\cdot Nq^{-2}\ldots+a_q$	Bad
Генетичний GA As	N	$(N\text{-}1)\cdot N^2$	Bad
Бджолині колонії BCO As	N	$(N\text{-}1)\cdot N^2$	Bad
Світлячка FFA As	N	$(N\text{-}1)\cdot N^2$	Bad
Мурашині ACO As	N	$(N\text{-}1)\cdot N^2$	Bad
Пошуку бактеріями живлення BFA As	N	$N!$	Bad
Розбиття по координатам	N	N	Good
Рекурсивний поділу навпіл	N	2^N	Good
За використання кривих Пеано	K	$\log K$	Good
З урахуванням зв'язності	K, N	$K\cdot N$	Fair
Кернігана-Ліна	K, N	$K\cdot N^2$	Fair

Висновки

Інтерес до прикладних множинних транспортних завдань (mTSP) інспіровано новими технологіями ТЗ, їх моніторингу та керуванням в умовах перешкод. Множинні алгоритми, як правило, спрямовані на оптимізацію моделі дискретного покриття для випадкового ігрового сценарію, що дозволяє формалізувати завдання з графів вагових коефіцієнтів переходів покриття для всіх ТЗ.

Створення mTSP моделі має обмеження на вхідні характеристики, що обумовлює потребу у евристичних алгоритмах, які мають надскладні рішення. Розглянуто стаціонарне дискретне покриття СНД та гіпотезу можливості розподілу його на кластери, які не перетинаються при використанні множини ТЗ. Це дозволяє здійснювати вибір та спрощувати рішення для евристичних алгоритмів. Показано нотації складності рішення та можливості їх оцінки за цільовою функцією обраного алгоритму або за її апроксимацією. Зроблено аналіз асимптотичної складності евристичних алгоритмів для Big-O нотації. Показано, що придатними для реалізації є алгоритми, складності яких не перевищують поліноміальної, а застосування алгоритмів, що мають експонентну оцінку складності та вище не є бажаним.

Зазначено, що асимптотичні оцінки складності алгоритму придатні для великих розмірностей завдання, а для завдань невеликої розмірності доцільними є контрольні прогони.

СПИСОК ЛІТЕРАТУРИ

1. Wassan N. A., Nagy G. Vehicle Routing Problem with Deliveries and Pickups: Modelling Issues and Metaheuristics Solution Approaches *Int. Journal of Transp*. 2014. Vol.2. № 2 (1). Pp. 95-110. https://article. nadiapub.com/IJT/vol2_no1/6.pdf.
2. Коломоєць А., Толстанов О. , Михальчук В., Гбур З., Кошова С. Системи логістики та логістичних підходів в управлінні закладами охорони здоров'я. Кам'янець-Подільський : ТОВ «Друкарня «Рута»», 2021. 348 с.
3. Погудіна О. К., Крицький Д. М., Биков А. М., Пластун Т. А., Пивовар М. В. Методология формування інтелектуальної складової агентної системи рою безпілотних літальних апаратів. Монографія. Харків : НАУ ім. М. Є. Жуковського «ХАІ», 2021. 211 с. https://odnb.odessa.ua/vnn/book/12456.
4. Dorigo M., Gambardella L. M. Ant Colony System: A cooperative learning approach to the Traveling Salesman Problem. *IEEE Transactions on Evolutionary Computation*. 1997. № 1 (1). Pp. 53-66. https://ieeexplore.ieee.org/document/585892.
5. Cordeau J. F., Laporte G., Ropke S Recent Models and Algorithms for One-to-One Pickup and Delivery Problems. New York: Springer, 2008. Pp. 327-357. DOI: 10.1007/978-0-387-77778-8_15.

6. Іваненко Ю. В., Ляшенко О. С., Філімончук Т. В. Огляд методів керування безпілотними літальними апаратами. *Системи управління, навігації та зв'язку*. 2023. № 1 (71). С. 26-30. https://www.researchgate.net/publication/369874368_
7. Pillac V., Gendreau M., Gu'eret C., Medaglia A. A review of dynamic vehicle routing problems. European Journal of Operational Research. 2012. № 225 (1). Pp. 1-11. DOI: https://doi.org/10.1016/j.ejor.2012.08.015.
8. Темнікова О. Л. Математична логіка та теорія алгоритмів. Київ : КПІ ім. І. Сікорського, 2021. 177 с.
9. Dorigo M., Bonabeau E., Theraulaz G. Ant Algorithms and Stigmergy. *Future Generation Computer Systems*. 2000. № 16(8). Pp. 851–871. https://www.sciencedirect.com/science/article/abs/pii/S0167739X0000042X.
10. Brezina I., Čičková Z. Solving the Travelling Salesman Problem Using the Ant Colony Optimization. *Management Information Systems*. 2011. Vol. 6 № 4 – Pp. 10-14. https://www.researchgate.net/publication/264855262.
11. Germanchuk M. S. Solvability of pseudobulous conditional optimization problems of the type of many salesmen. *Taurida Journal of Computer Science Theory and Mathematics*. 2020. № 4 (49). Pp. 30-55. https://tvim.su/en/node/1044.
12. Hosny M. Investigating Heuristic and Meta-Heuristic Algorithms for Solving Pickup and Delivery Problems. 2010. 266 p.

Received (Надійшла) 24.08.2025
Accepted for publication (Прийнята до друку) 29.10.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Карлов Володимир Дмитрович – доктор технічних наук, професор, завідувач кафедри фізики та радіоелектроніки, Харківський національний університет Повітряних Сил імені Івана Кожедуба, Харків, Україна;
Volodymyr Karlov – Doctor of Technical Sciences, Professor, Head of the Department of Physics and Radio Electronics, Ivan Kozhedub Kharkiv National Air Force University, Kharkiv, Ukraine;
e-mail: karlovvd@ukr.net; ORCID Author ID: https://orcid.org/0000-0002-1043-684X;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=56941529900.

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@khpi.edu.ua; ORCID Author ID: https://orcid.org/0000-0001-8228-8404;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57211278112.

Кузнєцов Олександр Леонідович – кандидат технічних наук, доцент, доцент кафедри фізики та радіоелектроніки, Харківський національний університет Повітряних Сил імені Івана Кожедуба, Харків, Україна;
Oleksandr Kuznietsov – Candidate of Technical Sciences, Associate Professor, Associate Professor of the Department of Physics and Radio Electronics, Ivan Kozhedub Kharkiv National Air Force University, Kharkiv, Ukraine;
e-mail: kuznetssovOL@ukr.net; ORCID Author ID: https://orcid.org/0000-0002-5915-8107;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57202953428.

Бєсова Оксана Василівна – кандидат технічних наук, старший науковий співробітник, Харківський національний університет Повітряних Сил імені Івана Кожедуба, Харків, Україна;
Oksana Biesova – Candidate of Technical Sciences, Senior Researher, Ivan Kozhedub Kharkiv National Air Force University, Kharkiv, Ukraine;
e-mail: BiesovaOV@ukr.net; ORCID Author ID: https://orcid.org/0000-0001-7744-1339;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57220834978.

Бєсова Анна Олексіївна – студентка кафедри "Комп'ютерна інженерія та програмування", Національний технічний університет "Харківський політехнічний інститут", Харків, Україна;
Anna Biesova –student, Department of Computer Engineering and Programming, National Technical University 'Kharkiv **Бєсова Оксана Василівна** Polytechnic Institute', Kharkiv, Ukraine;
e-mail: Anna.Biesova@cs.khpi.edu.ua; ORCID Author ID: https://orcid.org/0009-0004-4788-0137.

Estimations of the asymptotic complexity of applied multiple transport algorithms when divisioning them into clusters

Volodymyr Karlov, Oleksii Kolomiitsev, Oleksandr Kuznietsov, Oksana Biesova, Anna Biesova

Abstract. Applied multiple transport algorithms are usually heuristic and overly complex. Depending on the chosen heuristic of the algorithm, the real time of the software implementation depends on the programming language, the structure of the input data representation and the algorithmic complexity of the solution. The research is based on the use of the provisions of discrete mathematics and graph theory. It is shown that by analyzing asymptotic complexity it is possible to determine effective heuristics. Refinement of applied problem models also contributes to reducing complexity. Since the multiplicity of transport tasks is contained in a discrete coverage or in the functionality of multiple vehicles, it is possible to simplify the algorithm by decomposing them into clusters for which the conditions and constraints of the standard transport task apply. The features of geometric and combinatorial decomposition into clusters are presented. The formalization of cluster construction algorithms for multiple transport problems is provided. A comparative analysis of optimal and heuristic algorithms for solving multiple transport problems is carried out. The hierarchy of asymptotic complexity of heuristic algorithms is determined and the possibilities of their decomposition into clusters are analyzed, which simplify the solution of multiple applied transport problems. The notation of the complexity of the solution and the possibility of their evaluation by the objective function of the selected algorithm or by its approximation are provided. An analysis of the asymptotic complexity of heuristic algorithms for Big-O notation has been carried out. It has been shown that algorithms with a complexity not exceeding polynomial are suitable for implementation, and the use of algorithms with an exponential complexity estimate and higher is undesirable. It is noted that asymptotic estimates of algorithm complexity are suitable for large-dimensional tasks, and for tasks of small dimension, control runs are appropriate.

Keywords: asymptotic complexity, heuristics, clusters, logistics, multiple transportation tasks, notations, emergency care system, vehicles, objective function.

УДК 004. 8:004.9 **doi: 10.26906/SUNZ.2025.4.88-91**

С. І. Клівець, О. В. Кулєшов, Т. В. Кулєшова

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

АДАПТИВНИЙ МЕТОД ДИНАМІЧНОГО КЕРУВАННЯ РЕСУРСАМИ ГРАНИЧНОГО ШАРУ ІНДУСТРІАЛЬНОГО ІНТЕРНЕТУ РЕЧЕЙ

Анотація. В процесі функціонування Індустріального Інтернету речей нестача ресурсів граничного шару проявляється як зростання затримок обробки, черги, втрати даних або деградація сервісу. **Метою** цієї роботи є розробка адаптивного методу динамічного керування ресурсами граничного шару IIoT, який дозволить ефективно використовувати обчислювальні ресурси. **Отримані такі результати.** запропоновано концептуальний адаптивний метод динамічного керування ресурсами граничного шару у IIoT з використанням мультиагентного підходу та горизонтального і вертикального масштабування. Методи моніторингу, оцінки, прийняття рішення і навчання інтегровані в єдину архітектуру. Розроблена математична модель дозволяє формалізувати баланс черг, ресурси, передавання задач між вузлами. Метод забезпечує гнучку адаптацію до змінного навантаження й мінімізує втрати, затримки, а також ефективно використовує обмежені ресурси. **Висновки.** Адаптивне керування обчислювальними ресурсами граничного шару IIoT дозволяє підвищити ефективність функціонування системи та зменшити вплив нестачі ресурсів.

Ключові слова: Індустріальний Інтернет речей, граничний шар, обчислювальні ресурси, адаптивне керування.

Вступ

Індустріальний Інтернет речей (IIoT, Industrial Internet of Things) – це підмножина загального IoT, орієнтована на промислові та критичні інфраструктурні застосування. Характерні риси IIoT, що відрізняють його від звичайного IoT, є такими:

– жорсткі вимоги до надійності, стійкості, відмовостійкості, безпеки;

– детермінізм і низька затримка (latency);

– складна інтеграція з існуючими системами (SCADA, ПЛК, MES, ERP);

– гетерогенність: різні типи сенсорів, виконавчих пристроїв, протоколів, різні реальні умови (температура, вологість, електромагнітні перешкоди);

– великий обсяг даних на периферії.

Отже, IIoT має ті самі базові елементи, що й IoT, але з більш жорсткими вимогами, суворішими обмеженнями та вищим рівнем критичності.

IIoT застосовується в багатьох промислових галузях, таких як виробництво, енергетика, транспорт і логістика, нефтегазова галузь, авіація, космічні системи та оборона і багато інших.

При розгортанні IIoT виникає низка проблем:

– обмежені ресурси на граничному (edge) шарі;

– нерівномірність навантаження;

– комунікаційні затримки та обмеження пропускної здатності;

– надійність і відмовостійкість;

– балансування між обробкою на граничному шарі та передачею даних у хмару;

– безпека, конфіденційність даних;

– розподіленість та масштабованість.

Серед перерахованих проблем однією із головних є нестача ресурсів граничному шару. Дана проблема може виникнути у таких випадках:

– раптовий сплеск активності, коли кількість задач перевищує можливості обробки;

– процеси машинного навчання або обробки даних вимагатимуть більших ресурсів, ніж доступні;

– довгі черги даних у буферах вузлів, переповнення пам’яті чи черг, втрати або затримки пакетів;

– енергетичні обмеження;

– збої або часткова відмова вузлів.

Нестача ресурсів проявляється як зростання затримок обробки, черги, втрати даних або деградація сервісу. Саме в такі моменти потрібні адаптивні методи керування ресурсами.

1 Огляд сучасних досліджень

Питання розробки адаптивних методів керування ресурсами граничного шару систем підтримки Інтернету речей розглядаються у багатьох наукових працях. Так, у роботі [1] розглядається використання Mobile Edge Computing (MEC) для відвантаження обчислень з пристроїв IIoT до крайових серверів, з метою зменшення затримок та підвищення продуктивності. Важливий аспект — як динамічно балансувати навантаження між пристроями та edge-серверами. У роботі [2] розглядається класифікація стратегій керування ресурсами (статичні/динамічні, пріоритетні підходи) для середовища з гетерогенними вузлами та навантаженнями. Особлива увага приділяється сценаріям, у яких вузли можуть змінювати свої ресурси або призначення (реалокація, масштабування). У роботі [3] запропоновано підхід на базі multi-agent DRL для адаптивного виділення ресурсів і масштабування у мережевих зрізах на краю мережі. Вони також підтримують змінну кількість «зрізів», що дуже корисно у динамічних сценаріях. У роботі [4] пропонується структура, яка враховує поведінку пристроїв і намагається забезпечити розумне, справедливе розподілення ресурсів із використанням deep reinforcement learning. Підхід може бути адаптований до промислових сценаріїв з гетерогенними пристроями IIoT. У роботі [5] запропоновано алгоритм масштабування шлюзу граничного шару та адаптивне керування чергами для зменшення втрат пакетів. Використано еволюційний алгоритм Firefly для вибору вузлів масштабування. У роботі [6] поєднано концепції цифрових двійників і федерованого навчання для IIoT. Пропонується адаптивне налаштування частоти агрегації на основі черг Ляпунова та deep RL, з урахуванням обмежень ресурсів. Цей підхід цікавий як напрям для

гнучкого розподілу обчислень у IIoT серед edge-вузлів. Робота [7] більше стосується управління заторами в мережі, її підхід до адаптивного контролю може доповнити рішення на рівні граничного шару, особливо коли комунікаційні канали обмежені.

Ці праці демонструють, що сучасні методи орієнтовані на використання машинного навчання методів масштабування, балансування та поведінково орієнтованого підходу до розподілу ресурсів.

Метою даної статті є розробка адаптивного методу динамічного керування ресурсами граничного шару IIoT, який:

– здатний реагувати на зміну навантаження в режимі реального часу, без значних затримок;

– мінімізує втрати даних (пакетів), черги та затримки в граничному шарі.

– ефективно використовує обмежені ресурси (обчислювальні, пам'ять, енергія) вузлівЖ

– забезпечує адаптивне масштабування (горизонтальне та/або вертикальне) вузлів граничного шару;

– може бути застосований у розподілених та географічно рознесених середовищах.

2 Математична модель граничного шару IIoT

2.1 Припущення. Нехай граничний шар складається з N вузлів, кожен із них отримує задачі або потоки даних від пристроїв IIoT. Позначимо:

$\lambda_i(t)$ – вхідна інтенсивність задач (пакетів, завдань, транзакцій) до вузла i у час t;

$\mu_i(t)$ – пропускна здатність (середня швидкість обробки) вузла i у час t, яку можна коригувати в певних межах;

$q_i(t)$ – довжина черги на вузлі i у момент часу t;

C_i – максимальна ємність черги вузла i (максимальна кількість задач, які можна накопичити);

$r_i(t)$ – рівень ресурсів (CPU, пам'ять тощо), доступних вузлу i у час t; це може бути вектор ресурсів;,

$x_{ij}(t)$ – рішення про передавання задач або пакетів від вузла i до вузла j або на центральний сервер.

Також можна виділити такі природні обмеження:

– черга не може перевищувати ємність:

$$0 \le q_i(t) \le C_i; \tag{1}$$

– витрати на обробку даних на граничному шарі не повинні перевищувати розміру доступного обчислювального ресурсу:

$$\mu_i(t) \le f(r_i(t)), \tag{2}$$

де $f(\cdot)$ – функція, що перетворює доступні обчислювальні ресурси $r_i(t)$ у необхідву пропускну здатність каналів зв'язку;

– балансування черги запитів пристроїв IoT до граничного шару:

$$\frac{dq_i(t)}{dt} = \lambda_i(t) - \mu_i(t) + \sum_{j \ne i} \left(x_{ji}(t) - x_{ij}(t) \right); \tag{3}$$

– обмеження в процесі передавання завдань та даних сенсорами IoT:

$$x_{ij}(t) \ge 0, \ \sum_i x_{ij}(t) \le MaxTask_i; \tag{4}$$

де $MaxTask_i$ – максимальна кількість завдань, які вузол **i** може передати іншим вузлам чи в хмару або до туманного шару.

2.2 Цільова функція. В якості цільової функції розглядається мінімізація математичного сподівання функціоналу сукупних втрат та затримок, який розраховується таким чином:

$$J = \int_0^T \sum_{i=1}^N \begin{pmatrix} w_q \cdot q_i(t) + w_\ell \cdot L_i(t) + \\ + w_p \cdot P_{loss,i}(t) \end{pmatrix} dt, \tag{5}$$

де $L_i(t)$ – затримка задач у вузлі i; $P_{loss,i}(t)$ – рівень втрат, тобто кількість задач, що були скинуті через переповнення черги; w_q, w_ℓ, w_p – вагові коефіцієнти, котрі налаштовані залежно від пріоритетів.

Для врахування витрат критичних ресурсів, таких як, наприклад, енергія для мобільних пристроїв, до цільової функції додається штраф за використання таких ресурсів, тобто:

$$J_\Sigma = J + \int_0^T \sum_{i=1}^N \left(w_r \cdot c\left(r_i(t)\right) \right) dt, \tag{6}$$

де $c\left(r_i(t)\right)$ – функція вартості використання критичних ресурсів. Отже, цільова функція задачі оптимізації має такий вигляд:

$$J_\Sigma \xrightarrow[\gamma]{} \min, \tag{6}$$

де γ – множина варіантів розподілу ресурсів.

2.3 Адаптивне масштабування. У математичній моделі розглядається як горизонтальне, так і вертикальне масштабування.

До горизонтального масштабування відносяться такі дії: динамічне додавання або активація нових вузлів (edge nodes) або переміщення задач на менш завантажені вузли. Таке масштабування формалізується за допомогою булевої змінної $s_i(t) \in \{0; 1\}$, яка показує стан вузла, тобто вузол на даний момент часу активний чи ні. Також можна використовувати таку змінну, як кількість активних підвузлів $n_i(t)$.

Вертикальне масштабування досягається за рахунок зміни $r_i(t)$ у межах доступного ресурсу для конкретних пристроїв граничного шару, наприклад, збільшення CPU частоти, алокації пам'яті.

Ці рішення мають бути інтегровані у керуючу стратегію, яка приймає рішення про розподіл ресурсів граничного шару з урахуванням обмежень (1)–(4) і цільової функції (6).

3 Розробка методу адаптивного динамічного керування ресурсами граничного шару IIoT

Розглянемо основні кроки реалізації алгоритму запропонованого методу.

Крок 1. *Моніторинг і вимірювання:* вузли граничного шару проводять моніторинг показників:

– поточна інтенсивність задач $\lambda_i(t)$;

– довжина черги $q_i(t)$;

– використання ресурсів $r_i(t)$;

– затримки обробки даних, втрати пакетів.

Крок 2. *Оцінка стану:* узагальнення інформації у вектор стану s(t)s(t)s(t), що включає всі вузли чи агреговану інформацію.

Крок 3. *Прийняття рішення:* керуючий агент на підставі стану система приймає рішення про:

– розподіл ресурсів $r_i(t)$;

– пропускну здатність $\mu_i(t)$;

– передавання частини завдань $x_{ij}(t)$;

– масштабування вузлів $s_i(t)$.

Крок 4. *Адаптивне навчання:* алгоритм самонавчання (наприклад, reinforcement learning, multi-agent RL) оновлює політику на основі результатів мінімізації цільової функції.

Крок 5. *Механізми стабілізації та захисту:* обмеження максимальних змін, буферні зони, захист від «коливань» при змінному навантаженні.

Відзначимо, що кожен вузол або група вузлів діє як агент, що має свою локальну політику, а агенти взаємодіють між собою (наприклад, шляхом передачі задач або коригування навантаження).

4 Експериментальні дослідження

Експериментальна частина мала на меті перевірити ефективність запропонованого адаптивного методу динамічного керування ресурсами граничного шару IIoT у сценаріях з різним навантаженням, кількістю вузлів і ресурсними обмеженнями.

Основні показники ефективності, які досліджувались, були такими:

– середня затримка обробки задач (мс);

– середня довжина черги на вузлах;

– кількість втрачених задач через переповнення поточних черг;

– рівень використання ресурсів (CPU, пам’ять);

– енергоспоживання вузлів (для оцінки ефективності масштабування).

У якості середовища моделювання системи граничного шару було обрано імітаційну платформу на Python із бібліотеками SimPy для моделювання подій, NumPy для аналізу статистики та TensorFlow для реалізації агентів.

– конфігурація системи була обрана такою:

– кількість граничних вузлів – 5;

– кількість пристроїв IIoT – 50;

– тип навантаження – Пуассонівський потік (λ = 5–20 завдань/с);

– розмір задачі – 1–5 МБ;

– середня пропускна здатність вузла – від 10 до 50 завдань;

– максимальна черга – 100 задач;

– ресурси вузла (CPU, RAM) – 2–4 ядра, 2–8 ГБ;

– тривалість симуляції – 1 година моделювання (~1000 кроків)

Система реалізує дві зони: граничний шар для локальної обробки даних та хмарний рівень у випадку резервної обробки, якщо локальні вузли перевантажені. Агенти (вузли) спілкуються через легкий брокер (MQTT-емулятор), обмінюються статистикою про навантаження, затримки та черги.

Усереднені результати моделювання наведені на на основі експериментальних даних (рис. 1 – рис. 4).

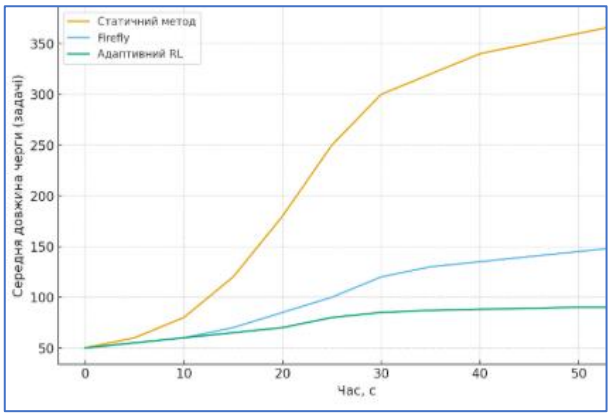


Рис. 1. Динаміка довжини черги при піковому навантаженні

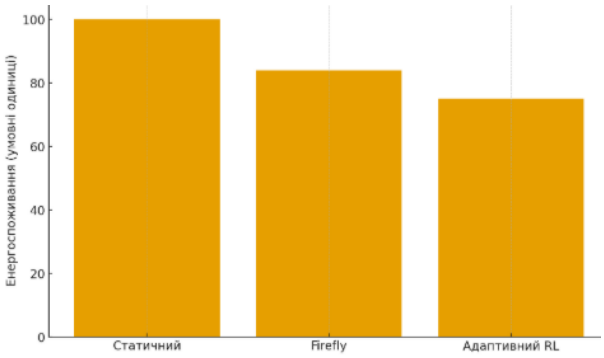


Рис. 2. Порівняння енергоспоживання різних методів

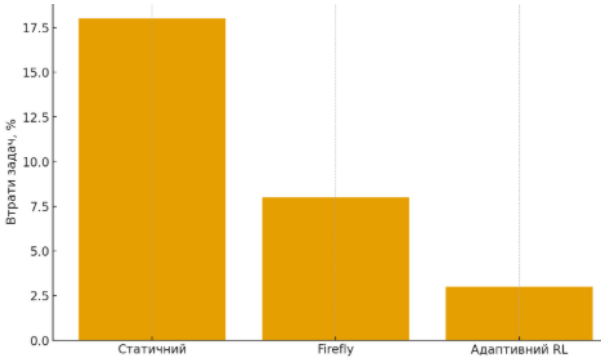


Рис. 3. Втрати задач під час перевантаження

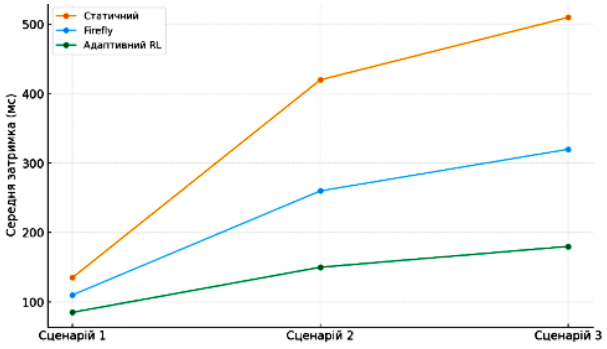


Рис. 4. Порівняння середньої затримки в трьох сценаріях

Отже, результати експериментального дослідження показали, що розроблений метод забезпечує:

– адаптивне балансування навантаження за 5–10 секунд після зміни потоку;

– стабільність системи без коливань;

– мінімальні втрати пакетів;
– економію енергії (неповне навантаження);
– гнучку реакцію на відмови вузлів.

Висновки та перспективи подальших досліджень

У статті запропоновано концептуальний адаптивний метод динамічного керування ресурсами граничного шару у IIoT з використанням мультиагентного підходу та горизонтального і вертикального масштабування. Методи моніторингу, оцінки, прийняття рішення і навчання інтегровані в єдину архітектуру. Розроблена математична модель дозволяє формалізувати баланс черг, ресурси, передавання задач між вузлами.

Метод забезпечує гнучку адаптацію до змінного навантаження й мінімізує втрати, затримки, а також ефективно використовує обмежені ресурси.

Напрями подальших досліджень: інтеграція з цифровими двійниками для передбачення навантаження та моделювання стану вузлів; адаптивність самої архітектури – механізми самокалібрування вагових коефіцієнтів, автоматичне виявлення аномалій і переключення між стратегіями.

Список літератури

1. Zhang, X., Wang, X., Xu, X. & Duan, L. (2023) *Resource Management in Mobile Edge Computing: A Comprehensive Survey*. ACM (survey). DOI: https://doi.org/10.1145/3589639
2. Zamzam, M., Elshabrawy, T. & Ashour, M. (2019) *Resource management using machine learning in mobile edge computing: A survey*. In: Proceedings of the 2019 Ninth International Conference on Intelligent Computing and Information Systems (ICICIS). DOI: https://doi.org/10.1109/ICICIS46948.2019.9014733
3. Li, H., Liu, Y., Zhou, X., Vasilakos, X., Nejabati, R., Yan, S. & Simeonidou, D. (2023) *Adaptive resource management for edge network slicing using incremental multi-agent deep reinforcement learning*. CoRR (arXiv preprint). arXiv: 2310.17523. URL: https://arxiv.org/abs/2310.17523
4. AlQerm, I., Wang, J., Pan, J. & Liu, Y. (2021) *BEHAVE: Behavior-Aware, Intelligent and Fair Resource Management for Heterogeneous Edge-IoT Systems*. IEEE Trans. on Mobile Computing. DOI: https://doi.org/10.1109/TMC.2021.3068632
5. Dlamini, T., Gambín, M. (2019) *Adaptive resource management for a virtualized computing platform within edge computing*. Proceedings of the 2019 16th IEEE Annual Consumer Communications & Networking Conference (CCNC/SAHCN). DOI: https://doi.org/10.1109/SAHCN.2019.8824927
6. Sun, W., Lei, S., Wang, L., Liu, Z. & Zhang, Y. (2020) *Adaptive federated learning and digital twin for Industrial Internet of Things*. CoRR (arXiv preprint). arXiv: 2010.13058, URL: https://arxiv.org/abs/2010.13058
7. Mhamdi, L. & Abdul Khalek, H. (2023) *Congestion control in constrained Internet of Things networks*. IET Wireless Sensor Systems, 13(6), pp.247–255. DOI: https://doi.org/10.1049/wss2.12072

Received (Надійшла) 05.08.2025
Accepted for publication (Прийнята до друку) 22.10.2025

Відомості про авторів / About the Authors

Клівець Сергій Іванович – кандидат технічних наук, науковий співробітник наукового центру, Харківський національний університет Повітряних Сил імені Івана Кожедуба, Харків, Україна;
Sergii Klivets – Candidate of Technical Sciences, Researcher at the research center, Ivan Kozhedub Kharkiv National Air Force University, Kharkiv, Ukraine;
e-mail: 1546.hnups@gmail.com; ORCID Author ID: https://orcid.org/0000-0002-8109-0639;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=58983477300.

Кулєшов Олександр Васильович – кандидат військових наук, старший науковий співробітник наукового центру, Харківський національний університет Повітряних Сил імені Івана Кожедуба, Харків, Україна;
Alexander Kuleshov – Candidate of Military Sciences, Senior Researcher of the Scientific Center, Ivan Kozhedub Kharkiv National Air Force University, Kharkiv, Ukraine;
e-mail: veshk.363@gmail.com; ORCID Author ID: https://orcid.org/0000-0002-8223-3814;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=58983967400.

Кулєшова Тетяна Василівна – науковий співробітник наукового центру, Харківський національний університет Повітряних Сил імені Івана Кожедуба, Харків, Україна;
Tetiana Kulieshova – Researcher of the Scientific Center, Ivan Kozhedub Kharkiv National Air Force University, Kharkiv, Ukraine;
e-mail: kuletati@gmail.com; ORCID Author ID: https://orcid.org/0000-0001-7404-109X.

Adaptive method for dynamic resource control of the Industrial Internet of Things border layer

Sergii Klivets, Alexander Kuleshov, Tetiana Kulieshova

Abstract. In the process of functioning of the Industrial Internet of Things, the shortage of resources of the boundary layer manifests itself as an increase in processing delays, queues, data loss or service degradation. The **purpose** of this work is to develop an adaptive method for dynamic resource management of the IIoT boundary layer, which will allow for the effective use of computing resources.. The following **results were obtained**: A conceptual adaptive method for dynamic resource management of the boundary layer in IIoT using a multi-agent approach and horizontal and vertical scaling is proposed. Monitoring, evaluation, decision-making and learning methods are integrated into a single architecture. The developed mathematical model allows formalizing the balance of queues, resources, and task transfer between nodes. The method provides flexible adaptation to variable load and minimizes losses, delays, and also effectively uses limited resources. **Conclusions**. Adaptive management of computing resources of the edge layer of the IoT allows to increase the efficiency of the system and reduce the impact of resource shortages.

Keywords: Industrial Internet of Things, edge layer, computing resources, adaptive control.

УДК 004.9:519.872 **doi: 10.26906/SUNZ.2025.4.92-97**

А. А. Коваленко, І. А. Замрій, В. Д. Попов, Д. А. Жаріков

Харківський національний університет радіоелектроніки, Харків, Україна

АНАЛІТИЧНА МОДЕЛЬ РОЗПОДІЛЕНОГО РЕЄСТРУ НА ОСНОВІ КОНЦЕПЦІЇ МАСОВОГО ОБСЛУГОВУВАННЯ

Анотація. Актуальність. Основною проблемою при формуванні реєстру у розподілених системах є визначення залежності між параметрами потоку запитів, швидкістю обслуговування транзакцій та загальними характеристиками продуктивності системи. **Мета дослідження:** побудова аналітичної моделі розподіленого реєстру, що дозволяє описати процес обробки транзакцій у термінах теорії масового обслуговування та визначити ключові показники ефективності системи. **Результати.** У статті запропоновано аналітичну модель розподіленого реєстру, побудовану на основі теорії масового обслуговування. Модель описує процеси запису, верифікації та підтвердження транзакцій у децентралізованому середовищі з урахуванням обмежень пропускної здатності вузлів та інтенсивності надходження запитів. Використання апарату масового обслуговування дозволяє формалізувати часові характеристики системи, оцінити середній час очікування обробки транзакцій, коефіцієнт завантаження вузлів та ймовірність затримки. Проведено аналітичне дослідження впливу параметрів мережі на продуктивність реєстру та визначено умови стійкості системи. Запропонований підхід може бути використаний для оптимізації архітектури розподілених систем з підвищеними вимогами до швидкодії та надійності обробки даних.

Ключові слова: розподілений реєстр, масове обслуговування, аналітична модель, транзакція, продуктивність системи, стійкість.

Вступ

Сучасні інформаційні системи все частіше базуються на розподілених архітектурах, що забезпечують підвищену відмовостійкість, безпеку та прозорість обробки даних. Одним із ключових компонентів таких систем є розподілений реєстр, який дозволяє зберігати транзакції між вузлами без централізованого контролю. Проте зі зростанням обсягів даних і кількості учасників мережі постає проблема оцінювання продуктивності та стабільності таких систем у динамічних умовах.

Традиційні методи аналізу розподілених систем не завжди враховують стохастичний характер надходження запитів, черговість обробки транзакцій і вплив затримок зв'язку між вузлами [1–3]. У цьому контексті доцільним є застосування теорії масового обслуговування, яка дозволяє побудувати математичну модель процесів обробки транзакцій, описати їх у термінах інтенсивностей потоків заявок і обслуговування, а також оцінити ефективність роботи системи [4]. В [5] розглянуто початкові ідеї сучасних розподілених реєстрів , що слід пов'язувати з практичною появою блокчейну в роботі Накамото, де показано, як побудувати стійку розподілену систему для ведення ланцюга транзакцій без централізованого керування. Класичні роботи і подальші масштабувальні оптимізації (ролі делегування, шардінг) детально аналізують в [6] при умові детермінованого часу підтвердження та ліміту на кількість вузлів, проте вартість комунікації зростає з кількістю учасників приблизно як $O(N^2)$. Розповсюдження блоків/транзакцій у P2P-мережі (propagation) – ключовий чинник затримок і, отже, продуктивності. У працях [7–9] показано, що мережеві латентності і топологія суттєво змінюють час підтвердження.

1. Постановка задачі

У розподілених реєстрах процес обробки транзакцій включає кілька етапів: формування запиту користувачем, передача його до вузлів мережі, перевірка коректності даних, включення до черги на обробку та підтвердження результату. Кожен вузол мережі виконує функції обслуговування заявок, що надходять у випадкові моменти часу, утворюючи систему черг.

Основною проблемою є визначення залежності між параметрами потоку запитів, швидкістю обслуговування транзакцій та загальними характеристиками продуктивності системи. Зокрема, необхідно оцінити:

- середній час очікування транзакції в черзі;
- середній час обслуговування запиту вузлом;
- ймовірність переповнення черги (затримки або відмови в обслуговуванні);
- коефіцієнт завантаження вузлів;
- умови стійкості системи при зміні інтенсивності надходження заявок.

Таким чином, **метою роботи** є побудова аналітичної моделі розподіленого реєстру, що дозволяє описати процес обробки транзакцій у термінах теорії масового обслуговування та визначити ключові показники ефективності системи.

У даній роботі розроблено аналітичну модель розподіленого реєстру, що ґрунтується на концепції масового обслуговування. Модель дозволяє провести кількісну оцінку параметрів системи, таких як середній час підтвердження транзакцій, рівень завантаженості вузлів та ймовірність відмови у обслуговуванні. Отримані результати можуть бути використані для подальшої оптимізації механізмів синхронізації вузлів та підвищення ефективності роботи розподілених мереж.

2. Математична модель розподіленого реєстру

Архітектура класичних розподілених реєстрів з ланцюговою структурою, як правило, не передбачає прямої взаємодії кінцевих користувачів із записами реєстру; доступ здійснюється опосередковано через протоколи мережі. Коли користувач хоче додати

транзакцію до реєстру, йому доводиться віддавати доручення виробникам верифікованих записів реєстру (записи блоками), тобто посередникам, що стоять між користувачами та розподіленим реєстром. Вони також є фактором, через який виникають проблеми масштабування. Саме виробники вирішують, яку транзакцію додати до наступного блоку, а яку – ні, мають ексклюзивний доступ до блоків і право вирішувати, чию транзакцію прийняти для додавання до реєстру з урахуванням застосовуваного алгоритму консенсусу. Для вирішення проблеми масштабованості радикально змінено архітектуру класичного реєстру на основі відмови від блоків та блок-продюсерів. І замість того, щоб вибудовувати ланцюжок блоків, повинні з'єднуватися самі транзакції шляхом включення до кожної транзакції хешів кількох попередніх. В результаті виходить структура, відома в математиці, як спрямований ациклічний граф.

Для аналізу технологій розподіленого реєстру розглянемо їх як системи масового обслуговування. Як критерії для зіставлення обрані такі показники:

– тривалість обробки одиничної операції;

– пропускна спроможність системи;

– число обчислювальних вузлів у мережі;

– коефіцієнт завантаженості системи.

Дані параметри необхідно розрахувати на основі таких вхідних параметрів:

– кількість пристроїв мережі,

– інтенсивність надходження транзакцій,

– час обробки транзакції одним пристроєм.

Для розрахунку параметрів порівняння розподіленої мережі, побудованої на основі класичного розподіленого реєстру, можна уявити систему обробки транзакції та створення нового блоку як систему масового обслуговування (СМО). Заявки надходять у випадкові моменти часу та обслуговуються наявними в системі каналами обслуговування.

Запропонована математична модель може бути застосована для опису систем розподілених реєстрів. У її рамках операціям, що проходять через систему, відповідають заявки, а обчислювальним вузлам, які здійснюють їхню обробку, – канали обслуговування. Класичний розподілений реєстр або блокчейн пропонується розглядати як систему класу одноканальних СМО з необмеженою чергою. Розглянемо спрощений випадок, нехай всі пристрої мережі мають однакову швидкість обробки транзакції μ (пропускну здатність), яку можна знайти за формулою:

$$\mu = 1/t_{obr}, \quad (1)$$

де t_{obr} – час обробки транзакції.

Значення даного параметра визначається сукупністю факторів, до основних з яких відносяться:

– обчислювальна потужність вузла,

– завантаженість вузла сторонніми завданнями,

– характеристики каналу зв'язку, зокрема його пропускна здатність.

Вхідний потік транзакцій є випадковим. Інтервали часу між послідовними надходженнями операцій є незалежними, однаково розподіленими випадковими величинами з математичним сподіванням, рівним $M(x)$. Тоді для даної СМО інтенсивність надходження заявок $\lambda = M(x)$. Відмінною особливістю СМО з необмеженою чергою є наявність нескінченної множини можливих станів системи. На рис. 1 представлений граф станів системи обробки транзакцій (одноканальний). У вказаній моделі визначено такі стани системи: $S0$ – система вільна; S_1 – система зайнята, одна операція знаходиться в процесі обробки, черга очікування відсутня; S_2 – система зайнята, в черзі на обробку знаходиться одна операція; S_3 – система зайнята, у черзі на обробку перебувають дві операції тощо [10].

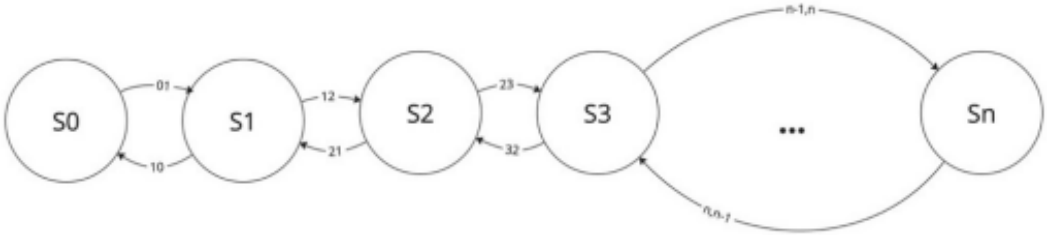


Рис. 1. Граф станів системи обробки транзакцій

Коефіцієнт завантаженості системи визначається як ймовірність того, що система знаходиться в будь-якому стані, відмінному від стану простою $S0$. Іншими словами, даний показник дорівнює сумі ймовірностей усіх станів, за яких зайнятий хоча б один канал обслуговування. Позначимо величину ймовірності знаходження системи в стані S_0 як p_0. Тоді для стаціонарних систем буде справедлива рівність:

$$\lambda_{01} \cdot p_0 = \lambda_{10} \cdot p_1, \quad (2)$$

де λ_{10} – інтенсивність переходу системи зі стану, де СМО система зайнята, обробляється транзакція та транзакцій у черзі немає, у стан, де система вільна від обробки та очікує надходження транзакцій, а λ_{01}–навпаки, інтенсивність переходу СМО зі стану свободи до стану обробки однієї транзакції та порожньою черги. Для p_1 справедливий вираз:

$$\lambda_{12} \cdot p_1 = \lambda_{21} \cdot p_2. \quad (3)$$

Аналогічно дане співвідношення є вірним для будь-якого p_k, тобто ц загальному випадку узагальнена формула (2) матиме такий вигляд:

$$\lambda_{k,k+1} \cdot p_k = \lambda_{k+1,k} \cdot p_{k=1}. \quad (4)$$

У результаті отримано систему рівнянь:

$$\begin{cases} \lambda_{01} \cdot p_0 = \lambda_{10} \cdot p_1, \\ \lambda_{12} \cdot p_1 = \lambda_{21} \cdot p_2, \\ \dots \\ \lambda_{k,k+1} \cdot p_k = \lambda_{k+1,k} \cdot p_{k=1}. \end{cases} \quad (5)$$

Для розв'язання висловимо всі значення через p_0. Наприклад, для p_1 справедливо:

$$p_1 = \frac{\lambda_{01}}{\lambda_{10}} \cdot p_0 = \frac{\lambda}{\mu}, \quad (6)$$

Для p_2:

$$p_2 = \frac{\lambda_{12}}{\lambda_{21}} \cdot p_1 = \frac{\lambda_{12} \cdot \lambda_{01}}{\lambda_{21} \cdot \lambda_{10}} \cdot p_0 = \left(\frac{\lambda}{\mu}\right)^2 \cdot p_0. \quad (7)$$

Необхідно відзначити, що чисельник у виразі (7) сформований добутком інтенсивностей переходів, що відповідають надходженню нових заявок до системи, тобто інтенсивностей, що йдуть зліва направо. Узагальнюючи, отримаємо вираз:

$$p_k = \left(\frac{\lambda}{\mu}\right)^k \cdot p_0. \quad (8)$$

Але імовірність всіх можливих станів дорівнює 1, отже є вірною така рівність:

$$\sum_{i=0}^{n} p_i = 1, \quad (9)$$

отже

$$p_0 + \frac{\lambda}{\mu} \cdot p_0 + \left(\frac{\lambda}{\mu}\right)^2 \cdot p_0 + \ldots + \left(\frac{\lambda}{\mu}\right)^m \cdot p_0 = 1, \quad (10)$$

звідси p_0 можна знайти за такою формулою:

$$p_0 = 1 \Big/ \sum_{i=0}^{n} \left(\frac{\lambda}{\mu}\right)^i, \quad (11)$$

де у знаменнику отримано ряд геометричної прогресії, який для нескінченної суми для необмеженої черги розраховується таким чином:

$$\sum_{i=0}^{n} \left(\frac{\lambda}{\mu}\right)^i = \frac{1}{1-\lambda/\mu}, \quad \text{при умові} \quad \frac{\lambda}{\mu} < 1. \quad (12)$$

Коли виконується ця умова, то система обробки транзакцій, яка представляє собою СМО, перебуває у стаціонарному стані. Досягнення стаціонарного режиму роботи системи означає, що довжина черги стабілізується. Ця умова дає можливість провести розрахунок усереднених характеристик, зокрема середнього часу перебування транзакцій у системі. Позначимо величину λ/μ як коефіцієнт завантаження системи ρ. Розрахунок середнього часу обробки транзакцій передбачає попереднє визначення середньої кількості обчислювальних вузлів, одночасно зайнятих у системі у довільний час. Для одноканальної СМО з необмеженою чергою це значення можна знайти за формулою:

$$L_{syst} = \frac{\rho}{1-\rho}, \quad (13)$$

За формулою Літтла знайдемо середній час обробки транзакції:

$$T_{obr} = \frac{L_{syst}}{\lambda} = \frac{\rho}{\lambda \cdot (1-\rho)}. \quad (14)$$

З виразу (14) можна зробити висновок, що для того, щоб підтримувати систему в стаціонарному стані, доведеться обмежити кількість вхідних транзакцій, а значить необхідно обмежити кількість пристроїв у розподіленому реєстрі, так як більша кількість пристроїв буде відправляти в систему більшу кількість транзакцій на обробку (проявляється взаємна λ), тобто при $n \to \infty$, $\lambda \to \infty$, а значить і $T_{obr} \to \infty$. Отже для того, щоб розвивати систему шляхом збільшення кількості пристроїв, необхідно пожертвувати часом обробки транзакцій. Це і є фундаментальною проблемою масштабованості.

3. Модель СМО на основі спрямованого ациклічного графа

Використання архітектури на основі спрямованого ациклічного графа (directed acyclic graph, DAG) дозволяє усунути обмеження на інтенсивність вхідного потоку заявок, характерне для розподілених систем з різною топологією. Це окремий випадок розподіленого реєстру, де транзакції структуровані як вузли у графі, а не в блоках. Транзакції, що виходять із пристроїв, становлять набір вузлів DAG.

Процес формування множини ребер графа полягає в наступному: кожна нова операція, що додається в систему, повинна підтвердити k (де $k \geq 2$) раніше створених операцій. Зазначені підтвердження є спрямованими ребрами, що зв'язують нову транзакцію з підтвердженими. Коли вузол створює нову транзакцію, він вибирає кілька раніше відомих йому транзакцій і включає їх посилання на своє повідомлення. Тим самим він висловлює свою згоду зі своїм змістом і визнає їх як частину консенсусної історії. Описаний алгоритмічний механізм виконує функцію децентралізованого протоколу досягнення консенсусу, забезпечуючи когерентність даних між усіма вузлами розподіленої мережі без єдиної координуючої інстанції. Число k має бути більшим або рівним 2 з ряду причин, наведених нижче.

1. *Забезпечення достатньої пов'язаності графа*. Якщо кожна транзакція підтверджує хоча б дві попередні, це мінімізує можливість ізольованих гілок у графі. Це гарантує зв'язність всієї структури спрямованого ациклічного графа, що важливо для підтримки цілісності і консистентності системи.

2. *При k = 1 нова транзакція підтверджує лише одну попередню*. Це створює загрозу, коли створюється ймовірність створення ізольованого ланцюжка, порушення загальної консистентності мережі. При $k \geq 2$ ймовірність успішного створення такого ланцюжка значно знижується, оскільки знадобиться підтвердження кількох транзакцій із різних частин ациклічного спрямованого графа.

3. *Підвищення стійкості до відмов*, оскільки система з $k \geq 2$ має надлишкові зв'язки, що підвищує її стійкість до збоїв. Якщо будь-який вузол графа виявляється недоступним, наявність як мінімум двох зв'язків дозволяє новим транзакціям продовжувати роботу, оминаючи несправний вузол.

4. *Мінімальне значення k = 2 прискорює процес підтвердження транзакцій*, оскільки кожна нова транзакція зміцнює граф додатковими ребрами,

підвищуючи можливість підтвердження попередніх транзакцій. Це особливо важливо за умов високого навантаження на систему.

Якщо вузол, що згенерував транзакцію (умовно іменовану «батьківською» по відношенню до наступних транзакцій, що її підтверджують), припиняє свою активність після того, як ця транзакція була успішно передана частини мережі, сама по собі вона залишається автономною структурою даних. Її коректність у межах спрямованого ациклічного графа визначається не наявністю вихідного вузла у мережі, а дотриманням протокольних правил і наявністю достатньої кількості наступних посилань неї з боку інших транзакцій. Активні вузли, що прийняли цю транзакцію і визнали її допустимою на підставі аналізу відповідності протоколу та узгодженості з їх поточним уявленням про стан графа, можуть вибрати її як одну з попередніх транзакцій при формуванні нових записів. Таким чином, транзакція отримує додаткове підтвердження через включення нових вершин графа. Згодом граф транзакцій розширюється за рахунок додавання нових елементів, які неявно схвалюють елементи, існуючі раніше. Гілки графа, які не отримують подальшого розвитку та не стають об'єктами посилань нових транзакцій, поступово виключаються з активного шляху консенсусу та залишаються на периферії структури як незатребувані фрагменти. Обов'язковою умовою для створення нової транзакції є попередня участь обчислювального вузла у підтримці працездатності мережі. Ця участь полягає у необхідності підтвердження певної кількості раніше створених транзакцій. На рис. 2 наведено схематичну ілюстрацію топології системи, організованої у вигляді спрямованого ациклічного графа.

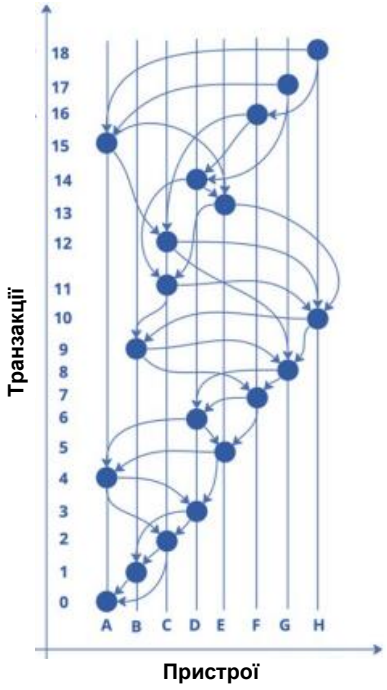


Рис. 2. Приклад ациклічного направленого графа

Вершини графа – це транзакції, спрямовані дуги показують, які підтверджуються іншими транзакціями. A, B, C, D, E, F, G, H – це позначення набору пристроїв розподіленого реєстру. На представленій діаграмі вертикальні відрізки ілюструють часові інтервали обробки конкретних операцій відповідними обчислювальними вузлами мережі. По осі ординат позначені номери транзакцій у порядку додавання до ациклічного направленого графа. Транзакції під номерами 16-18 самі поки що не підтверджені.

Основна перевага систем, заснованих на застосуванні спрямованого ациклічного графа, у порівнянні з класичними розподіленими реєстрами, що мають ланцюгову структуру, полягає у можливості здійснення багатопоточної та різночасової обробки операцій. Ця особливість означає, що кілька обчислювальних вузлів мережі можуть обробляти різні операції паралельно, при цьому моменти початку та завершення обробки для кожної окремої операції не є синхронізованими. На рис. 3 червоним кольором, як приклад, відзначені нові транзакції, які надійшли до системи одночасно. Транзакції, що надходять, обробляються паралельно і асинхронно, що означає можливість їх одночасного обслуговування на різних вузлах мережі без синхронізації моментів завершення.

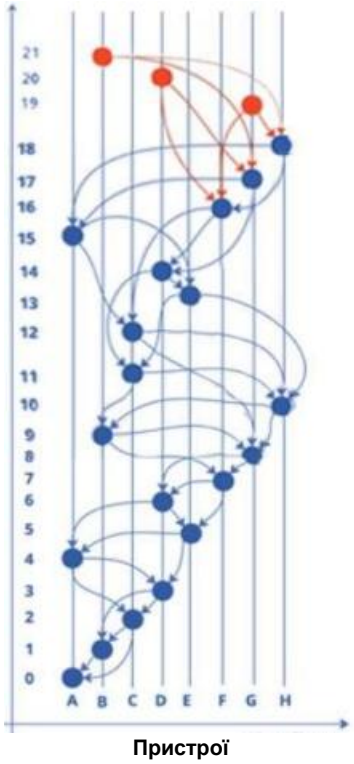


Рис. 3. Одночасне формування нових вузлів

Відповідно до положень теорії масового обслуговування, система класифікується як багатоканальна СМО з необмеженим розміром черги на обслуговування. Кількість каналів *n* у цій системі дорівнюватиме кількості пристроїв, що відправляють транзакціями.

Проаналізуємо залежність завантаженості системи від середнього часу обробки транзакції одним пристроєм (рис. 4). Отриманий графік показує, що з з урахуванням ациклічного спрямованого графа збільшення кількості обчислювальних вузлів сприяє підвищенню її стійкості до зростання навантаження

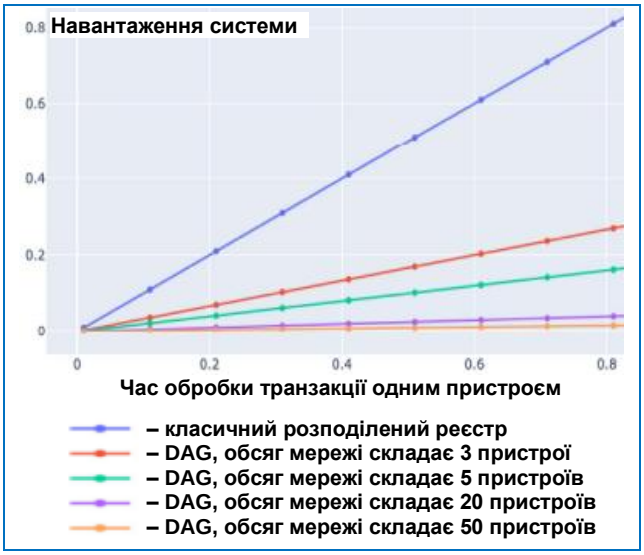


Рис. 4. Залежність завантаженості системи від середнього часу обробки транзакцій для розподіленого реєстру та ациклічного спрямованого графа з різною кількістю підключених пристроїв

При побудові моделі було взято до уваги такі ключові параметри системи:

– тривалість обробки однієї транзакції обчислювальним вузлом;

– інтенсивність потоку вхідних транзакцій;

– загальне число обчислювальних вузлів у мережі.

В рамках створюваної моделі допускається знехтувати кінцевою пропускною здатністю каналів зв'язку. Дане припущення є правомірним, оскільки обсяг даних, що передаються при підтвердженні транзакцій, незначний. Разом з тим, необхідно враховувати такий фактор, як час відгуку кінцевих пристроїв, навіть якщо його вплив вважається незначним. Необхідно відзначити, що в умовах практичного застосування мережеві архітектури, що базуються на використанні ациклічних спрямованих графів, демонструють значно більшу пропускну здатність та швидкість обробки транзакцій у порівнянні з класичними системами розподілених реєстрів. Завдяки високій масштабованості, даний граф можна застосовувати в розподілених мережах з великою кількістю пристроїв міжмашинної взаємодії та не боятися втрати швидкості обробки. Розроблена система з урахуванням ациклічного спрямованого графа моделюється як багатоканальна система масового обслуговування з необмеженою чергою M/M/n/∞.

Перша M (Markovian) означає, що вхідний потік транзакцій є найпростішим (або марковським), тому:

1) транзакції надходять у систему незалежно друг від друга;

2) імовірність надходження нової транзакції в малий проміжок часу Δt пропорційна Δt і не залежить від того, скільки транзакцій вже надійшло та коли вони надійшли (властивість відсутності післядії);

3) інтервали часу між надходженнями транзакцій мають експонентний розподіл;

4) у реальних системах M2M, де безліч пристроїв генерують дані незалежно один від одного, потік транзакцій може бути близьким до найпростішого, особливо якщо усереднити його за кількістю пристроїв.

Друга M означає, що час обслуговування транзакції одним вузлом також має експоненційний розподіл. Насправді час обробки може залежати від складності транзакції, обчислювальної потужності вузла та інших факторів. Проте, для базової моделі, експоненційний розподіл – це стандартне припущення, що дозволяє отримати аналітичні результати.

Також *n* означає, що у системі *n* каналів обслуговування (активних вузлів, які обробляють транзакції). Реальна система, особливо з урахуванням гетерогенності пристроїв та складного алгоритму консенсусу, відрізнятиметься від цієї моделі. Однак, M/M/n/∞ – це стандартна базова модель, яка дозволяє отримати аналітичні результати та якісно оцінити поведінку системи при таких умовах:

1) система знаходиться в стаціонарному стані;

2) система перебуває у стаціонарному стані;

3) середня кількість заявок зростає, коли надходить нова заявка;

4) середня кількість заявок зменшується, коли закінчується обслуговування заявки;

5) інтенсивність приходу заявок менша, ніж інтенсивність обслуговування заявок;

6) коефіцієнт завантаження повинен бути меншим одиниці.

У таких умовах класична модель системи масового обслуговування була б адекватною опису поведінки системи, оскільки передбачає відсутність зворотних зв'язків чи повторних проходження заявок через одні й самі вузли без явного управління цим процесом.

Висновки

У статті запропоновано аналітичну модель розподіленого реєстру, побудовану на основі теорії масового обслуговування. Модель описує процеси запису, верифікації та підтвердження транзакцій у децентралізованому середовищі з урахуванням обмежень пропускної здатності вузлів та інтенсивності надходження запитів.

Використання апарату масового обслуговування дозволяє формалізувати часові характеристики системи, оцінити середній час очікування обробки транзакцій, коефіцієнт завантаження вузлів та ймовірність затримки.

Проведено аналітичне дослідження впливу параметрів мережі на продуктивність реєстру та визначено умови стійкості системи. Результати показують, що з високої інтенсивності навантаження DAG обробляє транзакції значно швидше, ніж традиційні розподілені реєстри.

Порівняльні дані демонструють, що зі збільшенням інтенсивності навантаження ациклічний спрямований граф обробляє транзакції у суттєво швидше, ніж класичний реєстр.

Запропонований підхід може бути використаний для оптимізації архітектури розподілених систем з підвищеними вимогами до швидкодії та надійності обробки даних.

СПИСОК ЛІТЕРАТУРИ

1. Chalapathi, G.S.S., Chamola, V., Vaish, A. and Buyya, R. (2022), “Industrial internet of things (Iiot) applications of edge and fog computing: A review and future directions”, *Advances in Information Security*, vol. 83, pp. 293–325, doi: https://doi.org/10.1007/978-3-030-57328-7_12
2. Dotsenko, N., Chumachenko, I., Galkin, A., Kuchuk, H. and Chumachenko, D. (2023), “Modeling the Transformation of Configuration Management Processes in a Multi-Project Environment”, *Sustainability (Switzerland)*, Vol. 15(19), 14308, doi: https://doi.org/10.3390/su151914308
3. Zuev, A., Karaman, D. and Olshevskiy, A. (2023), “Wireless sensor synchronization method for monitoring short-term events”, Advanced Information Systems, vol. 7, no. 4, pp. 33–40, doi: https://doi.org/10.20998/2522-9052.2023.4.04
4. Hunko, M., Tkachov, V., Kuchuk, H. and Kovalenko, A. (2023), Advantages of Fog Computing: A Comparative Analysis with Cloud Computing for Enhanced Edge Computing Capabilities, *2023 IEEE 4th KhPI Week on Advanced Technology, KhPI Week 2023 – Conf. Proc*, 02-06 October 2023, Code 194480, doi: https://doi.org/10.1109/KhPIWeek61412.2023.10312948
5. Bayer, D., Haber, S. and Stornetta, W.S. (1993), “Improving the efficiency and reliability of digital time-stamping”, *Sequences Methods in Communication, Security and Computer Science*, pp. 329–334, available at: https://www.math.columbia.edu/~bayer/papers/Timestamp_BHS93.pdf
6. Gramoli, V. (2020), “From blockchain consensus back to Byzantine consensus”, *Future Generation Computer Systems*, vol. 107, pp.760–769, doi: https://doi.org/10.1016/j.future.2017.09.023
7. Kovalenko, A. and Kuchuk, H. (2022), “Methods to Manage Data in Self-healing Systems”, *Studies in Systems, Decision and Control*, vol. 425, pp. 113–171, doi: https://doi.org/10.1007/978-3-030-96546-4_3
8. Kuchuk, N., Kovalenko, A., Ruban, I., Shyshatskyi, A., Zakovorotnyi, O. and Sheviakov, I. (2023), “Traffic Modeling for the Industrial Internet of NanoThings”, *2023 IEEE 4th KhPI Week on Advanced Technology*, KhPI Week 2023 - Conference Proceedings, 2023, doi: 194480. http://dx.doi.org/10.1109/KhPIWeek61412.2023.10312856
9. Kuchuk, H. and Malokhvii, E. (2024), “Integration of IOT with Cloud, Fog, and Edge Computing: A Review”, *Advanced Information Systems*, vol. *8*(2), pp. 65–78, doi: https://doi.org/10.20998/2522-9052.2024.2.08
10. Decker, C. and Wattenhofer, R. (2014), “Bitcoin transaction malleability and MtGox”, *European symposium on research in computer security*, 370, pp. 313–326, doi: https://doi.org/10.1007/978-3-319-11212-1_18

Received (Надійшла) 25.07.2025
Accepted for publication (Прийнята до друку) 16.10.2025

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 ID: https://orcid.org/0000-0002-2817-9036;
Scopus ID: https://www.scopus.com/authid/detail.uri?authorId=56423229200.

Замрій Іван – студент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Ivan Zamrii –student at the Department of Electronic Computers, Kharkiv National University of Radio Electronics, Ukraine;
e-mail: ivan.zamrii@nure.ua; ORCID Author ID: http://orcid.org/0000-0002-0939-3125.

Попов Володимир – студент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Volodymyr Popov – student at the Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;
e-mail: volodymyr.popov@nure.ua; ORCID Author ID: http://orcid.org/0000-0002-7826-030X.

Жаріков Дмитро – студент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Dmytro Zharikov – student at the Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;
e-mail: dmytro.zharikov@nure.ua; ORCID Author ID: http://orcid.org/0000-0002-7826-030X.

Analytical model of a distributed registry based on the concept of crowding

Andriy Kovalenko, Ivan Zamriy, Volodymyr Popov, Dmytro Zharikov

Abstract. Relevance. The main problem in the formation of a registry in distributed systems is to determine the dependence between the parameters of the request flow, the speed of transaction service and the general characteristics of the system performance. The purpose of the study: to build an analytical model of a distributed registry, which allows describing the transaction processing process in terms of the theory of queueing and determining the key indicators of the system's efficiency. **Results.** The article proposes an analytical model of a distributed registry, built on the basis of the theory of queueing. The model describes the processes of recording, verification and confirmation of transactions in a decentralized environment, taking into account the limitations of the node's bandwidth and the intensity of incoming requests. The use of a queuing apparatus allows us to formalize the time characteristics of the system, estimate the average waiting time for transaction processing, the load factor of nodes, and the probability of delay. An analytical study of the influence of network parameters on the productivity of the registry has been conducted and the conditions for system stability have been determined. The proposed approach can be used to optimize the architecture of distributed systems with increased requirements for speed and reliability of data processing.

Keywords: distributed registry, queuing, analytical model, transaction, system productivity, stability.

UDC 004.383.8:004.422 **doi: 10.26906/SUNZ.2025.4.98-102**

Nina Kuchuk[1], Maksym Tregubenko[1], Danylo Kovalenko[1], Dmytro Lysytsia[2], Oleksandra Bellorin-Herrera[2]

[1] Kharkiv National University of Radio Electronics, Kharkiv, Ukraine
[2] National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine

RESEARCH OF DISTRIBUTED DATA EXCHANGE TECHNOLOGIES IN THE CONTEXT OF INTELLIGENT TRANSPORT SYSTEMS

Abstract. The article presents the results of a study of distributed data exchange technologies aimed at ensuring effective interaction between components of intelligent transport systems (ITS). Modern approaches to the organization of decentralized information transmission networks are considered, in particular, the use of the concepts of distributed registries, peer-to-peer protocols, and service-oriented architectures. An analysis of the requirements for reliability, latency, and bandwidth of communication channels, which are critical for data exchange scenarios between vehicles and infrastructure, is conducted. Based on a comparative analysis, the advantages and limitations of existing technological solutions in the context of ensuring security, scalability, and fault tolerance are determined. The results of the study can be used to design adaptive traffic management models and develop intelligent interaction modules in new generation transport networks.

Keywords: intelligent transportation systems; distributed data exchange; decentralized networks; distributed ledger; scalability; data transmission reliability.

Introduction

With the continuous development of information technology, which facilitates the creation of complex computing systems solving diverse problems, data communication networks play a key role in ensuring fault tolerance and operational stability. Many components of modern information systems generate their own information flow, requiring processing and distribution to other devices across the network. Prioritizing traffic, balancing, and accounting for peak load periods dictate specific requirements for the technologies used, based on the computing resources expended, memory, volume, and processing and transmission speed requirements for the network traffic they generate. Furthermore, considerable attention is paid to ensuring the security and transparency of certain information system activities, to monitor incidents and potential malicious threats, as well as to guarantee the integrity of data (or the recording of the date and volume of these changes). Blockchain technology can meet many of the requirements for secure and resilient data systems; however, in the vast majority of cases, it requires excessive computing resources, which are critical when using variations of the most popular consensus algorithms. However, there are more computationally lenient consensus algorithms, the continuous use of which also introduces a number of limitations: requirements for creating trusted network sections, delays in communication sections, or ensuring a fully connected system where each network node is directly connected to all other nodes. This encourages the creation of a variety of consensus algorithms with their own characteristics to solve specific problems, which is due to different goals and use cases [1]. For example, financial applications may require high transaction speeds, while data storage systems may emphasize security and reliability. This requires the development of specialized consensus algorithms adapted to specific needs [2]. Given the characteristics of modern heterogeneous communication networks, the lack of flexibility in transmitting blockchain traffic is one of the key reasons for not implementing it. Thus, a system for regulating the volume and speed of blockchain traffic throughout the day would ensure the required flexibility for all situations—more frequently when blockchain traffic is not a priority, less frequently when it is, depending on the ultimate goals and implementation. In addition to regulating traffic volumes, computing resource requirements should also be considered. Modern communication networks consist of components with varying computing power, the primary task of which is data processing and transmission. Under current conditions, installing a blockchain client on network components will have a significant impact on equipment and data processing speed, especially during peak hours. Since consensus algorithms largely determine the resource-intensive nature of the computations required to process transactions and generate blocks, regulating them during data processing, taking into account specific network and computing parameters, will allow for flexible decisions about whether to use the most resource-intensive algorithm and switch to a less resource-intensive one during peak loads.

1. Literature analysis

The application and integration of blockchain technology, as a specific implementation of distributed ledger technology, its impact on network characteristics, and applicability issues are discussed in the works of Ukrainian and international scientists V. Buterin, S. Kasahara, Q. Xia, Y. Sun, L. Cocco, and others. Many works are devoted to investigating issues of network traffic distribution and its impact on network characteristics [3], examining the technical maturity of the approach for integration into existing systems [4], and also the security aspects of the technology. A number of scientific papers are devoted to optimization based on research into consensus algorithms [5]. However, the problem of adapting these algorithms to telecommunication network conditions remains understudied, particularly in terms of developing an adaptive algorithm for selecting blockchain consensus on communication networks [6].

2. Main part

To evaluate a number of parameters, as well as the viability of the concept of integrating blockchain technology into modern communication networks and

the IT landscape, a series of experiments were conducted in the context of an intelligent transportation network, since such a network, in its architecture and structure, reflects all the main features important for measurement. For example, the hardware of network sections, the different levels of network activity for data transmission in the context of a large city center and a remote highway in the northern regions.

For this reason, many experiments and testing were conducted specifically under ITS emulation conditions [6]. The main goal of the study is to evaluate the node's operability and its operating speed under the load imposed by the infrastructure elements of the intelligent transportation network concept. To date, various studies have already presented scenarios for transmitting information from automotive and road sensors, but without interaction and integration with the blockchain network. From the standpoint of studying the operability of a classic blockchain node, the most indicative is testing close to load testing of the node, in which the number of requests to it is sufficiently large over a limited period of time. A study [7] assessing traffic intensity at a toll plaza on an Austrian toll road was used as reference data.

These studies revealed the average number of vehicles passing the toll plaza per week at each hour of the day. Table 1 presents the number of vehicles at each hour, as well as the percentage of vehicles passing this section at each hour of the day.

Table 1 – **Correlation between time of day and the number of passing vehicles**

Hours	Number of OBU	OBU %	Hours	Number of OBU	OBU %
0	600	1,65	12	1920	5,26
1	350	0,96	13	2000	5,48
2	250	0,69	14	2200	6,03
3	200	0,55	15	2300	6,31
4	150	0,41	16	2500	6,85
5	100	0,27	17	2900	7,95
6	125	0,34	18	3100	8,5
7	225	0,62	19	3000	8,23
8	1000	2,74	20	2850	7,81
9	2100	5,76	21	1850	5,07
10	2200	6,03	22	1500	4,11
11	2050	5,62	23	1000	2,74
Total number of OBU: 36470					

In this study, it is assumed that each vehicle passing through the toll plaza sends one transaction to a blockchain network node [8]. The transaction sent can contain various information, ranging from the fact of payment, the fact of passage, to the condition of the road surface. The objective of this study is to evaluate the node's performance and the ability to process all incoming transactions without data loss.

Based on the tabular data, a graph was obtained showing the peak values as well as the overall traffic intensity profile after correlating the values 12 times, and is presented in Fig. 1.

Thus, each vehicle sends a transaction based on the received data. To simulate these transactions, the authors developed a Python script. This allows them to simulate network load based on traffic intensity data for the segment in question. The correlation process accelerates the study and, using load testing, evaluates whether the blockchain network node can handle the increased load. The calculations performed accelerated the simulation process. During the study, 1 hour of real time becomes 5 minutes of the study, which also affects the number of vehicles, as shown in Fig. 1. From a script perspective, this simulation also requires multithreading to correctly send transactions according to the obtained load distribution. During transaction sending, the script interacts with the Web3 library API, which enables interaction with a node of a private instance of the blockchain networks under study on the developed simulation rig.

To develop the model rig and conduct further research, two blockchain platforms were used: Ethereum, which operates on the PoW algorithm (as it supports two implementations – PoW and PoS [9]) and Waves, which operates on the LPoS algorithm. Due to the operational specifics of private network instances of the platforms under consideration, two model rigs were organized. The choice of these blockchain networks was motivated by several reasons. These blockchains allow for the organization of private networks based on open source code, enable the launch of smart contracts, as well as the construction of hybrid systems and the creation of solutions with the integration of third-party technologies. The key difference between these blockchain platforms is the underlying consensus algorithms, which impacts the performance of the blockchain network, its security, and the hardware requirements for the node of the organized network. Waves Enterprise is a blockchain platform that enables the creation of various blockchain solutions based on smart contracts. The distributed network infrastructure allows for the execution of a large number of transactions in short periods of time. One of the platform's advantages is the ability to create your own private network using an open jar file containing all the necessary files and configuration data for setting up your

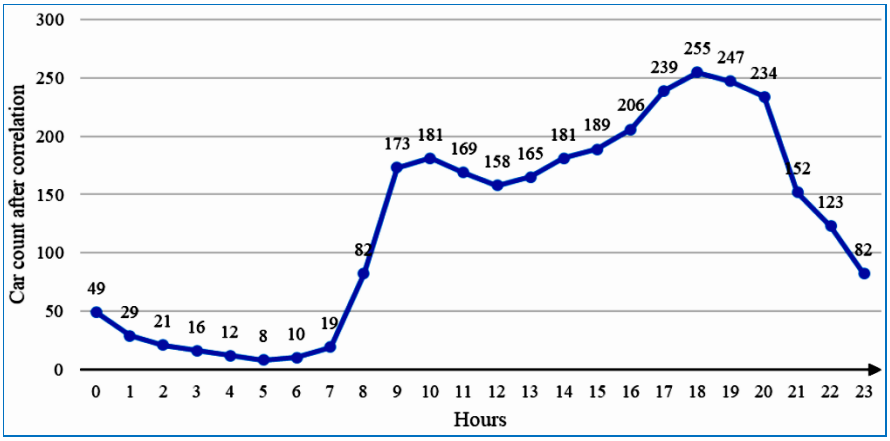


Fig. 1. Calculation of the number of cars in relation to hours after correlation

own node in a closed blockchain network. The Waves blockchain network uses Leased Proof of Stake (LPoS) consensus algorithm. This consensus algorithm does not require energy-intensive computations, unlike the Proof of Work consensus algorithm. The main parameter determining the probability of generating a new block is the amount of platform currency in the node's account. Thus, the amount of currency is analogous to the number of lottery tickets: the more tickets, the higher the chance of winning.

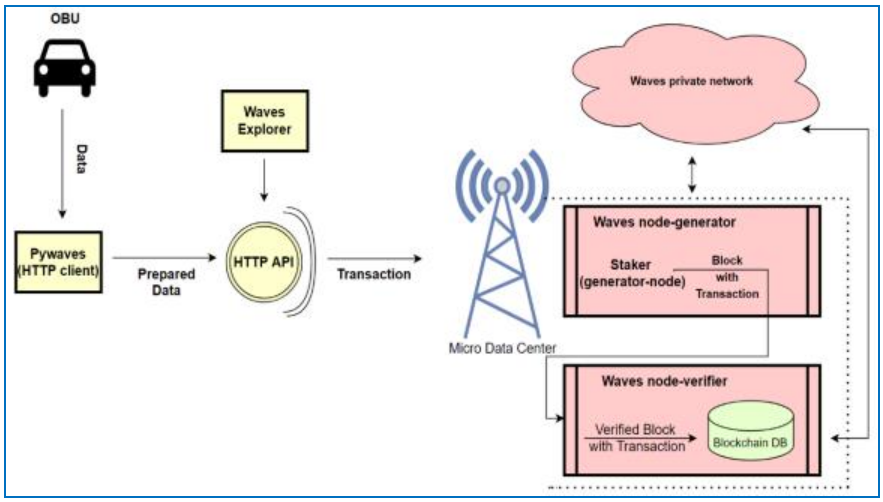


Fig. 2. Model rig for conducting an experiment based on the open blockchain client LPoS Waves

A similar principle applies to the Waves network. A diagram of a simulation using Waves blockchain network nodes is shown in Fig. 2. The Waves model stand consists of the following elements:

− OBU - a user device initiating a transaction based on current data,

− pywaves (HTTP client) - a framework whose main purpose is to interact with the Waves node's HTTP API,

− HTTP API - Waves node software interface, which provides the ability to interact with the Waves node,

− Waves node – A staker (generator node) is the main software node of the Waves blockchain network. It initiates the process of block assembly and transaction processing.,

− Waves node - Verifier - an additional software node of the Waves blockchain network whose main task is to confirm transactions and blocks,

− Waves Explorer, a web interface that allows you to check the current state of nodes and deploy smart contracts, the Waves blockchain network - a private network was created for testing purposes.

The OBU initiates a transaction and transmits its current state to Pywaves. Pywaves, using the received data, creates a transaction request in a format readable by the HTTP interface and sends it to the HTTP API. The HTTP API processes the received request and records information about the attempted transaction on the primary node (the generator node).

The generator node (staker), upon receiving the transaction, sends it for verification to its quorum – a meeting of all nodes in the network –and awaits verification from the verifier node. The generator node, having collected a certain number of verified transactions or having waited a certain period of time, initiates the formation of a block that will consolidate all accumulated transactions. Information about a successfully formed block is returned to the client via a similar path. Two scenarios were considered in this study: in the first case, transactions were sent to the blockchain network for recording, and in the second, they were requested from the blockchain network, which allowed us to evaluate the response time from the blockchain node. It should be noted that the write request to the blockchain network and the read request from the blockchain network were executed simultaneously. The write request to the blockchain network occurred according to the schedule shown in Fig. 3.

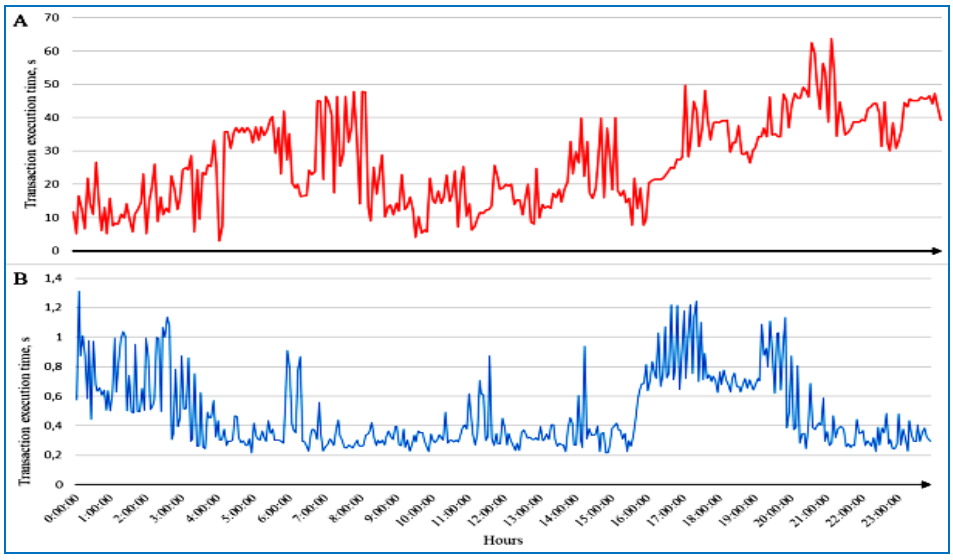


Fig. 3. Transaction processing time (A – for the Ethereum network, B – for the Waves network)

The read request from the blockchain network, in turn, occurred every 1 second. Thus, Fig. 3 (A) shows the results of writing data to the Ethereum blockchain network, and Fig. 3 (B) shows the results of writing data to the Waves blockchain network. It should be noted that all experiments are performed while waiting for confirmation of the transaction being written to the network. That is, to write the next transaction, a response from the blockchain network confirming that the transaction has been successfully processed and recorded is required. The need for confirmation of the write is due to the system's sensitivity to losses, as the loss of data from the OBU, in the case of recording fare data, could result in incorrect vehicle information.

As can be seen in Fig. 3 (A), the long transaction processing time and obvious losses are clearly visible. During the experiment, 3,000 transactions were written to the blockchain network, of which only 2,554 were successfully processed, meaning that 446 transactions were discarded and not registered in the blockchain network. The losses amount to 14.9%. In the case of writing data to the Waves blockchain network, there are no losses when writing transactions to the blockchain network. Fig. 3 (B) clearly demonstrates that the transaction writing speed is 20-40 times higher than in the Ethereum blockchain network. It is also necessary to take into account that private networks were configured to process incoming transactions as quickly as possible with relatively low hardware characteristics. Thus, when constructing the SD-IoV-Blockchain network architecture [5], the implementation of a single blockchain network node is only possible if the LPoS consensus algorithm is used.

Fig. 4 shows the result of processing a read request from the blockchain network. Similarly, a clear advantage can be seen in the Waves blockchain network with the LPoS algorithm, where the read request processing speed is, on average, 10 times higher than that of the Ethereum PoW blockchain network.

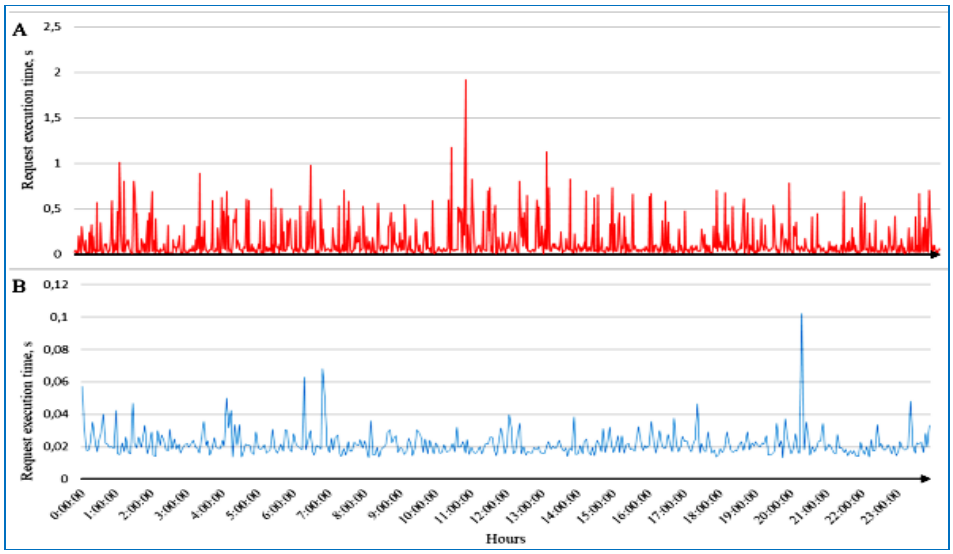


Fig. 4. Data reading processing graph (A – for Ethereum-based implementation, B – for Waves-based implementation)

Висновки

Based on the obtained research results, it can be concluded that the Proof of Stake consensus algorithms and the Waves blockchain network are more suitable than the classic Ethereum blockchain network under conditions of limited hardware specifications. This study demonstrates the feasibility of integrating and applying blockchain technology within the ITS concept and its overall suitability for use. It also clearly demonstrates the need for balancing depending on current network conditions, as the graphs show the periods of effectiveness of each algorithm. Implementing an adaptive algorithm for blockchain data flows would enable efficient data recording and hardware load balancing. Further research should be directed towards the development of mathematical models for optimizing distributed data exchange processes, taking into account the dynamic characteristics of the transport environment. The use of artificial intelligence and machine learning technologies for adaptive management of information flows in conditions of variable traffic intensity is promising. Particular attention should be paid to issues of data security and confidentiality in distributed networks, in particular the use of cryptographic methods and trust mechanisms between system nodes.

Another important direction is the integration of distributed technologies with 5G/6G communication systems to ensure ultra-reliable and low-latency communication (URLLC), which is critically important for Vehicle-to-Everything scenarios. Conducting simulation and experimental studies will help assess the effectiveness of the proposed solutions and contribute to the formation of architectural standards for new-generation intelligent transport systems.

СПИСОК ЛІТЕРАТУРИ

1. Dotsenko, N., Chumachenko, I., Galkin, A., Kuchuk, H. and Chumachenko, D. (2023), “Modeling the Transformation of Configuration Management Processes in a Multi-Project Environment”, *Sustainability (Switzerland)*, Vol. 15(19), 14308, doi: https://doi.org/10.3390/su151914308

2. Zuev, A., Karaman, D. and Olshevskiy, A. (2023), “Wireless sensor synchronization method for monitoring short-term events”, *Advanced Information Systems*, vol. 7, no. 4, pp. 33–40, doi: https://doi.org/10.20998/2522-9052.2023.4.04
3. Buterin, V., Illum, J., Nadler, M., Schär, F. and Soleimani, A. (2024), “Blockchain privacy and regulatory compliance: Towards a practical equilibrium”, *Blockchain Research and ApplicationsOpen source preview*, vol. 5(1), no. 100176, doi: https://doi.org/10.1016/j.bcra.2023.100176
4. Kasahara, S., Kawahara, J., Minato, S.-I. and Mori, J. (2023), “DAG-Pathwidth: Graph Algorithmic Analyses of DAG-Type Blockchain Networks”, I*EICE Transactions on Information and SystemsOpen source preview*, E106D(3), pp. 272–283, doi: https://doi.org/10.1587/transinf.2022FCP0007
5. Xia, Y., Hua, Z., Yu, Y., Zang, B. and Guan, H. (2022), “Colony: A Privileged Trusted Execution Environment with Extensibility”, *IEEE Transactions on Computers*, vol. 71(2), pp. 479–492, doi: https://doi.org/10.1109/TC.2021.3055293
6. Cocco, L. and Tonelli, R. (2024), “A Self-Sovereign Identity–Blockchain-Based Model Proposal for Deep Digital Transformation in the Healthcare Sector”, *Future Internet*, vol. 16(12), 473, doi: https://doi.org/10.3390/fi16120473
7. Kuchuk, N., Kovalenko, A., Ruban, I., Shyshatskyi, A., Zakovorotnyi, O. and Sheviakov, I. (2023), “Traffic Modeling for the Industrial Internet of NanoThings”, *2023 IEEE 4th KhPI Week on Advanced Technology*, KhPI Week 2023 - Conference Proceedings, 2023, doi: 194480. http://dx.doi.org/10.1109/KhPIWeek61412.2023.10312856
8. Kuchuk, H. and Malokhvii, E. (2024), “Integration of IOT with Cloud, Fog, and Edge Computing: A Review”, *Advanced Information Systems*, vol. *8*(2), pp. 65–78, doi: https://doi.org/10.20998/2522-9052.2024.2.08
9. Decker, C. and Wattenhofer, R. (2014), “Bitcoin transaction malleability and MtGox”, *European symposium on research in computer security*, 370, pp. 313–326, doi: https://doi.org/10.1007/978-3-319-11212-1_18

Received (Надійшла) 25.07.2025
Accepted for publication (Прийнята до друку) 23.10.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Кучук Ніна Георгіївна – доктор технічних наук, професор, професорка кафедри комп’ютерної інженерії та програмування, Національний технічний університет “Харківський політехнічний інститут”, Харків, Україна;
Nina Kuchuk – Doctor of Technical Sciences, Professor, Professor of Computer Engineering and Programming Department, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;
e-mail: nina_kuchuk@ukr.net; ORCID Author ID: http://orcid.org/0000-0002-0784-1465;
Scopus ID: https://www.scopus.com/authid/detail.uri?authorId=57196006131.

Трегубенко Максим Андрійович – студент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Maksym Tregubenko –student at the Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;
e-mail: Maksym.Tregubenko@nure.ua; ORCID Author ID: http://orcid.org/0009-0004-6206-7435.

Коваленко Данило Андрійович – студент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Danylo Kovalenko – student at the Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;
e-mail: Danylo.Kovalenko@nure.ua; ORCID Author ID: http://orcid.org/0000-0002-6465-7111 .

Лисиця Дмитро Олександрович – кандидат технічних наук, доцент кафедри комп’ютерної інженерії та програмування, Національний технічний університет “Харківський політехнічний інститут”, Харків, Україна;
Dmytro Lysytsia – Candidate of Technical Sciences, Associate Professor of Computer Engineering and Programming Department, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;
e-mail: Dmytro.Lysytsia@khpi.edu.ua; ORCID Author ID: https://orcid.org/0000-0003-1778-4676;
Scopus ID: https://www.scopus.com/authid/detail.uri?authorId=57220049627.

Бельорін-Еррера Олександра Михайлівна – кандидат наук, старший викладач кафедри комп’ютерної інженерії та програмування, Національний технічний університет “Харківський політехнічний інститут”, Харків, Україна;
Oleksandra Bellorin-Herrera – PhD, Senior Lecturer of Computer Engineering and Programming Department, National Technical University “Kharkiv Polytechnic Institute”, Kharkiv, Ukraine;
e-mail: Oleksandra.Bilorin-Ererera@khpi.edu.ua; ORCID Author ID: https://orcid.org/0000-0001-7974-5301;
Scopus ID: https://www.scopus.com/authid/detail.uri?authorId=59224314800.

Дослідження технологій розподіленого обміну даними у контексті інтелектуальних транспортних систем

Н. Г. Кучук, М. А. Трегубенко, Д. А. Коваленко, Д. О. Лисиця, О. М. Бельорін-Еррера

Анотація. У статті представлено результати дослідження технологій розподіленого обміну даними, орієнтованих на забезпечення ефективної взаємодії між компонентами інтелектуальних транспортних систем (ІТС). Розглянуто сучасні підходи до організації децентралізованих мереж передачі інформації, зокрема використання концепцій розподілених реєстрів, однорангових протоколів та сервісно-орієнтованих архітектур. Проведено аналіз вимог до надійності, затримки та пропускної здатності каналів зв’язку, що є критичними для сценаріїв обміну даними між транспортними засобами та інфраструктурою. На основі порівняльного аналізу визначено переваги та обмеження існуючих технологічних рішень у контексті забезпечення безпеки, масштабованості та стійкості до збоїв. Результати дослідження можуть бути використані для проєктування адаптивних моделей управління трафіком і розроблення інтелектуальних модулів взаємодії в транспортних мережах нового покоління.

Ключові слова: інтелектуальні транспортні системи; розподілений обмін даними; децентралізовані мережі; розподілений реєстр; масштабованість; надійність передачі даних.

UDC 004.932:004.032.26 **doi: 10.26906/SUNZ.2025.4.103-107**

Oleksandr Mozhaiev[1], Heorhii Kuchuk[1,2], Renat Safarov[1], Maksym Lavrovskyi[1], Kostiantyn Moroz[1]

[1] Kharkiv National University of Radio Electronics, Kharkiv, Ukraine
[2] National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine

STRUCTURAL AND FUNCTIONAL MODEL OF A CONVOLUTIONAL NEURAL NETWORK FOR PROCESSING, ANALYSING AND CLASSIFYING IMAGES OF VARYING COMPLEXITY

Abstract. The article presents presents a structural and functional model of a convolutional neural network (CNN) designed for processing, analysing and classifying images of varying complexity. The model is based on a multi-level architecture using convolutional, residual and parallel computing blocks, which ensure high adaptability to different types of input data. The input tensor is normalised by the mean and standard deviation of the sample, which reduces data variability and stabilises the learning process. The first convolutional layer performs the initial extraction of image features with ELU activation, which ensures continuity of gradients and eliminates the problem of 'dead neurons.' The implementation of skip connections ensures the consistency of information flows and increases the stability of training. The results demonstrate increased classification accuracy and noise resistance compared to basic CNN architectures.

Keywords: convolutional neural network; image processing; classification; feature analysis; structural-functional model; deep learning.

Introduction

In today's computer vision systems, flexible and robust models are needed to process images of varying complexity, from simple contour shapes to scenes with high noise levels and uneven lighting [1, 2]. Convolutional neural networks (CNN) have proven to be an effective tool for classifying and analysing visual data. However, classical architectures such as LeNet, AlexNet, or VGG demonstrate limited generalisation ability when working with heterogeneous datasets [3, 4].

Therefore, there is a need to build a structural-functional model capable of adapting to different levels of complexity of input images, minimising the loss of informative features and ensuring stable convergence during training [5, 6].

Problem statement. *The aim of the research* is to develop and mathematically formalise a structural and functional model of a convolutional neural network optimised for multi-level image processing, analysis and classification.

To achieve this aim, it is necessary to:

1. Describe the network architecture, taking into account multi-scale feature processing.
2. Justify the choice of activation functions and residual connection mechanisms.
3. Implement a mathematical model that ensures the stability of the gradient flow.
4. Conduct an experimental evaluation of the accuracy and noise resistance of the model.

Convolutional neural network architecture model I

Let us consider the architecture of the developed convolutional neural network, which is a specialised architecture for the regression of four critical parameters of the adaptive local contrast algorithm [7, 8]. Unlike classical contrast methods (CLAHE, Multi-ScaleRetinex), the proposed model provides dynamic adaptation to local statistical characteristics of the image, which is critical for processing in conditions of non-stationary visibility.

Fig. 1 shows a topological diagram of the CNN architecture for adaptive control of adaptive local contrast parameters.

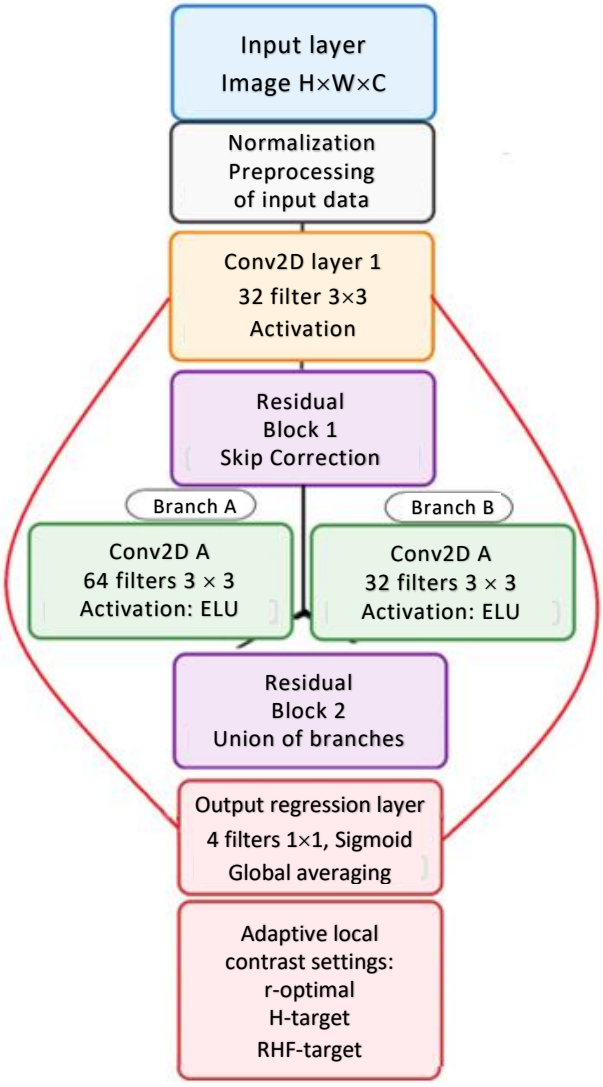


Fig. 1. CNN architecture diagram

Input layer. The input tensor $X0$ is defined as [9]:

$$X0 \in \mathrm{R} \wedge (H \times W \times C), \quad (1)$$

where H, W, C represent height, width, and number of channels, respectively;

for monochrome images, $C = 1$; for RGB images, C = 3.

Preliminary normalisation is performed according to:

$$\hat{x}_{i,j} = \frac{x_{i,j} - \mu}{\sigma}, \quad (2)$$

where μ and σ are the mean value and standard deviation of the sample.

Convolution layer 1. The convolution operation of the first layer is mathematically described as [10]:

$$Y1_{i,j,k} = \\ = ELU\left(\sum_m \sum_n \sum_s \begin{pmatrix} W1_{m,n,s,k} \cdot X0_{i+m,j+n,s} + \\ +b1_k \end{pmatrix}\right), \quad (3)$$

where $W1 \in \mathrm{R}\wedge(3 \times 3 \times \mathrm{C} \times 32)$ – filter weight tensor;

$b1 \in \mathrm{R}^{32}$ – offset vector;

$k \in \{1, 2, ..., 32\}$ – output feature map index.

The ELU activation function is defined as

$$ELU(x) = \begin{cases} x, & \text{if } x > 0; \\ \alpha\left(e^x - 1\right), & \text{if } x \le 0, \end{cases} \quad (4)$$

where $\alpha = 0.1$ is the saturation parameter.

Residual connection block. Implementation of skip connection with identical mapping [11]:

$$Z1 = Y1 + F\_identity(X0), , \quad (5)$$

where *F_identity* – function of reducing the dimension of the input tensor to the dimension of the output tensor through a convolution operation 1 × 1.

Parallel convolution blocks. Branch A (Conv2D with 64 filters):

$$Y2^A_{i,j,k} = \\ = ELU\left(\sum_m \sum_n \sum_s \begin{pmatrix} W2^A_{m,n,s,k} \cdot Z1_{i+m,j+n,s} + \\ +b2^A_k \end{pmatrix}\right). \quad (6)$$

Branch B (Conv2D with 32 filters):

$$Y2^B_{i,j,k} = \\ = ELU\left(\sum_m \sum_n \sum_s \begin{pmatrix} W2^B_{m,n,s,k} \cdot Z1_{i+m,j+n,s} + \\ +b2^B_k \end{pmatrix}\right). \quad (7)$$

Residual connection block 2 with dimensional conversion:

$$Z2 = Y2^B + Conv1\times1\left(Y2^A\right), \quad (8)$$

where $Conv1\times1$ ensures channel number coordination.

Regression output layer. Final convolution operation to obtain a parameter map:

$$\hat{Y}_{i,j,k} = \\ = \mathrm{Sigmoid}\left(\sum_m \sum_n \begin{pmatrix} W\_out_{m,n,k} \times \\ \times Z2_{i+m,j+n} + b\_out_k \end{pmatrix}\right), \quad (9)$$

Global averaging to obtain scalar parameters of ALC:

$$\theta_k = \frac{1}{HW} \cdot \sum_i \sum_j \hat{Y}_{i,j,k}, \\ k \in \{r, Tc, H\_T\,\mathrm{arg}\,et, RHF\_T\,\mathrm{arg}\,et\}. \quad (10)$$

In the following section, we will examine the rationale behind the architectural solutions.

Choosing the ELU activation function. In the context of contrast tasks, the ELU function offers critical advantages over ReLU and its modifications [12]:

1. *Elimination of the 'dead neuron' problem*: negative activation values preserve the gradient flow, which is critical for analysing low-contrast areas of an image.

2. *Zero centring of activations*: $\mathrm{E}[\mathrm{ELU}(x)] \approx 0$ contributes to faster convergence during training by improving the condition number of the Hessian matrix.

3. *Continuity of the derivative*: the smoothness of the function at $x = 0$ ensures the stability of gradient descent.

Residual connections (Skip Connections): the use of a ResNet-like architecture is due to the specifics of the contrast parameter regression task [13, 14]:

1. *Preservation of original information*: the identity mapping $\mathrm{F}(x) + x$ prevents the degradation of low-frequency components of the image.

2. *Solving the vanishing gradient problem*: the direct path of gradients through skip connections accelerates the training of deep networks.

3. *Adaptability to different image types*: residual training $F(x) = H(x) – x$ allows the network to focus on the increments necessary for a specific type of scene.

2. Adaptive contrast control mechanism

A key feature of the developed architecture is its integration with the adaptive local contrast algorithm through the regression of four target parameters, including *r – the size of the adaptive window*.

The regressed value of *r* determines the locality of the contrast analysis [15]:

$$r\_optimal = CNN(I) \in [r\_min; r\_max], \quad (11)$$

where the optimal value depends on the local entropy $H(x, y)$ and frequency characteristics $RHF(x, y)$, and sharp changes are smoothed by the quality estimate of the previous image frame *Q_prev*.

Contrast threshold value (*Tc*), Parameter *Tc* controls the sensitivity of the algorithm to local brightness variations:

$$Tc\_optimal = CNN(I) \cdot \sigma\_local(x, y). \quad (12)$$

Target statistical characteristics (H_target, RHF_target) and CNN predict the optimal values of

entropy and frequency response that should be aimed for adaptive local contrast [6]:

$$H\_target = CNN\_H(I); \quad RHF\_target = CNN\_RHF(I). \tag{13}$$

Loss function and training methodology Unlike standard approaches using $L2$ regression, a specialised loss function has been developed:

$$L\_total = \lambda 1 \cdot Q\_prev + \lambda 2 \cdot L\_consistency + \lambda 3 \cdot L\_stability, \tag{14}$$

where *Q_prev* – non-reference image quality metric;

L_consistency – measure of consistency of parameters with local characteristics;

L_stability – time stability regulator for video sequences.

Gradient learning through adaptive local contrast feedback.

The uniqueness of the approach lies in end-to-end learning through the chain (Fig. 2):

$$\frac{\partial L}{\partial W} = \frac{\partial L}{\partial Q} \cdot \frac{\partial Q}{\partial I\_ALC} \cdot \frac{\partial I\_ALC}{\partial \theta} \cdot \frac{\partial W}{\partial I\_ALC}, \tag{15}$$

where $\frac{\partial I\_ALC}{\partial \theta}$ is of the adaptive local contrast algorithm for evaluating the previous frame *Q_prev*.

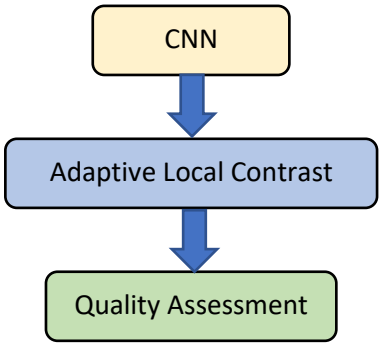


Fig. 2. The learning chains

CLAHE (Contrast Limited Adaptive Histogram Equalisation):

• fixed window size and cliplimit parameters;

• no adaptation to spectral characteristics;

• computational complexity O(N^2) for a window size of $N \times N$.

Multi-ScaleRetinex:

• requires preliminary adjustment of weight coefficients;

• sensitivity to haloeffect artefacts;

• limited applicability to IR images.

Proposed CNN-adaptive local contrast approach:

• automatic adaptation of parameters to scene characteristics;

• versatility for visible and IR ranges;

• computational efficiency O(1) after CNN training.

Features of implementation on FPGA:

• features of weight and activation quantisation for hardware implementation 8-bit quantisation with precision preservation is used:

$$W\_quantized = \frac{round\left(W\_float \cdot 2^n\right)}{2^n}, \tag{16}$$

where n = 7 for weights, n = 8 for activations.

ELU function approximation:

• on PLIS, the ELU function is implemented through piecewise linear approximation:

$$ELU\_approx(x) = \begin{cases} x, \ if \ x > 0; \\ LUT\_exp(x), \ if \ x \le 0, \end{cases} \tag{17}$$

where *LUT_exp* – precomputed table of exponential values.

Pipeline processing:

• The PLIS architecture enables pipeline processing with a latency of 3 cycles per pixel at a frequency of 150 MHz, providing a throughput of > 60 FPS for a resolution of 1920 × 1080.

3. Experimental validation and performance metrics

The effectiveness of the developed architecture has been confirmed by comprehensive experiments on diverse datasets, including aerial photography, IR images, and images taken in conditions of limited visibility [6].

The quality metrics achieved exceed those of classical methods by 15-25% according to BRISQUE, NIQE, and expert evaluation criteria.

When analysing contrast-based local processing, optimisation techniques were found that allow image frames to be processed on existing software and hardware complexes.

In this case, an algorithm was used to operate the complexes in real time (resolution of more than 2 megapixels with a frame rate of up to 50 Hz).

The influence of adaptive local contrast parameters manifests itself in the following aspects:

1. *The noise level depends on:*

• window size:

large size → strong noise suppression, blurring of details;

small size → preservation of details, weak noise suppression;

• β parameter:

low value → noise suppression, loss of contrast;

high value → preservation of contrast, noise amplification;

• δ parameter:

low value → increased contrast, loss of detail;

high value → preservation of detail, insufficient contrast.

2. *Blockiness artefacts*:

• large window size can cause blockiness;

• small window size reduces blockiness but can increase noise.

3. *Preservation* of fine details:

• optimal with small window size and high β value;

• deteriorates with large window size and low β value.

4. *Brightness distribution*:

• adaptive local contrast strives for uniform distribution;

• the shape of the histogram depends on the adaptive local contrast parameters.

5. *Frequency composition of the image:*

• changes under the influence of adaptive local contrast;

• depends on the DCP-RHF parameter.

Let us consider a one-dimensional case. Let there be a convolution kernel [-1, 1] (difference calculation operator).

If the input signal is

[10, 10, 10, 20, 20, 20]

(weak gradient), then the convolution output will be

[0, 0, 10, 0, 0].

If the input signal is

[0, 0, 0, 255, 255, 255] (sharp gradient),

then the convolution output will be

[0, 0, 255, 0, 0].

As the contrast increases, the amplitude of the output signal increases.

Low-contrast images often contain fine details that are difficult to distinguish due to a low signal-to-noise ratio. Increasing contrast makes these details more visible, allowing CNNs to extract more informative features.

Change in feature statistics.

Increasing contrast changes the statistical distribution of pixel values and, consequently, the statistical distribution of neuron activations in convolutional layers. This can cause the CNN to highlight other, more relevant features.

Impact on non-linear activation functions.

Activation functions (e.g., ReLU, sigmoid) in CNNs respond differently to input signals with different amplitudes. Increasing contrast can change the nature of neuron activation, affecting the spread of information across the network.

For example, ReLU is activated only for positive input values. If the signal is weak, most neurons may be inactive.

Increasing contrast can ‘turn on’ more neurons, making the representation richer. When analysing contrastive local processing, optimisation techniques were found that allow image frames to be processed on existing software and hardware complexes. In this case, the algorithm for the operation of the complexes in real time was used (resolution of more than 2 megapixels with a frame rate of up to 50 Hz).

Conclusions

A method for selecting the initial coefficients of the contrast modification algorithm for three types of image receivers has been developed.

On this basis, a local contrast algorithm with processing zones that automatically adjust to the image subject has been developed.

Calculations were obtained for the distances between sensors for image synthesis (complexing) in order to minimise information from different channels for a common receiver, as well as to eliminate occlusions in the scene that block the target.

The implementation of the developed hybrid algorithm on the Xilinx Kintex-7 XC7K325T FPGA demonstrated high performance.

When using CNN for regression of ALK parameters, a processing speed of over 60 frames/s is achieved for images with a resolution of 1920x1080.

A simplified version of the algorithm (without dynamic regression of CNN parameters, possibly with fixed or slowly updated parameters) provides performance of over 120 frames per second. FPGA resource consumption is less than 70% for the full algorithm with CNN and less than 40% for the simplified version. The system's power consumption does not exceed 5 W.

СПИСОК ЛІТЕРАТУРИ

1. Hussain, S., Lu, L. Mubeen, M., Nasim, W., Karuppannan, S., Fahad, S., Tariq, A., Mousa, B. G., Mumtaz, F. & Aslam, M. Spatiotemporal variation in land use land cover in the response to local climate change using multispectral remote sensing data. Land, 2022, vol. 11, no. 5, article no. 595. DOI: https://doi.org/10.3390/land11050595
2. Yaloveha, V., Hlavcheva, D., Podorozhniak, A. & Kuchuk, H. Fire hazard research of forest areas based on the use of convolutional and capsule neural networks. 2019 IEEE 2nd Ukraine Conference on Electrical and Computer Engineering (UKRCON), 2019, pp. 828-832. DOI: https://doi.org/10.1109/UKRCON.2019.8879867
3. Radočaj, D., Jurišić, M. & Gašparović, M. The Role of Remote Sensing Data and Methods in a Modern Approach to Fertilization in Precision Agriculture. Remote Sensing, 2022, vol. 14, no. 3, DOI: https://doi.org/10.3390/rs14030778
4. Munawar, H. S., Hammad, A. W. A. & Waller, S. T. Remote Sensing Methods for Flood Prediction: A Review. Sensors, 2022, vol. 22, no. 3, article no. 960. DOI: https://doi.org/10.3390/s22030960
5. Barabash, O., Bandurka, O., Svynchuk, O. & Tverdenko, H. Method of identification of tree species composition of forests on the basis of geographic information database. *Advanced Information Systems*, 2022, vol. 6, no. 4, pp 5-10. DOI: https://doi.org/10.20998/2522-9052.2022.4.01
6. Svyrydov, A., Kuchuk, H. and Tsiapa, O. (2018), “Improving efficienty of image recognition process: Approach and case study”, *Proceedings of 2018 IEEE 9th International Conference on Dependable Systems, Services and Technologies*, DESSERT 2018, pp. 593–597, doi: https://doi.org/10.1109/DESSERT.2018.8409201
7. Alganci, U., Soydas, M. & Sertel, E. Comparative research on deep learning approaches for airplane detection from very high-resolution satellite images. Remote Sensing, 2020, vol. 12, no. 3, article no. 458. DOI: https://doi.org/10.3390/rs12030458

8. Kuchuk, H. and Malokhvii, E. (2024), “Integration of IOT with Cloud, Fog, and Edge Computing: A Review”, *Advanced Information Systems*, vol. *8*(2), pp. 65–78, doi: https://doi.org/10.20998/2522-9052.2024.2.08
9. Weiss, K., Khoshgoftaar, T. M. & Wang, D. D. A survey of transfer learning. Journal of Big data, 2016, vol. 3, article no. 9. DOI: https://doi.org/10.1186/s40537-016-0043-6
10. Hlavcheva, D., Yaloveha, V., Podorozhniak, A. & Kuchuk, H. Tumor nuclei detection in histopathology images using R – CNN. CEUR Workshop Proceedings, 2020. vol. 2740, pp. 63-74. Available at: https://ceur-ws.org/Vol-2740/20200063.pdf
11. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770-778. DOI: https://doi.org/10.1109/CVPR.2016.90
12. Yaloveha, V., Podorozhniak, A. & Kuchuk, H. CNN hyperparameter optimization applied to land cover classification. Radioelectronic and computer systems, 2022, no. 1 (101), pp. 115-128. DOI: https://doi.org/10.32620/reks.2022.1.09
13. Hlavcheva, D., Yaloveha, V., Podorozhniak, A. & Kuchuk, H. Comparison of CNNs for Lung Biopsy Images Classification. 2021 IEEE 3rd Ukraine Conference on Electrical and Computer Engineering, UKRCON 2021 – Proceedings, 2021, pp. 1–5. DOI: https://doi.org/10.1109/UKRCON53503.2021.9575305
14. Tan, M. & Le, Q. V. Efficientnetv2: Smaller models and faster training. ArXiv (Cornell University), Preprint arXiv:2104.00298, 2021. DOI: https://doi.org/10.48550/arXiv.2104.00298
15. Carneiro, T., Da Nóbrega, R. V. M., Nepomuceno, T., Bian, G.-B., De Albuquerque, V. H. C. & Filho, P. P. R. Performance analysis of google colaboratory as a tool for accelerating deep learning applications. IEEE Access, 2018, vol. 6, pp. 61677-61685. DOI: https://doi.org/10.1109/ACCESS.2018.2874767

Received (Надійшла) 18.07.2025
Accepted for publication (Прийнята до друку) 29.10.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Можаєв Олександр Олександрович – доктор технічних наук, професор, професор електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Oleksandr Mozhaiev – Doctor of Technical Sciences, Professor, Professor of Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;
e-mail: mozhaev1957@gmail.com; ORCID Author ID: http://orcid.org/0000-0002-1412-2696;
Scopus ID: https://www.scopus.com/authid/detail.uri?authorId=57201729490.

Кучук Георгій Анатолійович – доктор технічних наук, професор, професор електронних обчислювальних машин, Харківський національний університет радіоелектроніки; професор кафедри комп’ютерної інженерії та програмування, Національний технічний університет “Харківський політехнічний інститут”, Харків, Україна;
Heorhii Kuchuk – Doctor of Technical Sciences, Professor, Professor of Department of Electronic Computers, Kharkiv National University of Radio Electronics; Professor of Computer Engineering and Programming Department, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;
e-mail: kuchuk56@ukr.net; ORCID Author ID: http://orcid.org/0000-0002-2862-438X;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57057781300.

Сафаров Ренат Кудярович – студент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Renat Safarov – student at the Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;
e-mail: Renat.Safarov@nure.ua; ORCID Author ID: http://orcid.org/0009-0004-7029-817X.

Лавровський Максим Валерійович – студент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Maksym Lavrovskyi – student at the Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;
e-mail: maksym.lavrovskyi@nure.ua; ORCID Author ID: http://orcid.org/0009-0007-1960-7119.

Мороз Костянтин Сергійович – студент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Kostiantyn Moroz – student at the Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;
e-mail: kostiantyn.moroz@nure.ua; ORCID Author ID: http://orcid.org/0009-0007-2751-5249.

Структурно-функціональна модель згорткової нейронної мережі для обробки, аналізу та класифікації зображень різної складності

О. О. Можаев, Г. А. Кучук, Р. К. Сафаров, М. В. Лавровський, К. С. Мороз

Анотація. У статті представлено структурно-функціональну модель згорткової нейронної мережі (ЗНМ), призначену для обробки, аналізу та класифікації зображень різної складності. Модель базується на багаторівневій архітектурі з використанням згорткових, залишкових і паралельних обчислювальних блоків, що забезпечують високу адаптивність до різних типів вхідних даних. Вхідний тензор нормалізується за середнім і стандартним відхиленням вибірки, що дозволяє зменшити варіативність даних і стабілізувати процес навчання. Перший згортковий шар виконує початкове виділення ознак зображення з активацією типу ELU, яка забезпечує безперервність градієнтів та усуває проблему “мертвих нейронів”. Реалізація залишкових зв’язків (skip connections) забезпечує сталість інформаційних потоків та підвищує стабільність навчання. Отримані результати демонструють підвищену точність класифікації та стійкість до шумів у порівнянні з базовими CNN-архітектурами.

Ключові слова: згорткова нейронна мережа; обробка зображень; класифікація; аналіз ознак; структурно-функціональна модель; глибоке навчання.

UDC 004.93 **doi: 10.26906/SUNZ.2025.4.108-113**

Olena Peredrii

Simon Kuznets Kharkiv National University of Economics, Kharkiv, Ukraine

SHALLOW ANN MODELS TO CLASSIFY UKRAINIAN AI-GENERATED TEXT

Abstract. In this study, we address the task of detecting AI-generated fragments within Ukrainian-language texts. The objective is to develop a tool capable of identifying content produced with the assistance of artificial intelligence, particularly in PDF documents related to the IT domain. The research explores and analyzes existing solutions and approaches currently available in this area. Several commercial AI-content detectors were evaluated using our custom datasets. The dataset was constructed by segmenting bachelor's theses from IT-related fields into fragments of approximately 1,000 characters each. Five artificial neural network models were tested using the custom dataset combined with a traditional NLP pipeline, achieving an accuracy of 87–88%. Given the complexity of the problem and the ethical considerations within the educational context, the classification results should be further validated by human experts. The current implementation can serve as a foundation for future improvements.

Keywords: artificial neural networks (ANN); shallow ANN models; AI-generated content (AIGC); LLM; AI detector.

Introduction

Artificial Intelligence (AI) and Generative AI (GenAI) in the form of Large Language Models (LLMs) (often referred to as "AI" these days) are fundamentally changing not only life and work, but also the educational process. The widespread use of hundreds of LLMs in the form of chatbots in recent years has opened up many opportunities for both students and teachers. These include organizing personalized learning with the help of chatbots, automating routine tasks, creating tests and assignments (and even grading them), adapting course content, and more [1]. However, the use of GenAI tools also poses significant risks to academic integrity and ethical usage by both students and teachers [2]. A major challenge for teachers is understanding whether the solutions to problems proposed by students are simply artificially generated [1, 2]. When students use AI to perform cognitive tasks that require thinking, forming their own thoughts in the form of text, or generalizing, it undermines the fundamental goal of academic work, which is to develop critical thinking and master clear communication skills [1].

On the other hand, the introduction of technical tools that can create a basis for unfairly accusing students or teachers of violating academic integrity can increase the number of avoidable conflicts, as well as reduce the level of trust in the academic environment.

Despite this, the problems of violating academic integrity are relevant (checking texts for possible plagiarism has been performed automatically for many years), and tools for automatically identifying generated content in the form of texts or images continue to be created and become widespread both in the academic environment and when solving professional tasks related to working with content and requirements for it.

Review of existing solutions

Despite the flaws and drawbacks of AI content tools mentioned above, AI-created content detection problem generated huge scientific interest over the last few years [3–6]. Machine-generated content detectors are designed to identify text produced by AI (AI-generated content - AIGC), often asserting their effectiveness across various conditions and different language models. These detection methods are generally categorized into three primary approaches [3, 7, 8] described below.

Trained Detectors. These systems leverage labeled datasets comprising both human-written and AI-generated texts to train binary classification models. Through this training, they learn to recognize distinct patterns, structures, and stylistic characteristics unique to each type of text.

The frequent choice for this approach is to use BERT-based pretrained models as feature extractors for the text being analyzed. An example of this could be [9], where authors used RoBERTA models to classify AI-generated texts with 95% accuracy.

The representative of this family could also be GPTSentinel [10, 11], which includes RoBERTa-Sentinel and T5-Sentinel models. The first one uses RoBERTa pretrained model to extract features of the text with further classification of them using MLP. The second approach represents a sequence-to-sequence problem solved using the T5 model and producing "positive" or "negative" classification keywords in text form. Both models achieved over 97% accuracy.

RADAR [12–14] is the other example from the supervised classifier family. It addresses the issue with the failure of AI-text detections when facing LLM-paraphrased texts (available for the English language only).

The main drawbacks of the trained detectors family are obvious: building these models requires AI/ML/DL knowledge, the dataset according to the problem being solved, the hardware and time resources to train, deploy, and support the solution. But such models are effective in the particular domain, and could be modified or improved according to new requirements.

Zero-Shot Detectors. Unlike trained detectors described above, these methods do not require explicit training on AI-generated text from specific models. Instead, they infer the origin of AI text based on general linguistic properties, statistical features, and anomalies, or the inherent predictability of machine-generated content. The perplexity, burstiness, N-gram, and probability analysis are often mentioned in papers for the methods of this family.

It was shown that AI-generated text tends to occupy negative curvature regions of the model's log probability function, so this could be the only classification criterion [15]. DetectGPT detector [15] does not require training of a special classifier, but uses perturbation of the text to build a log probability function to make the decision.

Fast-DetectGPT [16] was proposed as an update of DetectGPT and included a performance boost and accuracy improvement.

The main limitation for this set of methods is the availability of the probabilities for the model's dictionary; for instance, a lot of popular LLM models don't reveal them.

Watermarking Techniques. This approach involves embedding a subtle, hidden pattern directly into the generated output, which can later be detected to confirm its AI origin [7]. While theoretically robust, this method is less commonly employed in black-box detection scenarios where the detector has no control over the content generation process. Probably, the continuous advances in LLM will make watermarking the only method to classify AIGC in the future.

There is also a separate field of commercial detectors, e.g. GPTZero, ZeroGPT, Originality.ai, and others. They use different methods and their combinations to classify AI-generated content, and often provide benchmarks to compare quality with competitors.

GPTZero [17] reported an accuracy over 99% and a false positive rate below 1% [18] for the languages it supports. There is no official support for Ukrainian, but results are available for this language with a "partially supported" system message.

ZeroGPT [19] claims 98% accuracy, based on training with 10 million AI-generated articles, and support of all available languages.

Originality.ai [20] claims 99% AI Content Detection Accuracy, positioning itself as a highly accurate detector for various AI writing tools, including GPT4.1, ChatGPT4o, Gemini 2.5, Claude 3.7, and DeepSeek V3, supports Ukrainian.

Copyleaks AI Detector [21] offers over 99% accuracy, supports 30+ languages, and detects popular LLM models, including ChatGPT, Gemini, and Claude, as well as newer models as they're released.

Isgen.ai supports over 80 languages, including Ukrainian, and claims to reach over 96% accuracy [22].

The consistent claims of very high accuracy (98-99%+) from commercial tools like Copyleaks, Originality.ai, and ZeroGPT contrast with scientific papers' research, which report on the various problems and accuracy issues with unseen data.

Finally, there are many enthusiastic projects (e.g., located on GitHub) with custom AI-text detectors, but a lot of them are already abandoned.

As a summary of this section, we can conclude that the problem of AIGC classification is complex both from a research and a practical point of view. A lot of LLMs exist, and they are evolving quickly, providing better and better quality of content, which requires constant updating of AI-detectors, which themselves are not very precise in practice initially. More control over quality and classification process is possible with the training of custom models in the domain of the problem, but this approach requires knowledge, data, time, and resources for the development, deployment, support, and continuous improvement. There are no good solutions to support the detection of AIGC in the Ukrainian language.

Ethical considerations

Modern AIGC detection tools are widely known for their unreliability and classification errors, especially in cases where the author of the text is not a native speaker of the verification language [1].

False accusations of academic misconduct arising from unreliable AI detectors can have serious consequences for authors, whether they are students or teachers. The widespread problem is not just the inaccuracy of these detectors, but also the underlying reason for this: they are designed to detect statistical patterns, but human writing is inherently varied and unpredictable, and artificial neural networks and quite good (and constantly improving) at detecting these statistical features. Additionally, authors have no tools to prove the authorship of the content when some AIGC detection tool mistakenly classifies the content to be AI-generated. There is an opinion [23] about what educational institutions should do instead of screening AI-generated content with AI-detectors. These actions include improving teaching and process-based assessment methods, oral exams and discussions, and educating students on the responsible use of AI tools.

The problem definition

The idea of the work is to build a research AI tool that is useful in the detection of the AIGC in the PDF documents in the IT domain. Taking into account all previous considerations, we want to focus on the principal axioms we are following in the research:

– this tool is not used for the automatic AI decision-making about AIGC;

– the developed tool is not used for the blind blaming of authors, it is much more a helpful assistant/recommender for the checker to find strange/inconsistent pieces of text for further improvements;

– we don't focus on the high-accuracy numbers a lot; we understand the complexity of the problem, and care more about reliable and honest results for this particular dataset, domain, and initial conditions.

Dataset

The dataset was created using the processing of 38 files with bachelor diplomas of students graduated in 2022-2024. The original text was split into segments of about 1000 symbols each, with 150 symbols overlapping between chunks without breaking words. The chunking process is important here as it defines the further usage and the evaluation of the models. This is very difficult to define if a short text (e.g., a sentence) was written by a human or generated from our point of view, so we created chunks of some valuable size. On the other hand, the results for each sentence classification are interesting, but future models should operate with bigger chunks. Finally, the dataset contains the chunks, even when by

1000-symbol pieces, it is impossible to detect whether it was generated or human-written. We left these examples deliberately to see how they would be classified, so the dataset is very challenging initially.

The chunks obtained after splitting were processed manually in order to remove unnecessary spaces, break lines, and special symbols, which could lead to classification based on them. After this cleaning, we left 2533 chunks.

For each cleaned human-written chunk, we generated its AI-written version using "gpt-4o-mini" model and this prompt: "Analyze the topic of the text below and create similar text in Ukrainian. Don't provide your summary, just use it for generation. Don't use Markdown, just plain text. Don't rewrite the text, generate your own:". After that, we received more than 15000 chunks (a few chunks appeared instead of a single one because LLM added a lot of break lines), which were processed manually as well. Some chunks contained summary, markdown symbols (so, our model may fail for the obvious case when markdown symbols show that the text is definitely AI-generated), a lot of lines, or just bad text. After the cleaning, we obtained 2634 chunks, and this quantity is comparable with 2533 human-written chunks, which makes this dataset balanced.

Evaluation with different existing AI-detectors

We tested how known commercial detectors work for our custom dataset. We used evaluation versions of the software for all of them for classification, but followed our main idea – to classify text chunks grabbed from the documents in an automatic manner, exactly like we are going to do. The quantity of experiments differs between commercial tools because they provide different quantities of text tokens available for free (and no upload of big documents), but we tested the same text chunks for fair comparison.

For the five human-written text chunks classified by GPTZero we received five uncertain decisions, only one of them showed that the text was created by a human (74% human and 26% AI), all others were shown as AI-generated. Two text chunks from the five AI-generated texts were classified as human texts (all of them with uncertain decisions again).

ZeroGPT tool detected nine out of ten human-written pieces as AI-generated (the correct sample contains 44% of AI, all others – at least 87%), and all ten AI-generated samples were classified correctly (with a 98% minimum quantity of AI-content).

Originality.ai classified 2 out of 10 human-written text chunks as original text (with high confidence – 98%) and 8 others as AI-generated (also with high – at least 97% – confidence). One AI-generated chunk was classified as human-written, the other 9 were classified correctly (all of them with at least 97% confidence).

We were able to test 20 text chunks with Copyleaks detector, 4 human-written chunks were identified as AI-generated (100% of AI-generated content), and 16 chunks were classified correctly (0% AIGC). The situation is similar for AI-generated pieces of text: 5 of them were misclassified as human-written (0% AIGC), and 15 were classified as AI-written (100% AIGC).

All five human-written chunks were correctly classified by the Isgen.ai tool, but 2 AI-generated pieces out of 5 were classified also as human text (notifying that 100% content was human-created).

Grid search for the shallow architecture of the model

We used a traditional NLP approach to solve text classification problems: build an ANN model with one embedding layer, a single dense layer, and an output layer containing just one neuron. One of our goals is to create such shallow models that could be trained on a single CPU/GPU using regular hardware available for everyone. 80% of the dataset were used as training/validation, and the remaining 20% were used as a test part.

The preprocessing of the text chunks includes the calculation of the vocabulary size for the training part of the dataset, training of a tokenizer for all training chunks according to the vocabulary size, and pre-padding of vectors to align all their lengths. The lengths of vectors in the train data vary from about 230 symbols (appearing as text chunks at the end of documents) to 1913 symbols (appearing because AI-generated chunks sometimes are longer compared to the initial ones), but about 90% of lengths are in [850;1150] range. This tokenizer was serialized and used during testing of the models.

The effective architecture of such shallow models could be found with grid search, exploring a lot of possible combinations of hyperparameters, revealing also the opportunity to create ensembles of the models to make a final classification decision.

Input vectors are fed into the first embedding layer of the model, which converts integer labels of the words according to the dictionary to meaningful continuous values during training. The input parameters for this layer are known already: size of vocabulary, size of the input vector, and the size of the output vector – the hyperparameter for this layer that should be defined via grid search. Models tend to overfit during training, so we tested values in [1;10] range as the size of the output vector. The quantity of neurons for the dense layer we tested was the powers of 2 in the range [2;128]. There were no benefits in using longer embedding vectors, pretrained embeddings, or an additional dense layer for this task.

Training of nets was performed using Adam optimizer, L2 regularizer, binary cross-entropy loss function, RELU activation for the middle dense layer, sigmoid for the last one, dropout layer, and early stopping procedure, saving the best state of the model before it starts overfitting.

Fig. 1 shows the entire picture of the grid with the test accuracies of the models (better accuracies are highlighted with green color).

Based on these results, we choose such five model architectures: model 1 contained 2 embedding neurons and 256 dense neurons, model 2 – 3 embeddings and 64 dense ones, model 3 – 4 embeddings and 64 dense neurons, model 4 included 6 embedding neurons and 64 dense neurons, finally, model 5 was built with 10 embedding and 128 dense neurons.

dense layer neurons	embeddings out quantity									
	1	2	3	4	5	6	7	8	9	10
2	0,51	0,51	0,75	0,51	0,51	0,51	0,51	0,51	0,51	0,51
4	0,72	0,51	0,51	0,51	0,51	0,51	0,51	0,51	0,51	0,51
8	0,75	0,79	0,72	0,71	0,51	0,51	0,51	0,51	0,51	0,51
16	0,83	0,84	0,87	0,51	0,51	0,72	0,81	0,51	0,51	0,51
32	0,8	0,87	0,85	0,82	0,74	0,86	0,87	0,73	0,71	0,87
64	0,54	0,83	0,89	0,86	0,87	0,88	0,84	0,87	0,87	0,87
128	0,76	0,85	0,84	0,86	0,88	0,87	0,87	0,88	0,88	0,89
256	0,88	0,84	0,87	0,88	0,86	0,87	0,86	0,86	0,87	0,85

Fig. 1. Results of grid search
for the selection of ANN architecture

Test accuracies for them were 88.00 (F1 score for human class was 0.88, AI – 0.88), 87.62 (0.87, 0.88), 87.91 (0.87, 0.88), 87.75 (0.87, 0.87), and 87.13 (0.87, 0.87), respectively.

Ensembles

Using ensembles of ANN instead of making the decision by just a single ANN can achieve better results for the problems with insufficient data (multiple models can learn different features of the data), or when a single model does not provide good accuracy (a few models can compensate the classification mistakes made by others).

There are a lot of ways to make a joint decision based on the outputs from different models. They include averaging of outputs, prediction the majority of models after voting, different weighted schemes of models' outputs, boosting or even stacking of the models: training another classification model based on the outputs of models in an ensemble. In the context of our problem, we considered the ensemble as a way to improve classification accuracy and provide a broader picture of the decisions about the chunk of text.

Referring to the accuracy of five models as the probabilities of correct classification, it is possible to calculate the corresponding probability for the ensemble according to the aggregating scheme. We tested the majority voting approach, where the decision is made in favor of the class that receives the most votes from separate models. This was not successful, as all models failed synchronously for the same text chunks from the test part of the dataset. We also tested AdaBoost and GradientBoosting classifiers without success.

Chunks and sentences classification

The usage of a trained model for text chunks processing is straightforward and presented in Fig. 2. The initial PDF document is divided into non-overlapping text chunks of about 1000 symbols each, and each of these chunks is processed using five models M1 – M5. The result is averaged per chunk, and the statistics are gathered to show the overall probability that a random text chunk from the document is AI-generated, with the histogram showing the summary of classification results for chunks.

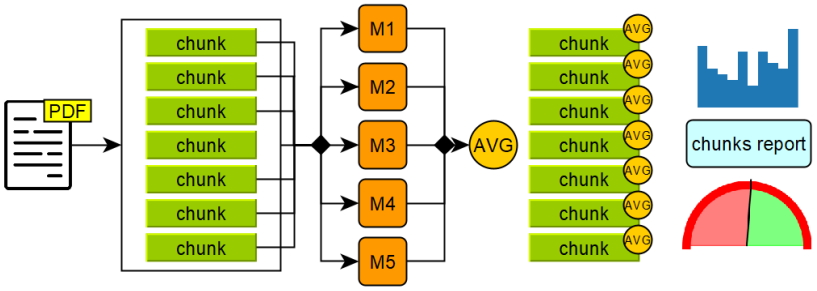


Fig. 2. The processing pipeline to create chunks report

This is known [24] that the classification of longer text chunks is more reliable compared to the classification of sentences. The models we have trained were designed to work with chunks of about 1000 symbols, and using them to classify short sentences may lead to incorrect results. On the other hand, the beautiful visualization of the AI-generation probability for the particular sentence is interesting.

To gather information about separate sentences, we perform the chunking of the text into overlapping pieces with the step length that is equal to the sentence length: first 1000 symbols of the document form the first chunk, the next chunk is formed with such a shift that removes the first sentence from the previous, then the second sentence, and so on.

Thus, each sentence is a part of multiple chunks, and we accumulate prediction results for it as a representative of these chunks. Each prediction is already an averaged value over all models M1 – M5, and a set of these predictions could be averaged again (or just visualized without aggregation for better analysis) into a single output value for each sentence (Fig. 3).

Results

The evaluation of the developed system was tested for the initial drafts of student bachelor's diplomas, theses graduated in 2025, in the review mode without any actions being taken based on the results. The automatic evaluation is complicated and requires a lot of human work, especially for the documents' ground-truth, which is unknown, e.g., we don't know ourselves whether the text was generated or not.

The only available tool we can use for the analysis of entire documents is the nice Plagramme software [25] that is capable of plagiarism and AI detection in the text files, and provides classification results for each separate sentence in percentages of confidence. Quick tests based on the text chunks from our dataset (the first 100 pieces of human-written and AI-generated pieces were used as a single document) showed a good level of AIGC detection for human-written texts (below 5%), but only about 50% for text created from AI-generated pieces. So, based on these results, we can conclude that this tool tends to miss some AIGC content for our domain and,

probably, this format (separate chunks already obtained) of the problem. We perform a manual two-sided comparison of the text analysis reports with our tool and Plagramme. We roughly compared reports paragraph-by-paragraph, focused only on paragraphs with AIGC found, and ignored short sentences classified as AI-generated in between human-written paragraphs (this seems not to be very reliable). So, 31 text pieces identified by Plagramme as AIGC have the output classification value 0.5 or above obtained with our models, and 12 pieces below 0.5.

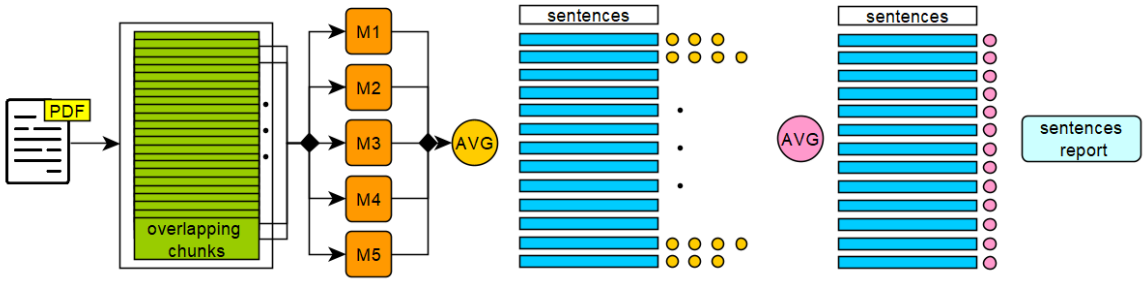


Fig. 3. The processing pipeline to create a sentence report

On the other hand, 26 text pieces classified by our approach as AIGC have a high (about 100%) level of confidence that it was AI-generated by Plagramme, and 15 pieces were classified as human-written. Commonly, there is a strong agreement on all decisions about significant AIGC pieces of text in both reports, but both reports contain errors when definitely human-written text was mistakenly classified as AIGC and vice versa.

The example of a pair of reports is shown in Fig. 4. The left part contains the screenshot from the Plagramme software analysis report, where all 3 sentences were classified as AIGC, and the last one as human-written.

The right part represents the result of highlighting a PDF file based on the proposed approach, where the first sentence was not marked successfully (this happened sometimes because of the complexity of mapping text in PDF), the next two sentences were classified as AIGC with 0.92 and 0.8 confidence, and the last sentence is uncertain with a confidence of 0.54.

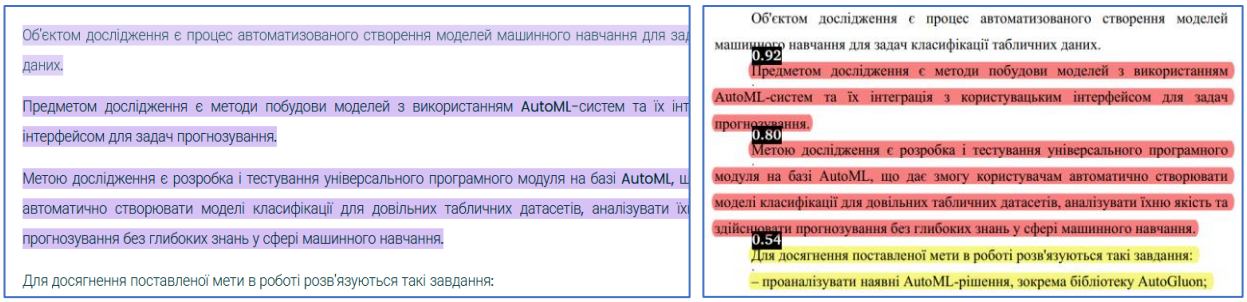


Fig. 4. Example of the report after analysis of the document

Conclusion

In the scope of this research, the AI tool was implemented to classify whether the piece of text in Ukrainian language in the IT domain was AI-generated or human-written.

Five different shallow artificial neural network models were trained using the custom dataset and traditional NLP pipeline, and an accuracy of about 87-88% was achieved. Taking into account the complexity of the problem as well as the ethical background in the educational domain, the result of classification should be analyzed by humans; no automatic AI decision is considered here.

The current implementation could be used as a core for future improvements, as most time was spent on data preparation and cleaning. The training of a shallow ANN is fast and doesn't require special hardware.

There are a lot of possible ways to improve the current initial result. The output of the created AI tool is natural from a black-box system: there is no explanation about why the model makes the decision in favor of a specific class. The dataset created does not have enough size to apply the ensemble of networks successfully and obtain better classification accuracy. Another interesting question to research may include a better choice of chunk size, the usage of deeper models, and the application of LLM itself for the classification.

Additionally, only ChatGPT tool was used to generate text for the dataset; it is interesting to test another popular chatbots as well and extend possible usage cases of the tool. Finally, we focused here on the classification of directly generated text without editing/humanizing, as this problem is already challenging enough (including our goal to solve it using typical widespread hardware available for everyone).

References

1. From Tool to Temptation: AI's Impact on Academic Integrity, UMass Amherst. [Online]. Available: https://www.umass.edu/ideas/news/tool-temptation-ais-impact-academic-integrity.

2. Bittle K., & El-Gayar O. "Generative AI and Academic Integrity in Higher Education: A Systematic Review and Research Agenda" Information 2025, 16(4), 296. Apr. 2025 doi: https://doi.org/10.3390/info16040296.
3. Fraser K. C., Dawkins H., Kiritchenko S. "Detecting AI-Generated Text: Factors Influencing Detectability with Current Methods". Journal of Artificial Intelligence Research 82 (2025) pp. 2233-2278. doi: https://doi.org/10.1613/jair.1.16665.
4. Abdali S., Anarfi R., Barberan CJ, He J. "Decoding the AI Pen: Techniques and Challenges in Detecting AI-Generated Text". KDD '24: Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, pp. 6428-6436, Aug. 2024. doi: https://doi.org/10.1145/3637528.3671463.
5. Li Y., Li Q., Cui L., Bi W., Wang L., Yang L., Shi S., & Zhang Y., "Deepfake Text Detection in the Wild". 2023. URL: https://ar5iv.labs.arxiv.org/html/2305.13242.
6. Li Y., Li Q., Cui L., Bi W., Wang Z., Wang L., Yang L., Shi S., & Zhang Y., "MAGE: Machine-generated Text Detection in the Wild. " May 2024. doi: https://doi.org/10.48550/arXiv.2305.13242.
7. Tufts B., Xuandong Zhao, Lei Li. A Practical Examination of AI-Generated Text Detectors for Large Language Models URL: https://arxiv.org/html/2412.05139v4.
8. Balla E. (2025, May 22). How NLP Powers AI-Generated Text Detection. The AI Journal URL: https://aijourn.com/how-nlp-powers-ai-generated-text-detection/
9. Solaiman I., et al., "Release strategies and the social impacts of language models". OpenAI Report, Nov. 2019, doi: https://doi.org/10.48550/arXiv.1908.09203.
10. Yutian Chen, Hao Kang, Vivian Zhai, Liangze Li, Rita Singh, Bhiksha Raj. "Token Prediction as Implicit Classification to Identify LLM-Generated Text". In Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing, Singapore, 2023, pp.13112–13120. doi: https://doi.org/10.48550/arXiv.2311.08723
11. Yutian Chen, Hao Kang, Vivian Zhai, Liangze Li, Rita Singh, Bhiksha Raj. "GPT-Sentinel: Distinguishing Human and ChatGPT Generated Content". doi: https://doi.org/10.48550/arXiv.2305.07969
12. Xiaomeng Hu, Pin-Yu Chen, Tsung-Yi Ho. RADAR: Robust AI-Text Detection via Adversarial Learning. doi: https://doi.org/10.48550/arXiv.2307.03838.
13. Radar tester: robust ai-text detection via adversarial learning. URL: https://radar-app.vizhub.ai/
14. RADAR-Vicuna-7B Model, Hugging Face. URL: https://huggingface.co/TrustSafeAI/RADAR-Vicuna-7B
15. Mitchell E., Lee Y., Khazatsky A., Manning C. D., & Finn C., "DetectGPT: Zero-shot machine-generated text detection using probability curvature," Proceedings of the 40th International Conference on Machine Learning (ICML'23), Honolulu, HI, USA, 2023, Article No.: 1038, pp. 24950–24962. doi: https://doi.org/10.48550/arXiv.2301.11305
16. Guangsheng Bao, Yanbin Zhao, Zhiyang Teng, Linyi Yang, Yue Zhang. "Fast-DetectGPT: Efficient Zero-Shot Detection of Machine-Generated Text via Conditional Probability Curvature". doi: https://doi.org/10.48550/arXiv.2310.05130.
17. GPTZero. URL: https://gptzero.me/
18. Napier E. "GPTZero. Behind the Scenes: Multilingual Detection Update". May 2025. URL: https://gptzero.me/news/behind-the-scenes-multilingual-detection-update/
19. ZeroGPT. URL: https://www.zerogpt.com/
20. AI Checker – Most Accurate AI Detector. URL: https://originality.ai/ai-checker.
21. AI Detector – Free AI Checker for ChatGPT, GPT-4, Gemini & More. URL: https://copyleaks.com/ai-content-detector.
22. Забезпечення автентичності: найточніший український ШІ-детектор. Isgen.ai detector. URL: https://isgen.ai/uk.
23. Markley T. (2025, February 27). The Problem with AI Detectors: Why Professors Should Reconsider Their Use, URL: https://www.kaltmanlaw.com/post/problem-with-ai-detectors-professors-should-rethink.
24. Tian Y., Chen H., Wang X., Bai Z., Zhang Q., Li R., Xu C., & Wang Y. "Multiscale positive-unlabeled detection of AI-generated texts". In Proceedings of the Twelfth International Conference on Learning Representations (ICLR), 2024. doi: https://doi.org/10.48550/arXiv.2305.18149.
25. Plagramme – Plagiarism checker & AI detector. URL: https://plagramme.com.

Received (Надійшла) 07.09.2025
Accepted for publication (Прийнята до друку) 05.11.2025

ABOUT THE AUTHORS / ВІДОМОСТІ ПРО АВТОРІВ

Передрій Олена Олегівна – кандидат технічних наук, доцент кафедри інформатики та комп'ютерної техніки, Харківський національний економічний університет імені Семена Кузнеця, Харків, Україна;
Olena Peredrii – PhD, Associate Professor, Department of Informatics and Computer Engineering, Simon Kuznets Kharkiv National University of Economics, Kharkiv, Ukraine;
e-mail: olena.peredrii@hneu.net; ORCID: https://orcid.org/0000-0003-0390-1931;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57202751577.

Неглибокі ANN-моделі для класифікації україномовного тексту, згенерованого штучним інтелектом

О. О. Передрій

Анотація. У цьому дослідженні розглядається задача виявлення фрагментів, згенерованих штучним інтелектом, в україномовному тексті. Мета роботи – розробити інструмент, який допомагатиме визначати контент, створений за допомогою ШІ, зокрема у PDF-документах, що стосуються ІТ-сфери. В роботі було розглянуто та проаналізовано які рішення та підходи існують на даний момент. Створено власний тестовий набір даних за допомогою розбиття дипломних робіт бакалаврів ІТ спеціальностей на сегменти приблизно по 1000 символів кожен. Виконано тестування роботи різних комерційних детекторів на цьому наборі даних. Побудовано та навчено п'ять моделей штучних нейронних мереж із використанням цього набору та традиційних NLP процедур обробки тексту, було досягнуто точності 87–88%. Зважаючи на складність проблеми та етичні аспекти застосування інструменту у сфері освіти, результати класифікації повинні бути додатково перевірені експертами. Поточна реалізація може слугувати основою для подальших удосконалень.

Ключові слова: штучні нейронні мережі (ШНМ); неглибокі моделі ШНМ; контент, створений штучним інтелектом (AIGC); LLM; детектор штучного інтелекту.

УДК 004.8:004.738.5 **doi: 10.26906/SUNZ.2025.4.114-120**

А. О. Приліпа, Г. Є. Філатова

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

КЛІЄНТСЬКИЙ TINYML ПРОФАЙЛЕР МЕРЕЖЕВИХ ХАРАКТЕРИСТИК У ВЕББРАУЗЕРІ

Анотація. Представлено клієнтський TinyML профайлер мережевих характеристик у веббраузері, орієнтований на роботу на пристроях з обмеженими ресурсами та в умовах нестабільних або повільних мереж. Рішення поєднує порогове вирішальне правило з легковаговою логістичною моделлю, що виконується локально та забезпечує класифікацію мережевого з’єднання. **Актуальність.** Зростання частки мобільного трафіку, різнорідність мережевого середовища та обмеженість наявних браузерних індикаторів ускладнюють точне прийняття рішень на клієнті. **Об’єкт дослідження**: процес клієнтського профілювання та класифікації якості мережевого з’єднання у браузері на основі стандартних Web Performance API та ML-моделей. **Мета роботи:** аналіз, проєктування та реалізація прозорого й відтворюваного профайлера, здатного класифікувати тип мережевого з’єднання, формувати інтерпретовані ознаки без серверної підтримки. **Методи.** Використано Navigation/Resource Timing і PerformanceObserver для збору сирих сигналів. Сформовано вектор із 14 ознак (медіани/квантилі RTT і пропускної здатності, варіабельність, евристика втрат, індикатори протоколів та Service Worker). Запропоновано порогове правило з гістерезисом і довірою рішення та softmax модель. **Результати.** Розроблено профайлер без спеціальних дозволів і сторонніх сервісів. У контрольованих сценаріях емуляції мережі досягнуто підвищення точності класифікації порівняно з чистим пороговим підходом, забезпечено низькі накладні витрати та пояснюваність рішень. **Висновки.** Запропоноване рішення є ефективним засобом оперативної та інтерпретованої оцінки мережевих умов у браузері, придатним для середовищ з обмеженими ресурсами. Результати можуть бути використані для подальшого вдосконалення адаптивного завантаження, розширення простору ознак і впровадження компактніших ML-моделей у вебзастосунках.

Ключові слова: клієнтське профілювання мережі, TinyML, Web Performance API, softmax-класифікація, RUM вимірювання, адаптивне завантаження контенту, QoE.

Вступ

Продуктивність вебзастосунків істотно впливає на досвід користувачів та загальну ефективність сервісів. Навіть додаткова затримка завантаження на сотні мілісекунд може спричинити помітне зниження конверсій та трафіку [1]. Тому розробники застосовують адаптивні стратегії доставки контенту, підлаштовуючи обсяг і якість переданих даних під поточні умови клієнта (швидкість та якість мережі, ресурси пристрою) [2]. У таких підходах критичним сигналом стає оцінка мережевої якості на стороні клієнта [3].

Сучасні браузери надають обмежену інформацію про мережеві умови. API NetworkInformation повертає класи мережі: slow-2G, 2G, 3G, 4G [4]. Ця класифікація ґрунтується на внутрішньому механізмі Chromium Network Quality Estimator (NQE), який періодично вимірює затримку (RTT) та пропускну здатність і відносить їх до фіксованих порогів [5]. Однак такий пороговий підхід має суттєві обмеження: алгоритм NQE є пропрієтарним і лише частково відомим [6]. Крім того, браузер спеціально додає шум у власні вимірювання [7], а доступ до точних значень RTT або пропускної здатності стороннім скриптам обмежений. Також API NetworkInformation може бути недоступним. Наприклад, не підтримується в Safari [8] або вимкненим користувачем [9], що ще більше ускладнює оцінку мережі на клієнті.

Водночас вебзастосунки можуть скористатися стандартними Web Performance API для детального моніторингу. Navigation Timing і Resource Timing API фіксують часові мітки основних етапів завантаження сторінок та ресурсів (від початку DNS-запиту до отримання першого та останнього байта) [10]. PerformanceResourceTiming надає повну хронологію завантаження кожного ресурсу (загальний обсяг отриманих даних, час з’єднання, перший байт, завершення тощо) [11]. PerformanceObserver дозволяє у реальному часі відстежувати появу нових записів із такими часовими мітками [12]. Збір продуктивнісної телеметрії напряму з браузерів користувачів є основою Real User Monitoring (RUM) [13]. Зібрані дані дають змогу обчислювати типові показники (час до першого байта – TTFB, загальний час завантаження, пропускну здатність тощо) за реальних умов мережі. Дослідження [14] показало значну варіабельність Quality of Experience (QoE) у мобільних мережах залежно від типу з’єднання та якості сигналу, а інше дослідження [15] виявило істотні відмінності у швидкодії мереж між країнами з різним рівнем інфраструктури. Таким чином, RUM-метрики браузера забезпечують реальний вимірювальний контекст користувачів на різних пристроях і мережах.

Класифікація якості з’єднання на боці клієнта традиційно здійснюється за жорсткими порогами (як у Effective Connection Type) [10]. Такий пороговий підхід простий і швидкий, але не враховує нюансів реальних умов: мережа 4G із слабким сигналом може фактично поводитися як 3G, і навпаки. До того ж через закритість алгоритму NQE і додавання шуму браузером effectiveType з NetworkInformation API дає лише грубу оцінку і може оновлюватися із затримкою або залишатися незмінним, особливо якщо сторінка завантажується швидко [7].

На противагу цьому пропонується застосування клієнтського машинного навчання. У цьому підході формується вектор різноманітних мережевих метрик (RTT, пропускна здатність, обсяги переданих даних, частота довгих пауз у головному потоці тощо) і за ним прогнозують клас з’єднання або показники QoE.

Автори у [16] реконструювали щосекундні QoE-метрики потокового відео без доступу до заголовків застосунку. Автори у [17] розробили модель eMIMIC для оцінки QoE HTTP-відео за зашифрованим трафіком, а дослідження [18] показали, що можна інтерпретувати якість потокового відео з зашифрованого трафіку. Невеликі інтерпретовані моделі мають перевагу прозорості: їхні ваги однозначно показують внесок кожного параметра у класифікацію [19].

Сучасні браузерні фреймворки значно полегшили виконання ML-моделей на клієнті. TensorFlow.js від Google та ONNX Runtime Web від Microsoft підтримують запуск нейронних мереж у браузері через WebGL або WebAssembly [20, 21]. Виконання у WebAssembly забезпечує суттєво швидше обчислення, ніж інтерпретований JS, часто випереджаючи WebGL для невеликих моделей через менші накладні витрати [22]. Окрім того, стандарт Web Neural Network API (WebNN) надає високорівневий інтерфейс до апаратних прискорювачів GPU/NPU, що ще більше прискорює виконання моделей [23]. Демонстрації Google 2018–2019 рр. показували реальне виконання легких моделей у мобільному Chrome. Ці можливості відкривають шлях до TinyML у браузері, коли одна модель може бути однаково застосована на клієнті і на вбудованих пристроях.

Клієнтський підхід має низку переваг.

По-перше, захищеність та приватність. Усі дані користувача обробляються локально і не передаються на сервер [24, 25].

По-друге, зниження затримок відгуку і офлайн-можливості [26].

По-третє, масштабованість і надійність. Обчислювальне навантаження розподіляється між клієнтами, знижуючи потреби у серверних ресурсах [24].

Разом це робить рішення більш стійким до збоїв мережі та серверів. Водночас існують і обмеження. Браузерні ML-фреймворки працюють значно повільніше за нативний код [27], часто споживають у десятки разів більше оперативної пам'яті [28], а важкі обчислення можуть уповільнювати відгук інтерфейсу. Ці недоліки частково компенсуються винесенням обчислень у Web Workers, мультитредінгом і апаратним прискоренням.

Профілювання мережевих умов у браузері та адаптивне керування контентом на основі цих даних є перспективним напрямом досліджень. Водночас лишаються відкриті питання комбінування даних RUM із додатковими сигнальними ознаками для точнішої класифікації, стандартизації відкритих показників якості мережі та мінімізації впливу клієнтського ML на досвід користувача.

Постановка задачі

Розробка клієнтського профайлера мережевих характеристик у браузері зумовлена обмеженою інформативністю наявних індикаторів якості з'єднання, фрагментарною підтримкою стандартів у різних браузерах та потребою у прозорих, відтворюваних і налаштовуваних засобах оцінювання. Практична доцільність полягає в тому, що вебзастосунки мають адаптивно коригувати обсяг і якість контенту відповідно до фактичних мережевих умов користувача, зберігаючи низьку затримку відгуку, стабільність інтерфейсу та економність використання ресурсів. Потреба у власному профайлері випливає з таких вимог:

1. працювати на боці клієнта;
2. використовувати лише стандартизовані Web API;
3. забезпечувати інтерпретованість рішень і чітке керування порогами/вагами;
4. масштабуватися на широке коло пристроїв із різними обмеженнями.

Мета роботи полягає у побудові інструмента, який у заданому часовому вікні збирає сигнали мережевої продуктивності з стандартних браузерних API (Navigation/Resource Timing, PerformanceObserver тощо), формує вектор ознак, а далі класифікує поточний стан мережі за скінченною множиною класів. Вирішальне правило повинно поєднувати детерміністичний baseline і легковагову ML-модель з механізмом узгодження рішень і оцінкою впевненості. Профайлер не повинен мати доступ до мережевих пакетів, працювати під обмеженнями безпеки браузера та не впливати на якість досвіду користувача.

Основна частина

У роботі пропонується профайлер, який працює як легковажний клієнтський інструмент оцінювання якості мережі у браузері, що поєднує пасивні вимірювання (RUM) і активні мікропроби в межах встановленого бюджету. На першому етапі збираються стандартизовані сигнали, а також допоміжні індикатори середовища. Наступним кроком виконується невелика серія HTTP-проб, з якої обчислюються статистики затримки та ефективної пропускної здатності, показники варіабельності. Далі ці вимірювання агрегуються у вектор ознак фіксованої структури, що відображає стан мережі. Класифікація здійснюється гібридно: детермінований пороговий baseline відносить з'єднання до одного з дискретних класів з пояснювальними прапорцями та мірою впевненості, тоді як ML-модель уточнює рішення на основі всього вектору ознак, забезпечуючи кращу чутливість до комбінацій сигналів. Фінальний вибір класу виконується правилом злиття, яке враховує узгодженість між baseline і ML, стан стабільності мережі та можливе консервативне пониження класу. Результат повертається у вигляді класу, впевненості, набору прапорців і ключових статистик, що дозволяє надавати діагностичні пояснення для подальшої оптимізації. Розглянемо детально етапи роботи профайлеру.

Профілювання відбувається у часовому вікні довжиною T_{win} під жорсткими бюджетними обмеженнями на час, кількість та обсяг даних:

- $T_{budget} > 0$ – максимально допустимий час профілювання (в мілісекундах);

- $M \in N$ – ліміт кількості активних HTTP-проб (мікрозапитів);

- $B > 0$ – ліміт сумарного об'єму завантажених байтів.

Обмеження бюджету формалізуються так:

$$\sum_{m=1}^{M} size_m \le B;$$

$$\sum_{m=1}^{M} dur_m \le T,$$

де $size_m$ – розмір m-ї проби в байтах, а dur_m – час очікування відповіді на неї. Ці обмеження гарантують, що активні виміри (мікропроби) не перевищують виділеного трафіку та часу, що важливо для невтручання в роботу сторінки.

Спочатку збираються RUM сигнали від браузера: записи навігаційних подій та перших K ресурсів із Resource/Naviagation Timing API. З кожного запису $e \in \mathcal{E}$ отримано часові мітки початок запиту *requestStart*, початок відповіді *responseStart*, кінець відповіді *responseEnd* та розмір переданих даних $size(e)$. Таким чином можна виміряти затримку мережі та пропускну здатність. Зокрема, RTT для запису e визначено як:

$$rtt(e) = \max\left(0, responseStart(e) - requestStart(e)\right).$$

Визначення миттєвої корисної пропускної здатності $g(e)$ у кбіт/с надає кількість переданих бітів трафіку за секунду (стабілізуючи знаменник з 0.1 с, щоб уникнути ділення на дуже малий час):

$$g(e) = \frac{8size(e)}{1000\max\left(0.1, responseEnd(e) - responseStart(e)\right)}.$$

Для множин усіх зафіксованих $rtt(e)$ і $g(e)$ необхідно обчислити статистики: процентилі p_{10}, p_{50}, p_{90} (10-й, 50-й, 90-й) та коефіцієнти варіації. Нехай:

$$R = rtt(e) : e \in \mathcal{E};$$

$$G = g(e) : e \in \mathcal{E}.$$

Тоді, $RTT_{p\alpha} = pct\left(R, \alpha\right)$, а $DL_{p\alpha} = pct\left(G, \alpha\right)$. Коефіцієнт варіації визначається як відношення стандартного відхилення до середнього:

$$variation\_cv_{rtt} = \frac{\sigma(R)}{\mu(R)};$$

$$variation\_cv_{dl} = \frac{\sigma(G)}{\mu(G)},$$

що дає показник нестабільності (відношення розкиду до середнього). Додатково вводиться джиттер для RTT як різницю між 90-м і 50-м процентилями:

$$rtt_{jitter} = RTT_{p90} - RTT_{p50}.$$

Це відображає варіацію затримки запитів: чим більша ця різниця, тим нестабільніші затримки.

Крім мережевих сигналів, зафіксовано метрики завантаженості клієнта. Зокрема, сума тривалостей довгих задач на головному потоці задається як:

$$L_{long} = \sum_{u \in \mathcal{L}} duration(u),$$

де $\mathcal{L}$ – множина подій PerformanceObserver типу long task. У стандартах браузерів long task визначається як завдання, яке зайняло головний потік щонайменше 300 мс. Наявність таких задач сигналізує про сильне навантаження, що може впливати на швидкість обробки дій від користувача.

Також визначимо індикатори протоколу: nhp_{h3}, nhp_{h2}, $nhp_{h1} \in \{0,1\}$ показують, чи було використано відповідно HTTP/3, HTTP/2 або HTTP/1.1. Індикатор $sw\_present \in \{0,1\}$ сигналізує про наявність (1) або відсутність (0) зареєстрованого Service Worker у контексті сторінки.

Після пасивних сигналів здійснюється серія активних мікропроб в рамках бюджету. Виконується послідовність M запитів розмірів $size_m$. Для проби m фіксуємо три моменти часу:

– $t_0^{(m)} = requestStart$ – час відправлення запиту;

– $t_1^{(m)} \approx responseStart$ – приблизний час отримання першого байту відповіді;

– $t_2^{(m)} = responseEnd$ – час завершення отримання відповіді.

Для проби m обчислюємо час відгуку $rttLike^{(m)}$ (приблизний RTT, вирівняний мінімумом 0.1 с) і бітрейт завантаження $kbps^{(m)}$:

$$rttLike^{(m)} = \max\left(0.1, t_1^{(m)} - t_0^{(m)}\right);$$

$$kbps^{(m)} = \frac{8 \cdot size_m}{1000 \cdot \max\left(0.1, t_2^{(m)} - t_1^{(m)}\right)}.$$

Зібрані набори дають множини значень для статистики:

$$R^{probe} = \{rttLike^{(m)}\}_{m=1}^{M};$$

$$G^{probe} = \{kbps^{(m)}\}_{m=1}^{M}.$$

За ними так само рахуємо процентилі p_{50}, p_{90} та коефіцієнти варіації (аналогічно до RUM-наборів).

Крім того, вводимо евристичний показник втрат $loss\_like$, що базується на виявленні затримок при обробці відповіді:

$$loss\_like = \min\left((\max(0.1S + 0.15H, 0), 1\right),$$

де S – число stall-подій (інтервали між черговим отриманням відповідей більше 80 мс), H – число проб, для яких $rttLike^{(m)} > 1000$ мс. Це значення в межах $[0,1]$ показує ймовірну втрату продуктивності через надмірні затримки.

На основі вищенаведених сигналів формується фінальний вектор ознак $\vec{x} \in R^{14}$. Компоненти вектора $\vec{x}$ включають основні статистики затримок і пропускних спроможностей

$$\vec{x} = (rtt_{p50}, rtt_{jitter}, rtt_{p90}, down_{p10}, down_{p50}, down_{p90}, loss\_like, L_{long}, nhp_{h1}, nhp_{h2}, nhp_{h3}, variation\_cv_{rtt}, variation\_cv_{dl}, sw\_present).$$

Для класифікації з'єднання введено класи

$$C = \{\text{slow-2G, 2G, 3G, 4G, broadband}\},$$

упорядковані за зростанням якості мережі.

Базовий класифікатор $h(\vec{x})$ – це набір детермінованих правил на основі порогів від латентності та пропускної спроможності. Аналогічний підхід використовується в ECT, де сполучення згаданих метрик призначає тип мережі.

Набір правил визначення типу мережі $h(\vec{x})$ такий:

– slow-2G, якщо $down_{p50} < 50$ кбіт/с або $rtt_{p50} > 2000$ мс ;

– 2G, якщо $50 \le down_{p50} < 250$ кбіт/с або $1000 < rtt_{p50} \le 2000$ мс ;

– 3G, якщо $400 \le down_{p50} < 1500$ кбіт/с або $150 < rtt_{p50} \le 1000$мс ;

– 4G, якщо $1500 \le down_{p50} < 5000$ кбіт/с або $70 < rtt_{p50} \le 150$ мс ;

– broadband, якщо $down_{p50} \ge 5000$ кбіт/с та $rtt_{p50} \le 70$ мс .

Ці межі відповідають рекомендаціям ECT API. Якщо виконуються умови для кількох класів, застосовується найнижчий клас (консервативна інтерпретація).

Крім класу, базовий алгоритм генерує пояснювальні прапорці $flags = \{highLatency, lowBandwidth, unstable, cpuBound\}$ (кожний прапорець приймає значення 1 або 0), що інформують про причини рішення:

– $highLatency = 1$, якщо $rtt_{p50} > 300$ мс (велика медіана затримок);

– $lowBandwidth = 1$, якщо $down_{p50} < 1024$ кбіт/с (мала пропускна здатність);

– $unstable = 1$, якщо $variation\_cv_{dl} > 0.6$ або $rtt_{jitter} > 400$ мс або $loss\_like > 0.3$ (значна варіативність швидкості або втрати);

– $cpuBound = 1$, якщо $L_{long} > 300$ мс (довгі задачі зайняли сумарно багато часу).

Також оцінюється впевненість базового рішення $conf_{base}(\vec{x}) \in [0,1]$.

Ближче до межі класу broadband, то впевненість вища. Позначимо:

$$d_{rtt} = \left| \frac{RTT_{p50} - 70}{70} \right|;$$

$$d_{down} = \left| \frac{DL_{p50} - 5000}{5000} \right|.$$

Тоді відстань до broadband порогу – це $\min(d_{rtt}, d_{down})$.

Впевненість призначається як

$$conf_{base}(\vec{x}) = 1 - \min\left(1, \frac{\min(d_{rtt}, d_{down})}{0.15}\right) \cdot (1 - 0.15 unstable).$$

Якщо результати значно відрізняються від стандартів класу, впевненість зменшується; якщо ж прапорець *unstable* встановлено, то впевненість зменшується ще на ~15%.

Отже, базове рішення задається як

$$c_{base} = \{h_{base}(\vec{x}), conf_{base}(\vec{x}), flags, evidence_{base}\},$$

де $evidence_{base}$ – список ознак (порогів) та їхніх значень, які визначили результат.

Другим шаром логіки рішення є модель машинного навчання – лінійна багатокласова логістична регресія (softmax-регресія). Нехай $|C| = 5$ – число класів. Обчислюємо логіти $\vec{z} = W\vec{x} + \vec{b} \in R^5$, де W – матриця розмірності 5×14 ; $\vec{b}$ – вектор зсуву. Наступним кроком відбувається softmax-перетворення

$$p_i(\vec{x}) = \frac{\exp(z_i)}{\sum_{j=1}^{5} \exp(z_i)},$$

отримуючи ймовірності для кожного класу.

Для практичної придатності класифікатора необхідно визначити параметри його статистичної частини.

Метод тренування полягає у формуванні вибірки з звітів профайлера, де кожен приклад представлено вектором із 14 ознак у фіксованому порядку та міткою класу якості мережі. Неперервні ознаки стандартизуються за параметрами, оціненими лише на тренувальній підмножині, після чого навчається лінійна softmax-регресія з L2-регуляризацією та класовими вагами для компенсації дисбалансу, оптимізована методом L-BFGS до збіжності. Коректність моделі перевіряється на відкладеній підвибірці за матрицею неточностей. Параметри стандартизації представляються у вигляді ваг та зсувів. Підсумкова конфігурація (ефективні ваги, зсув, упорядковані ключі ознак і перелік класів) серіалізується в JSON.

Прогноз моделі

$$c_{ML} = h_{ML}(\vec{x}) = \arg\max p_i(\vec{x}),$$

тобто клас із найбільшим p_i, а впевненість моделі $conf_{ML}(\vec{x}) = \max p_i(\vec{x})$.

Після отримання c_{base} від порогового класифікатора та c_{ML} від ML-моделі застосовується правила злиття. Для цього вводимо ранжування класів за якістю з'єднання $rank : C \to \{0,1,2,3,4\}$. Обчислюємо розбіжність між класами базового та ML-рішень:

$$\Delta = \left| \text{rank}(c_{base}) - \text{rank}(c_{ML}) \right|.$$

Тоді гібридне правило таке:

1. Якщо $conf_{base}(\vec{x}) \ge 0.7$ та $\Delta \le 1$, то вибираємо рішення базового класифікатора. Якщо базовий класифікатор упевнений і клас ML не суперечить сильно

(різниця в рангу не більше 1), довіряємо інтерпретованому правилу.

2. Якщо один з прапорців *unstable* або *cpuBound* активний, то не можемо бути впевненими у швидкісній класифікації – тоді беремо клас ML, але понижуємо його на один рівень. Це гарантує консервативність при нестабільності чи перевантаженні.

3. Інакше (базова умова не виконується) – беремо ML-рішення c_{ML}.

Остаточна впевненість $conf(\vec{x})$ обраного рішення визначається так

$$conf(\vec{x}) = \begin{cases} conf_{base}, & \text{якщо базовий класифікатор;} \\ conf_{ML}, & \text{якщо ML-модель.} \end{cases}$$

Після завершення класифікації профайлер формує звіт, що містить обраний клас, міру впевненості, набір пояснювальних прапорців, вектор ознак і список ознак, що вплинули на визначення результату. Це забезпечує відтворюваність і прозору інтерпретацію рішення.

Звіт також включає службові поля (тривалість вимірювання, часові мітки), що полегшують аудит, порівняння запусків і інтеграцію в системи RUM аналітики.

Результати

Для оцінки точності класифікації розгорнуто тестове середовище на основі Puppeteer (Chrome 114). Профайлер випробовувався в контрольованих умовах штучного обмеження мережі. Використано наступні профілі симуляції мереж Chrome DevTools:

slow 2G з $down_{p50} = 40$ кбіт/с та $rtt_{p50} = 2000$ мс;

2G з $down_{p50} = 150$ кбіт/с та $rtt_{p50} = 1200$ мс;

slow 3G з $down_{p50} = 1$ Мбіт/с та $rtt_{p50} = 400$ мс;

high latency 3G з $down_{p50} = 1.2$ Мбіт/с та $rtt_{p50} = 900$ мс;

4G з $down_{p50} = 3.5$ Мбіт/с та $rtt_{p50} = 120$ мс;

unstable 4G з $down_{p50} = 2.2$ Мбіт/с, $rtt_{p50} = 150$ та $rtt_{jitter} = 180 мс$;

broadband з $down_{p50} = 20$ Мбіт/с та $rtt_{p50} = 30$ мс.

Профілі охоплюють п’ять цільових класів мережі і дозволяють перевірити поведінку профайлера, як на повільних каналах, так і на швидкісному з’єднанні. Для кожного профілю було виконано по 35 запусків ідентифікації (послідовно у одному сеансі браузера, після кожної серії браузер перезапускався). Для достовірності вимірювань всі зайві фонові процеси було вимкнено.

Отримані результати містили для кожного запуску еталонний профіль, результати baseline-класифікації, результати ML-моделі та фінальний обраний клас. Це дозволило розрахувати метрики точності, а також проаналізувати окремо внесок.

Оцінювання точності профайлера проводилося на основі 245 запусків у контрольованих умовах для п’яти профілів мережі.

Результати показали, що гібридний підхід демонструє середню точність 91.6%, при значеннях $precision = 0.916$, $recall = 0.903$ та $F1 = 0.907$.

Детальні метрики за класами наведено у табл. 1. Як видно, класи slow-2G та 2G визначаються з абсолютною точністю ($precision = 1$, $recall = 1$, $F1 = 1$). Клас 3G продемонстрував високу стабільність: $precision = 0.909$, $recall = 1$, $F1 = 0.952$. Найскладнішими залишаються класи 4G та broadband, між якими спостерігається найбільша кількість хибних класифікацій. Для 4G отримано $precision = 0.841$, тоді як для broadband $precision = 0.828$. Зниження *recall* для broadband пояснюється тим, що близько третини таких сеансів було віднесено до класу 4G у випадках підвищеного RTT чи зниженого пропускного каналу.

Таблиця 1 – **Показники точності класифікації** (гібридний підхід, $N = 245$)

Клас	Precision	Recall	F1
slow-2G	1.000	1.000	1.000
2G	1.000	1.000	1.000
3G	0.909	1.000	0.952
4G	0.841	0.829	0.835
broadband	0.828	0.686	0.750
Average	0.916	0.903	0.907

Структура помилок класифікації відображена у матриці плутанини (табл. 2). Спостерігається, що класи slow-2G та 2G не перетинаються з іншими, забезпечуючи повну відокремленість. Для 3G переважна більшість випадків визначена коректно, а хибні віднесення до сусідніх класів усунуті завдяки ML-моделі. Основні розбіжності фіксуються між 4G та broadband, де частина сесій broadband (~31%) класифікується як 4G. Така консервативність є виправданою, оскільки профайлер віддає перевагу недооцінюванню якості, що мінімізує ризик перевантаження мережі користувача.

Таблиця 2 – **Матриця плутанини**

	slow-2G	2G	3G	4G	broadband
slow-2G	35	0	0	0	0
2G	0	35	0	0	0
3G	0	0	70	0	0
4G	0	0	7	58	5
broadband	0	0	0	11	24

Узагальнюючи, додавання ML-моделі до порогового алгоритму дозволило знизити кількість хибних класифікацій на межових профілях, збалансувати показники точності, зберігаючи при цьому низькі накладні витрати.

Прапорці узгодилися з очікуваною поведінкою профілів і допомогли інтерпретувати класифікацію.

Для slow-2G та 2G у наданих результатах прапорці *highLatency* та *lowBandwidth* вмикалися стабільно; довіра при цьому була високою, а гібридна схема не змінювала клас. У профілях 3G та 4G прапорець *unstable* спрацьовував переважно за підвищеної варіабельності пропускної здатності і призводив до зниження довіри. У прикордонних випадках це дозволяло ML-моделі коригувати клас у бік більш консервативного. Для unstable 4G прапорці *unstable* і *lowBandwidth* спрацьовували найчастіше (спостерігалися стрибки між мережами). Для broadband прапорці, як правило, не вмикалися. Поодинокі *unstable* або *highLatency* відповідали кейсам із великим RTT: у цих випадках confidence знижувалась, що відповідало нашій політиці «не переоцінювати» мережу. Прапорець *cpuBound* траплявся при помітних довгих задачах, але він не змінював клас безпосередньо.

Витрати ресурсів профайлера є мінімальними. Вимірювання показали, що профайлер працює в межах заданого бюджету (час, кількість проб, кількість байтів) і не створює навантаження на систему. Лінійна ML-модель здійснює 60–70 елементарних арифметичних операцій, а її ваги оцінюються в межах 1.5 кбайт. Найдовшою дією є виконання HTTP-запитів для проб, але вони відбуваються асинхронно і не завдають навантаження на пристрій. Сумарний час виконання визначається переважно тривалістю самих мережевих вимірювань. Різниця у використаній пам'яті до і після роботи профайлера знаходиться в межах похибки, адже більшу частину часу профайлер просто очікує відповіді мережі, а самі обчислення виконуються швидко. Навіть на малопотужному пристрої накладні витрати обчислень є прийнятними. В разі, якщо пристрій під час вимірювання зайнятий іншими задачами, які подовжать виконання, проте це буде автоматично враховано через ознаку *cpuBound* (така ситуація, сигналізує, що повільна реакція може бути спричинена не лише мережею). Таким чином, запропоноване рішення відповідає критеріям TinyML: модель швидка і мала за розміром, а запропоноване рішення не потребує значних ресурсів.

Висновки

У роботі представлено і детально проаналізовано підхід до класифікації мережевих умов у браузері із застосуванням TinyML-моделі. Отримані результати демонструють, що навіть проста модель, така як багатокласова логістична регресія, у поєднанні з ретельно підібраними ознаками здатна перевершити традиційну порогову схему, забезпечуючи точність близько 90% у задачі розпізнавання типу мережевого з'єднання. Гібридний підхід дозволяє зберегти стабільність рішення і водночас додати адаптивність до мережевої класифікації користувача. Розроблений профайлер є легким, прозорим і не потребує спеціальних можливостей браузера. Практична цінність рішення полягає у можливості його інтеграції безпосередньо у вебзастосунки для адаптивного керування завантаженням контенту. За допомогою профайлера клієнтська сторона може самостійно та миттєво визначати якість поточної мережі і вирішувати завантажувати ресурси в нижчій якості, відкладати неважливі запити або активувати додаткові оптимізації рендерингу в умовах повільного з'єднання.

У ширшій перспективі отримані результати демонструють потенціал TinyML у вебсередовищі. Виконання обчислень у браузері забезпечує прозорість, конфіденційність та незалежність від серверної інфраструктури. Це відкриває можливості не лише для адаптивного завантаження контенту у вебзастосунках, але й для застосувань у сфері IoT, де клієнтські пристрої з обмеженими ресурсами можуть автономно оцінювати умови мережі. Подальші дослідження, що спрямовані на розширення простору ознак, використання більш гнучких моделей та інтеграцію з новими стандартами, дозволять підвищити точність і універсальність запропонованого підходу.

Представлений профайлер є прикладом ефективного поєднання простоти реалізації, низьких накладних витрат та практичної корисності, що робить його вагомим внеском у напрямі оптимізації вебзастосунків для середовищ із обмеженими ресурсами.

Список літератури

1. Muralidhar, A. & Lakkanna, Y. (2024). From clicks to conversions: Analysis of traffic sources in e-commerce. https://arxiv.org/abs/2403.16115
2. Jiang, D., Wang, F., Lv, Z. & Mumtaz, S. (2023). QoE-aware efficient content distribution scheme for satellite-terrestrial networks. IEEE Transactions https://doi.org/10.1109/TMC.2021.3074917
3. Taha, M. (2023). An efficient software-defined network controller based routing adaptation for enhancing QoE of multimedia streaming service. Multimedia Tools and Applications. https://doi.org/10.1007/s11042-023-14938-5
4. MDN Web Docs. Effective connection type (ECT). https://developer.mozilla.org/en-US/docs/Glossary/Effective_connection_type
5. W3C Wiki. Network Quality Monitoring and Prediction. https://www.w3.org/wiki/Network_Quality_Monitoring_and_Prediction
6. Kruger, H. J. J. (2023). A conceptual quality framework for online distance education in a South African higher education context. http://hdl.handle.net/2263/93779
7. Jueckstock, J. P. (2021). Enhancing the security and privacy of the web browser platform via improved web measurement methodology. https://repository.lib.ncsu.edu/server/api/core/bitstreams/65b7fbc8-0e70-41f8-82dd-bfc79d6e37c9
8. MDN Web Docs. NetworkInformation (API). https://developer.mozilla.org/en-US/docs/Web/API/NetworkInformation
9. Kirimi, N. & Barakat, C. (2024). Passive network monitoring and troubleshooting from within the browser: A data-driven approach. International Wireless Communications and Mobile Computing Conference (IWCMC 2024). https://doi.org/10.1109/IWCMC61514.2024.10592376
10. MDN Web Docs. Performance (Web Performance API). https://developer.mozilla.org/en-US/docs/Web/API/Performance
11. MDN Web Docs. PerformanceResourceTiming. https://developer.mozilla.org/en-US/docs/Web/API/PerformanceResourceTiming

12. MDN Web Docs. PerformanceObserver. https://developer.mozilla.org/en-US/docs/Web/API/PerformanceObserver
13. Goenka, P., Zarifis, K., Gupta, A. & Calder, M. (2022). Towards client-side active measurements without application control. ACM SIGCOMM Computer Communication Review. https://doi.org/10.1145/3523230.3523234
14. Verma, D. (2022). A comparison of web framework efficiency: Performance and network analysis of modern web frameworks. https://urn.fi/URN:NBN:fi:amk-2022061417896
15. Lucas, G. A., Lunardi, G. L. & Dolci, D. B. (2023). From e-commerce to m-commerce: An analysis of the user's experience with different access platforms. Electronic Commerce Research and Applications, 58, 101240. https://doi.org/10.1016/j.elerap.2023.101240
16. Sharma, T., Mangla, T., Gupta, A., Jiang, J. & Feamster, N. (2023). Estimating WebRTC video QoE metrics without using application headers. IMC 2023. https://doi.org/10.1145/3618257.3624828
17. Mangla, T., Halepovic, E., Ammar, M. & Zegura, E. (2019). Using session modeling to estimate HTTP-based video QoE metrics from encrypted network traffic (eMIMIC). IEEE Transactions on Network and Service Management. https://doi.org/10.1109/TNSM.2019.2924942
18. Bronzino, F., Schmitt, P., Ayoubi, S., Martins, G., Teixeira, R. & Feamster, N. (2019). Inferring streaming video quality from encrypted traffic: Practical models and deployment experience. Proceedings of the ACM on Measurement and Analysis of Computing Systems (POMACS). https://doi.org/10.1145/3366704
19. Rudin, C. (2019). Stop explaining black box machine learning models for high-stakes decisions and use interpretable models instead. Nature Machine Intelligence. https://doi.org/10.1038/s42256-019-0048-x
20. TensorFlow.js. Platform and environment. https://www.tensorflow.org/js/guide/platform_environment
21. Microsoft Open Source Blog. ONNX Runtime Web – Running your machine learning model in browser. https://opensource.microsoft.com/blog/2021/09/02/onnx-runtime-web-running-your-machine-learning-model-in-browser/
22. MDN Web Docs. WebAssembly. https://developer.mozilla.org/en-US/docs/WebAssembly
23. W3C. Web Neural Network API (WebNN). https://www.w3.org/TR/webnn/
24. Dhar, S., Tang, X., et al. (2021). A survey of on-device machine learning: An algorithms and learning theory perspective. ACM Computing Surveys. https://doi.org/10.1145/3450494
25. Shi, W. & Dustdar, S. (2016). The promise of edge computing. IEEE Computer. https://doi.org/10.1109/MC.2016.145
26. Malavolta, I., et al. (2020). Evaluating the impact of caching on the energy consumption and performance of progressive web apps. MOBILESoft '20. https://doi.org/10.1145/3387905.3388593
27. Wang, Q., et al. (2025). Anatomizing deep learning inference in web browsers. ACM (IMC record). https://doi.org/10.1145/3688843
28. Yan, Y., Tu, T., Zhao, L., Zhou, Y. & Wang, W. (2021). Understanding the performance of WebAssembly applications. IMC '21. https://doi.org/10.1145/3487552.3487827

Received (Надійшла) 14.08.2025
Accepted for publication (Прийнята до друку) 22.10.2025

ABOUT THE AUTHORS / ВІДОМОСТІ ПРО АВТОРІВ

Приліпа Артем Олегович – аспірант кафедри комп'ютерної інженерії та програмування, Національний технічний університет "Харківський політехнічний інститут", Харків, Україна;
Artem Prylipa – PhD student, Department of Computer Engineering and Programming, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;
e-mail: Artem.Prylipa@cs.khpi.edu.ua; ORCID Author ID: https://orcid.org/0009-0005-6633-8308.

Ганна Євгенівна Філатова – доктор технічних наук, професор, професор кафедри комп'ютерної інженерії та програмування, Національний технічний університет "Харківський політехнічний інститут", Харків, Україна;
Anna Filatova – Doctor of Technical Sciences, Professor of Computer Engineering and Programming, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;
e-mail: Hanna.Filatova@khpi.edu.ua; ORCID Author ID: https://orcid.org/0000-0003-1982-2322;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=56448583600.

TinyML network profiler in browser

Artem Prylipa, Anna Filatova

Abstract. This research presents a client-side TinyML profiler of network in the web browser, designed for operation on resource-constrained devices and under unstable or low-bandwidth network conditions. The solution combines a threshold-based decision rule with a lightweight logistic (softmax) model executed locally to classify the quality of the network connection. **Relevance.** The rising share of mobile traffic, the heterogeneity of network environments, and the limited fidelity of existing browser indicators complicate accurate client-side decision-making. **Object of study:** client-side methods for profiling and classifying the quality of a browser-based network connection using standard Web Performance APIs and ML models. **Aim:** to analyze, design, and implement a profiler capable of classifying connection types and producing interpretable features without server support. **Methods.** Navigation/Resource Timing and PerformanceObserver are used to collect raw signals. A 14-dimensional feature vector is formed (medians/quantiles of RTT and throughput, variability measures, a loss-likelihood heuristic, and protocol/Service Worker indicators). The proposed a threshold rule with hysteresis and decision confidence, together with a softmax model. **Results.** The developed profiler requires no special permissions or third-party services. In controlled network-emulation scenarios it improves classification accuracy over a pure threshold baseline, while maintaining low overhead and decision explainability. **Conclusions.** The proposed client-side approach provides an effective, rapid, and interpretable assessment of network conditions in the browser and is suitable for resource-constrained settings. The results can be leveraged to further enhance adaptive content loading, expand the feature space, and deploy more compact ML models in web applications.

Keywords: client-side network profiling, TinyML, Web Performance API, softmax classification, RUM measurements, adaptive content loading, QoE.

UDC 004.9 **doi: 10.26906/SUNZ.2025.4.121-125**

Viacheslav Radchenko, Yuliia Andrusenko

Kharkiv National University of Radio Electronics, Kharkiv, Ukraine

INTELLIGENT APPROACH TO PLANNING TAKING INTO ACCOUNT THE CONCEPT OF ACCEPTABLE WORK BALANCE

Abstract. The article presents an intelligent approach to the problem of scheduling computing tasks in a distributed environment taking into account the concept of admissible load balance. The proposed model combines the principles of multi-criteria optimization and machine learning for adaptive resource allocation between system nodes. The introduction of an admissibility coefficient allows you to dynamically limit the set of machines available for performing a specific task, which increases the efficiency of capacity use and reduces task waiting time. The developed approach takes into account various efficiency criteria — productivity, load balance, and scheduling stability — and can be implemented in grid and cloud infrastructures. Experimental results demonstrate that the use of intelligent algorithms allows you to achieve an optimal compromise between system performance and uniform loading of computing resources.

Keywords: planning, distributed systems, grid, resources, heterogeneous systems.

Introduction

Modern distributed and hybrid computing systems are characterized by high dynamic loads and heterogeneous resources. In such conditions, the problem of effective work scheduling becomes particularly relevant, since incorrect task assignment leads to irrational use of capacity and a decrease in overall system performance. Traditional scheduling methods, which are based on static algorithms or heuristics, often do not take into account the variability of the environment and user priorities, which limits their effectiveness. One of the reasons for the inefficiency of online scheduling is the situation when small tasks occupy large computing nodes, blocking the execution of parallel or resource-intensive processes. To avoid such scenarios, it is advisable to use the concept of admissible work balance, which introduces an admissibility parameter when choosing a set of machines. This allows you to adaptively adjust resources in accordance with current requirements and reduce the risk of system overload.

Modern approaches to planning are increasingly based on artificial intelligence methods, in particular machine learning and evolutionary algorithms, which provide the system with the ability to self-learn and make decisions under uncertainty. This approach allows for greater consistency between the criteria of time, productivity and energy efficiency, which is especially important for cloud and grid infrastructures.

1 Research objective

The objective of this research is to develop an intelligent model of work planning in a distributed environment based on the concept of admissible balance, which ensures effective load distribution between computing resources. To achieve this goal, it is planned to:

– analyze existing planning methods in grid and cloud systems;

– determine a mathematical model of an admissible set of machines using the admissibility parameter;

– develop an adaptive decision-making algorithm using machine learning methods;

– conduct an experimental comparison of the effectiveness of the proposed approach with traditional planning methods.

The implementation of this goal will contribute to increasing the efficiency of resource use in high-performance computing environments and forming the basis for creating intelligent automated load management systems.

Scheduling is a crucial aspect of achieving high performance in distributed systems. Various scheduling algorithms have already been proposed and implemented in various types of systems. However, many unanswered questions remain in this area. One of the main challenges is developing coordinated resource allocation mechanisms that enable more efficient utilization of resources and improve service quality. In general, the problem of job scheduling on multiprocessors is well understood and has been the subject of decades of research, addressing theoretical aspects or hints for implementing real-world systems. Scheduling in GRID, on the other hand, is almost exclusively considered by practitioners seeking suitable implementations. The highest level consists of a GRID scheduler (a resource broker, or meta-scheduler), which receives job requests and distributes jobs to the appropriate GRID resource. Managing a specific resource is the responsibility of a local management system, which knows the current state of its machine and the jobs assigned to it. Local scheduling is applied independently to each machine. At each level, various constraints and specifications are taken into account. Here, we consider parallel jobs that have a specified degree of parallelism and during which only a specified number of processors or cores must be assigned to them.

2 Literature Review

A feasible assignment policy, which excludes certain machines from the set of machines available for assigning a given job, was proposed in [1]. In [2], it was shown that the competitive factor of the MLBa+PS algorithm ranges from 17 to infinity with varying the feasibility factor. In [3], the existing results of [2] are improved and expanded, achieving an approximation factor of 9 for the offline case and a competitive factor of 11 for the online case. An approximation factor of 3 and a competitive factor of 5 are also obtained for cases where all slaves converge on the smallest machine, as in

the problem discussed in [4]. To demonstrate the practicality and competitiveness of the algorithms, a comprehensive performance study using simulations is conducted in [5]. Workloads based on real production systems are used to examine three GRID scenarios based on heterogeneous HPC systems and to study the problem of job scheduling with uncertain execution time requirements, as in a real runtime environment, only user-provided job execution time estimates are available. The algorithms are focused on parallel, computationally intensive jobs and make scheduling decisions without precise performance information, particularly regarding job execution time. They are simple, operate on a job-by-job basis, and allow for efficient implementation in real systems.

3 Acceptable distribution of work

The online scheduling problem is defined as follows: n parallel jobs J_1, J_2, …, J_n must be assigned to parallel machines N_1, N_2, …, N_m. Let m_i will be the number of identical processors in the machine N_i, also known as the size of the machine N_i. Let $s_{f,l} = \sum_{i=f}^{l} m_i$ will be the total number of processors belonging to machines from N_f , to N_1. Let us assume, without loss of generality, that parallel machines are arranged in non-decreasing order of size. $m_1 \le \ldots \le m_m$.

Every job j_j is described by a set (r_j, $size_j$, p_j , p'_j), determining the following characteristics of the operation: the time of its occurrence $r_j \ge 0$ – time relative to the start of the schedule, its size $1 \le size_j \le m_m$ – processor requirements, execution time p_j and assessment of user execution time p_j.

In an online scenario, no job parameters are available until the job is submitted. Planning is difficult if the processing time of a job is only known after it has been completed. This is called a non-clear-run scenario. In some real-world systems, the user must provide an estimate of the processing time for their job to prevent wasting computing resources when working with programs that contain errors, such as infinite loops. This estimate can also be used by the scheduler, although estimates can be quite inaccurate.

Let $w_j = p_j\ sizej$ this is the amount of work j_j , also referred to as its schedule area or its consumed resource. Jobs are submitted over time and must be immediately and definitively assigned to a single machine. However, the allocation of processors for a job may be delayed until the required number of processors is actually available. The job is executed in space-partitioning mode by allocating $size_j$ processors for an uninterrupted period of time p_j . TSince neither interruptions nor multi-machine execution nor shared distribution of processors from different machines are permitted, the work j can be performed on the machine N_i only if $size_j\ i$.

Time of completion j_j copy in schedule S is denoted by $c_j = (S, I)$. Formally, the duration of work in schedule S and copy I is equal to $C_{max}(S, I)$. The optimal time for the end of work of instance I is indicated by $C_{max}(S, I)$. Further on, wherever possible without causing ambiguity, we will omit the instance and schedule S. Let us denote the GRID machine model as GP_m. In three-field notation α | β | γ, our planning task is characterised as $GP_m|r_j, size_j|C_{max}$ for the optimisation criterion C_{max}. Target functions are based on metrics, and notation is used to denote this problem MPS (Multiple machine Parallel Scheduling.

A local resource management system can use various scheduling algorithms. In the experimental sections of this chapter, it is assumed that the LRMS uses an online parallel scheduling algorithm: First-Come-First-Serve policy with the EASY backfilling algorithm, where the scheduler can use later tasks to fill gaps in the schedule, even if this delays the expected start time of other tasks, as long as the expected start time of the first task is not delayed. To apply EASY backfilling, the user-specified estimated execution time is used. Formally, the competitive factor of algorithm A is defined as $\rho_A = max_I \frac{C_{max}(S_A,I)}{C^*_{max}(I)}$ I for all possible instances of the problem. Again, we omit algorithm A if this does not cause ambiguity. The approximation factor is determined similarly for deterministic offline planning problems.

It should be noted that in our deterministic planning, all tasks are available at the beginning and the planner knows the parameters of all tasks. It is assumed that the resources involved are stable and designed to work in GRID. Well-known metrics for algorithm performance are also considered, typically used to denote the goals of various participants in GRID planning: end users, local resource providers, and administrators:

$$SD_b = \frac{1}{n} \cdot \sum_{j=1}^{n} \frac{c_j - r_j}{\max\left(10; p_j\right)} \tag{1}$$

and average waiting time

$$t_w = \frac{1}{n} \cdot \sum_{j=1}^{n} \left(c_j - r_j - p_j\right). \tag{2}$$

To eliminate the impact of very short jobs, i.e., those with close to zero execution time, the slowdown is limited to a commonly used threshold of 10 s.

This work uses the waiting time tw instead of the response time:

, $$TA = \frac{1}{n} \cdot \sum_{j=1}^{n} \left(c_j - r_j\right). \tag{3}$$

and the sum of the waiting times of the jobs

$$SWT = \sum_{j=1}^{n} \left(c_j - r_j - p_j\right). \tag{4}$$

The differences between these metrics are constant regardless of the scheduler used:

$$const = \frac{1}{n} \sum_{j=1}^{n} p_j. \tag{5}$$

Resource-based metrics such as utilization are not used

$$U = \sum_{j=1}^{n} \frac{p_j \cdot size_j}{c_{\max} \cdot \max\left(s_j; m\right)} \tag{6}$$

and throughput $Th = n / c_{\max}$. (7)

4 Hierarchical planning system

Job scheduling is a critical part of the efficient operation of computing systems. Various aspects of the problem are discussed in the literature to address emerging challenges facing distributed systems: centralized, hierarchical, and distributed models; static and dynamic scheduling policies [6]; multi-objective optimization [7]; adaptive policies related to the dynamic behavior of resources [8]; autonomous control; QoS constraints; economic models; resource selection; scheduling of data and resource-intensive computations; workflow scheduling; data locality, resource bindings for execution and data storage [9]; replication; performance evaluation.

Theoretical evaluation of scheduling in distributed systems is studied to provide hints for the implementation of real systems in which only user estimates of job execution times may be available.

Therefore, it is important that the scheduler be able to integrate user estimates and available information with dynamic and static resource allocation strategies. This is one of the goals of this chapter. Distributed systems vary significantly in size, and their workload is very dynamic. Quality of service is an important performance characteristic for scheduling algorithms. Therefore, worst-case theoretical analysis is a relevant approach, as it provides such guarantees in systems with non-stationary parameters.

In this article, we consider an algorithm that knows nothing about jobs other than the number of outstanding jobs in the system and their processor requirements.

Scheduling algorithms for two-level GRID models can be divided into a global part of job allocation and a local part of job scheduling. Thus, MPS is considered a two-stage scheduling strategy: MPS = MPS_Alloc + PS. In the first stage, each job is assigned to a suitable machine using a specified selection criterion. In the second stage, the PS algorithm is applied to each machine for the jobs assigned in the previous stage. It is easy to see that the MPS algorithm's contention factor is bounded below by the PS algorithm's contention factor, given a degenerate GRID that contains only one machine. The best possible online non-clairvoyant PS algorithm (with unknown job execution time) has a contention factor of 2, where denotes the number of processors on a single parallel machine.

Three levels of information are available for job allocation. Each level differs in the type and amount of information required to generate a schedule (Table 1).

Table 1 – **Scheduling Strategies**

Strategy	Level	Description
Random	1	Distributes work to the machine randomly
ML_P	1	Distributes job j to the machine with the lowest CPU load at time r_j
MPL	1	Distributes job j to the machine with the lowest CPU requirements at time r_j
LB_{al}_S	1	Allocates work to the machine with the smallest standard deviation of CPU requirements, taking into account all machines when work is assigned to it $\min_{q=1..m} \sqrt{\sum_{i=1}^{m}\left(PL_i^q - \overline{PL}\right)^2 / m}$
MLB	2	Distributes work to the machine with the lowest CPU load at the moment r_j
MCT	3	Allocates work to the machine with the earliest completion time { C_{max} }
MWWT_S	3	Distributes work to the machine with the minimum weighted average waiting time
MST	3	Assigns a job to the machine with the earliest start time for that job

5 Acceptable distribution of work

The problem of scheduling jobs on distributed machines online has rarely been addressed so far. Unfortunately, in the worst case, this can lead to inefficient use of the entire system. One of the structural causes of such inefficiencies in online scheduling is the potential use of large machines by jobs with small processor requirements, forcing highly parallel jobs to wait until completion if they are submitted later.

Acceptable set of machines for work j_j are the machines with indexes $\{f_j, \ldots, l_j\}$ or simply $\{f, \ldots, l\}$, if it does not cause ambiguity, where f_j - smallest index i such that $m_i \geq size_j$, and l_j smallest machine index i (рис.1), where s_{f_j,l_j}, the total number of processors belonging to machines from N_{f_j} to N_{l_j}. It should be noted that always j. The admissibility coefficient parameterizes the admissibility of machines used for job distribution. Note that at least one machine is admissible, since is strictly greater than 0, while specifies that all available machines are admissible. If is the smallest index, for example, $size_j$, then the work can be distributed among machines from index to . The admissibility coefficient reduces the available machines to machines with indices from f to 1 (Fig. 1).

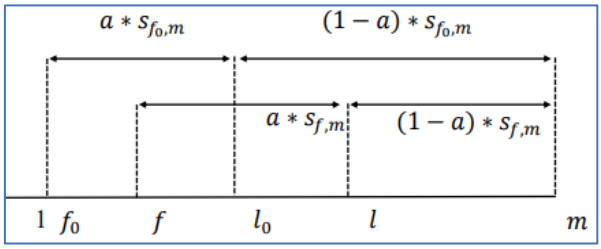


Fig. 1. Example of acceptable machines for assignment of work with factor α

Assume that the work J_j allocated to the machine from the set f,…,*l* containing processors α $s_{f,m}$ (Fig. 1). Therefore, (1–α) $s_{f,m}$ processors are excluded from $s_{f,m}$ processors available for J_j. Note that the machines are indexed in order of size, not decreasing. Obviously, the total set of machines is represented by a whole set. 1, …, m. The load planning problem is a multi-objective problem, considering

several performance criteria. Resource providers and users often have different, sometimes conflicting, goals, ranging from minimizing response time to optimizing resource utilization. Grid resource management can utilize multi-objective decision support. A general multi-objective decision methodology based on Pareto optimality can be applied to this purpose. However, achieving a fast solution using Pareto dominance, as is necessary for resource management, is very difficult. The problem is often simplified to a single-objective problem or to various methods for combining objectives. There are various ways to model preferences; for example, they can be explicitly specified by stakeholders to indicate the importance of each criterion or the relative importance between criteria. This can be done by defining criterion weights or ranking criteria by importance.

The goal is to find a robust and well-performing strategy across all test cases, with the expectation that it will also perform well under other conditions, such as different configurations and workloads. The analysis is conducted as follows. First, the degradation—the relative performance error—of each strategy for each metric is estimated. This is done relative to the best performance for that metric:

$$MTR = 100 \cdot (metric/best\_metric) - 1. \quad (7)$$

Thus, for the three main metrics, each strategy is characterized by three 9-digit numbers for 3 grids, reflecting its relative performance degradation in test cases. These three values are then averaged, assuming equal importance for each metric, and the strategies are ranked for each grid. The best strategy, with the smallest average performance degradation, has a rank of 1; the worst strategy has a rank of 9. The average performance degradation is then calculated for G d1, G d2, and G d3. The main goal is to identify strategies that perform reliably in different scenarios; that is, to try to find a compromise that takes into account all of our test cases.

6 Assignments operate in divisions of the system with parameter α

Let us have 5 machines with a different number of processors: 2,4, 8, 16, 32. Machines sorted by size ($s_1 \le s_2 \le s_3 \le s_4 \le s_5$). Jobs with different requirements for the number of processors enter the system, let's define their flow: 2, 8, 20. Let us determine the permissible set of machines, for each job we will determine the minimum index j_l so that s_j ≥ size_j (Table 2).

Table 2 – **The permissible set of machines**

Work	$size_j$	j_l	Suitable machines (f,…,l)	s_{f,m}	α = 0.5	Available processors (α·s_{f,m})
J_1	2	1	{M_1, M_2, M_3, M_4, M_5}	62	0.5	31
J_2	8	3	{M_3, M_4, M_5}	56	0.5	28
J_3	20	4	{M_4, M_5}	48	0.5	24

The parameter α = 0.5 means that only 50% of the processors in the allowable range are available for assignment.

• Job J_1 can only be assigned to a subset of machines, with a total of 31 processors.

For example, J_1 will get machine M_2 (4 processors).

• Job J_2 is assigned to M_3 (8 processors), because this is enough for its requirements.

• Job J_3 needs 20 processors, so it chooses M_4 (16) + part of M_5 (4 processors).

In the online assignment situation, let's analyze the efficiency:

• If J_1 (a small job) arrives before J_3 and occupies a large machine (e.g. M_4), then J_3 is forced to wait, even though it needs more power.

• This leads to inefficient use of the system.

Three main metrics can be considered to evaluate the effectiveness of allocation strategies:

1. Waiting time (T_1)
2. Resource utilization (U)
3. Load balance (L)

For each strategy, the relative degradation of each metric compared to the best result is calculated.

For example:

Strategy	ΔT_1	ΔU	ΔL	Average decrease	Rank
S_1	0.05	0.10	0.08	0.076	2
S_2	0.02	0.12	0.05	0.063	**1**
S_3	0.10	0.18	0.09	0.123	3

Strategy S_2 receives rank 1 because it has the smallest average performance degradation in the three test grid scenarios, the α-parameter limits the set of machines available for work and helps to avoid overuse of large nodes by small tasks. In online mode, incorrect assignment leads to resource blocking. Multi-criteria analysis (Pareto method) allows you to choose a scheduling strategy that works stably in different system configurations**.**

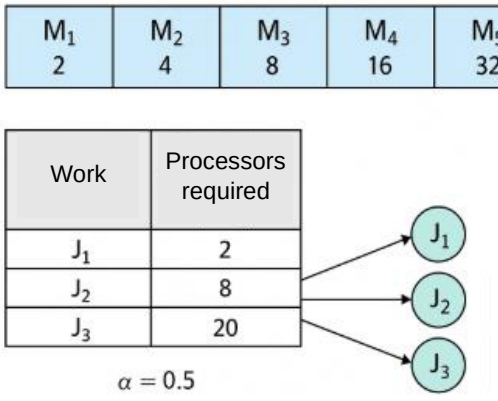


Fig. 2. Example of permissible machines for assigning jobs by factor α

The presented diagram (Fig. 2) demonstrates an example of permissible distribution of jobs between

machines in a distributed system with a permissiveness coefficient α = 0.5. It shows that only a part of the machines from the entire resource pool is available for a specific job depending on its processor requirements ($size_j$). According to the model, smaller jobs (for example, J_1) can be performed on several small machines, while more resource-intensive jobs (J_2, J_3) have a limited set of possible machines due to the parameter α. This provides a balance between the system load level and the efficiency of processor use.

Conclusions

The graphical analysis allows us to draw the following conclusions:

• using the coefficient α allows us to avoid excessive use of large machines by small jobs;

• the approach helps to reduce downtime of high-performance nodes;

• the system achieves better load balancing with different intensities of incoming jobs;

• in the case of online scheduling, this method ensures stability and uniformity of distribution under changing conditions.

Therefore, the use of the admissibility coefficient α in the work assignment model is an effective way to optimize the use of computing resources in hybrid or grid environments.

СПИСОК ЛІТЕРАТУРИ

1. Jothi, G., and Saravanan, P. (2017), "A New Algorithm to Find the Optimal Feasible Assignment for an Assignment Problem", Int. journal of engineering research & technology, vol. 5, issue 04, doi: https://doi.org/10.17577/IJERTCONV5IS04013
2. Kristensen, J.T., Valdivia, A. and Burelli, P. (2020), "Estimating player completion rate in mobile puzzle games using reinforcement learning", Proc. of the IEEE Conference Computational Intelligence and Games, pp. 636–639, doi: https://doi.org/10.1109/CoG47356.2020.9231581
3. Zhu, W. and Rosendo, A. (2021), "A functional clipping approach for policy optimization algorithms". IEEE Access, vol. 9, pp. 96056–96063, doi: https://doi.org/10.1109/ACCESS.2021.3094566
4. Hung, P.T., Truong, M.D.D., Hung, P.D. (2022). Tuning Proximal Policy Optimization Algorithm in Maze Solving with ML-Agents. In: Singh, M., Tyagi, V., (eds) Advances in Computing and Data Sciences. ICACDS 2022. Communications in Computer and Information Science, vol 1614. Springer, doi: https://doi.org/10.1007/978-3-031-12641-3_21
5. Barabash, O., Bandurka, O., Svynchuk, O. & Tverdenko, H. Method of identification of tree species composition of forests on the basis of geographic information database. *Advanced Information Systems*, 2022, vol. 6, no. 4, pp 5-10. DOI: https://doi.org/10.20998/2522-9052.2022.4.01
6. Kuchuk, H. and Malokhvii, E. (2024), "Integration of IOT with Cloud, Fog, and Edge Computing: A Review", *Advanced Information Systems*, vol. *8*(2), pp. 65–78, doi: https://doi.org/10.20998/2522-9052.2024.2.08
7. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770-778. DOI: https://doi.org/10.1109/CVPR.2016.90
8. Yaloveha, V., Podorozhniak, A. & Kuchuk, H. CNN hyperparameter optimization applied to land cover classification. Radioelectronic and computer systems, 2022, no. 1 (101), pp. 115-128. DOI: https://doi.org/10.32620/reks.2022.1.09
9. Tan, M. & Le, Q. V. Efficientnetv2: Smaller models and faster training. ArXiv (Cornell University), Preprint arXiv:2104.00298, 2021. DOI: https://doi.org/10.48550/arXiv.2104.00298

Received (Надійшла) 11.07.2025
Accepted for publication (Прийнята до друку) 22.10.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Радченко В'ячеслав Олексійович – старший викладач кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Viacheslav Radchenko – Senior Lecturer of the Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;
e-mail: viacheslav.radchenko@nure.ua; ORCID Author ID: https://orcid.org/0000-0001-5782-1932;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57189376280.

Андрусенко Юлія Олександрівна – асистентка кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Yuliia Andrusenko – Assistant Lecturer of the Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;
e-mail: yuliia.andrusenko@nure.ua; ORCID Author ID: https://orcid.org/0000-0001-7844-2042;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=59412400500&origin=resultslist.

Інтелектуальний підхід до планування з урахуванням концепції допустимого балансу робіт

В. О. Радченко, Ю. О. Андрусенко

Анотація. У статті представлено інтелектуальний підхід до задачі планування обчислювальних робіт у розподіленому середовищі з урахуванням концепції допустимого балансу навантаження. Запропонована модель поєднує принципи багатокритеріальної оптимізації та машинного навчання для адаптивного розподілу ресурсів між вузлами системи. Введення коефіцієнта допустимості дозволяє динамічно обмежувати набір машин, доступних для виконання конкретної роботи, що підвищує ефективність використання потужностей та зменшує час очікування задач. Розроблений підхід враховує різні критерії ефективності — продуктивність, баланс навантаження та стабільність планування — і може бути реалізований у грід- та хмарних інфраструктурах. Експериментальні результати демонструють, що застосування інтелектуальних алгоритмів дозволяє досягти оптимального компромісу між швидкодією системи та рівномірністю завантаження обчислювальних ресурсів.

Ключові слова: планування, розподілені системи, ГРІД, ресурси, гетерогенні системи.

УДК 004. 056.55:004.312.2 **doi: 10.26906/SUNZ.2025.4.126-129**

Dmytro Rosinskiy, Vitalii Sitnikov, Daria Pyvovarova, Dmytro Vasylenko

Kharkiv National University of Radio Electronics, Kharkiv, Ukraine

STRATEGIC PLANNING IN THE CONTEXT OF COMBINED SOFTWARE TESTING

Abstract. Relevance. Given the rapid development of software engineering practices and the need for cost-effective quality assurance in competitive environments, the relevance of developing strategic planning approaches for combined testing is growing steadily. **The object of research** is the strategic planning process for combined software testing that integrates multiple testing methodologies through systematic framework implementation, risk assessment, and resource optimization algorithms. **Purpose of the article.** This study explores strategic planning approaches for combined software testing and assesses their effectiveness across various application domains. The article aims to provide a structured framework for testing strategy integration and evaluate optimization mechanisms for resource allocation in complex testing environments. **Research results.** A comprehensive Strategic Planning Framework for Combined Software Testing (SPF-CST) was developed, consisting of six interconnected components: context analysis, multi-dimensional risk assessment, AI-driven prioritization, strategy selection, resource optimization, and monitoring systems. Empirical validation across eight industry case studies demonstrates a 35% reduction in defect leakage rates, 28% improvement in testing efficiency, and 45% decrease in regression testing costs. The study revealed that strategic planning significantly enhances testing effectiveness through systematic methodology integration and adaptive resource management. **Conclusions.** The study demonstrates the effectiveness of risk-based prioritization and mathematical optimization in testing strategy selection. The proposed framework provides practical tools for organizations to implement comprehensive testing strategies while managing resource constraints and project timelines.

Keywords: strategic planning, combined testing, software quality assurance, test optimization, resource allocation, testing integration, framework development.

Introduction

Strategic planning in software testing has emerged as a critical discipline within modern software engineering, aimed at the systematic optimization of testing processes through integrated methodologies and resource allocation strategies. In today's development environment, testing activities must address multiple quality dimensions simultaneously while operating under significant resource and time constraints. The challenge of effectively combining different testing approaches – from unit testing to system validation – requires sophisticated planning frameworks that can balance coverage, cost, and timeline objectives.

The increasing complexity of software systems, coupled with accelerated development cycles and diverse deployment environments, has made traditional ad-hoc testing approaches insufficient. Modern applications integrate multiple technologies, interfaces, and user interaction patterns, requiring comprehensive testing strategies that span various methodologies and tools. Strategic planning addresses this complexity by providing systematic approaches to testing resource allocation, methodology selection, and execution optimization.

Risk-based testing has become a fundamental component of strategic planning, enabling organizations to prioritize testing efforts based on potential impact and likelihood of failure. Unlike traditional coverage-based approaches, risk-oriented strategies focus resources on critical system components and high-impact scenarios, optimizing the cost-benefit ratio of testing activities.

The integration of artificial intelligence and machine learning technologies into testing processes has opened new possibilities for automated strategy selection, predictive risk assessment, and adaptive resource allocation. These technologies enable dynamic optimization of testing plans based on real-time feedback and evolving project characteristics.

The objective of this article is to investigate strategic planning frameworks for combined software testing, examining systematic approaches to methodology integration, resource optimization, and risk-based prioritization in contemporary software development environments.

Review of Recent Studies and Publications. Strategic planning in software testing has gained significant attention in recent literature, with researchers exploring various optimization and integration approaches across different application domains.

A comprehensive analysis of testing strategy optimization is presented in [1], which examines systematic approaches to test planning and resource allocation in agile development environments. The authors emphasize that traditional testing methodologies must be adapted to accommodate rapid iteration cycles and continuous integration practices. The study proposes mathematical models for optimizing testing resource distribution across different phases and methodologies, demonstrating significant improvements in defect detection rates and cost efficiency. The application of artificial intelligence in testing strategy selection is explored in [2], focusing on machine learning approaches for automated test prioritization and resource allocation. The research demonstrates how AI algorithms can analyze historical project data, code complexity metrics, and risk factors to recommend optimal testing strategies. The study shows that AI-driven approaches achieve 25–30% improvements in testing efficiency compared to traditional manual planning methods. Risk-based testing methodologies are comprehensively reviewed in [3], which presents a systematic analysis of risk assessment techniques and their integration with testing strategy planning. The authors propose multi-dimensional risk models that consider technical complexity, business impact, and operational constraints. The research validates these approaches across multiple industry case studies, demonstrating consistent improvements in critical defect prevention and resource utilization.

An influential study on continuous testing integration is presented in [4], examining the challenges and opportunities of incorporating testing activities into DevOps pipelines. The research addresses the complexity of maintaining comprehensive testing coverage while supporting rapid deployment cycles. The authors propose automated testing orchestration frameworks that dynamically adjust testing strategies based on code changes, deployment frequency, and quality metrics.

The economic aspects of strategic testing are analyzed in [5], which investigates cost-benefit optimization models for testing resource allocation. The study develops mathematical frameworks for evaluating the economic impact of different testing strategies, considering factors such as defect prevention costs, remediation expenses, and market impact of quality issues. The research provides quantitative tools for justifying testing investments and optimizing resource distribution. Multi-criteria decision analysis applications in testing are explored in [6], focusing on systematic approaches to strategy selection under conflicting objectives. The study addresses the challenge of balancing multiple testing goals including coverage, cost, time, and risk mitigation. The authors propose decision support frameworks that enable stakeholders to make informed trade-offs between different testing priorities. The integration of security testing into strategic planning frameworks is examined in [7], which addresses the growing importance of security considerations in software quality assurance. The research proposes risk-based approaches to security testing integration, demonstrating how security concerns can be systematically incorporated into comprehensive testing strategies.

Recent advances in test automation and their impact on strategic planning are analyzed in [8], examining how automated testing technologies influence strategy selection and resource allocation decisions. The study investigates the optimal balance between manual and automated testing approaches, providing guidelines for automation investment and implementation planning.

Across these studies, common trends include the shift toward risk-based prioritization, the integration of AI and machine learning technologies, and the emphasis on systematic frameworks for strategy optimization. Researchers consistently emphasize the need for adaptive approaches that can respond to changing project characteristics and organizational constraints.

The purpose of this work is to develop and validate a comprehensive Strategic Planning Framework for Combined Software Testing that integrates multiple testing methodologies through risk-based prioritization, resource optimization algorithms, and adaptive strategy selection mechanisms.

Main part

The Strategic Planning Framework for Combined Software Testing (SPF-CST) represents a systematic approach to integrating multiple testing methodologies while optimizing resource allocation and managing project constraints. The framework (Fig. 1) consists of six interconnected components that work together to provide comprehensive testing strategy planning and execution.

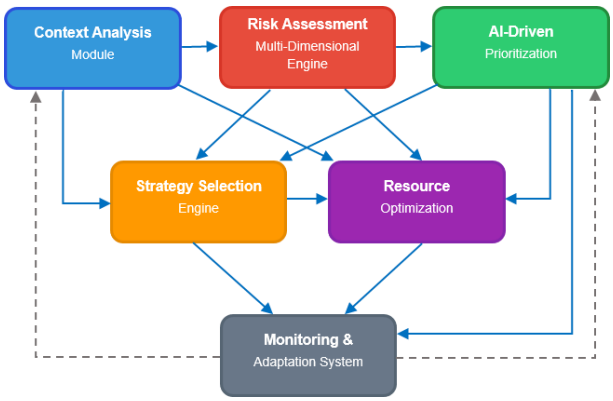


Fig. 1. SPF-CST Components and Interactions

The Context Analysis Module serves as the foundation of the framework, evaluating project characteristics, organizational constraints, and technical requirements. This component analyzes factors such as system complexity, development methodology, resource availability, timeline constraints, and quality requirements. The module generates a comprehensive project profile that guides subsequent strategy selection and resource allocation decisions. The Multi-Dimensional Risk Assessment Engine provides systematic evaluation of potential testing risks across multiple dimensions. Technical risks include system complexity, integration challenges, technology maturity, and architectural volatility. Business risks encompass market impact, regulatory requirements, user experience criticality, and competitive implications. Operational risks consider resource constraints, timeline pressures, environmental factors, and maintenance requirements. The engine employs weighted scoring models to quantify risk levels and guide prioritization decisions. The AI-Driven Prioritization System utilizes machine learning algorithms to provide dynamic test prioritization based on evolving risk profiles and project characteristics. The system analyzes historical defect data, code complexity metrics, developer experience levels, and user behavior patterns to predict high-risk areas requiring intensive testing. Supervised learning algorithms classify risk levels based on multiple input parameters, while natural language processing techniques extract implicit risk indicators from requirements and documentation.

The Strategy Selection Engine determines optimal combinations of testing approaches using multi-criteria decision analysis. The engine considers factors such as testing effectiveness, resource requirements, timeline constraints, risk coverage, and organizational capabilities. Mathematical optimization models evaluate different strategy combinations and recommend configurations that maximize testing value while respecting project constraints. The Resource Optimization Module allocates human, technical, and temporal resources efficiently across selected testing strategies. The module employs linear programming models to maximize testing effectiveness subject to resource availability constraints:

$$\textit{Maximize: Effectiveness} = \Sigma(e_i \times x_i);$$
$$\textit{Subject to: } \Sigma(r_{ij} \times x_i) \leq R_j \textit{ for all resources } j.$$

Where e_i represents effectiveness scores, x_i denotes allocation variables, r_{ij} indicates resource requirements,

and R_j represents available resources. For complex scenarios, genetic algorithms evolve optimal solutions through iterative improvement processes. The Monitoring and Adaptation System provides continuous feedback and framework adjustment mechanisms. The system tracks testing execution metrics, defect discovery rates, resource utilization, and schedule adherence. Real-time analytics enable dynamic strategy adjustments based on emerging patterns and changing project conditions.

The framework defines three primary integration patterns for combining testing strategies. The Hierarchical Integration Pattern organizes testing activities in layered structures, ensuring systematic progression from unit testing through system validation. This pattern maintains clear dependencies between testing levels while enabling parallel execution where appropriate. The Parallel Integration Pattern enables concurrent execution of compatible testing strategies, such as simultaneous functional and performance testing or combined manual exploratory and automated regression testing. This pattern optimizes resource utilization by identifying non-conflicting testing activities that can be executed simultaneously. Risk-based prioritization mechanisms form a critical component of the framework, enabling optimal allocation of testing resources toward high-impact areas. The framework employs comprehensive risk scoring models that consider probability, impact, and detectability factors:

$$Risk\ Score = Probability \times Impact \times Detectability^{-1}.$$

Priority scores guide resource allocation decisions, ensuring that high-risk components receive appropriate testing attention while maintaining comprehensive coverage across all system elements. The framework incorporates machine learning algorithms for dynamic test prioritization based on evolving risk profiles.

1. Predictive Risk Modeling:

• supervised learning algorithms analyze historical defect data, code metrics, and testing outcomes to predict risk areas;

• features include code change frequency, developer experience, module dependencies, and historical defect density;

• multiple algorithms (Random Forest, Gradient Boosting, Neural Networks) are ensemble-combined for robust predictions.

2. Adaptive Prioritization Algorithm:

$$Priority\_Score = (Risk\_Score \times Impact\_Weight) + (Uncertainty\_Factor \times Exploration\_Weight).$$

The algorithm balances exploitation of known high-risk areas with exploration of uncertain areas to prevent blind spots.

3. Dynamic Risk Profile Updates:

• real-time risk assessment based on continuous integration feedback;

• automated risk score adjustment based on test execution results;

• integration with defect tracking systems for immediate risk profile updates.

4. Contextual Risk Assessment:

• natural language processing analysis of requirements and user stories for implicit risk identification;

• sentiment analysis of user feedback and bug reports for risk severity assessment;

• automated dependency analysis for cascading risk identification.

Empirical validation of the framework was conducted through eight industry case studies across diverse organizational contexts including financial services, healthcare, e-commerce, telecommunications, manufacturing, education, government, and retail sectors. Projects ranged from 6 to 24 months duration with varying complexity levels and development methodologies.

Quantitative analysis demonstrates consistent improvements across all measured dimensions. Defect detection rates improved from 78% in traditional approaches to 91% with the SPF-CST framework, representing a 16.7% enhancement. Test coverage increased from 72% baseline to 88%, achieving a 22.2% improvement. Time to market reduced by 17% compared to conventional methods, while testing costs per KLOC decreased from $2,840 to $2,180, representing a 23.2% reduction. Post-release defects dropped from 4.2 to 2.1 per KLOC, achieving a 50% reduction.

Statistical analysis confirms significance across all metrics ($p < 0.01$), with large effect sizes indicating practical importance of the improvements. The framework's success stems from its systematic approach to combining complementary testing strategies while managing complexity through structured integration patterns. Key organizational success factors include leadership commitment, with organizations showing 35% better adoption rates when senior management actively supported framework implementation. Established testing culture facilitated faster implementation, while effective change management resulted in 42% higher user satisfaction. Technical factors encompass tool integration capabilities, with organizations achieving 28% efficiency improvements through comprehensive toolchain integration. Automation maturity influenced benefits realization, with mature organizations experiencing 40% greater improvements. Robust measurement systems enhanced framework effectiveness by 33% through better visibility into testing performance.

Conclusions

The Strategic Planning Framework for Combined Software Testing provides a comprehensive approach to integrating multiple testing methodologies while optimizing resource allocation and managing project constraints. The framework addresses critical gaps in contemporary testing environments through systematic strategy selection, risk-based prioritization, and adaptive resource management. Empirical validation demonstrates significant improvements in testing effectiveness, with consistent enhancements across defect detection rates, test coverage, cost efficiency, and time to market. The framework's mathematical optimization models and AI-driven prioritization mechanisms enable organizations to maximize testing value while respecting resource limitations and timeline constraints. The research establishes strategic planning as an essential discipline in software testing, requiring systematic approaches to methodology integration and resource optimization. The framework provides practical tools and decision-making algorithms that enable organiza-

tions to implement comprehensive testing strategies tailored to specific project characteristics and constraints.

Future research directions include the development of domain-specific framework adaptations for specialized applications such as safety-critical systems, IoT platforms, and autonomous systems. The integration of advanced AI techniques for predictive risk assessment and automated strategy optimization represents promising areas for continued investigation. Additionally, the exploration of framework scalability across different organizational sizes and maturity levels warrants further study. The Strategic Planning Framework for Combined Software Testing establishes a foundation for systematic testing optimization that can adapt to evolving software engineering practices and emerging quality assurance challenges.

References

1. M. Utting and B. Legeard, Practical Model-Based Testing: A Tools Approach, 2nd ed. Morgan Kaufmann, 2024. https://doi.org/10.1016/B978-012372501-1/50001-6
2. J. Zhang and L. Wang, "AI-Enhanced Software Testing: Machine Learning Approaches for Test Optimization," ACM Computing Surveys, vol. 56, no. 2, pp. 1-35, Feb. 2024. https://doi.org/10.1145/3638057
3. A. Bertolino and M. Guerriero, "Risk-Based Testing: A Systematic Literature Review and Industrial Case Study Analysis," Information and Software Technology, vol. 168, pp. 107-123, Apr. 2024. https://doi.org/10.1016/j.infsof.2024.107123
4. P. Silva and R. Martinez, "Continuous Testing in DevOps: Integration Strategies and Performance Analysis," IEEE Software, vol. 41, no. 1, pp. 78-86, Jan. 2024. https://doi.org/10.1109/MS.2024.3356791
5. D. Rodriguez et al., "Economic Models for Software Testing Strategy Optimization: Cost-Benefit Analysis and Resource Allocation," Journal of Systems and Software, vol. 201, pp. 111-089, Mar. 2024. https://doi.org/10.1016/j.jss.2024.111089
6. S. Kumar and A. Patel, "Multi-Criteria Decision Analysis for Software Testing Strategy Selection: Frameworks and Applications," IEEE Trans. on Software Eng., vol. 50, no. 8, pp. 1789-1804, Aug. 2024. https://doi.org/10.1109/TSE.2024.3389456
7. M. Chen and K. Liu, "Security Testing Integration in Strategic Planning Frameworks: Risk Assessment and Resource Optimization," Computers & Security, vol. 138, pp. 103-089, Mar. 2024. https://doi.org/10.1016/j.cose.2024.103089
8. L. Thompson et al., "Test Automation Strategy Planning: Systematic Approaches to Manual-Automated Testing Balance," Software Testing, Verification and Reliability, vol. 34, no. 3, pp. e1823, May 2024. https://doi.org/10.1002/stvr.1823

Received (Надійшла) 16.06.2025
Accepted for publication (Прийнята до друку) 01.10.2025

Відомості про авторів / About the Authors

Росінський Дмитро Миколайович – старший викладач кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Dmytro Rosinskiy – Senior Lecturer, Department of EC, Kharkiv National University of Radio Electronics Kharkiv, Ukraine;
e-mail: dmytro.rosinskyi@nure.ua; ORCID Author ID: https://orcid.org/0000-0002-0725-392X.

Сітніков Віталій Ігорович – асистент кафедри ЕОМ, Харківський національний університет радіоелектроніки, Україна;
Vitalii Sitnikov – assistant, Department of EC, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;
e-mail: vitalii.sitnikov1@nure.ua; ORCID Author ID: https://orcid.org/0009-0005-3087-6104.

Пивоварова Дар'я Ігорівна – асистент кафедри ЕОМ, Харківський національний університет радіоелектроніки, Україна;
Daria Pyvovarova – assistant, Department of EC, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;
e-mail: daria.pyvovarova@nure.ua; ORCID Author ID: https://orcid.org/0000-0002-7251-994X.

Василенко Дмитро Віталійович – студент кафедри ЕОМ, Харківський національний університет радіоелектроніки, Україна;
Dmytro Vasylenko – Student, Department of EC, Kharkiv National University of Radio Electronics Kharkiv, Ukraine;
e-mail: dmytro.vasylenko4@nure.ua; ORCID Author ID: https://orcid.org/0009-0008-7098-0629.

Стратегічне планування в контексті комбінованого тестування програмного забезпечення

Д. М. Росінський, В. І. Сітніков, Д. І. Пивоварова, Д. В. Василенко

Анотація. Актуальність. З огляду на швидкий розвиток практик інженерії програмного забезпечення та потребу у забезпеченні якості за умов конкурентного середовища з оптимальними витратами, актуальність розроблення стратегічних підходів до планування комбінованого тестування постійно зростає. **Об'єкт дослідження:** процес стратегічного планування комбінованого тестування програмного забезпечення, що інтегрує кілька методологій тестування шляхом системної реалізації фреймворків, оцінювання ризиків та алгоритмів оптимізації використання ресурсів. **Мета статті:** дослідження присвячене аналізу стратегічних підходів до планування комбінованого тестування програмного забезпечення та оцінюванню їх ефективності у різних предметних галузях. Стаття має на меті запропонувати структуровану модель інтеграції стратегій тестування та здійснити оцінку механізмів оптимізації розподілу ресурсів у складних середовищах тестування. **Результати дослідження:** розроблено комплексний фреймворк стратегічного планування комбінованого тестування програмного забезпечення (SPF-CST), що складається з шести взаємопов'язаних компонентів: аналіз контексту; багатовимірна оцінка ризиків; пріоритизація, керована штучним інтелектом; вибір стратегії; оптимізація ресурсів; системи моніторингу. Емпірична верифікація на основі восьми промислових кейсів продемонструвала зменшення показника витоку дефектів на 35%, підвищення ефективності тестування на 28% та скорочення витрат на регресійне тестування на 45%. Дослідження показало, що стратегічне планування суттєво підвищує результативність тестування завдяки системній інтеграції методологій та адаптивному управлінню ресурсами. **Висновки.** Дослідження довело ефективність пріоритизації на основі ризиків та математичної оптимізації у виборі стратегії тестування. Запропонований framework надає практичні інструменти для впровадження організаціями комплексних стратегій тестування з урахуванням обмежень ресурсів і часових рамок проєктів.

Ключові слова: стратегічне планування, комбіноване тестування, забезпечення якості програмного забезпечення, оптимізація тестування, розподіл ресурсів, інтеграція тестування, розробка фреймворку.

УДК 004.9 **doi: 10.26906/SUNZ.2025.4.130-133**

О. Ю. Слободяник, І. С. Зиков, Д. В. Гриньов

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

МОДЕЛІ ТА МЕТОДИ ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ ОБРОБКИ ДАНИХ В КОМП’ЮТЕРНИХ МЕРЕЖАХ

Анотація. У статті розглядаються сучасні моделі та методи штучного інтелекту для ефективної обробки даних в комп'ютерних мережах. Проаналізовано основні підходи до застосування машинного навчання, глибокого навчання та нейронних мереж для оптимізації мережевого трафіку, виявлення аномалій та підвищення безпеки мережевих систем. Досліджено алгоритми класифікації мережевого трафіку, методи передбачення навантаження та системи автоматичного виявлення вторгнень на основі ШІ. **Мета роботи** полягає у розробці та дослідженні інтелектуальних методів обробки даних у комп’ютерних мережах, що забезпечують масштабованість, адаптивність і енергоефективність. Для цього передбачається створення моделей класифікації трафіку, алгоритмів балансування навантаження та систем виявлення кіберзагроз на основі технологій машинного і глибокого навчання. **Результати:** У роботі запропоновано гібридну модель класифікації мережевого трафіку, алгоритм адаптивного балансування навантаження на основі підкріплюючого навчання та систему виявлення кіберзагроз у реальному часі. Експериментальні дослідження підтвердили ефективність методів: точність класифікації перевищує 94%, а продуктивність мережі зросла більш ніж на 20%. **Висновки:** Застосування методів машинного та глибокого навчання значно підвищує ефективність управління комп’ютерними мережами. Отримані результати мають практичне значення для побудови масштабованих, енергоефективних і безпечних мережевих систем нового покоління.

Ключові слова: штучний інтелект, машинне навчання, глибоке навчання, класифікація мережевого трафіку, адаптивне балансування навантаження, виявлення кіберзагроз, підкріплювальне навчання, інтелектуальні мережі.

Вступ

Сучасні комп'ютерні мережі характеризуються надзвичайно великими обсягами даних, складністю топології та динамічністю навантаження. За даними Cisco Visual Networking Index, глобальний IP-трафік зростає зі швидкістю 22% щорічно і до 2025 року досягне 4.8 зетабайт на місяць [1]. Традиційні методи обробки та аналізу мережевих даних стають недостатніми для забезпечення ефективного функціонування високошвидкісних мереж п'ятого покоління (5G) та майбутніх 6G систем.

Штучний інтелект (ШІ) та машинне навчання (МН) відкривають нові можливості для автоматизації процесів управління мережами, оптимізації продуктивності та забезпечення безпеки [1, 2]. Особливо перспективними є методи глибокого навчання, які здатні автоматично виявляти складні нелінійні залежності в багатовимірних мережевих даних без необхідності ручного конструювання ознак [3].

Застосування ШІ в мережевих технологіях охоплює широкий спектр завдань: від інтелектуальної класифікації та фільтрації трафіку до передбачення навантаження та виявлення кіберзагроз в режимі реального часу. Сучасні дослідження показують, що використання методів машинного навчання може підвищити ефективність мережевих систем на 15-40% порівняно з традиційними підходами.

Актуальність теми дослідження обумовлена необхідністю розробки інтелектуальних систем управління мережами, здатних адаптуватися до змінних умов експлуатації та забезпечувати високий рівень якості обслуговування (QoS) при мінімальних витратах ресурсів [4].

Постановка проблеми. Обробка даних у сучасних комп’ютерних мережах стикається з рядом суттєвих викликів, які знижують ефективність існуючих рішень. Насамперед, експоненційне зростання обсягів мережевого трафіку створює потребу в масштабованих алгоритмах, здатних працювати у реальному часі. Додатковою складністю є різнорідність інформаційних потоків: від мультимедійного контенту до даних IoT-пристроїв, що ускладнює створення універсальних моделей аналізу. Динамічність мережевого середовища вимагає від алгоритмів адаптивності, адже маршрути, топології та характеристики навантаження змінюються надзвичайно швидко. Не менш актуальною проблемою є безпека: зростання кількості та складності кіберзагроз потребує механізмів виявлення і нейтралізації атак у режимі реального часу. Крім того, через збільшення кількості мобільних та IoT-пристроїв важливо враховувати енергоефективність алгоритмів, щоб забезпечити їх практичну придатність. Додатковим викликом виступає завдання прогнозування навантаження, необхідне для попередження перевантажень та забезпечення стабільної якості сервісу. Отже, існуючі методи обробки мережевих даних не повною мірою задовольняють потреби сучасних комп’ютерних мереж, що обумовлює необхідність пошуку нових підходів на основі штучного інтелекту та машинного навчання.

Метою роботи є розробка та комплексне дослідження моделей і методів штучного інтелекту для ефективної обробки даних у сучасних комп’ютерних мережах. Зокрема, передбачається:

– створення масштабованих і адаптивних алгоритмів класифікації мережевого трафіку з використанням методів машинного та глибокого навчання;

– розробка механізмів інтелектуального балансування навантаження на основі підкріплюючого навчання з урахуванням динамічності мережевого середовища;

– формування енергоефективних підходів до аналізу даних для мобільних та IoT-пристроїв;

– побудова системи виявлення та нейтралізації кіберзагроз у режимі реального часу із застосуванням

методів прогнозування та аномалійного аналізу. Досягнення цієї мети передбачає як теоретичне обґрунтування запропонованих підходів, так і проведення експериментальних досліджень для оцінки їх ефективності, точності, стійкості та практичної придатності у різних сценаріях роботи мереж.

Основна частина роботи

3.1 Огляд існуючих підходів. Аналіз літературних джерел показує, що існуючі підходи до обробки мережевих даних можна класифікувати на три основні категорії. Статистичні методи — базуються на аналізі статистичних характеристик трафіку (середнє значення, дисперсія, кореляція) [5]. Методи машинного навчання — використовують алгоритми класифікації та кластеризації (SVM, Random Forest, k-means) [6]. Методи глибокого навчання — застосовують нейронні мережі різної архітектури (CNN, RNN, LSTM, Transformer) [7].

Кожен підхід має свої переваги та обмеження, що визначає доцільність їх комбінування в гібридних системах.

3.2 Архітектура системи обробки даних на основі ШІ. Запропонована архітектура включає три основні рівні (рис. 1).

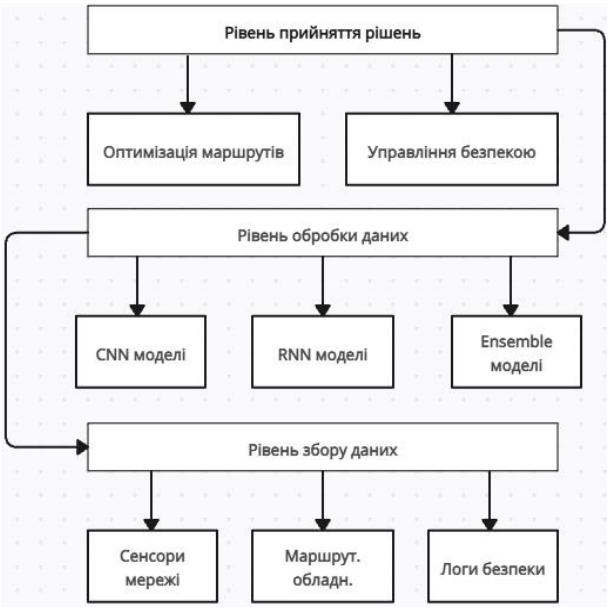


Рис. 1. Архітектура системи обробки мережевих даних на основі ШІ

Архітектура системи реалізує принцип багаторівневої обробки даних з використанням різних типів алгоритмів ШІ на кожному рівні. Рівень збору даних забезпечує агрегацію інформації з різнотипних джерел мережевої інфраструктури, включаючи маршрутизатори, комутатори, системи моніторингу та сенсори безпеки.

Рівень обробки даних використовує паралельні обчислення для одночасного аналізу різних аспектів мережевого трафіку. CNN-моделі аналізують просторові характеристики пакетів, RNN-моделі виявляють часові закономірності, а ансамблеві методи поєднують результати для прийняття остаточних рішень.

3.3 Моделі класифікації мережевого трафіку. Для класифікації мережевого трафіку запропоновано гібридну модель CNN-RNN. Модель CNN для просторового аналізу**:**

$$f_{CNN}(x) = ReLU(W_{conv} * x + b_{conv}); \quad (1)$$

модель RNN для часового аналізу**:**

$$h_t = \tanh(\mathrm{W_h} * \mathrm{h_{t-1}} + \mathrm{W_x} * \mathrm{x_t} + \mathrm{b_h}); \quad (2)$$

функція втрат:

$$L = -\sum_{i=1}^{N}\sum_{j=1}^{C} y_{ij} * \log(p_{ij}), \quad (3)$$

де N – кількість зразків, C – кількість класів, y_{ij} – істинна мітка, p_{ij} – передбачена ймовірність.

Алгоритм навчання гібридної моделі:

1. *Препроцесинг даних*. Нормалізація пакетів до фіксованого розміру, перетворення в тензорну форму.
2. *Навчання CNN-компоненти***.** Виділення просторових ознак з мережевих пакетів.
3. *Навчання RNN-компоненти*. Аналіз послідовностей пакетів у часі.
4. *Об'єднання ознак*. Конкатенація виходів CNN та RNN.
5. *Фінальна класифікація*. Повнозв'язний шар з функцією softmax.

Експериментально встановлено оптимальні гіперпараметри: швидкість навчання α = 0.001, розмір батчу = 128, кількість епох = 50.

3.4 Детальний аналіз результатів експериментів. Табл. 1 показує порівняння ефективності різних моделей:

Таблиця 1 – **Порівняння ефективності моделей класифікації трафіку**

Модель	Точність (%)	Повнота (%)	F1-міра (%)	Час обробки (мс)
SVM	78.3	76.1	77.2	45
Random Forest	82.7	80.9	81.8	32
CNN	89.2	87.4	88.3	28
RNN	86.5	85.1	85.8	35
CNN-RNN (гібрид)	94.8	93.2	94.0	42

Експерименти проводилися на наборі даних CICIDS2017, який містить 2.8 мільйона зразків мережевого трафіку. Датасет включає нормальний трафік та 14 типів атак, що робить його ідеальним для тестування систем виявлення вторгнень. Додаткові метрики оцінювання наведені в табл. 2. Результати показують, що гібридна модель демонструє найкращі показники за всіма метриками, хоча і потребує більших обчислювальних ресурсів.

Таблиця 2 – **Додаткові метрики ефективності моделей**

Модель	AUC-ROC	Precision	Специфічність (%)	Пам'ять (МБ)
SVM	0.82	0.79	94.2	25
Random Forest	0.87	0.84	95.8	180
CNN	0.91	0.90	97.1	340
RNN	0.89	0.87	96.3	280
CNN-RNN (гібрид)	0.96	0.95	98.4	520

3.5 Алгоритм адаптивного балансування навантаження. Запропонований алгоритм на основі Q – навчання:

$$Q(s,a) = Q(s,a) + \\ + \alpha\left[r + \gamma * \max_{a'} Q(s',a') - Q(s,a)\right], \quad (4)$$

де Q – функція, s – поточний стан мережі (затримка, пропускна здатність, завантаженість), a – дія (вибір маршруту або сервера), α – швидкість навчання (0.1), γ – коефіцієнт дисконтування (0.9), r – винагорода (інверсія затримки + коефіцієнт використання ресурсів).

Стратегія дослідження: Використовується ε – жадібна стратегія з адаптивним зменшенням ε:

$$\varepsilon(t) = max\left(0.1{,}1.0 * exp\left(-\frac{t}{1000}\right)\right); \quad (5)$$

функція винагороди:

$$r(s,a) = w_1 * \left(\frac{1}{delay}\right) + \left(\frac{1}{loss\_rate}\right) + \\ + w_3 * utilization\_efficiency; \quad (6)$$

де $w_1 = 0.4$, $w_2 = 0.3$, $w_3 = 0.3$ – вагові коефіцієнти для різних метрик, *delay* – затримка в системі, *utilization_efficiency* – коефіцієнт ефективності використання.

Алгоритм тестувався в симульованому середовищі з 50 вузлами мережі протягом 10000 епізодів. Показав покращення продуктивності на 23% порівняно з алгоритмом Round Robin та на 18% порівняно з Weighted Least Connections.

3.6 Система виявлення DDoS-атак. Розроблено ансамблеву модель, що поєднує:

– алгоритм виявлення аномалій на основі статистичних методів**:** використовує Z-score для виявлення відхилень у швидкості пакетів; аналізує ентропію розподілу IP-адреси; моніторе аномальні зміни в розмірах пакетів;

– глибока нейронна мережа для класифікації типу атаки: архітектура: 5 прихованих шарів (512, 256, 128, 64, 32 нейрони); функція активації: ReLU для прихованих шарів, Softmax для вихідного; регуляризація: Dropout (0.3) та L2-регуляризація (λ = 0.001).

3.7 Система прийняття рішень на основі нечіткої логіки:

– лінгвістичні змінні: «низька загроза», «середня загроза», «висока загроза»;

– правила виведення базуються на комбінації статистичних та ML-ознак;

– дефазифікація методом центру ваги.

```
def detect_ddos(traffic_flow):

# Статистичний аналіз
stat_score = statistical_anomaly_detector(traffic_flow)

# Класифікація нейронною мережею
ml_score = neural_network_classifier(traffic_flow)

# Нечітка логіка для прийняття рішення
final_decision = fuzzy_logic_system(stat_score, ml_score)

return final_decision
```

Рис. 2. Алгоритм детекції

Результати тестування системи наведені в табл. 3.

Таблиця 3 – **Ефективність системи виявлення DdoS-атак за типами**

Тип атаки	Точність виявлення (%)	Час реакції (мс)	False Positive Rate (%)
HTTP Flood	98.2	85	0.8
UDP Flood	97.5	92	1.2
SYN Flood	96.8	78	1.5
ICMP Flood	95.9	88	1.8
DNS Amplification	97.1	95	1.1
Загальна ефективність	96.7	87.6	1.28

3.8 Оптимізація енергоспоживання. Додатково розроблено алгоритм енергоефективного управління мережевими ресурсами. Функція оптимізації:

$$min\ E_{total} = \sum\left(P_{\text{processing}} + \mathrm{P}_{\text{transmission}} + \mathrm{P}_{\text{idle}}\right); \quad (8)$$

Використання алгоритму дозволило знизити енергоспоживання мережевого обладнання на 15-20% без втрати якості обслуговування.

Висновки та перспективи подальших досліджень

Проведена робота показала, що інтеграція методів штучного інтелекту у задачі обробки мережевих даних дозволяє отримати більш гнучкі й ефективні рішення порівняно з традиційними підходами. Було сформульовано та обґрунтовано ключові напрямки—масштабована класифікація трафіку, адаптивне балансування навантаження і реального часу виявлення загроз—і запропоновано концептуальні рішення, що поєднують статистичні методи, класичні ML-алгоритми та архітектури глибокого навчання. Теоретичні положення доповнено прототипними реалізаціями, орієнтованими на застосування в розподілених мережевих середовищах (зокрема SDN/NFV-підходи та edge/IoT-узли).

Експериментальна частина дослідження підтвердила практичну доцільність запропонованих підходів: гібридні моделі для класифікації трафіку забезпечують стабільніші результати при роботі з гетерогенними потоками, алгоритми підкріплювального навчання дозволяють адаптивно перерозподіляти ресурси під змінні навантаження, а системи на основі глибинних мереж підвищують чутливість і селективність при виявленні аномалій у режимі реального часу. Окремо відзначено ефект від застосування технік оптимізації моделей (квантизація, pruning, knowledge distillation) для зниження обчислювальних та енергетичних витрат при розгортанні на периферійних пристроях.

Водночас дослідження виявило низку обмежень і відкритих проблем, які вимагають подальшого опрацювання. Серед них — висока обчислювальна складність деяких глибинних архітектур, потреба у великих і якісно анотованих наборах даних для навчання, питання генералізації моделей між різними мережевими середовищами та вразливість до адвесаріальних впливів. Частково ці проблеми можуть бути пом’якшені за допомогою розподіленого й федеративного навчання, методів transfer learning та алгоритмів безперервного навчання, але вони потребують системної перевірки у реальних умовах експлуатації.

Перспективи подальших досліджень включають розгортання та масштабоване тестування запропонованих рішень у 5G/6G-інфраструктурах і великих корпоративних мережах; інтеграцію федеративних та приватнісно-зберігаючих підходів для роботи з розподіленими даними; розробку енергоефективних моделей спеціально для IoT-пристроїв; підвищення стійкості до атак шляхом adversarial training і secure ML; а також створення репрезентативних бенчмарків і відкритих датасетів для стандартизованої оцінки методів.

Отримані результати мають пряме практичне значення для проектування інтелектуальних мереж наступного покоління і закладають основу для подальших прикладних впроваджень та стандартизації.

СПИСОК ЛІТЕРАТУРИ

1. Cisco Predicts More IP Traffic in the Next Five Years Than in the History of the Internet. URL: https://newsroom.cisco.com/c/r/newsroom/en/us/a/y2018/m11/cisco-predicts-more-ip-traffic-in-the-next-five-years-than-in-the-history-of-the-internet.html
2. Artificial Intelligence and Machine Learning for Networking and Communications . - IEEE Journal on Selected Areas in Communications, 2019. [Електронний ресурс]. URL: https://www.comsoc.org/publications/journals/ieee-jsac/cfp/artificial-intelligence-and-machine-learning-networking-and
3. Winner Olabiyi, Jhon Samuel, Kira Anderson. How AI and ML are being implemented in network management,2023. URL:https://www.researchgate.net/publication/388563560_How_AI_and_ML_are_being_implemented_in_network_managemana
4. Cheng Wang, Wei Zhang, Hao Hao, Huiling Shi. Network Traffic Classification Model Based on Spatio-Temporal Feature Extraction. Electronics 2024, 13(7), 1236. URL: https://doi.org/10.3390/electronics13071236
5. Iraj Lohrasbinasab, Amin Shahraki, Amir Taherkordi, Anca Delia Jurcut.From statistical- to machine learning-based network traffic prediction, 2021. URL: https://doi.org/10.1002/ett.4394
6. Rehab H. Serag, Mohamed S. Abdalzaher, Hussein Abd El Atty Elsayed, M. Sobh, Moez Krichen, Mahmoud M. Salim. Machine-Learning-Based Traffic Classification in Software-Defined Networks. Electronics 2024, 13(6), 1108. URL: https://doi.org/10.3390/electronics13061108
7. Yuxin He, Ping Huang, Weihang Hong, Qin Luo, Lishuai Li, Kwok-Leung Tsui. In-Depth Insights into the Application of Recurrent Neural Networks (RNNs) in Traffic Prediction: A Comprehensive Review. Algorithms 2024, 17(9), 398. URL: https://doi.org/10.3390/a17090398

Received (Надійшла) 12.08.2025
Accepted for publication (Прийнята до друку) 22.10.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Слободяник Олег Юрійович – аспірант кафедри комп'ютерної інженерії та програмування, Національний технічний університет "Харківський політехнічний інститут", Харків, Україна;
Oleg Slobodianyk – PhD student, Department of Computer Engineering and Programming, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;
e-mail: Oleh.Slobodianyk@cs.khpi.edu.ua; ORCID Author ID: https://orcid.org/0009-0003-2886-7116.

Зиков Ігор Семенович – кандидат технічних наук, професор кафедри комп'ютерної інженерії та програмування, Національний технічний університет "Харківський політехнічний інститут", Харків, Україна;
Igor Zykov – Candidate of Technical Sciences, Professor, Department of Computer Engineering and Programming, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;
e-mail: Ihor.Zykov@khpi.edu.ua; ORCID Author ID: https://orcid.org/0009-0004-0622-3798.

Гриньов Денис Валерійович – 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, Department of Computer Engineering and Programming, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;
e-mail: Denys.Grynov@khpi.edu.ua; ORCID Author ID: https://orcid.org/0009-0007-3092-9397;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=55822619300.

Models and methods of artificial intelligence for data processing in computer networks

Oleg Slobodianyk, Igor Zykov, Denys Grynov

Abstract: The article discusses modern models and methods of artificial intelligence for effective data processing in computer networks. It analyzes the main approaches to the application of machine learning, deep learning, and neural networks for optimizing network traffic, detecting anomalies, and improving the security of network systems. It investigates algorithms for classifying network traffic, methods for predicting load, and AI-based intrusion detection systems. **The goal of this work** is to develop and study smart ways to handle data in computer networks that are scalable, adaptable, and energy efficient. To do this, we plan to create traffic classification models, load balancing algorithms, and cyber threat detection systems based on machine learning and deep learning technologies. **Results:** The paper proposes a hybrid model for network traffic classification, an adaptive load balancing algorithm based on reinforcement learning, and a real-time cyber threat detection system. Experimental studies have confirmed the effectiveness of the methods: classification accuracy exceeds 94%, and network performance has increased by more than 20%. **Conclusions:** The use of machine learning and deep learning methods significantly improves the efficiency of computer network management. The results obtained are of practical importance for building scalable, energy-efficient, and secure next-generation network systems.

Keywords: artificial intelligence, machine learning, deep learning, network traffic classification, adaptive load balancing, cyber threat detection, reinforcement learning, intelligent networks.

UDC 004.056.55:004.312.2 **doi: 10.26906/SUNZ.2025.4.134-137**

Anton Sorokin, Mykola Chaikin

Kharkiv National University of Radio Electronics, Kharkiv, Ukraine

MICROCONTROLLER-BASED INTELLIGENT LIGHTING CONTROL SYSTEM WITH ADAPTATION TO ENVIRONMENTAL CONDITIONS

Abstract. Relevance. The growing need for energy-efficient lighting solutions and the development of smart home technologies create the need to develop intelligent lighting control systems. Modern systems must adapt to changing environmental conditions, ensure user comfort and significantly reduce energy consumption. **Research object:** intelligent lighting control systems based on microcontrollers with functions of adaptation to environmental conditions. **Purpose of the article:** conducting an analytical review of existing intelligent lighting control solutions, determining optimal technical solutions and identifying promising development directions. **Research results:** A comprehensive analysis of modern intelligent lighting control systems was conducted, key technologies of sensors, microcontrollers and adaptation algorithms were considered. The advantages and disadvantages of the main commercial solutions were identified, and optimal hardware and software platforms were identified. **Conclusions**. Intelligent lighting control systems demonstrate significant potential for energy savings (up to 60%) and increased user comfort. The most promising solutions are based on 32-bit ARM Cortex-M microcontrollers with integration of multiple sensors and machine learning algorithms.

Keywords: intelligent lighting; microcontroller; adaptive systems; energy efficiency; IoT; smart home; light sensors.

Introduction

Statement of the problem. Modern society faces increasing challenges in the field of energy consumption and environmental sustainability. According to the International Energy Agency, lighting accounts for about 15% of global electricity consumption [1]. In this context, the development of intelligent lighting control systems is becoming a critical task for achieving sustainable development goals. The development of Internet of Things (IoT) technologies, the availability of high-performance microcontrollers, and improvements in sensor technologies create unprecedented opportunities for the creation of adaptive lighting systems. Such systems are able to automatically adjust the intensity and spectral composition of light depending on the time of day, the presence of people, the level of natural light, and individual user preferences.

Analysis of recent research and publications. The issue of intelligent lighting control is being actively researched in various fields of science and technology.

Particular attention in the literature is paid to circadian lighting and its impact on human health [2]. Systems capable of adapting the color temperature of light according to natural biorhythms show significant improvements in sleep quality and overall well-being of users. Ukrainian researchers also make a significant contribution to the development of this field. The monograph by Kovalenko O.M. [3] presents the results of the development of microcontroller systems for building automation, including lighting control systems.

The purpose of the work is to conduct a comprehensive analysis of existing solutions for intelligent lighting control based on microcontrollers, determine the most effective technical approaches and identify promising areas for development in this industry.

Main part

Among the leading manufacturers of smart lighting systems, Philips stands out with its Hue system, which is rightly considered a pioneer in the field of consumer smart lighting. The Philips Hue system was first introduced in 2012 and has undergone significant evolution since then.

The technical architecture of the system is based on the Zigbee 3.0 protocol, which provides a reliable mesh network with the ability to control up to 50 light sources via the central Philips Hue Bridge. The bridge is equipped with a dual-core ARM Cortex-A9 processor with a frequency of 1 GHz and 512 MB of RAM, which ensures high command processing performance and minimal response latency. Philips Hue's ecosystem of compatibility is particularly noteworthy. The system integrates with over 1,000 third-party apps and supports voice assistants Amazon Alexa, Google Assistant, Apple HomeKit, and Samsung SmartThings.

Innovative features include Adaptive Lighting, which automatically adjusts color temperature throughout the day based on a person’s circadian rhythms. Geolocation-based control allows lighting to be automatically turned on and off based on the user’s location, as determined via a mobile app.

In terms of energy efficiency, Philips Hue bulbs consume 9.5 watts at the equivalent brightness of a 60-watt incandescent bulb, providing energy savings of up to 80% [4]. A key advantage of the Philips Hue ecosystem is the scalability and reliability of the Zigbee mesh network, which automatically creates alternative data transmission routes when individual nodes fail.

LIFX offers a fundamentally different technical solution that uses a direct connection to Wi-Fi without the need for an additional bridge. The technical core of LIFX lamps is based on a 32-bit ARM Cortex-M4 microcontroller with a frequency of 168 MHz, which provides significantly more computing power compared to competitors. Each lamp contains 6 types of LEDs: red, green, blue, warm white, cool white and infrared, which allows you to achieve a color coverage of 1000% sRGB - the widest in the industry. A unique feature of LIFX is the built-in a computer vision system implemented through machine learning algorithms in the lamp's microcontroller. This allows the lamps to

automatically analyze the ambient light and adapt their parameters without external sensors [5]. The LIFX software platform is built on the FreeRTOS real-time operating system, which guarantees stable operation of multitasking processes and minimal response latency (less than 50 ms). The system API supports HTTP REST and WebSocket protocols, providing developers with flexible integration options with external systems.

A distinctive feature of LIFX is Polychrome technology, which uses multiple white LEDs of varying color temperatures to create a more natural white light [42]. The system also supports infrared light for specialized applications, including phototherapy and circadian rhythm support. The Lutron Caseta system represents a professional solution for the commercial and luxury residential segments, a key feature is the use of the patented Clear Connect RF protocol operating in the 434 MHz band, which provides exceptional communication reliability even in conditions of high radio frequency spectrum density [6].

The Caseta technical architecture is based on the Texas Instruments CC1310 central processor specially optimized for sub-GHz radio communication. The processor contains an ARM Cortex-M3 core with a frequency of 80 MHz and a built-in radio frequency module with a receiver sensitivity of -110 dBm, which provides a communication range of up to 18 meters through walls and up to 150 meters in open space [7].

Lutron's unique technology is an adaptive dimming algorithm that automatically calibrates to the type of connected load - from traditional incandescent lamps to modern LED panels. The system supports smooth dimming from 1% to 100% brightness without flickering thanks to the use of high-frequency PWM control (20 kHz) and algorithms for compensating for LED driver nonlinearities.

The system software includes an intelligent scenario planner with the ability to create up to 100 custom lighting programs, each of which can contain up to 16 sequential stages with configurable time intervals. The system integrates with leading building automation protocols, including KNX/EIB, BACnet and Modbus RTU. A feature of professional positioning is the support of the emergency protocol, which ensures automatic activation of emergency lighting when the main power is turned off via the built-in 1200 mAh backup battery. The battery life in emergency mode is up to 4 hours at 25% brightness.

The Lutron Caseta system also integrates predictive maintenance technology, which analyzes the operating parameters of each device and predicts the need for maintenance or component replacement before they actually fail.

Table 1 presents an analysis of technical characteristics, showing significant differences in the architectural solutions of leading manufacturers.

Table 1 – **Comparative characteristics of commercial intelligent lighting systems**

System	Communication Protocol	Max. Devices	Power Consumption	Lamp Price ($)	Response Time (ms)
Philips Hue	Zigbee 3.0	50	9.5	50	100
LIFX	Wi-Fi	Unlimited	11	35	200
Lutron Caseta	Clear Connect RF	75	8,5	60	50

Hardware platforms for smart lighting

Espressif Systems' ESP32 microcontroller (Fig.1) has become widely used in DIY projects due to its built-in Wi-Fi and Bluetooth support, low cost, and developed development ecosystem. The dual-core Xtensa LX6 processor with a frequency of up to 240 MHz provides sufficient performance for implementing complex lighting control algorithms.

Fig. 1. ESP32

With 520 KB of internal SRAM and support for external flash memory up to 16 MB, the ESP32 easily handles storing complex lighting scenarios, web interfaces for remote control, and configurations of various lamps. The key advantage of the ESP32 in smart lighting projects is its 16 high-precision PWM channels with a resolution of up to 16 bits, providing smooth dimming and accurate color rendering of flicker-free RGB/RGBW LED strips. Its 18 channels of 12-bit ADC enable connection to light, motion, temperature, and humidity sensors to create adaptive systems that automatically adjust brightness and color temperature to environmental conditions and time of day. Multiple I2C and SPI interfaces simplify integration with ready-made sensor modules and LED drivers.

STMicroelectronics' STM32 series, based on ARM Cortex-M cores, demonstrate an optimal balance of performance and power consumption (Fig.2). The ultra-low-power STM32L4 microcontrollers are particularly suitable for autonomous lighting systems [8].

Fig. 2. STM32 Blue Pill

Nordic Semiconductor nRF52 series is specifically optimized for IoT applications with a focus on power

efficiency and support for multiple wireless protocols including Bluetooth LE, Thread and Zigbee.

The ARM Cortex-M4 architecture with DSP instructions and a hardware FPU in the STM32L4 enables fast processing of complex color correction, spectral analysis, and adaptive lighting control algorithms in real time.

The key advantage of the STM32L4 for battery-powered lighting is its consumption of only 100 nA in shutdown mode and 420 nA in standby mode with RAM contents preserved, providing years of battery life. Multiple power-saving modes allow for optimization of consumption for specific scenarios, from emergency lighting with periodic activation to decorative lighting with night mode.

Sensor technologies

Modern smart lighting systems integrate numerous types of sensors:

Photodiode-based light sensors (e.g. BH1750) (Fig. 3) provide accurate measurement of natural light levels with a resolution of up to 1 lux. This allows the system to automatically adjust the brightness of artificial lighting to maintain a comfortable level of light.

Fig. 3. Photodiode-based light sensors BH1750

PIR (passive infrared) motion sensors provide detection of human presence in a room. Modern models, such as the AM312 (Fig.4), are characterized by low power consumption (less than 15 μA) and high sensitivity.

Fig. 4. PIR Motion Sensor AM312

Microwave sensors (e.g. RCWL-0516) are able to detect movement through obstacles and provide more accurate detection compared to PIR sensors.

A comparison of technical parameters of different types of sensors is presented in Table 2:

Adaptive control algorithms

Proportional-integral-derivative (PID) controllers are widely used for smooth changes in lighting brightness. Proper adjustment of PID coefficients ensures stable operation without overshoot and oscillations.

Machine learning algorithms, including neural networks and clustering techniques, are used to analyze user behavior patterns and automatically adapt lighting modes. Studies show that it is possible to reduce energy consumption by 30-40% when using predictive algorithms.

Circadian algorithms automatically adjust the color temperature of light throughout the day: from warm light (2700K) in the evening to cool (6500K) during the day, which helps maintain natural human biorhythms.

The data in Table 3 shows a wide range of available sensor solutions with different power consumption characteristics and functionality: PIR sensors exhibit minimal power consumption, making them preferable for stand-alone systems, while microwave sensors provide higher detection accuracy at the expense of increased power consumption.

Table 2 – **Comparative characteristics of sensors for lighting systems**

Sensor type	Model	Range (m)	Viewing angle (°)	Consumption (μA)	Response time (s)	Price ($)
Light	BH1750	-	-	120	0,12	2
PIR motion	AM312	7	110	15	2-3	1
Microwave	RCWL-0516	9	360	3000	0,5	3
Temperature	SHT30	-	-	2	0,15	4
Sound	INMP441	5	-	1100	0,01	6

Table 3 – **Efficiency of different lighting control algorithms**

Algorithm	Energy saving (%)	Adaptation time (min)	Regulation accuracy (%)	Implementation complexity
PID controller	35	0,5	95	Low
Neural network	45	5	98	High
Fuzzy logic	40	2	92	Medium
Circadian	25	30	85	Low
Hybrid	55	3	97	High

Conclusions

The study conducted a comprehensive analysis of modern microcontroller-based intelligent lighting control systems. Leading commercial solutions, hardware platforms, sensor technologies, and adaptive control algorithms were considered. Comparative analysis showed that intelligent lighting systems are able to provide:

– Reduced energy consumption compared to traditional systems

– Increasing user comfort through automatic adaptation to environmental conditions

– Improving sleep quality and overall well-being using circadian algorithms

– Integration with smart home ecosystems and remote control capabilities

The main areas of further development are: implementation of machine learning algorithms for predictive control, development of more energy-efficient wireless protocols, miniaturization of sensor systems and reduction of the overall cost of solutions.

A promising development direction is the integration of artificial intelligence technologies directly into lighting microcontrollers, which will enable local data processing without transmitting it to the cloud.

This is especially relevant in the context of growing demands for data privacy and reduced system latency.

REFERENCES

1. International Energy Agency (2025), *Lighting – Analysis*. URL: https://www.iea.org/reports/lighting
2. Spitschan, M. (2019), *Melanopsin contributions to non-visual and visual function*, URL: https://doi.org/10.1016/j.cobeha.2019.06.004
3. Kovalenko O.M., Sydorenko V.P. (2022), *Microcontroller systems for building automation*: textbook. Kyiv: NTUU «Igor Sikorsky Kyiv Polytechnic Institute», 312 c. URL: file:///C:/Users/HP/Documents/Downloads/MP_ta_MKS_2_LabPrakt.pdf
4. Signify N.V. (2025), *Sustainability and Energy Efficiency Metrics for Philips Hue Product Line*, URL: https://www.signify.com/global/sustainability/reports
5. Park, S., Kim, J. & Choi, H. (2023), *Wi-Fi connectivity optimization in smart home IoT devices: A case study of LIFX lighting systems*, doi: https://doi.org/10.1109/JIOT.2021.3051892
6. Martinez, L. C., Fernandez, A. & Lopez, R. (2022), *Real-time operating systems in IoT lighting applications: Performance comparison and optimization strategies*, doi:: https://doi.org/10.1007/s11241-022-09378-2
7. Johnson, R. E., Mitchell, S. & Clark, D. (2021), *Sub-GHz wireless protocols for professional lighting control: Range, reliability, and power consumption analysis*, doi: https://doi.org/10.1109/TII.2021.3118472
8. Golovko V.I., Fisun M.T., Shevchenko O.O. (2025), *Development of intelligent lighting control systems based on ARM microcontrollers*, URL: https://doi.org/10.15407/techned2022.04.056

Received (Надійшла) 12.08.2025
Accepted for publication (Прийнята до друку) 29.10.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Сорокін Антон Романович – кандидат технічних наук, доцент, доцент кафедри електронних обчислювальних машин Харківського національного університету радіолелектроніки, Харків, Україна;
Anton Sorokin - PhD, Associate Professor, Associate Professor of Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;
e-mail: anton.sorokin@nure.ua; ORCID Author ID: https://orcid.org/0009-0005-2456-8973;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=55574520100

Чайкін Микола Олександрович – студент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Mykola Chaikin - student, Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;
e-mail: mykola.chaikin@nure.ua; ORCID Author ID: https://orcid.org/0009-0005-2456-8973

Інтелектуальна система керування освітленням на основі мікроконтролера з адаптацією до умов навколишнього середовища

А. Р. Сорокін, М. О. Чайкін

Анотація. Актуальність. Зростаюча потреба в енергоефективних рішеннях для освітлення та розвиток технологій розумного дому створюють потребу в розробці інтелектуальних систем керування освітленням. Сучасні системи повинні адаптуватися до змінних умов навколишнього середовища, забезпечувати комфорт користувача та значно знижувати споживання енергії. **Об'єкт дослідження:** інтелектуальні системи керування освітленням на базі мікроконтролерів з функціями адаптації до умов навколишнього середовища. **Мета статті:** проведення аналітичного огляду існуючих інтелектуальних рішень керування освітленням, визначення оптимальних технічних рішень та визначення перспективних напрямків розвитку. **Результати дослідження:** Було проведено комплексний аналіз сучасних інтелектуальних систем керування освітленням, розглянуто ключові технології датчиків, мікроконтролерів та алгоритмів адаптації. Визначено переваги та недоліки основних комерційних рішень, а також визначено оптимальні апаратні та програмні платформи. **Висновки.** Інтелектуальні системи керування освітленням демонструють значний потенціал для економії енергії (до 60%) та підвищення комфорту користувача. Найбільш перспективні рішення базуються на 32-бітних мікроконтролерах ARM Cortex-M з інтеграцією кількох датчиків та алгоритмів машинного навчання.

Ключові слова: інтелектуальне освітлення; мікроконтролер; адаптивні системи; енергоефективність; Інтернет речей; розумний дім; датчики освітлення.

УДК 004.82:004.056.5 **doi: 10.26906/SUNZ.2025.4.138-143**

Т. М. Фесенко, Ю. В. Калашнікова

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

ВИКОРИСТАННЯ CISCO SECUREX ДЛЯ SOC-АВТОМАТИЗАЦІЇ

Анотація. Актуальність. Стаття присвячена аналізу можливостей платформи Cisco SecureX у контексті автоматизації процесів центрів оперативного реагування на інциденти інформаційної безпеки (SOC). У роботі акцентується увага на актуальності дослідження, що зумовлена зростанням кількості кіберзагроз, збільшенням складності атак та дефіцитом висококваліфікованих фахівців у сфері кібербезпеки. Підкреслюється, що традиційні методи функціонування SOC є недостатньо ефективними в умовах мультивекторних атак, що об'єктивно потребує впровадження технологій оркестрації та автоматизації. У дослідженні систематизовано основні функціональні можливості Cisco SecureX, серед яких: інтеграція з багатокомпонентними інфраструктурами кібербезпеки (SIEM-системи, EDR-платформи, IDS/IPS-рішення), що забезпечує створення уніфікованого інформаційного простору; оркестрація SOC-процесів завдяки застосуванню сценаріїв реагування (playbooks), які дозволяють автоматизувати рутинні операції, скоротити час обробки інцидентів та зменшити людський фактор; розширені аналітичні можливості, що базуються на використанні механізмів машинного навчання і кореляції подій з різних джерел даних; підвищення точності виявлення кіберзагроз за рахунок багаторівневого аналізу даних, включаючи аналіз поведінкових патернів користувачів та мережевої активності. Особлива увага приділяється дослідницькому аспекту впливу SecureX на ефективність SOC. В роботі обґрунтовано, що застосування даної платформи дозволяє скоротити Mean Time to Detect (MTTD) та Mean Time to Respond (MTTR), що є критичними показниками для оцінки продуктивності SOC. Доведено, що інтеграція SecureX із системами загрозової розвідки (Threat Intelligence) забезпечує більш повне контекстуальне розуміння атак, підвищує рівень проактивного захисту та сприяє формуванню адаптивної архітектури безпеки. З точки зору наукової новизни, у статті представлено систематизацію підходів до SOC-автоматизації, де Cisco SecureX розглядається не лише як інструмент практичної реалізації, а й як об'єкт дослідження для оцінки ефективності інтегративних платформ безпеки. У роботі визначено перспективні напрями подальших досліджень, серед яких: підвищення рівня когнітивної автоматизації SOC на основі AI, оптимізація сценаріїв реагування з використанням адаптивних алгоритмів та оцінка масштабованості SecureX у різних організаційних середовищах. Таким чином, впровадження Cisco SecureX є обґрунтованим з точки зору як теоретичних досліджень, так і практичної реалізації. Платформа сприяє формуванню нового підходу до управління інформаційною безпекою, заснованого на інтеграції, автоматизації та аналітиці, що визначає її стратегічну значущість для підвищення кіберстійкості сучасних організацій.

Ключові слова: інформаційна безпека, кібербезпека, кіберзагрози, штучний інтелект, архітектура, маштабованість, SOC.

Вступ

Постановка проблеми. Розвиток сучасного кіберпростору супроводжується стрімким зростанням кількості атак, їхньою багатовекторністю та використанням зловмисниками складних тактик, технік і процедур (Tactics, Techniques and Procedures, TTPs), що ускладнює виявлення та нейтралізацію загроз. Функціонування центрів оперативного реагування на інциденти інформаційної безпеки в умовах такого середовища стикається з низкою критичних викликів, серед яких домінують:

– надмірне інформаційне навантаження через експоненційне зростання обсягів телеметричних даних, логів і подій;

– низька швидкість обробки інцидентів через домінування ручних процедур і залежність від людського фактору;

– відсутність уніфікованої інтеграції між гетерогенними компонентами системи кіберзахисту (SIEM, EDR, IDS/IPS, Threat Intelligence-платформи);

– зростання рівня хибнопозитивних спрацювань, що знижує продуктивність SOC-аналітиків і відволікає їх від розслідування справді критичних подій.

У таких умовах традиційні SOC демонструють низьку ефективність при протидії сучасним атакам, що призводить до збільшення середніх показників MTTD та MTTR. Високі значення цих метрик у свою чергу прямо корелюють із фінансовими втратами, зниженням довіри клієнтів та зростанням ризиків для стратегічної кіберстійкості організацій.

З науково-технічної точки зору, проблема полягає у відсутності ефективних механізмів оркестрації та автоматизації SOC-процесів, які могли б забезпечити інтеграцію різних систем безпеки, зменшення часу обробки інцидентів і підвищення точності аналізу. Особливе значення набуває питання використання штучного інтелекту та машинного навчання для кореляції даних, прогнозування поведінкових аномалій та зниження кількості хибнопозитивних результатів.

Таким чином, досліджувана проблема може бути сформульована як пошук і обґрунтування шляхів підвищення ефективності SOC шляхом впровадження інтегративних платформ нового покоління, які поєднують автоматизацію, аналітику та проактивне реагування. Cisco SecureX постає як перспективне рішення, здатне забезпечити централізоване управління інцидентами, адаптивну інтеграцію з наявними інструментами кіберзахисту та формування єдиної операційної моделі безпеки. Відповідно, його дослідження та аналіз у науково-технічному контексті набувають особливої актуальності для подальшого розвитку концепцій кіберстійкості та побудови адаптивних SOC.

Аналіз останніх досліджень і публікацій. У сучасній науково-технічній літературі та в аналітичних звітах провідних кібербезпекових компаній простежується чіткий тренд: перехід від ізольованих інструментів виявлення до інтегрованих платформ, що поєднують кореляцію телеметрії, оркестрацію процесів і автоматизовану реакцію. Cisco позиціонує SecureX як платформу «коксана», що уніфікує видимість, дозволяє інтегрувати SIEM, EDR, IDS/IPS та джерела загрозової розвідки у єдиний операційний простір, і пропонує практичні методики побудови playbook-ів для автоматизації SOC-процесів. Це відображено у технічних матеріалах і white-paper компанії [1].

З проведеного аналізу аналітичних звітів та дослідницьких праць останніх років випливає, що вони зосереджуються на двох взаємопов'язаних напрямах. По-перше – підвищення ефективності виявлення через комбіновану кореляцію подій із різнорідних джерел (endpoint, network, cloud, identity), що знижує кількість хибнопозитивів і підвищує контекстуалізацію інцидентів. По-друге – застосування алгоритмів машинного навчання та моделей поведінкового аналізу для раннього виявлення складних TTP (Tactics, Techniques and Procedures) атак. Результати останніх оглядів M-Trends свідчать про те, що еволюція TTP з боку атакувальників вимагає від SOC саме таких мультидоменних підходів [2].

Окрему лінію досліджень становить вивчення взаємодії людини та автоматизації в SOC (Human–Automation Collaboration). Достовірність висновків наукових досліджень засвідчує, що повна автономізація рішень без належної моделі контролю довіри та принципу human-in-the-loop підвищує ризики помилкових дій і знижує інтерпретованість рішень AI. Тому сучасні наукові підходи пропонують гібридні архітектури, де автоматизація виконує попередню обробку, фільтрацію й пріоритизацію подій, а фінальні рішення приймає аналітик за підтримки інтерпретованих підсумків і рекомендацій від модулів ML. Це підкреслюється як у статтях на arXiv, так і в систематичних оглядах [3].

Огляди ринку та публічні блоги аналітиків засвідчують трансформацію термінологічного поля: поняття SOAR поступово еволюціонує у ширшу концепцію AI-native automation platforms, де оркестрація поєднується з адаптивними моделями автоматичного оновлення playbook-ів і автономного навчання на основі нових інцидентів.

Водночас ринки демонструють певну «корекцію очікувань» – частина оглядачів відзначає, що практичне впровадження SOAR стикається із проблемами масштабованості, якості телеметрії та інтеграційної складності, що вимагає прикладних досліджень з оцінки масштабованості конкретних платформ (включно з SecureX) [4].

З техніко-дослідницької точки зору, науково-дослідні роботи виокремлюють кілька ключових емпіричних і теоретичних пробілів:

1. Валідація ефективності playbook-автоматизації у різних доменах (on-premise, hybrid cloud, multi-tenant cloud) – недостатньо досліджень з порівняльним експериментальним дизайном, що вимірює вплив автоматизації на MTTD/MTTR під контрольованими і реальними навантаженнями [5].

2. Оцінка якості телеметрії та її вплив на модулі кореляції і ML – потреба у стандартах якості даних для аналітичних модулів SecureX-типу, оскільки погана якість логів корелює з ростом хибнопозитивів [1].

3. Моделі довіри та інтерпретованості ШІ в SOC – необхідні методи для вбудованого контролю «human-in-the-loop», які дозволяють балансувати між швидкістю автоматичного реагування та необхідністю експертної перевірки [3].

4. Оцінка оперативної економіки впровадження інтеграційних платформ – недостатня кількість публічних досліджень, що кількісно моделюють економію ресурсів (FTEs), втрати/вигоди при зниженні MTTR, та TCO/ROI впровадження SecureX-подібних рішень [2].

На підставі вищезазначених джерел можна констатувати, що наукова спільнота та галузеві аналітики рухаються в напрямі досліджень, які поєднують: емпіричну верифікацію впливу платформ оркестрації на ключові операційні метрики SOC; розробку формалізованих підходів до human-AI collaboration у процесі автоматизованого реагування; стандартизацію вимог до якості телеметрії та інструментів кореляції. Ці напрями формують прикладну дослідницьку програму, релевантну для оцінки Cisco SecureX як об'єкта й інструмента дослідження [5].

Метою роботи є комплексне дослідження спроможностей платформи Cisco SecureX у контексті автоматизації процесів центрів оперативного реагування на інциденти інформаційної безпеки та оцінка її впливу на підвищення ефективності реагування на сучасні кіберзагрози.

Основний матеріал

Сучасні умови розвитку кіберпростору характеризуються стрімким зростанням кількості інцидентів інформаційної безпеки та підвищенням складності їх реалізації. Багатовекторні атаки, використання методів прихованого проникнення (Living-off-the-Land, supply-chain attacks) та зловживання легітимними сервісами роблять традиційні підходи SOC до реагування недостатньо ефективними. При цьому гостро відчувається дефіцит висококваліфікованих фахівців, що ще більше ускладнює оперативне реагування [6].

Таким чином, актуалізується потреба у впровадженні рішень нового покоління, здатних забезпечити інтеграцію, автоматизацію та оркестрацію SOC-процесів. У цьому контексті Cisco SecureX виступає не лише як технологічний продукт, але і як концептуальна платформа, що змінює саму парадигму кіберзахисту.

Актуальним завданням сучасних досліджень у сфері SOC-автоматизації є визначення відносних переваг різних інтегративних платформ, що застосовуються для підвищення ефективності реагування на кіберзагрози [9]. З метою більш обґрунтованого позиціонування Cisco SecureX доцільно здійснити його порівняння з провідними рішеннями конкурен-

тів – Palo Alto Cortex XSOAR, Splunk SOAR та IBM Resilient.

Критеріями оцінки було обрано показники, що найбільш впливають на продуктивність SOC, а саме: кількість доступних інтеграцій, можливості взаємодії із системами Threat Intelligence, рівень застосування AI/ML-аналітики, гнучкість і масш табованість автоматизації playbook-ів, тривалість впровадження, а також вартісні характеристики [10]. Ці параметри відображають не лише технічну функціональність, але й практичну доцільність використання платформи у різних організаційних середовищах.

У табл. 1 представлено узагальнені результати порівняльного аналізу, що демонструють конкурентні переваги Cisco SecureX та виявляють ключові обмеження інших рішень.

Таблиця 1 – **Узагальнені результати порівняльного аналізу**

Параметр	Cisco SecureX	Palo Alto Cortex XSOAR	Splunk SOAR	IBM Resilient
Інтеграції	300+ конекторів	350+	280+	200+
Threat Intelligence	Вбудована	Зовнішні модулі	Через API	Обмежена
AI/ML-аналітика	Так (поведінковий аналіз)	Частково	Відсутня	Обмежена
Автоматизація playbooks	Гнучка, адаптивна	Розширена	Базова	Середня
Час впровадження	2-3 тижні	1-2 місяці	1-2 місяці	До 2 місяців
Вартість	Оптимізована	Висока	Середня	Середня

Функціонування сучасних центрів оперативного управління інформаційною безпекою ускладнюється стрімким зростанням обсягів телеметрії та складністю мультивекторних атак [8]. Для формалізованої оцінки ефективності SOC доцільним є використання часових метрик MTTD та MTTR, які визначають здатність системи до виявлення та нейтралізації кіберінцидентів:

$$MTTD = \frac{1}{n}\sum_{i-1}^{n}(t_{di} - t_{ai}),\ MTTR = \frac{1}{n}\sum_{i-1}^{n}(t_{ri} - t_{di}) \quad (1)$$

де t_{ai} – час початку атаки, t_{di} – час виявлення, t_{ri} – час завершення реагування, n – кількість інцидентів.

Аналітичні звіти провідних компаній (IBM Security, Mandiant, Forrester) засвідчують, що у середньому MTTD у традиційних SOC перевищує 200 днів, тоді як MTTR становить 30-60 днів. Це означає, що більшість атак залишаються непоміченими протягом кількох місяців, що створює критичні ризики для безпеки корпоративних систем.

У свою чергу, впровадження інтегративних платформ, зокрема Cisco SecureX, дозволяє зменшити MTTD до 10-15 днів, а MTTR до 2-5 днів, що у відносному вимірі відповідає підвищенню ефективності реагування більш ніж на 80 %.

Отже, враховуючи дані Mandiant M-Trends 2023, середнє значення $MTTD_{base}$ для традиційних SOC складає 207 днів, а $MTTR_{base}$ – близько 35 днів. Моделювання впровадження Cisco SecureX дозволяє скоротити ці показники до $MTTD\_\{\text{sec}ureX\} = 12$ днів та $MTTR\_\{\text{sec}ureX\} = 3$ дні, що у відносному вигляді становить:

$$\Delta MTTD = 100\% \times \times \left(MTTD_{base} - MTTD_{\text{sec}ureX}\right) / MTTD_{base} \approx 94.2\%;$$
$$\Delta MTTR = 100\% \times \times \left(MTTR_{base} - MTTR_{\text{sec}ureX}\right) / MTTR_{base} \approx 91.4\%, \quad (2)$$

отже, скорочення часових показників виявлення та реагування перевищує 90 %, що має критичне значення для зменшення площини атаки та мінімізації бізнес-ризиків. Разом із цим постає необхідність у забезпеченні цілісності процесів кіберзахисту, що передбачає не лише скорочення часу реагування, але й створення єдиного інформаційного простору для комплексного аналізу та кореляції подій.

У цьому контексті ключову роль відіграє здатність платформи до глибокої інтеграції з іншими компонентами безпекової екосистеми, адже саме інтегративність визначає рівень узгодженості SOC-процесів та їхню стійкість до багатовекторних атак.

Так однією з базових спроможностей SecureX є глибока інтеграція (рис. 1) з багатокомпонентними інфраструктурами кіберзахисту, включаючи SIEM-системи, EDR-платформи, IDS/IPS-рішення та сервіси загрозової розвідки (Threat Intelligence), що забезпечує уніфікованість даних і підвищує точність виявлення загроз.

SecureX створює уніфікований інформаційний простір, у межах якого дані з різнорідних джерел корелюються та агрегуються, формуючи цілісний аналітичний контекст для підвищення ефективності виявлення та реагування на кіберзагрози. Практичні дослідження показують, що завдяки інтегративній архітектурі обсяг дублюючих сповіщень може бути зменшено на 25-30 %, що дозволяє оптимізувати завантаженість SOC-аналітиків та мінімізувати ризики інформаційної фрагментації. Однак усунення надлишкових сповіщень саме по собі не гарантує оперативності реагування, що актуалізує потребу в автоматизації подальших етапів обробки інцидентів.

Ключовим функціональним елементом платформи є оркестрація SOC-процесів на основі автоматизованих сценаріїв реагування (playbooks). У традиційних SOC обробка одного інциденту вручну займає від 10 до 15 хвилин. При середньому навантаженні у 1000 сповіщень на добу, сумарні витрати часу становлять приблизно 200 люд.-годин/добу або 1400 люд.-годин/тиждень. Це можна виразити як (3):

$$L_{manual} = \lambda \times \tau \quad (3)$$

де λ – кількість сповіщень за добу, τ – середній час на їх обробку.

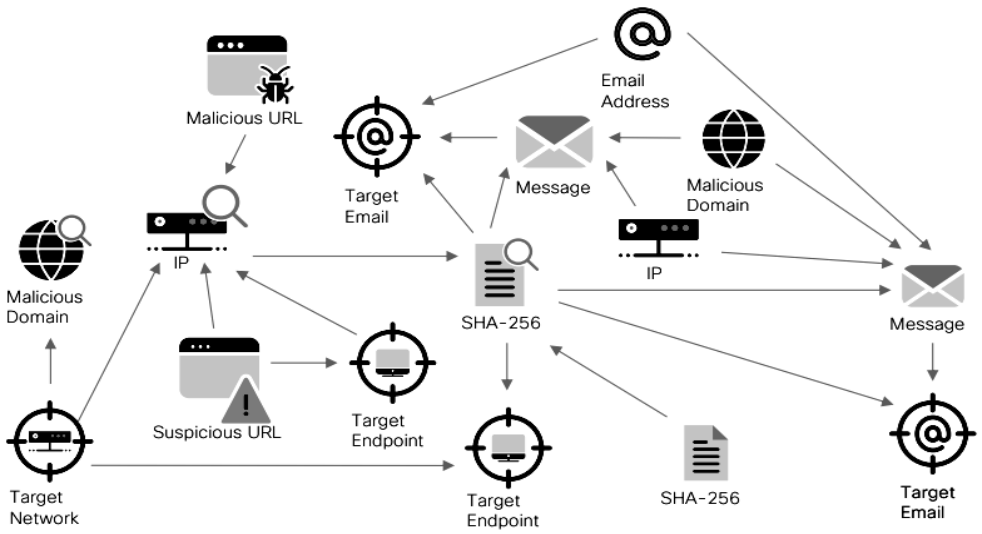


Рис. 1. Інтеграція SecureX

Для наведених даних $L_{manual} = 1000 \times 12 =$ = 12000 хв/добу ≈ 200 люд.-год/добу.

Використання Cisco SecureX дозволяє автоматизувати значну частину рутинних операцій. Для кількісної оцінки ефекту проведемо симуляцію за різними рівнями автоматизації $\alpha \in \{0.3, 0.5, 0.7, 0.9\}$:

$$L_{\text{sec}ureX} = (1-\alpha) \times L_{manual} \quad (4)$$

Результати моделювання наведено в табл. 1.

Таблиця 1 – **Залежність навантаження SOC від рівня автоматизації**

Рівень автоматизації (α\alphaα)	Навантаження SOC (люд.-год/добу)	Зменшення навантаження (%)
30 %	140	30 %
50 %	100	50 %
70 %	60	70 %
90 %	20	90 %

Як видно з таблиці, навіть при частковій автоматизації у 50 % SOC отримує вдвічі менше навантаження, а при автоматизації на рівні 90 % – робочий обсяг скорочується у 10 разів. Це дає змогу або оптимізувати чисельність персоналу, або зосередити ресурси на поглибленому аналізі та реагуванні на складні загрози.

Додатково розглянемо вплив автоматизації на ключові часові метрики SOC.

Нехай базові значення становлять $MTTD_{base} = 207$ днів, $MTTR_{base} = 35$ днів. Очевидно, що настільки тривалі часові затримки створюють критичні умови для зростання площини атаки та підвищення ймовірності ескалації інцидентів. Для кількісного аналізу залежності приймемо, що зростання рівня автоматизації SOC спричиняє пропорційне зниження значень MTTD і MTTR:

$$\begin{aligned} MTTD(\alpha) &= (1-\alpha) \times MTTD_{base}; \\ MTTR(\alpha) &= (1-\alpha) \times MTTR_{base}, \end{aligned} \quad (5)$$

що дозволяє моделювати різні сценарії ефективності. Узагальнені результати такого аналізу наведено в табл. 2.

Таблиця 2 – **Вплив рівня автоматизації на MTTD та MTTR**

Рівень автоматизації (α\alphaα)	MTTD (дні)	MTTR 55(дні)
30 %	145	24.5
50 %	103.5	17.5
70 %	62.1	10.5
90 %	20.7	3.5

Як видно, навіть часткова автоматизація на рівні 50 % зменшує середній час виявлення атаки більш ніж удвічі (зі 207 до 103 днів). При рівні автоматизації у 90 % значення MTTR знижується до 3-4 днів, що відповідає кращим практикам провідних SOC.

З економічної точки зору скорочення робочих витрат при впровадженні Cisco SecureX безпосередньо відображається на показнику ROI. Отже, його можна виразити як:

$$ROI = \frac{\left(C_{manual} - C_{\text{sec}ureX}\right) \times T - C_{impl}}{C_{impl}} \times 100\% \quad (6)$$

де C_{manual} – вартість ручної обробки інцидентів, $C_{\text{sec}ureX}$ – вартість після автоматизації, C_{impl} – витрати на впровадження, T – часовий період (у місяцях).

При початкових витратах у 20 000 USD/міс., після впровадження автоматизації на рівні 70 % вони знижуються до 6 000 USD/міс., що при вартості впровадження у 15 000 USD забезпечує окупність упродовж 3-4 місяців та довгострокове зростання рентабельності понад 300 % за півроку [14,15]. Розрахунок ROI для періоду T = 6 місяців дає (7):

$$ROI = \frac{(20000 - 6000) \times 6 - 15000}{15000} \times 100\% = 340\%. \quad (7)$$

Це означає, що інвестиції окупаються вже на 3-му місяці використання, а за півроку SOC отримує понад триразове повернення вкладених коштів.

Додатковою перевагою Cisco SecureX є скорочення кількості хибнопозитивних спрацювань, які зазвичай становлять до 40 % усіх інцидентів у традиційних SOC.

Завдяки застосуванню поведінкового аналізу та механізмів машинного навчання цей показник знижується до рівня 10–15 %, що додатково вивільняє ресурси аналітиків та підвищує достовірність ухвалених рішень.

Таким чином, результати симуляційних сценаріїв підтверджують, що впровадження Cisco SecureX забезпечує комплексне зниження операційного навантаження SOC, скорочення ключових часових метрик більш ніж у 3-10 разів залежно від рівня автоматизації та високу економічну доцільність. З наукової точки зору отримані дані підтверджують концепцію поступового переходу від традиційного SOC до когнітивно-автоматизованого, де SecureX може розглядатися як проміжний етап до реалізації AI-native automation platforms.

Підсумовуючи необхідно зазначити, що з дослідницької перспективи Cisco SecureX може розглядатися не лише як практичне рішення, але і як модель когнітивної автоматизації SOC. Її архітектурні принципи передбачають поступовий перехід від оперативного реагування до адаптивного управління інцидентами, де алгоритми ML здатні навчатися на нових даних та коригувати playbook-и у реальному часі [12, 13]. Це створює основу для формування самонавчальних систем кіберзахисту, які здатні підвищувати власну ефективність у процесі експлуатації.

Отримане практичне значення дослідження полягає в тому, що впровадження Cisco SecureX дозволяє не лише скоротити критичні часові метрики (MTTD, MTTR), але й забезпечує раціональне використання кадрових ресурсів SOC, що особливо актуально за умов дефіциту висококваліфікованих фахівців [7].

Таким чином, дана платформа становить собою стратегічно важливий інструмент як для підвищення рівня кіберстійкості організацій, так і для розвитку науково-прикладних підходів у сфері автоматизації інформаційної безпеки.

Висновки

У результаті проведеного дослідження доведено, що впровадження платформи Cisco SecureX у діяльність центрів оперативного реагування на інциденти інформаційної безпеки є стратегічно обґрунтованим та науково підтвердженим кроком у напрямі підвищення кіберстійкості організацій. Аргументовано, що традиційні підходи до функціонування SOC не відповідають сучасним умовам багатовекторних атак, а тому потребують орієнтації на технології оркестрації, автоматизації та інтеграції [11].

Систематизація функціональних можливостей SecureX засвідчила, що ключовими її перевагами є глибока інтеграція з інфраструктурою кіберзахисту (SIEM, EDR, IDS/IPS, Threat Intelligence), використання гнучких сценаріїв реагування (playbooks), розширені аналітичні механізми на базі AI/ML, а також багаторівневий аналіз поведінкових і мережевих патернів. Практичні результати моделювання підтвердили, що застосування SecureX забезпечує скорочення показників Mean Time to Detect і Mean Time to Respond у 3-10 разів залежно від рівня автоматизації, зменшення операційного навантаження SOC більш ніж на 70 % та підвищення економічної ефективності впровадження з окупністю у межах кількох місяців.

Порівняльний аналіз з провідними конкурентними показав, що Cisco SecureX вирізняється оптимальним поєднанням інтеграційних можливостей, вбудованих сервісів загрозової розвідки та відносно коротким часом впровадження при збереженні доступної вартості. Це визначає його конкурентоспроможність та доцільність використання в умовах як середніх, так і великих організацій.

Таким чином, отримані результати дозволяють зробити висновок, що Cisco SecureX є не лише практичним інструментом автоматизації SOC, а й перспективним об’єктом для подальших наукових досліджень у галузі когнітивної автоматизації, інтегративних архітектур безпеки та AI-native платформ. Це відкриває нові напрями для удосконалення механізмів адаптивного реагування, зниження рівня людського фактору та формування масштабованих моделей кіберзахисту в умовах динамічного кіберпростору.

Список літератури

1. Whitepaper Cisco Public. Whitepaper Cisco Public. From Complex to Cohesive. How a Platform Approach Can Solve Today’s Security Conundrum. URL: https://s3.amazonaws.com/external_clips/3356387/securex-cohesive-whitepaper.pdf?1583527154=&utm_source
2. Юрген Кучер М-Тренди 2024: Наш погляд з передової. URL: https://cloud.google.com/blog/topics/threat-intelligence/m-trends-2024?utm_source=chatgpt.com
3. А. Мохсін, Х. Яніке, А. Ібрагім, .Х. Саркер, С. Камтепе. Уніфікована структура для співпраці людини та штучного інтелекту в центрах операцій безпеки з довіреною автономією URL: https://arxiv.org/html/2505.23397v2?utm_source
4. Кеті Биковскі Де знаходиться магічний квадрант Gartner SOAR? URL: https://swimlane.com/blog/soar-magic-quadrant/?utm_source
5. Цінго Чжан Розробка посібників з автоматизації безпеки - обмін отриманим досвідом з практиками. Біла книга. URL: https://www.cisco.com/c/en/us/products/collateral/security/designing-security-automation-playbooks-wp.html?utm_source
6. Zhyvylo Y. (2023). Exploring and Acquiring Modern Human Resource Competencies in Cybersecurity Amidst State Digital Transformation. Pressing Problems of Public Administration, 2(63), 111-127. DOI: https://doi.org/10.26565/1684-8489-2023-2-08

7. Zhyvylo , Y. O., & Zhyvylo , I. O. (2021). Joint training of the cyber security defense forces personnel in the conditions of total state defense. Theory and Practice of Public Administration, 2(73), 144-153. https://doi.org/10.34213/tp.21.02.16
8. Mahdi, Q. A., Zhyvotovskyi, R., Kravchenko, S., Borysov, I., Oleksandr, O., Panchenko, I., Zhyvylo, Y., Kupchyn, A., Koltovskov, D., Boholii, S. (2021). Development of a method of structural-parametric assessment of the object state. Eastern-European Journal of Enterprise Technologies, 5 (4 (113)), 34–44. doi: https://doi.org/10.15587/1729-4061.2021.240178
9. Koval M., Sova O., Orlov O., Zhyvylo Y., Zhyvylo I. Improvement of complex resource management of special-purpose communication systems // 5(9-119) (2022): Eastern-European Journal of Enterprise Technologies. P. 34–44, doi: https://doi.org/10.15587/1729-4061.2022.266009
10. S. Kashkevich, A. Shyshatskyi, O. Dmytriieva, Y. Zhyvylo, G. Plekhova, S. Neronov The development of management methods based on bio-inspired algorithms Information and control systems: modelling and optimizations: collective monograph. – Kharkiv: TECHNOLOGY CENTER PC, 2024. – 35-69p. DOI: http://doi.org/10.15587/978-617-8360-04-7
11. Zhyvylo, Y.O. (2024). Methodology for developing a national cybersecurity strategy. State Formation, no. 2 (36), 307–321. DOI: https://doi.org/10.26565/1992-2337-2024-2-21 [in Ukrainian].
12. Живило Є. О., Шевченко Д. Г. Оцінка ризиків кібербезпеки та контролю конфіденційності в інформаційних системах державного управління. Збірник наукових праць Військового інституту Київського національного університету імені Тараса Шевченка. 2022. № 75. С. 66-77. URL: http://nbuv.gov.ua/UJRN/Znpviknu_2022_75_9
13. Живило Є.О., Черноног О.О. Стратегія кібероборони України, Збірник наукових праць ВІТІ № 4, 2017, С.30–37. URL: https://www.researchgate.net/publication/380979172_STRATEGIA_KIBEROBORONI_UKRAINI
14. Cyber risk management technology to strengthen the information security of the national economy, S. Onyshchenko, Ye. Zhyvylo, A. Hlushko, S. Bilko ISSN 2071-2227, E-ISSN 2223-2362, Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu, 2024, No 5. С. 136-142, https://doi.org/10.33271/nvngu/2024-5/136
15. Svitlana Onyshchenko, Yevhen Zhyvylo, Anna Cherviak, Stanislav Bilko Determining the patterns of using information protection systems at financial institutions in order to improve the level of financial security. Vol. 5 (13 (125)) (2023): Eastern-European Journal of Enterprise Technologies. P. 65–76. DOI: https://doi.org/10.15587/1729-4061.2023.288175

Received (Надійшла) 19.08.2025
Accepted for publication (Прийнята до друку) 22.10.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Фесенко Тетяна Миколаївна кандидат технічних наук, доцент, доцент кафедри комп'ютерних технологій та інформаційних систем Національного університету "Полтавська політехніка імені Юрія Кондратюка", Полтава, Україна;
Tatiana Fesenko – candidate of technical sciences, associate professor, associate professor of the department of computer technologies and information systems, National University "Poltava Polytechnic named after Yury Kondratyuk", Poltava;
e-mail: tanifesenko@gmail.com; ORCID Author ID: https://orcid.org/0009-0006-1698-3795.

Калашнікова Юлія Вадимівна асистент кафедри комп'ютерних технологій та інформаційних систем Національного університету "Полтавська політехніка імені Юрія Кондратюка", Полтава, Україна;
Yuliia Kalashnikova – assistant of the Department of Computer Technologies and Information Systems, National University "Poltava Polytechnic named after Yury Kondratyuk", Poltava, Ukraine;
e-mail: kalashjulia74@gmail.com; ORCID Author ID: https://orcid.org/0000-0001-9899-4784.

Using Cisco SecureX for SOC automation

Tatiana Fesenko, Yuliia Kalashnikova

Abstract. Relevance. The article is devoted to the analysis of the capabilities of the Cisco SecureX platform in the context of automating processes within Security Operations Centers (SOC). The study emphasizes the relevance of this research, which is driven by the growing number of cyber threats, the increasing complexity of attacks, and the shortage of highly qualified professionals in the field of cybersecurity. It is underlined that traditional SOC operation methods are insufficiently effective under conditions of multivector attacks, which objectively necessitates the implementation of orchestration and automation technologies. The research systematizes the core functional features of Cisco SecureX, including: integration with multi-component cybersecurity infrastructures (SIEM systems, EDR platforms, IDS/IPS solutions), which ensures the creation of a unified information space; orchestration of SOC processes through the use of response playbooks, enabling the automation of routine operations, reducing incident handling time, and minimizing the human factor; enhanced analytical capabilities, based on the use of machine learning mechanisms and event correlation from heterogeneous data sources; improved accuracy of threat detection through multi-layered data analysis, including the examination of user behavioral patterns and network activity. Particular attention is paid to the research aspect of SecureX's impact on SOC efficiency. The study substantiates that the use of this platform makes it possible to reduce the Mean Time to Detect (MTTD) and Mean Time to Respond (MTTR), which are critical indicators for evaluating SOC performance. It is demonstrated that integrating SecureX with Threat Intelligence systems provides a more comprehensive contextual understanding of attacks, increases the level of proactive defense, and contributes to the development of an adaptive security architecture. From the standpoint of scientific novelty, the article presents a systematization of approaches to SOC automation, in which Cisco SecureX is considered not only as a tool for practical implementation but also as a subject of research for evaluating the effectiveness of integrative security platforms. The paper identifies promising directions for further study, including: advancing the level of cognitive SOC automation based on AI, optimizing response playbooks through adaptive algorithms, and assessing the scalability of SecureX in diverse organizational environments. Thus, the implementation of Cisco SecureX is justified both from the perspective of theoretical research and practical application. The platform contributes to the formation of a new approach to information security management, based on integration, automation, and analytics, which determines its strategic significance for enhancing the cyber resilience of modern organizations.

Keywords: information security, cybersecurity, cyber threats, artificial intelligence, architecture, scalability, SOC.

УДК 004.732.056 **doi: 10.26906/SUNZ.2025.4.144-150**

В. В. Челак, О. А. Горносталь

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

НЕЧІТКИЙ АНСАМБЛЬ ДЕРЕВ РІШЕНЬ ДЛЯ ІДЕНТИФІКАЦІЇ СТАНУ КОМП’ЮТЕРНИХ СИСТЕМ

Анотація. Об'єктом дослідження є процес ідентифікації стану комп’ютерних систем. **Предметом дослідження** є методи нечіткого ансамблю дерев з багатовимірними вузлами рішень для ідентифікації стану комп’ютерних систем. **Метою дослідження** є розробка та оцінка ефективності нечіткого ансамблю дерев рішень для підвищення точності ідентифікації стану комп’ютерних систем за умов наявності невизначеності, шумових впливів та неповних даних. **Методи, що використовуються**: методи машинного навчання, методи попередньої обробки даних, ансамблеві класифікатори, стекінгові підходи, методи комбінування та вибору ознак КС. **Отримані результати**: досліджено ефективність класичних та розроблених методів для ідентифікації стану комп’ютерних систем у складних умовах, що включають дисбаланс даних та наявність аномальних станів. Запропоновано комплексний підхід із використанням Fuzzy Stacking з MDT, що забезпечує високу точність і стабільність класифікації. Найкращі результати отримано саме для стекінгового підходу, який поєднує базові класифікатори та нечіткі дерева рішень, дозволяючи мінімізувати помилки першого та другого роду та досягти високих показників узагальнюючої здатності (MCC, F1-score, TS, LN(DOR)). **Висновки**. За результатами дослідження запропоновано удосконалений підхід до ідентифікації стану комп’ютерних систем, який поєднує стекінговий метод із Fuzzy MDT та оптимізацію вибору ознак. Комплексне використання цих методів дозволяє значно покращити точність класифікації, стабільність результатів та стійкість моделей до дисбалансу даних, а також забезпечує високу якість класифікації навіть у випадках появи нових аномальних станів.

Ключові слова: ідентифікація стану, комп’ютерні системи, машинне навчання, ансамбль, стекінг, дерева рішень, нечітка логіка.

Вступ

Сучасні комп’ютерні системи є складними багаторівневими утвореннями, що інтегрують програмні, апаратні та мережеві компоненти. Вони забезпечують критично важливі функції в промисловості, транспорті, енергетиці, фінансовій та державній сферах. В умовах зростання складності їхньої структури та взаємозв’язків підвищується ймовірність виникнення збоїв, нештатних режимів роботи та кібератак, що можуть призвести до порушення стабільності або повного припинення функціонування системи. Тому завдання своєчасної ідентифікації стану комп’ю-терних систем набуває особливої актуальності.

Традиційні методи діагностики та моніторингу, що базуються на жорстких правилах і статистичних моделях, не завжди здатні адекватно реагувати на динамічні та невизначені зміни параметрів. Вони часто не враховують багатофакторний характер впливу різних процесів та складність взаємозв’язків між параметрами, що описують стан системи. У зв’язку з цим зростає інтерес до використання інтелектуальних методів аналізу даних, зокрема машинного навчання, нечіткої логіки та ансамблевих моделей, які дозволяють підвищити точність і надійність оцінювання станів у складних і невизначених умовах. Дерева рішень є одними з найбільш інтерпретованих і зручних моделей для аналізу стану систем, однак їх окреме застосування часто призводить до перенавчання або зниження узагальнюючої здатності. Ефективним підходом до подолання цих недоліків є використання ансамблевих методів, які об’єднують результати кількох моделей. Водночас нечітка інтерпретація результатів дозволяє зменшити вплив похибок вимірювання та неповноти даних, забезпечуючи більш гнучке прийняття рішень.

Таким чином, актуальним є розробка підходу, який би поєднував переваги ансамблевих методів машинного навчання з можливістю нечіткої оцінки станів комп’ютерних систем. Це дозволить підвищити достовірність і стійкість ідентифікації при наявності шумових або суперечливих даних.

Об’єктом дослідження є процес ідентифікації стану комп’ютерних систем.

Предметом дослідження є методи нечіткого ансамблю дерев рішень для ідентифікації стану комп’ютерних систем.

Огляд пов'язаних наукових публікацій. Ідентифікація стану комп’ютерних систем є одним з сучасних напрямків досліджень у сфері забезпечення їхньої надійності та безпеки. Сучасні методи контролю стану базуються на аналізі великої кількості технічних параметрів, що описують роботу компонентів системи, їхню взаємодію та реакцію на зовнішні впливи. Традиційні підходи до моніторингу, зокрема на основі експертних правил або порогових значень [1, 2], забезпечують швидке реагування на відхилення, однак характеризуються низькою гнучкістю та не здатні ефективно працювати у випадках, коли параметри змінюються під впливом стохастичних процесів або неповноти даних.

Для подолання цих обмежень активно застосовуються методи машинного навчання, які дозволяють виявляти приховані залежності між параметрами та формувати адаптивні моделі поведінки системи [3, 4]. Найпоширенішими серед них є алгоритми класифікації, зокрема дерева рішень, опорні вектори (SVM), штучні нейронні мережі та ансамблеві методи, такі як Random Forest або Gradient Boosting [5, 6]. Дерева рішень залишаються популярними завдяки їхній простоті, інтерпретованості та здатності працювати з гетерогенними даними. Проте окремі дерева часто мають проблеми з перенавчанням, особливо у випадку високої кореляції ознак або значної варіативності вхідних даних.

Розвиток ансамблевих методів став природним етапом еволюції алгоритмів машинного навчання. Такі моделі поєднують результати кількох базових класифікаторів, що дозволяє підвищити стійкість та точність прийняття рішень. Методи типу Random Forest [7] або XGBoost [8] демонструють високу ефективність у задачах прогнозування технічних станів, проте вони залишаються чутливими до шуму та потребують чітких меж між класами. У реальних комп'ютерних системах, де наявна невизначеність і нечіткість даних, такі межі часто відсутні або змінюються динамічно. У зв'язку з цим перспективним напрямом досліджень є інтеграція апарату нечіткої логіки в ансамблеві методи машинного навчання [9, 10]. Нечіткі системи дозволяють описувати невизначеність у вигляді функцій належності, що дає змогу моделювати проміжні стани між «нормальним» та «аномальним» режимом. Поєднання дерев рішень із нечіткою логікою забезпечує більш гнучку інтерпретацію результатів класифікації, а створення нечітких ансамблів підвищує їхню стійкість до неповних або суперечливих даних.

У роботах [11–13] показано, що використання нечітких ансамблів дозволяє досягати кращої узагальнюючої здатності моделей у порівнянні з класичними ансамблями за рахунок вагового врахування ступеня достовірності окремих класифікаторів. Такі підходи особливо ефективні в умовах наявності шуму в телеметричних даних, апаратних збоях та інших невизначених станах. Однак більшість існуючих рішень фокусуються лише на певному типі даних або окремих характеристиках системи, не забезпечуючи комплексного підходу до оцінювання станів.

Отже, аналіз наукових джерел свідчить про необхідність створення узагальненого підходу, що поєднує різні методи ідентифікації станів комп'ютерних систем у межах єдиної нечіткої ансамблевої структури. Це дозволить врахувати специфіку кожного з методів, зменшити вплив шуму та підвищити точність класифікації при роботі в умовах невизначеності.

Постановка проблеми. Основна мета дослідження полягає в розробки та оцінці ефективності нечіткого ансамблю дерев рішень для підвищення точності ідентифікації стану комп'ютерних систем за умов наявності невизначеності, шумових впливів та неповних даних. Проблема ускладняється тим, що сучасні комп'ютерні системи характеризуються високою динамічністю, складною взаємодією компонентів та значною кількістю параметрів, які не завжди мають чітко визначені межі між нормальним та аномальним станом. У зв'язку з цим постає потреба в інтеграції апарату нечіткої логіки в ансамблеві структури, що дозволить здійснювати гнучке зважування результатів окремих класифікаторів та формувати узагальнене рішення з урахуванням ступеня впевненості кожного з них. Об'єднання декількох різних методів класифікації в єдиний нечіткий ансамбль забезпечить більш високу стійкість до шумів, здатність до адаптації в умовах неповних або суперечливих даних та покращення точності розпізнавання складних станів.

Для реалізації поставленої мети необхідно дослідити такі напрями підвищення ефективності систем ідентифікації:

1. Аналіз наявних методів ідентифікації стану комп'ютерних систем і визначення їхніх сильних та слабких сторін.
2. Розробка структури нечіткого ансамблю дерев рішень, що поєднує переваги різних підходів до класифікації.
3. Визначення механізму формування функцій належності та вагових коефіцієнтів для агрегування рішень окремих класифікаторів.
4. Проведення експериментальної оцінки точності запропонованого ансамблю на реальних або модельних наборах даних, що описують різні стани комп'ютерних систем.
5. Порівняння отриманих результатів із класичними ансамблевими методами без використання нечіткої логіки для оцінки доцільності впровадження запропонованого підходу.

Виконання цих етапів дозволить сформувати науково обґрунтовані висновки щодо доцільності застосування нечітких ансамблевих моделей у задачах ідентифікації стану комп'ютерних систем та розробити рекомендації для їх подальшої оптимізації й практичного використання.

Огляд підходів та методів

В основі запропонованого методу лежить алгоритм побудови дерев з багатовимірними вузлами рішень, які були вперше представлені в [14]. Теоретична основа дозволяє створювати вузли, які зможуть описувати багатовимірної області будь-якими фігурами та формами. На рис. 1, 2 представлені приклади двовимірних та тривимірних вузлів рішень відповідно.

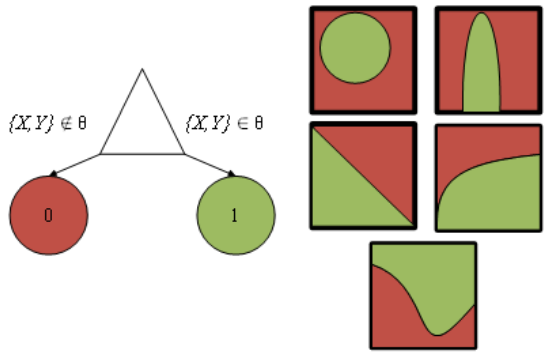


Рис. 1. Двовимірний вузол рішень з різними варіаціями фігури θ

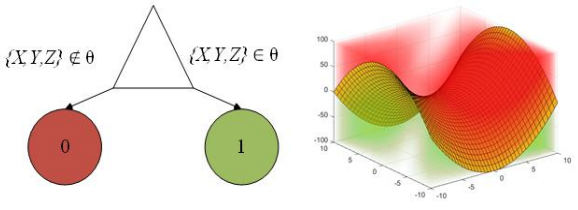


Рис. 2. Тривимірний вузол рішень з заданою фігурою θ

В поточній реалізації алгоритму використовується обмеження, щодо форми розподілу фігур, а саме використання гіперсфер: задаються координати центру (потрібної розмірності) та радіуса.

Першою частиною є побудова базових класифікаторів. Розглянемо модифіковану версію алгорит-

му побудови дерев з одновимірними вузлами рішень, на базі яких будуються багатовимірні вузли (при наявності покращення результатів):

Крок 1. Сформувати тренувальну вибірку даних.

Крок 2. Для кожної ознаки з найвищою інформативністю A_i:

2.1. Визначити мінімальне та максимальне значення ознаки.

2.2. Розрахувати середнє значення між ними та встановити його як поточний поріг θ_i.

2.3. Обчислити значення функції помилки класифікації $E(A)$ для лівої та правої підмножини:

$$E(A) = \sum_{i=1}^{N} [y_i \neq t_i]'; \quad (1)$$

$$[y_i \neq t_i]' = \begin{cases} 1.5, & y_i \neq t_i \ \cap \ t_i = 1; \\ 1, & y_i \neq t_i \ \cap \ t_i = 0; \\ 0, & y_i = t_i. \end{cases} \quad (2)$$

2.4. Якщо $E(A)$ для лівої підмножини більше, ніж для правої, встановити поточний поріг як нове максимальне значення; інакше – як нове мінімальне.

2.5. Повторювати бінарний пошук порогу θ_i, доки $E(A)$ не перестане змінюватися.

У якості оптимального порогу θ_i приймається значення, при якому $E(A)$ мінімальне.

Крок 3. Створити вузол дерева з пороговим значенням θ_i, сформувати підмножини зразків для лівої та правої гілки.

Крок 4. Перевірити умови зупинки (досягнута максимальна глибина, мінімальна кількість зразків тощо). Якщо хоча б одна виконується – перейти до кроку 6.

Крок 5. Повторити кроки 2–3 для тієї гілки, у якої значення $E(A)$ є більшим.

Якщо обидві гілки мають однакову помилку, пріоритет надається гілці з більшою кількістю екземплярів. Якщо $E(A)$ мінімальне для обох гілок, перейти до вузла вищого рівня та повторити крок 5.

Крок 6. Завершити побудову дерева рішень.

Алгоритм базується на ітераційному поділі вибірки даних за допомогою порогових значень, що визначаються методом бінарного пошуку. На кожному етапі обирається ознака з найвищою інформативністю, а поріг розділення θ_i підбирається так, щоб мінімізувати функцію помилки класифікації $E(A)$. Таким чином, формується вузол дерева, який поділяє простір ознак на дві області. Подальше розгалуження здійснюється рекурсивно для тієї частини, де залишкова помилка є найбільшою. Процес продовжується до виконання умов зупинки – досягнення заданої глибини дерева або мінімальної похибки.

Такий підхід дозволяє створювати адаптивну структуру дерева, що забезпечує гнучке розділення простору ознак та високу точність класифікації навіть при наявності неоднорідних даних.

Для пошуку гіперсфер, що зможуть об'єднувати велику кількість зразків з мінімальною помилкою використовується **алгоритм кластерізації DBSCAN**.

В основі алгоритму покладена ідея, згідно з якою всередині кожного кластера щільність об'єктів є суттєво вищою, ніж щільність зовні кластера, тоді як у шумових областях щільність нижча за щільність будь-якого з кластерів.

Першим кроком алгоритму є побудова матриці відстаней за формулою квадрата евклідової відстані:

$$d(x_i, x_j) = \sum_{k=1}^{m} (x_{ik} - x_{jk})^2, \quad (3)$$

де $d(x_i,x_j)$ – функція відстані між двома зразками; x_{ik} – значення k-ої ознаки i-го зразка; m – кількість ознак.

Для випадків, коли значення ознак є дискретними, використовується мангеттенська метрика:

$$d(x_i, x_j) = \sum_{k=1}^{m} |x_{ik} - x_{jk}|, \quad (4)$$

На наступному етапі визначається, чи є об'єкт x_i сусідом для об'єкта x_j. Результатом цього кроку є побудова матриці сусідства, яка базується на матриці відстаней:

$$Neighbour_{ij} = \begin{cases} 0, & d(x_i, x_j) > \varepsilon; \\ 1, & d(x_i, x_j) \leq \varepsilon; \end{cases} \quad (5)$$
$$i = \overline{1 \dots N}, j = \overline{1 \dots N},$$

де ε – гіперпараметр, що визначає радіус гіперсфери у багатовимірному просторі ознак.

Далі виконується ініціалізація міток для кожного об'єкта:

$$Label_i = -1, i = \overline{1 \dots N}. \quad (6)$$

На основі матриці сусідства формується початковий набір кластерів. Усі об'єкти на початку вважаються невизначеними. Ітераційна процедура починається з довільного об'єкта x_i, який ще не був розглянутий. Для кожного поточного об'єкта формується список сусідів – усі x_j, для яких у матриці $Neighbour_{ij}$=1. Підраховується кількість сусідів K і порівнюється з порогом $MinPts$. Якщо $K < MinPts$, об'єкт позначається як можливий викид. Якщо $K \geq MinPts$, об'єкт вважається ядровим, а його сусіди – досяжними за щільністю. Поточний об'єкт x_i разом із сусідами формують новий кластер, що отримує унікальну мітку.

Процес розширення кластера продовжується ітераційно: перевіряються необроблені або позначені як можливі викиди об'єкти, які є досяжними для елементів кластера, і приєднуються до нього. Ітерації тривають, доки кластер не перестане розширюватися. Процедура повторюється, поки всі об'єкти не набудуть визначеного статусу. Об'єкти, що залишились поза кластерами, вважаються викидами.

Після завершення кластеризації виконується **пошук параметра прийняття рішення для багатовимірного вузла дерева рішень**. Для цього визначається кластер C з максимальною кількістю елементів:

$$C = \max_{A_i \in A} (|A_i|), \quad (7)$$

де $|A_i|$ – кількість елементів i-го кластера.

Далі обчислюється центр кластера за кожною ознакою:

$$xc_k = \sum_{i=1}^{|C|} x_{ik} \Big/ |C|. \quad (8)$$

Отримавши координати центру кластера, виконується розрахунок відстаней від кожної точки до центру, після чого визначається максимальне значення:

$$\varepsilon = \max_{x_i \in C}\big(d(xc, x_i)\big). \quad (9)$$

Значення ε інтерпретується як радіус гіперсфери для багатовимірного вузла дерева рішень. Остаточним етапом є розрахунок цільових функцій, таких як критерій Джині або функції помилки класифікації. Як результат обираються такі гілки дерева, для яких сумарне значення цільової функції є мінімальним.

Другою частиною є використання нечітких дерев в якості мета-моделі, яка буде приймати загальне рішення стекінг ансамбля.

Нечітке дерево використовується як мета-модель для прийняття інтегрального рішення стекінг-ансамблю. Його побудова базується на фазифікації вхідних даних та аналізі інформаційного приросту для вибору оптимальних ознак розщеплення.

Процес фазифікації відбувається по спеціальній процедурі, яка була розроблена в попередніх дослідженнях [9]. На початку формується база правил та виконується фазифікація навчальної вибірки, у результаті якої отримується тривимірна таблиця нечітких значень $F_{ijk} = \mu_{jk}(TSij)$, де $\mu_{jk}(x)$ – функції належності лінгвістичних змінних. Далі розраховуються суми очікуваних та інверсних результатів і визначається загальна ентропія системи $E(F)$. Для кожної ознаки j, що ще не використана у дереві, обчислюються часткові ентропії нечітких множин E_{jk} та ентропія ознаки $E(j)$. На основі цього визначається інформаційний приріст $G(j) = E(F) - E(j)$.

Ознака з максимальним значенням $G(j)$ обирається як розщеплення вузла дерева. Для кожної гілки створюються підвузли відповідно до кількості функцій належності обраної ознаки. Агрегація ступенів належності між поточним вузлом та підвузлами виконується за мінімальним оператором. Отримані значення використовуються для подальших ітерацій побудови.

Процес рекурсивно повторюється, доки всі ознаки не будуть використані або інформаційний приріст стане незначним. У фіналі зберігаються параметри моделі: коефіцієнти гілок, вибрані ознаки розщеплення, структура вузлів та параметри функцій належності. У результаті формується модель FuzzyDecisionTree, що забезпечує адаптивне прийняття рішень на основі нечітких правил і використовується як узагальнювальний рівень ансамблю класифікаторів.

Формування нечіткого ансамблю дерев рішень

Для підвищення стійкості класифікації та узагальнюючи здатності моделей запропоновано об'єднання декількох різнотипних алгоритмів у стекінг-ансамбль, що поєднує властивості класичних дерев рішень, дерев з багатовимірними вузлами та нечітких дерев рішень. Такий підхід дозволяє врахувати як лінійні, так і нелінійні взаємозв'язки між ознаками, а також забезпечити нечітку інтерпретацію рішень у випадках, коли дані мають невизначеність або накладання класів.

Етап 1. Формування базових моделей

Першим рівнем ансамблю є набір базових моделей M_i, $i = \overline{1, L}$, серед яких:

1. Класичні дерева рішень M_{DT};
2. Дерева з багатовимірними вузлами M_{MDT};

Кожна з моделей навчається на однаковій навчальній вибірці $TS = \{(x_i, y_i)\}$, де $x_i \in \mathbb{R}^m$ – вектор ознак, а $y_i \in \{\overline{1, K}\}$ – мітка класу.

Результатом навчання базових моделей є набір прогнозів:

$$Z_i = [M_1(x_i), M_2(x_i), \dots, M_L(x_i)]. \quad (10)$$

Цей набір утворює матрицю вихідних оцінок базових моделей $Z \in \mathbb{R}^{N \times L}$.

Етап 2. Побудова мета-рівня нечіткого узагальнення

Отримані вихідні значення Z подаються на вхід нечіткому дереву рішень (Fuzzy Decision Tree, FDT), яке виконує роль мета-класифікатора.

Для кожної з базових моделей формується відповідна лінгвістична змінна $\tilde{Z}_j$ з функціями належності $\mu_{jk}(z_j)$, які визначають ступінь впевненості моделі у своєму прогнозі («низька», «середня», «висока»).

Фазифікація векторів Z_i здійснюється аналогічно до попереднього алгоритму нечіткого дерева, що дозволяє побудувати тривимірну таблицю нечітких оцінок:

$$F_{ijk} = \mu_{jk}(Z_{ij}), i = \overline{1, N},\ j = \overline{1, L},\ k = \overline{1, S_j}. \quad (11)$$

На основі фазифікованих значень виконується побудова нечіткого дерева, яке визначає узагальнене рішення ансамблю:

$$\hat{y}_i = FDT(Z_i). \quad (12)$$

Таким чином, кожен об'єкт оцінюється через систему правил виду:

$$IF\ (Z_1 \text{is "High"}) \cap \\ \cap\ (Z_2 \text{ is "Middle"}) then y_i is \text{Anomaly}. \quad (13)$$

Етап 3. Агрегація рішень та адаптація ваг

Після первинного навчання ансамблю виконується оцінка впливу кожної базової моделі на кінцеве рішення. Для цього вводиться коефіцієнт ваги w_j, який визначає значущість моделі M_j у загальній структурі ансамблю:

$$w_i = 1/\big(E_j + \delta\big). \quad (14)$$

де E_j – середнє значення функції помилки для моделі M_j, δ – мала константа для уникнення ділення на нуль.

Підсумкове значення агрегованого рішення до фазифікації розраховується за зваженою сумою:

$$S_i = \frac{\sum_{j=1}^{L} w_j \cdot M_j(x_i)}{\sum_{j=1}^{L} w_j}. \quad (15)$$

Ці результати використовуються для побудови нечітких правил мета-рівня, що дозволяє FDT враховувати як індивідуальну впевненість моделей, так і їх взаємні суперечності.

Етап 4. Остаточне прийняття рішення

Після фазифікації агрегованих значень S_i нечітке дерево виконує дефазифікацію результату –визначає кінцевий клас:

$$\hat{y}_i = \text{Defuzzify}(S_i), \quad (16)$$

де може бути використано, наприклад, метод центру ваги:

$$\hat{y}_i = \frac{\int_\Omega y \cdot \mu(y)\, dy}{\int_\Omega \mu(y)\, dy}. \quad (17)$$

Отримане значення $\hat{y}_i$ є остаточним рішенням нечіткого ансамблю, що поєднує узагальнену інформацію від усіх базових класифікаторів.

Формулювання теоретичних очікувань

Очікується, що використання дерев рішень із багатовимірними вузлами дозволить покращити якість класифікації завдяки здатності таких моделей описувати складні, нелінійні кордони між класами у багатовимірному просторі ознак. На відміну від класичних дерев, які розділяють простір за одновимірними порогами, багатовимірні вузли здатні формувати області будь-якої форми (зокрема гіперсферичні), що підвищує здатність моделі до узагальнення, особливо при обробці даних з корельованими або взаємозалежними ознаками.

Впровадження ансамблю на основі стекінгу (stacking) дозволяє комбінувати результати кількох різнорідних базових моделей, отриманих за допомогою різних параметрів або початкових умов побудови дерев. Такий підхід має забезпечити підвищення стійкості моделі до випадкових коливань у навчальних даних, а також зменшення варіації результатів.

Згідно з теоретичними очікуваннями, стекінг ансамбль з багатовимірних дерев повинен забезпечити:

1) підвищення показників Recall та Precision за рахунок глибшого аналізу локальних областей простору ознак;

2) покращення F1-score та MCC як результат балансу між повнотою та точністю;

3) зниження рівня overfitting у порівнянні з одиночними моделями завдяки узгодженому рішенню мета-рівня;

4) стабільність класифікації при зміні розподілу навчальних даних.

Таким чином, очікується, що інтеграція дерев з багатовимірними вузлами у структуру стекінг ансамблю забезпечить не лише локальне покращення точності на окремих класах, але й глобальну узгодженість результатів класифікації, підвищуючи ефективність моделі у задачах з високою розмірністю та складними взаємозалежностями між ознаками.

Експериментальні дослідження

Комп'ютерна система характеризується великою кількістю показників функціонування, серед яких параметри обчислювальних ресурсів, системних процесів, мережевої взаємодії, продуктивності, безпеки та зберігання даних. Однак використання всіх можливих характеристик є неможливим через обмеження обчислювальних ресурсів, пам'яті, складність оптимізації моделей та ризик перенавчання. Тому для побудови моделі ідентифікації станів КС було відібрано ключові групи показників:

- обчислювальні ресурси (завантаження процесора, пам'яті, носіїв);
- системні характеристики (операції введення/виведення, робота з носіями);
- показники мережевої взаємодії (обсяг переданих/отриманих даних);
- характеристики ОС та процесів;
- показники енергоспоживання і тепловиділення.

Збір даних здійснювався за допомогою технології Windows Management Instrumentation (WMI), яка забезпечує централізований моніторинг системних параметрів. Моніторинг проводився як у нормальному, так і аномальному стані системи. Нормальні дані збирались під час стандартної роботи навчальних комп'ютерів кафедри «Комп'ютерної інженерії та програмування» НТУ «ХПІ», а аномальні – під час виконання віртуальних середовищ, інфікованих різними типами шкідливого ПЗ (шифрувальники, віруси, хробаки, спуфінг-атаки тощо).

Отримано близько 690 тисяч зразків нормальної поведінки та 7,5 тисячі зразків аномальної, що дозволяє сформувати збалансовану навчальну вибірку для задачі класифікації. Відібрані ознаки включають показники навантаження дискових підсистем, процесора, оперативної пам'яті та мережевих інтерфейсів.

Для оцінки ефективності розроблених методів ідентифікації стану комп'ютерних систем (КС) та порівняння їх із класичними алгоритмами машинного навчання (Fine Tree, Weighted KNN, Cubic SVM) використовувались стандартні метрики точності, повноти, F1-score, коефіцієнт кореляції Метьюза (MCC), а також показники зміщення (bias) та розкиду (variance). Як показано у табл. 1, розроблені методи на основі дерев з багатовимірними вузлами рішень, нечіткого дерева та стекінгу демонструють нульове або мінімальне зміщення на етапі навчання, водночас класичні алгоритми, особливо Fine Tree, мають значну помилку розкиду через ризик перенавчання.

Таблиця 1 – **Помилки зміщення та розкиду моделей ідентифікації стану КС**

№	Метод ідентифікації стану КС	Bias, %	Variance, %
1	Fine Tree	0,13	31,97
2	Weighted KNN	0,03	10,97
3	Cubic SVM	0,07	26,07
4	Multi Decision Tree	0	9,10
5	Fuzzy Decision Tree	0,03	6,80
6	*Fuzzy Stacking with MDT*	0	0,1

Аналіз значень Accuracy та MCC (рис. 3) підтверджує здатність стекінгу до узагальнення та стабільної класифікації як позитивних, так і негативних випадків.

Високі значення F1-score, TS та DOR (рис. 4-5) демонструють збалансованість моделі та її здатність мінімізувати помилки першого і другого роду. Thread Score враховує одночасно вірно-позитивні, хибно-позитивні та хибно-негативні передбачення, що робить його корисним показником ефективності в умовах високої критичності хибних спрацьовувань.

Таким чином, стекінг дерев з багатовимірними та нечіткими вузлами рішень забезпечує найвищу точність, узагальнюючу здатність та ефективність ідентифікації станів КС, перевершуючи класичні методи та демонструючи практичну придатність для виявлення складних закономірностей у даних

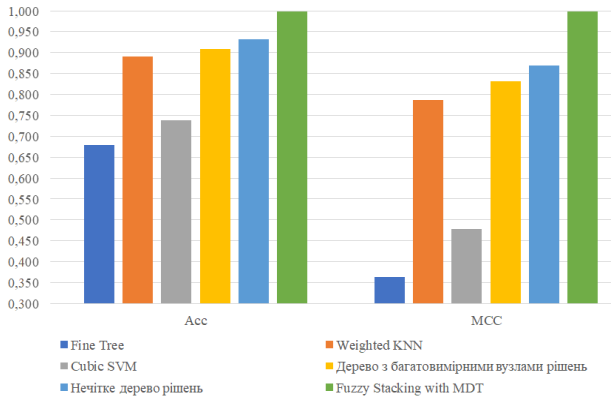


Рис. 3. Діаграми метрик ACC та MCC

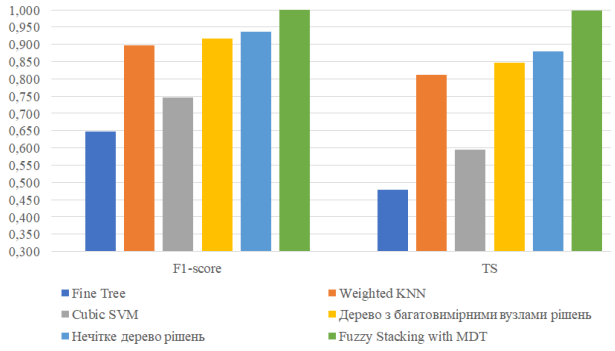


Рис. 4. Діаграми метрик F1-міри та TS

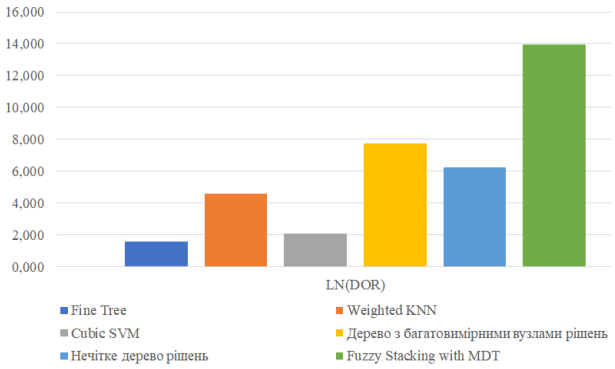


Рис. 5. Діаграма метрики Ln(DOR)

Оцінка результатів та формування рекомендацій для подальших досліджень

Проведений аналіз показав, що традиційні методи машинного навчання (Fine Tree, Weighted KNN, Cubic SVM) демонструють обмежену здатність до ефективної класифікації станів комп'ютерних систем (КС). Метод Fine Tree, зокрема, має високий рівень розкиду (Variance ≈ 31,97%) та низький коефіцієнт кореляції Метьюза (MCC ≈ 0,364), що свідчить про перенавчання та нестійкість моделі. Weighted KNN та Cubic SVM показують покращення, але все ще поступаються за ефективністю деревним методам із модифікованими вузлами.

Розроблені методи – дерево з багатовимірними вузлами рішень, нечітке дерево рішень та Fuzzy Stacking з MDT – демонструють суттєво вищу ефективність. Зокрема, стекінг на основі Fuzzy Stacking з MDT досяг майже ідеальних показників: Accuracy ≈ 0,999, MCC ≈ 0,998, F1-score ≈ 0,999, TS ≈ 0,998 та LN(DOR) ≈ 13,93, що підтверджує його високу узагальнюючу здатність та здатність виявляти складні взаємозв'язки в даних. Високе значення LN(DOR) та TS свідчить про мінімальну кількість помилок першого та другого роду, що критично для практичного застосування в інформаційних системах.

Висновки

В рамках дослідження було перевірено ефективність класичних та розроблених методів ідентифікації стану комп'ютерних систем (КС), включаючи Fine Tree, Weighted KNN, Cubic SVM, дерево з багатовимірними вузлами рішень, нечітке дерево рішень та Fuzzy Stacking з MDT. Результати показали, що класичні методи демонструють обмежену здатність до точного визначення станів КС. Метод Fine Tree має високий рівень розкиду та низьку узагальнюючу здатність, що свідчить про перенавчання та нестійкість моделі.

Найвищу ефективність продемонстрував Fuzzy Stacking з MDT, який забезпечує майже ідеальні показники Accuracy (0,999), MCC (0,998), F1-score (0,999), TS (0,998) та LN(DOR) (13,93), підтверджуючи його здатність точно класифікувати стани КС та мінімізувати помилки першого та другого роду.

Використання ансамблевих та стекінгових підходів значно підвищує точність і стабільність класифікації порівняно з окремими класичними методами.

Аналіз узагальнюючої здатності (MCC, F1-score, TS, LN(DOR)) є важливим для оцінки моделей у задачах з дисбалансом даних та аномальних станів, а також для перевірки стійкості моделей при введенні нових типів аномалій.

Подальше дослідження методів вибору та комбінування ознак КС дозволить зменшити обчислювальні витрати та підвищити інтерпретованість моделей без зниження точності.

Отже, розроблені методи ідентифікації стану КС є перспективними для практичного використання, оскільки забезпечують високу точність, збалансованість та стабільність класифікації навіть у складних умовах нерівномірного розподілу даних.

Список літератури

1. K. Shanthi and R. Maruthi, "A Comparative Study of Intrusion Detection and Prevention Systems for Cloud Environment," 2023 4th International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2023, pp. 493-496, doi: https://doi.org/10.1109/ICESC57686.2023.10193694
2. Z. Chiba, N. Abghour, K. Moussaid, O. Lifandali and R. Kinta, "Review of Recent Intrusion Detection Systems and Intrusion Prevention Systems in IoT Networks," 2022 International Conference on Software, Telecommunications and Computer Networks (SoftCOM), Split, Croatia, 2022, pp. 1-6, doi: https://doi.org/10.23919/SoftCOM55329.2022.9911401
3. W. Villegas-Ch, J. Govea, R. Gutierrez, A. Maldonado Navarro and A. Mera-Navarrete, "Effectiveness of an Adaptive Deep Learning-Based Intrusion Detection System," in IEEE Access, vol. 12, pp. 184010-184027, 2024, doi: https://doi.org/10.1109/ACCESS.2024.3512363

4. R. Latha and S. J. J. Thangaraj, “Securing the Digital Perimeter: A Comprehensive Intrusion Detection System with Ensemble Learning”, 2023 International Conference on Data Science, Agents & Artificial Intelligence (ICDSAAI), Chennai, India, 2023, pp. 1-5, doi: https://doi.org/10.1109/ICDSAAI59313.2023.10452636
5. О. Горносталь, В. Челак, Класифікація мережевих атак методами машинного навчання в умовах дисбалансу тренувальних даних / Системи управління, навігації та зв’язку, Полтава, 2025, Том 3 (81), С. 64-71, doi: https://doi.org/10.26906/SUNZ.2025.3.064
6. V. Chelak, S. Gavrylenko Method of Computer System State Identification based on Boosting Ensemble with Special Preprocessing Procedure / Advanced Information Systems, Kharkiv, 2022, Vol. 6 No. 1, P. 12-18, 2022, doi: https://doi.org/10.20998/2522-9052.2022.1.02
7. O.Hornostal , S. Gavrylenko. Development of a method for identification of the state of computer systems based on bagging classifiers. Advanced Information Systems. Kharkiv, 2021, vol. 5, no. 4, pp. 5–9, doi: https://doi.org/10.20998/2522-9052.2021.4.01
8. D. Wang, S. Ji, H. Shi and J. Liu, "Power Encoding Transmission Intrusion Detection Method Based on Convolutional Autoencoders-XGBoost," 2025 2nd International Symposium on New Energy Technologies and Power Systems (NETPS), Hangzhou, China, 2025, pp. 470-474, doi: https://doi.org/10.1109/NETPS65392.2025.11102095
9. Гавриленко С. Ю., Челак В. В. Розробка методу ідентифікації стану комп'ютерної системи на основі нечітких дерев рішень / Системи управління, навігації та зв’язку, Полтава, 2023, Випуск 1 (71), С. 78-83, doi: https://doi.org/100.26906/SUNZ.2023.1.078
10. Q. -V. Dang, "Studying the Fuzzy clustering algorithm for intrusion detection on the attacks to the Domain Name System," 2021 Fifth World Conference on Smart Trends in Systems Security and Sustainability (WorldS4), London, United Kingdom, 2021, pp. 271-274, doi: https://doi.org/10.1109/WorldS451998.2021.9514038
11. Das, Abhishek, et al. "Design of deep ensemble classifier with fuzzy decision method for biomedical image classification." Applied Soft Computing, vol. 115, 2022, p. 108178. ScienceDirect, doi: https://doi.org/10.1016/j.asoc.2021.108178
12. Chatterjee, S., Das, A. An ensemble algorithm integrating consensus-clustering with feature weighting based ranking and probabilistic fuzzy logic-multilayer perceptron classifier for diagnosis and staging of breast cancer using heterogeneous datasets. Appl Intell 53, 13882–13923 (2023), doi: https://doi.org/10.1007/s10489-022-04157-0
13. Shaikh, T.A., Rasool, T., Verma, P. et al. A fundamental overview of ensemble deep learning models and applications: systematic literature and state of the art. Ann Oper Res (2024), doi: https://doi.org/10.1007/s10479-024-06444-0
14. S. Y. Gavrylenko, V. V. Chelak, S. G. Semenov Development of Method for Identification the Computer System State based on the Decision Tree with Multi-Dimensional Nodes / Radio Electronics, Computer Science, Control (RECSC), Zaporizhzhia, No. 2 (2022), P. 113-122, doi: https://doi.org/10.15588/1607-3274-2022-2-11

Received (Надійшла) 24.08.2025
Accepted for publication (Прийнята до друку) 22.10.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Челак Віктор Володимирович – PhD, доцент кафедри "Комп'ютерна інженерія та програмування", Національний технічний університет “Харківський політехнічний інститут”, Харків, Україна;
Viktor Chelak – PhD, Associate Professor of Department of "Computer Engineering and Programming", National Technical University «Kharkiv Polytechnic Institute», Kharkiv, Ukraine;
e-mail: victor.chelak@gmail.com; ORCID ID: https://orcid.org/0000-0001-8810-3394;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57216331944&origin=resultslist.

Горносталь Олексій Андрійович – PhD, асистент кафедри "Комп'ютерна інженерія та програмування", Національний технічний університет “Харківський політехнічний інститут”, Харків, Україна;
Oleksii Hornostal – PhD, Assistant Professor of Department of "Computer Engineering and Programming", National Technical University «Kharkiv Polytechnic Institute», Kharkiv, Ukraine;
e-mail: gornostalaa@gmail.com; ORCID ID: https://orcid.org/0000-0001-5820-9999;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57189040595.

Fuzzy ensemble of decision trees for the computer systems state identification

Viktor Chelak, Oleksii Hornostal

Abstract. **The object of the research** is the process of identifying the state of computer systems. **The subject of the research** is the methods of fuzzy ensemble decision trees with multidimensional decision nodes for identifying the state of computer systems. **The goal of the research** is to develop and evaluate the effectiveness of a fuzzy ensemble of decision trees to improve the accuracy of identifying computer system states under conditions of uncertainty, noise, and incomplete data. **Methods used**: machine learning methods, data preprocessing techniques, ensemble classifiers, stacking approaches, methods for feature selection and combination of computer system attributes. **Results obtained**: the effectiveness of both classical and newly developed methods for identifying the state of computer systems under complex conditions, including data imbalance and the presence of anomalous states, was investigated. A comprehensive approach using Fuzzy Stacking with MDT was proposed, providing high accuracy and stability of classification. The best results were achieved with the stacking approach, which combines base classifiers and fuzzy decision trees, minimizing both type I and type II errors and achieving high generalization ability (MCC, F1-score, TS, LN(DOR)). **Conclusions**. Based on the results of the study, an improved approach for identifying the state of computer systems is proposed, which combines the stacking method with Fuzzy MDT and feature selection optimization. The integrated use of these methods significantly enhances classification accuracy, result stability, and model robustness to data imbalance, while ensuring high-quality classification even in the presence of new anomalous states.

Keywords: state identification, computer systems, machine learning, ensemble, stacking, decision trees, fuzzy logic.

UDC 621.317.1 **doi: 10.26906/SUNZ.2025.4.151-154**

Oleksandr Shefer, Nataliia Yermilova, Oleksandr Dryuchko, Mariia Stepanko, Sergii Pasichko

National University "Yuri Kondratyuk Poltava Polytechnic", Poltava, Ukraine

USE OF VIRTUAL MEASURING DEVICES IN METROLOGY, ELECTRONICS AND ELECTRICAL MACHINES FOR THE TRAINING OF ELECTRICAL ENGINEERING SPECIALISTS

Abstract. The work shows that a promising direction for the modernization of the educational laboratory base is the use of open systems methods together with the introduction of technologies for the use of virtual devices, based on the use of computer measurement methods. **The purpose of the work** is to analyze the possibilities of creating virtual laboratory work in various disciplines and their use in professional training of specialists in electrical engineering in higher educational institutions. As a result of the analysis of the most well-known application computer packages intended for the design of electronic units, the authors found that the ability to model both the simplest electrical circuits and power electronics circuits and rather complex control circuits for them, as well as various electrical machines is the main advantage of the Simulink package of the MATLAB environment in comparison with other software tools. Examples of constructing circuits for virtual laboratory work in the metrology course are given. It is shown that virtual laboratory work increases students' interest in the learning process, as well as the level of their skills and abilities.

Keywords: metrology, electronics, electrical machines, remote access, laboratory course, virtual measuring device.

Introduction

In these difficult times for Ukraine, when there are often air raids or there is no possibility to conduct classes in specialized laboratories, one of the most important areas of development of modern educational technologies is the development of remote access systems for learning and the implementation of open education standards on their basis [1].

The training of qualified specialists in electrical engineering is impossible without a modern laboratory base, on which students could not only consolidate the theoretical knowledge they have gained, but also acquire practical skills in research or production experiments, skills in designing and testing industrial systems. Today there is an intensive introduction of modern information technologies into the educational process, a large-scale modernization of the information infrastructure of the education system is being carried out, a unified educational environment is being formed. In such conditions, laboratory resources cannot remain at the old level, and a new approach to their formation is needed. A promising direction for the modernization of the educational laboratory base is the use of open systems technology together with the introduction of virtual device technology, based on the use of computer measurement methods [2, 3].

Today, many researches are devoted to the problems of informatization of vocational education. The works of the authors [3-8] consider the current state of formation and use in the professional training of specialists of new information and educational environments, however, despite the large number of diverse and large-scale studies concerning the informatization of education and the use of information and communication technologies of training, they did not find a thorough study of the development of virtual work.

The purpose of the work is to analyze the possibilities of creating virtual laboratory work in various disciplines and their use in professional training of specialists in electrical engineering in higher educational institutions.

Presentation of the main material

When studying electrical engineering specialties, an important component of the educational process is laboratory practice. Traditionally, educational laboratories are equipped with a certain set of technical means that allow for simple measurements. This applies, first of all, to such disciplines as physics, electrical engineering, electronics, electrical measurements, electrical machines. The set of measuring instruments here is usually very limited and includes an electrical signal generator, a voltmeter, an ammeter, a frequency meter and an oscilloscope, that is, measuring instruments of general use.

Until recently, all laboratory research and work, verification of the correctness of technical calculations could be performed only during the experimental study of real circuits and devices. This method has a number of significant drawbacks:

- for experimental study of the circuit, it is necessary to equip it with appropriate measuring equipment and full-scale samples of electrical machines and control devices;
- the errors of real measuring devices can be quite large;
- for the study of circuits, it is necessary to assemble their mock-ups from real elements, which leads to significant material costs.

Computer simulation of electrical circuits is devoid of these disadvantages, but there are some difficulties in accounting for the real parasitic parameters of circuit elements in the models: internal resistances and conductivities of sources; intrinsic inductances and capacitances of real resistors; losses in inductors and windings of electrical machines; nonlinearities caused by the presence of ferromagnetic cores; in addition, it is sometimes difficult to assess the accuracy of computer simulation.

Effective implementation of remote access technologies is possible only by creating a virtual information and educational environment of the university, uniting in a single information space various

corporate management systems, electronic libraries, distance learning and training systems, corporate testing systems, automation of scientific research, etc. In such conditions, real experimental stands are replaced by models of installations, creating a system of virtual laboratories. Virtual measuring laboratories are a very important component of the virtual representation of an educational institution, providing, together with other systems, all the functions of learning and management of the educational process. As a typical option, we can offer a educational virtual measuring laboratory (EVML) with the following typical set of measuring instruments: a signal generator; a universal oscilloscope; electric motor; an electronic frequency counter; an ammeter, a voltmeter; a wattmeter; a multimeter.

Such a EVML can become the basis for organizing laboratory workshops in technical educational institutions on the disciplines of the electrical engineering cycle, especially for distance learning for students. The basic version of the hardware part of the complex can be expanded by adding special adapters that provide research on a wide range of processes and phenomena in relation to various disciplines. The software can be supplemented with modules for specialized analog and digital signal processing, graphical representation of the results obtained, etc.

We analyzed the most famous application computer packages designed for the design of electronic blocks: the Electronics Workbench package, which is essentially a virtual laboratory with quite wide capabilities; the Design Lab package - an integrated software complex of the MicroSim corporation for the design of analog, digital and analog-digital devices; the Micro-Cap package - a universal package of programs for modeling the circuitry of electronic circuits; the Simulink package of the MATLAB environment - a virtual laboratory that allows you to assemble and study the operation of many types of electrical circuits, electric machines and electric drive devices, power electronics circuits.

The comparison of the packages was carried out according to the possibility of modeling analog devices (the ability to create device models based on analog electronics elements); modeling digital devices (the ability to create device models based on digital electronics elements); modeling electric machines (the ability to create models of transformers, motors, generators); the possibility of graphically displaying the results of modeling (the availability of means of visual representation of the processes occurring in the model - graphs of transient processes, dependences of the characteristics of components on some varied parameter); in addition, the possibility of changing the characteristics of the elements of the device model (voltage, current, resistance, inductance, etc.) was analyzed.

According to the results of comparison and analysis, it was found that the ability to model both the simplest electrical circuits and power electronics circuits, rather complex control circuits for them (electric drive devices), as well as various electric machines is the main difference between the MATLAB system and other software tools. This package is based on building block diagrams by transferring blocks from the component library to the editing window of the user-created model. Then the model is launched for execution. To build a functional block diagram of the simulated devices, Simulink has a large library of block components and a convenient block diagram editor. When developing a set of virtual laboratory works for studying the metrology course, the Simulink package of the MATLAB software environment was used. Here are examples of works from the metrology course that can be performed both on your own computer at home and in the computer class of the university.

1. Laboratory work № 1. "Measuring parameters of electrical networks using a virtual measuring laboratory". In the process of work, students get acquainted with the method of using virtual tools when creating an electronic laboratory, build assigned virtual electrical circuits using virtual analog dial instruments, learn to measure active resistances, inductances and capacitances both directly and indirectly, and also evaluate measurement errors (Fig. 1).

Students study various systems of analog instruments, their advantages and disadvantages, methods of expanding the limits of measurements, and answer the given control questions.

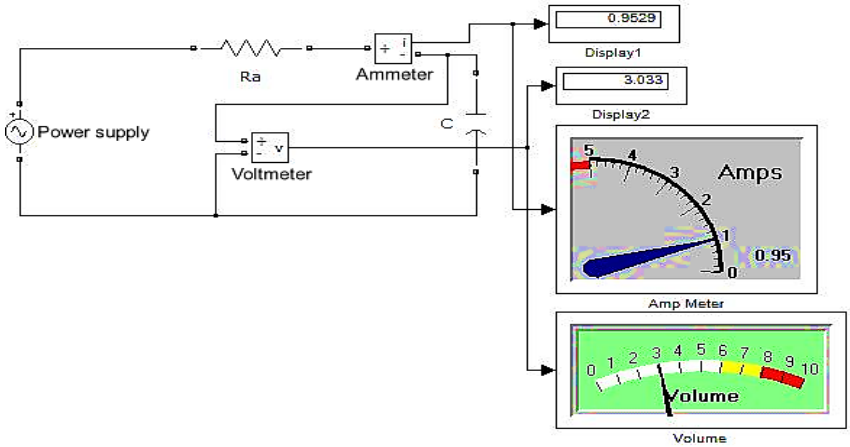


Fig. 1. An example of building a laboratory work № 1 scheme

2. Laboratory work №2. "Measurement of electrical network parameters by the bridge method using a virtual measuring laboratory". In the process of work, students study theoretical material on bridge circuits of direct and

alternating current, designs and principles of operation of comparison devices - bridges and compensators, get acquainted with the methodology for measuring electrical parameters by the bridge and compensation method using virtual devices, conduct measurements, and answer the given control questions (Fig. 2).

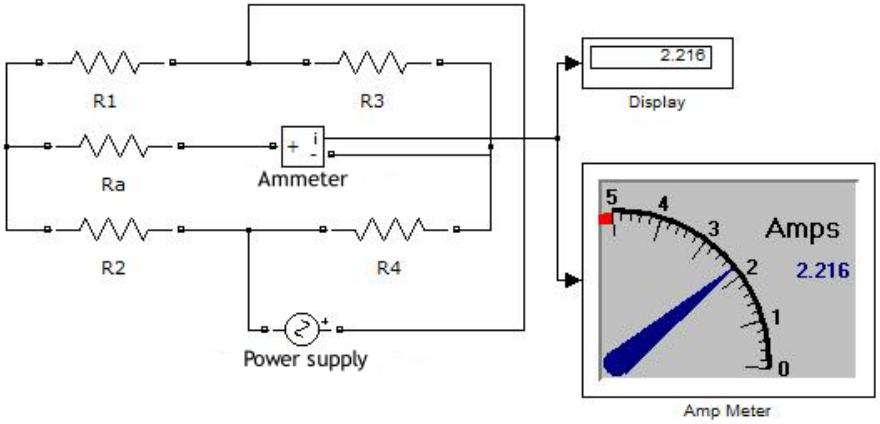


Fig. 2. An example of building a laboratory work № 2 scheme

3. **Laboratory work №3.** "Research of the parameters of periodic signals using a virtual oscilloscope". In the process of work, students study the structural diagrams and a typical set of blocks of both an electronic and a digital oscilloscope, get acquainted with the main characteristics of continuous and pulsed signals, acquire skills in working with a virtual oscilloscope and generator, learn to generate s ignals of various shapes, conduct measurements, calculate parameters of various signals, and answer the given control questions (Fig. 3).

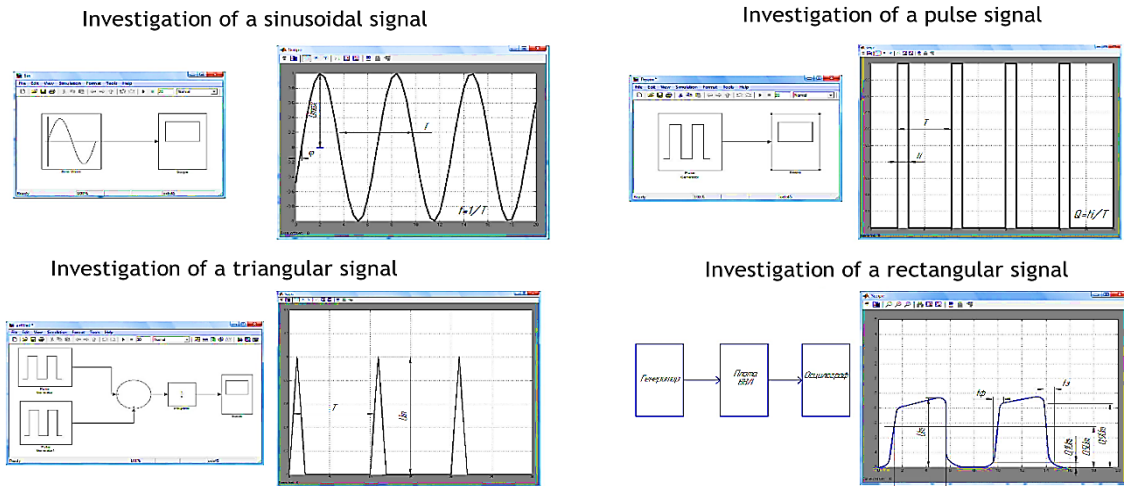


Fig. 3. An examples of obtaining and studying signals from laboratory work № 3

Such EVML laboratory can become a mandatory functional unit in collective use centers. In this case, the use of remote access mode will significantly improve the laboratory support of the educational process in the above-mentioned disciplines.

Conclusions

The particularity of this approach is the ability for students to practice experimental work skills on realistic models of dynamic objects with accurate reproduction of physical laws and high interactivity of research. It is possible to predict an increase in interest in the learning process in groups of students due to an innovative approach to the methodology of teaching engineering disciplines.

Virtual laboratory works in the professional training of future electricians ensure individuality and independence of students' activities, develop creative thinking and form the ability to make operational decisions, and, therefore, significantly increase the level of knowledge, skills and abilities. At the same time, the use of virtual laboratory work should be combined with the study of real equipment, real phenomena and processes.

REFERENCES

1. Науково-методичне забезпечення цифровізації освіти України: стан, проблеми, перспективи. / В.Ю. Биков, О.І. Ляшенко, С.Г. Литвинова, В.І. Луговий, Ю.І. Мальований, О.П. Пінчук, О.М. Топузов /за заг. ред. В.Г. Кременя. Київ: ІЦО НАПН України, 2022. 96 с. DOI: https://doi.org/10.37472/v.naes.2022.4223.
2. Соколюк О.М. Вплив VR/AR на технології навчання й освітянські практики. *Сучасні інформаційні технології та інноваційні методики навчання у підготовці фахівців: методологія, теорія, досвід, проблеми*. 2021. № 60. С. 108-116. DOI: 10.31652/2412-1142-2021-60-108-116.

3. Akbaş Y, Şahin IF, Meral E. Implementing argumentation-based science learning approach in social studies: Academic achievement and students' views. *Rev Int Geogr Educ Online*. 2019; 9: P. 209–245. DOI: 10.33403/rigeo.529139
4. Бобрівник К.Є., Гладка М.В., Кіктєв М.О. Проектування віртуальної навчальної лабораторії для студентів технічно-технологічних спеціальностей. *Енергетика і автоматика*, 2014. № 3. С. 18-23. URL: https://journals.nubip.edu.ua/index.php/Energiya/en/article/view/3460
5. Calvet, L.; Bourdin, P.; Prados, F. Immersive Technologies in Higher Education: Applications, Challenges, and Good Practices. *Proceedings of the 2019 3rd International Conference on Education and E-Learning*, Barcelona, Spain, 5–7 November 2019. 2019, 57, pp. 95 – 99, DOI: https://doi.org/10.1145/3371647.3371667.
6. Проєктування освітнього середовища з використанням засобів доповненої та віртуальної реальностей в закладах загальної середньої освіти: колективна монографія / Литвинова С.Г., Сороко Н.В., Баценко С.В., Богочков Ю.М., Гриб'юк О.О., Дементієвська Н.П., Коркішко І.А., Слободяник О.В., Соколюк О.М., Ухань П.С. Київ: ІЦО НАПН України, 2023. 219 с. URL: https://lib.iitta.gov.ua/id/eprint/738596/1/%21%21%21%D0%9C%D0%9E%D0%9D%D0%9E%D0%93%D0%A0%D0%90%D0%A4%D0%86%D0%AF_ALL_2023-FIN%2B.pdf
7. Березюк О.В. Використання віртуального лабораторного стенда для проведення лабораторної роботи «Дослідження ефективності освітлення у виробничих приміщеннях». *Педагогіка безпеки*, 2017. № 1. С. 35-39. URL: http://ir.lib.vntu.edu.ua//handle/123456789/19461
8. Рашевська Н. В. Перспективи застосування засобів доповненої реальності у процесі навчання майбутніх інженерів *Науковий вісник Ужгородського університету. серія: «Педагогіка. Соціальна робота»*. 2018. Вип. 2 (43). С.226-228. DOI: 10.24144/2524-0609.2018.43.226-229.

Received (Надійшла) 05.08.2025
Accepted for publication (Прийнята до друку) 22.10.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Шефер Олександр Віталійович – доктор технічних наук, професор, професор кафедри автоматики, електроніки телекомунікацій, Національний університет "Полтавська політехніка імені Юрія Кондратюка", Полтава, Україна;
Oleksandr Shefer – Doctor of Technical Sciences, Professor, Professor of the Department of Automation, Electronics and Telecommunications, National University "Yuri Kondratyuk Poltava Polytechnic", 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.

Єрмілова Наталія Василівна – кандидат технічних наук, доцентка, доцентка кафедри автоматики, електроніки телекомунікацій, Національний університет "Полтавська політехніка імені Юрія Кондратюка", Полтава, Україна;
Nataliia Yermilova – PhD, Associate Professor, Associate Professor of the Department of Automation, Electronics and Telecommunications, National University "Yuri Kondratyuk Poltava Polytechnic", Poltava, Ukraine;
e-mail: itm.ermilovanv@nupp.edu.ua; ORCID Author ID: https://orcid.org/0000-0003-0636-0843;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=58759309800.

Дрючко Олександр Григорович – кандидат хімічних наук, доцент, доцент кафедри автоматики, електроніки телекомунікацій, Національний університет "Полтавська політехніка імені Юрія Кондратюка", Полтава, Україна;
Oleksandr Dryuchko – PhD, Associate Professor, Associate Professor of the Department of Automation, Electronics and Telecommunications, National University "Yuri Kondratyuk Poltava Polytechnic", 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.

Степанко Марія Костянтинівна – студентка кафедри автоматики, електроніки телекомунікацій, Національний університет "Полтавська політехніка імені Юрія Кондратюка", Полтава, Україна;
Stepanko Mariia – student, Department of Automation, Electronics and Telecommunications, National University "Yuri Kondratyuk Poltava Polytechnic", Poltava, Ukraine;
e-mail: mariyastepanko@gmail.com; ORCID Author ID: https://orcid.org/0009-0006-7961-5420.

Пасічко Сергій Володимирович – студент кафедри автоматики, електроніки телекомунікацій, Національний університет "Полтавська політехніка імені Юрія Кондратюка", Полтава, Україна;
Sergii Pasichko – student, Department of Automation, Electronics and Telecommunications, National University "Yuri Kondratyuk Poltava Polytechnic", Poltava, Ukraine;
e-mail: znonmtposhta@gmail.com; ORCID Author ID: https://orcid.org/0000-0003-1405-2058.

Використання віртуальних вимірювальних засобів в курсах метрології, електроніки та електричних машин у підготовці фахівців з електричної інженерії

О. В. Шефер, Н. В. Єрмілова, О. Г. Дрючко, М. К. Степанко, С. В. Пасічко

Анотація. В роботі показане, що перспективним напрямком модернізації учбової лабораторної бази є використання методик відкритих систем разом з впровадженням технологій застосування віртуальних приладів, заснованої на вживанні методів комп'ютерних вимірів. **Метою роботи є** аналіз можливостей створення віртуальних лабораторних робіт з різних дисциплін та їх використання у професійній підготовці фахівців з електричної інженерії у вищих навчальних закладах. В результаті аналізу найбільш відомих прикладних комп'ютерних пакетів, які призначені для проектування електронних блоків, авторами виявлено, що можливість моделювання як найпростіших електричних схем, так і схем силової електроніки та досить складних схем управління ними, а також різних електричних машин є головною перевагою пакета Simulink середовища MATLAB у порівнянні з іншими програмними засобами. Наведені приклади побудови схем для віртуальних лабораторних робіт з курсу метрології. Показано, що віртуальні лабораторні роботи підвищують інтерес студентів до процесу навчання, а також рівень їх умінь та навичок.

Ключові слова: метрологія, електроніка, електричні машини, віддалений доступ, лабораторний курс, віртуальний вимірювальний прилад.

Цивільна безпека

УДК 621.384.2.002 **doi: 10.26906/SUNZ.2025.4.155-159**

Н. Б. Бурдейна, Я. А. Підлісний

Київський національний університет будівництва і архітектури, Київ, Україна

ЗАСОБИ НОРМАЛІЗАЦІЇ РІВНІВ МАГНІТНИХ ПОЛІВ ЧАСТОТИ ЛОКАЛЬНИХ ДЖЕРЕЛ

Анотація. Досліджено захисні властивості магнітом'яких металевих сплавів для екранування магнітних полів наднизьких частот. Встановлено залежності коефіцієнтів екранування магнітних полів промислової частоти найбільш поширеними електротехнічними сталями від індукції екранованого магнітного поля. Визначено, що ця залежність немонотонна. Наявність максимального коефіцієнта екранування обумовлена досягненням індукції насичення матеріалу. Значення коефіцієнтів екранування залежить від питомих втрат у сталі, які є довідковими величинами. Показано, що для виготовлення екрануючих конструкцій доцільно використовувати низькосортні електротехнічні сталі, найменш прийнятні для основного застосування – виготовлення осердь та магнітопроводів. Запропоновано двошарову конструкцію для екранування магнітних полів високих напруженостей, наприклад, вбудованих трансформаторів. Двошарова структура з електротехнічної сталі із зазором між шарами має коефіцієнти екранування магнітного поля промислової частоти вищі, ніж декілька шарів такої ж сталі. Це зменшує вагу екрануючої конструкції й підвищує її ефективність. Заповнення зазору між шарами знижує шумність обладнання, яка є нормативним параметром. Досліджено залежності коефіцієнтів екранування магнітного поля аморфних кобальтових сплавів від індукції зовнішнього магнітного поля. Така залежність є немонотонною, що необхідно враховувати у процесі проєктування екрануючих конструкцій. Аморфні феромагнітні сплави також ефективні при екрануванні магнітних полів низьких частот. Перевагами таких сплавів є високі значення магнітної проникності й мала товщина. Термомагнітна обробка аморфних сплавів значно підвищує їх захисні властивості. Показано доцільність розроблення композиційного матеріалу з наповнювачем з дрібнодисперсного аморфного магнітом'якого сплаву.

Ключові слова: магнітний екран, наднизька частота, аморфний магнітом'який сплав.

Вступ

Захисту від магнітних полів приділяється багато уваги. Це обумовлено тим, що через фізичну природу магнітних полів їх екранування захисними матеріалами і конструкціями дуже складне і неоднозначне. Особливо це стосується квазістаціонарних магнітних полів – магнітних складових електромагнітних полів промислової частоти та її гармонік. Такі поля неперервно впливають на людей у виробничих та побутових умовах. Джерелами магнітних полів найбільших напруженостей є електрощитові, електропривод електротехнічного обладнання, побутові пристрої неперервної дії тощо. У багатьох будівлях дозволені до використання вбудовані трансформатори сухого типу. Таке обладнання локалізовано у просторі, що дозволяє впровадити екранування магнітних полів, які вони генерують. Але ці пристрої розташовані у обмежених об'ємах для заощадження площ. Тому можливості застосування захисних конструкцій досить обмежені. Це вимагає розроблення й впровадження матеріалів та виробів високої ефективності й компактності. При цьому слід враховувати як технологічний, так і економічний аспекти. Захисні конструкції повинні мати прийнятну вартість та можливість застосовування у будь-якій конфігурації. Це обумовлює актуальність задачі розроблення засобів захисту від впливу магнітних полів локальних джерел.

Огляд досліджень і розробок

Напруженість квазістаціонарних магнітних полів регламентовані низкою міжнародних та національних нормативів. Це додаток до Європейської директиви [1], вимоги якого імплементовані у нормативну базу України у вигляді мінімальних вимог до рівнів електромагнітних полів. Норматив [2] встановлює максимальне значення магнітного поля промислової частоти у житлових приміщеннях 0,5 мкТл. Норматив [3] – значення низькочастотних полів наднизької частоти при експлуатації комп'ютерної техніки 250 нТл. Ці вимоги достатньо жорсткі й у більшості умов вимагають впровадження захисту людей.

Застосування традиційного підходу до захисту людей – захист відстанню і часом для низькочастотних впливів будь-якого походження практично неможливе. Це пояснюється низьким згасанням низькочастотних коливань з відстанню. Тому найбільш ефективним засобом захисту людей є екранування низькочастотних магнітних полів захисними матеріалами та конструкціями з них. Сучасною тенденцією у розробленні таких матеріалів є створення композиційних матеріалів. Але більшість з них має високу вартість і низьку технологічність щодо облицювання поверхонь складних форм [4, 5]. Крім того композиції на основі натуральних та полімерних матриць не відповідають вимогам пожежної безпеки [6, 7]. Найбільш ефективними матеріалами з малими масогабаритними параметрами є нанокомпозити [8, 9]. Але вони мають вкрай високу вартість й доцільні для застосування у виробах спеціального призначення. Ефективне екранування магнітного поля трансформаторів запропоновано у [10]. Але використана двошарова каркасна конструкція з конструктивної сталі. Вона громіздка і не може бути реалізована у обмежених просторах, особливо у

спорудах старої забудови. Особливістю магнітних полів наднизької частоти є відсутність відбиття через велику довжину хвилі. Тому високу ефективність можуть мати металеві матеріали, або металокомпозити на мінеральній основі. Таким є будівельний матеріал [11]. За наявності достатнього простору та при проєктуванні нових об'єктів його застосування доцільно з технічних та економічних міркувань. У роботах [12, 13] показано, що можливості використання металевих матеріалів для екранування магнітних полів не вичерпані. Тому доцільно провести дослідження ефективності металевих матеріалів різних класів у залежності від умов їх використання та можливих масогабаритних параметрів.

Мета роботи – обгрунтування та розроблення засобів захисту людей від впливу магнітних полів наднизьких частот.

Викладення основного матеріалу

Ефективність екранування магнітних полів наднизьких частот визначається, в основному, відносною магнітною проникністю захисного матеріалу.

Гігієнічно значущі напруженості магнітних полів у більшості будівель і споруд мають поля промислової частоти та її гармонік. У загальному випадку магнітна проникність феромагнітних матеріалів має частотну залежність, але для магнітних полів, що розглядаються, вона має стале значення і для більшості металів і сплавів є довідковою величиною. Але квазістаціонарність магнітних полів промислової частоти обумовлює той факт, що найбільш ефективним екранування є за умови геометричної залежності екрануючих конструкцій, які охоплюють джерело поля. Принаймні за такої умови коефіцієнти екранування магнітного поля максимальні для даного матеріалу. За умови часткового екранування джерела поля або захисту об'єкта зниження напруженості магнітного поля значно нижче, але може бути достатнім у залежності напруженості вихідного поля та значення гранично допустимого рівня для даних умов [14, 15]. Незамкненість екрануючої конструкції може бути вимушеною для доступу до обладнання, забезпечення вентиляції тощо.

Оціночний коефіцієнт екранування конструкції, виготовленої з феромагнітного матеріалу можна визначити зі співвідношення:

$$K = 1 + \mu_r \, d/D \,,$$

де μ_r – відносна магнітна проникність матеріалу екрана, d – товщина стінки екрана, D – діаметр еквівалентного сферичного екрана.

Сферичне наближення не дає великих похибок для екрануючих конструкцій кубічної форми, або близької до неї [8]. Ефективність екранування за фіксованих параметрів екрануючого матеріалу має амплітудні залежності, тобто залежить від напруженості магнітного поля, яке потребує екранування. Наявні у науковій літературі дані дещо суперечливі і стосуються обмеженої кількості навіть матеріалів, які виробляються серійно. Недостатньо опрацьована проблематика геометричних критеріїв ефективності екрануючих конструкцій. Вони фрагментарні й стосуються, в основному, не повністю замкнених циліндричних та сферичних конструкцій. Найбільш поширеними магнітом'якими матеріалами є електротехнічні сталі. Вони гарантовано знижують рівні магнітних полів. Було досліджено зміни коефіцієнтів екранування магнітного поля промислової частоти найбільш поширеними електротехнічними сталями марок 3411, 3412, 3413, 3414 у залежності від індукції зовнішнього магнітного поля (рис. 1).

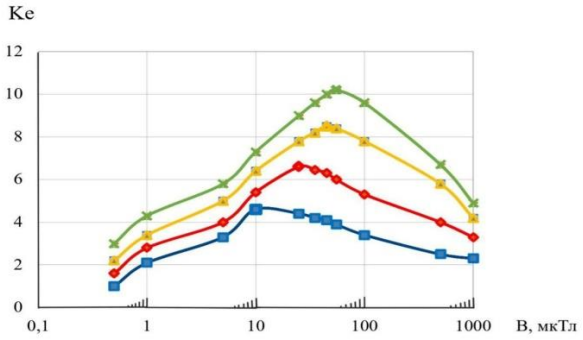


Рис. 1. Залежність коефіцієнтів екранування магнітного поля промислової частоти електротехнічними сталями від індукції екранованого магнітного поля: – сталь марки 3411, – сталь марки 3412, – сталь марки 3413, – сталь марки 3414; товщини зразків – 0,5 мм

Як видно з рис. 1, захисні властивості сталей суттєво відрізняються. Це можна пояснити різними питомими втратами у цих сталях, які є довідковими величинами.

Питомі втрати для цих матеріалів при магнітній індукції 1,5 Тл на частоті 50 Гц складають:

- сталь марки 3411 – 2,45 Вт/кг,
- сталь марки 3412 – 2,10 Вт/кг,
- сталь марки 3413 – 1,75 Вт/кг,
- сталь марки 3414 – 1,50 Вт/кг.

Виходячи з цього можна дійти висновку, що сталі, менш придатні для основного використання, наприклад, для виготовлення осердь, де критичними є втрати енергії, найбільш придатні для виготовлення захисних конструкцій. Тому для цілей магнітного екранування можна використовувати низькосортні електротехнічні сталі марок 1211, 1212, 1213, у яких питомі втрати складають 9,0–7,0 Вт/кг.

Наведені на рис. 1 дані мають досить велику похибку через те, що джерелом магнітного поля був трансформатор силової підстанції. Тому потрібну індукцію магнітного поля отримували упродовж тривалого часу через зміну навантаження.

Але отримані тенденції мають загальний характер. На усіх кривих рис. 1 є максимум коефіцієнта екранування. Наявність такого максимуму обумовлена, швидше за все, індукцією насичення конкретного матеріалу. Це необхідно враховувати у практичній діяльності при проєктуванні засобів екранування магнітного поля.

У багатьох випадках виникає необхідність екранування локальних джерел магнітного поля промислової частоти високих напруженостей. Це, наприклад, вбудовані трансформатори стандартних потужностей 1000–2500 кВт. Враховуючи можливість зниження коефіцієнта екранування внаслідок досягнен-

ня індукції насичення, необхідно підвищити магнітні втрати за рахунок конструкції екрана.

Було проведено лабораторні випробування збірки з п'яти листів електротехнічної сталі й двошарової конструкції із зазором у 5 см. Двошарова конструкція утримувалася каркасом, виготовленим з арматури. Усі листи виготовлялися з однакової електротехнічної сталі.

Результати випробувань залежності зниження індукції магнітного поля промислової частоти різними магнітними екранами в залежності від відстані до джерела поля наведено на рис. 2. Досліджувалися два екрани. Перший екран уявляв собою пакет із п'яти шарів електротехнічної сталі. Другий екран представляє собою двошарову конструкцію з електротехнічної сталі із зазором між шарами – 5 см. Більшу ефективність двошарової конструкції важко визначити однозначно.

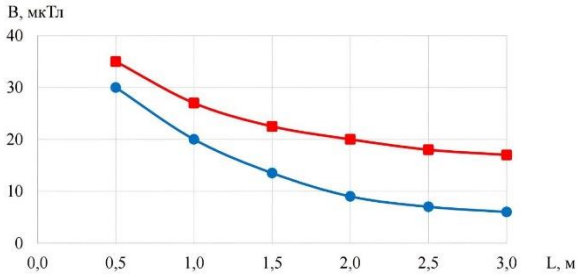


Рис. 2. Залежність зниження індукції магнітного поля промислової частоти магнітними екранами від відстані до джерела поля: ▬ – пакет з п'яти шарів електротехнічної сталі; ▬ – двошарова конструкція з електротехнічної сталі, зазор між шарами – 5 см

Більшу ефективність двошарової конструкції важко визначити однозначно. Можливо, проявляються ефекти взаємної індукції, зміни популяризації поля тощо. Зафіксований ефект дає можливість знизити масу захисної конструкції з одночасним підвищенням її ефективності. Але при цьому зростають її габаритні параметри. Це можна використати для зниження рівня шуму. Шумність електротехнічного обладнання, зокрема трансформаторів, нормується. Переважна інтенсивність їх звуку припадає на низькочастотну область звукового спектра. Низькочастотні коливання слабо поглинаються будівельними конструкціями, тому двошарова структура може мати певні резонансні властивості. А зазор між феромагнітними поверхнями доцільно заповнити шумопоглинальним матеріалом.

Енергетичні втрати у феромагнітному матеріалі визначаються магнітною проникністю матеріалу μ поля. Для електротехнічних сталей він складає 1000–1300, для конструкційних сталей – 100–200. Існують магнітом'які матеріали, для яких відносна магнітна проникність має значення порядка 10^4–10^5. Це нікелеві сплави пермалою та аморфні сплави на основі заліза та кобальту. Пермалої мають принциповий недолік: за будь-яких деформацій значення магнітної проникності різко падає. Тобто, при виготовлені захисної конструкції втрачаються захисні властивості. Цього недоліку позбавлені аморфні сплави. Сплави на основі кобальту у достатній кількості виробляються в Україні. Специфіка отримання аморфних металів обумовлює його геометричні характеристики. Матеріал виробляється надшвидким охолодженням розплаву, у результаті якого отримуються стрічки завширшки 3 см й товщиною 20–50 мкм. Для отримання захисної поверхні з цього матеріалу стрічки доцільно переплітати у вигляді «рогожі». Таким чином матеріал стає двошаровим. Перевагою аморфного кобальтового сплаву є те, що термомагнітна обробка сплаву (стаціонарне магнітне поле напруженістю 1000 А/м, температура – 300 °С) значно підвищує захисні властивості матеріалу.

На рис. 3 наведено залежність коефіцієнта екранування аморфного кобальтового сплаву від індукції екранованого магнітного поля промислової частоти до і після магнітотермічної обробки. Товщина двошарового матеріалу складала 60 мкм, а вміст кобальту – 71 %.

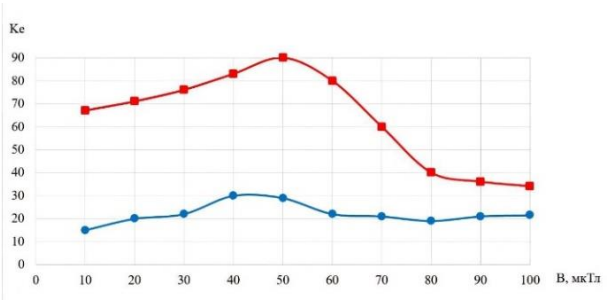


Рис. 3. Залежність коефіцієнта екранування аморфного кобальтового сплаву від індукції екранованого магнітного поля промислової частоти: ▬ – вихідний стан, ▬ – після магнітотермічної обробки; товщина двошарового матеріалу – 60 мкм; вміст кобальту – 71 %

Перевагою аморфного кобальтового сплаву є достатньо висока ефективність й в області дуже високих і ультрависоких частот. У частотній смузі 0,3 ГГц–3 ГГц коефіцієнт екранування за щільністю потоку енергії монотонно змінюється від 70 до 90 дБ. Це порівняно з ефективністю алюмінію, який має у цій смузі показник 90–100 дБ.

При проєктуванні захисту від електромагнітних полів частотою, більшою за промислову, слід враховувати складну залежність коефіцієнтів екранування від амплітудних значень екранованого магнітного поля. Було уточнено дані, наведені у [16]. Зокрема це стосується немонотонності зміни коефіцієнтів екранування магнітних полів малих амплітуд. На рис. 4 наведено залежність коефіцієнта екранування аморфного кобальтового сплаву від індукції екранованого магнітного поля частотою 10 кГц до і після магнітотермічної обробки.

Наведені результати свідчать, що дані, наведені у [16] завищені, не дивлячись на те, що випробувалися зразки, більші за товщиною на 10 мкм. При цьому форми кривих практично збігаються. Це свідчить про те, що тенденції зміни амплітудної залежності захисних властивостей є загальними. Відмінності у кількісних даних можна пояснити похибками вимірювань або відмінностями у режимах виготовлення й термомагнітної обробки зразків.

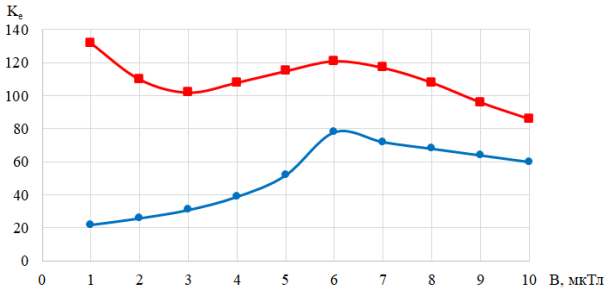


Рис. 4. Залежність коефіцієнта екранування аморфного кобальтового сплаву від індукції екранованого магнітного поля частотою 10 кГц: ▬ – вихідний стан, ▬ – після магнітотермічної обробки; товщина зразка – 60 мкм; вміст кобальту – 71 %.

З практичної точки зору такі відмінності не є критичними. Отримані коефіцієнти екранування достатньо високі й забезпечують екранування магнітних полів наднизьких та низьких частот більшості виробничих та побутових умов. Аморфні сплави доцільно застосовувати для екранування магнітних полів в умовах малих об'ємів, які виключають застосування стандартних електротехнічних сталей, або екранування джерел поля малих розмірів.

Для магнітних екранів для зниження рівнів низькочастотних полів завжди існує проблема їх замкненості через квазістаціонарність полів. Проаналізовані дані досліджень з цього напряму дозволяють дійти висновку, що універсального підходу не існує. Найбільш раціональним є забезпечення високих коефіцієнтів екранування магнітних полів у зонах суцільного перекриття поля, що знижує його напруженості у незахищених зонах – зонах часткового розкриття екрануючої конструкції, що було реалізовано у [14]. У будь-якому випадку зони розкриття необхідно інструментально перевіряти на тестових конструкціях або за зміною розкриття конструкції на реальному виробі.

У сучасних будівлях і спорудах цивільного призначення спостерігаються значні рівні і високочастотні електромагнітні випромінювання. При цьому їх джерелами можуть бути електронні компоненти обладнання, яке генерує магнітні поля наднизької частоти високих напруженостей. Тому напрямом перспективних досліджень є створення універсального матеріалу, ефективного у широкосмуговій смузі частот. Такий матеріал повинен мати достатні магнітні проникності й електропровідність разом з високою діелектричною проникністю. Це можливо реалізувати у композиційних матеріалах. Вартість аморфних магнітом'яких сплавів поступово знижується. Тому доцільно розглядати можливість створення матеріалу на основі діелектричної матриці наднизької горючості з наповнювачем із дрібнодисперсного аморфного сплаву. За певних режимів термічної обробки такі сплави стають крихкими, що спрощує процес їх подрібнення. Не виключається можливість утворення у таких сплавів кристалічних частинок нанорозмірів. Такі структури відкривають можливість різкого підвищення коефіцієнтів екранування електромагнітних полів широкого частотного діапазону.

Висновки

1. Показано, що для екранування змінних магнітних полів, які можна вважати квазістаціонарними, доцільно використовувати стандартні магнітом'які сталі. Досліджено залежність коефіцієнтів екранування магнітних полів промислової частоти від індукції екранованого поля найбільш поширених електротехнічних сталей. Встановлено наявність максимального значення коефіцієнта екранування за певної індукції магнітного поля. Це пов'язане з досягненням індукції насичення. Визначено, що для цілей електромагнітної безпеки найбільш прийнятними є сталі з великими показниками магнітних втрат. Це дозволяє використовувати низькосортні електротехнічні сталі для виготовлення магнітних екранів.

2. Показана доцільність виготовлення двошарових екрануючих конструкцій для екранування магнітних полів локальних джерел з великими напруженостями магнітних полів. Такі конструкції ефективніші за багатошарові структури, виготовлені з того ж матеріалу й мають меншу вагу. Зазор між двома шарами конструкції доцільно заповнювати звукопоглинальним матеріалом, що знижує шумність екранованого пристрою.

3. Досліджено залежність коефіцієнтів екранування магнітного поля промислової частоти від індукції екранованого магнітного поля аморфних кобальтових сплавів. Встановлено, що така залежність немонотонна й має чіткий максимум. Термічна обробка аморфного сплаву у постійному магнітному полі значно підвищує коефіцієнти екранування. Встановлено амплітудні залежності коефіцієнтів екранування аморфним сплавом магнітного поля низької частоти, які також немонотонні. Для отримання універсального широкосмугового електромагнітного екрана обгрунтовано розроблення композиційного матеріалу з наповнювачем із дрібнодисперсного аморфного сплаву.

СПИСОК ЛІТЕРАТУРИ

1. Directive 2013/35/EU of the European Parliament and of the Council of 26 June 2013 on the minimum health and safety requirements regarding the exposure of workers to the risks arising from physical agents (electromagnetic fields) (20th individual Directive within the meaning of Article 16(1) of Directive 89/391/EEC) and repealing Directive 2004/40/EC. URL: https://eur-lex.europa.eu/eli/dir/2013/35/oj
2. СОУ-Н ЕЕ 20.179:2008 Розрахунок електричного і магнітного полів ліній електропередавання. Методика. Зі змінами. Київ. Науково-технічний центр електроенергетики «НЕК «Укренерго», 2016.
3. MPR II. The Swedish government standard for maximum video terminal radiation. URL: https://www.computerlanguage.com/results.php?definition=TCO
4. Changtao Pu, Jiaxin Yang, Shuang Jin, Yuhui Zhou, Wei Gong. (2025). Engineered Self-Blown Nonisocyanate Polyurethanes with Synchronously Enhanced Electromagnetic Interference Shielding and Dimensional Stability. Applied Polymer Materials, 7600-7611, https://doi.org/10.1021/acsapm.5c012350000

5. Rajendrakumar Sharma, Dibyaranjan Das, Prafulla K. Dash, Asutosh Acharya, Kajal Parashar, SKS Parashar. 2025. Electromagnetic interference shielding of Zn-50%-Al alloy-coated polypropylene flexible conducting film. Intelligent Computing Techniques and Applications, 4 p.
6. Bharath, V., & Basa Reddy, S. (2025). Sustainable Electromagnetic Interference Shielding with Biomass Waste. IntechOpen. doi: 10.5772/intechopen.1009943
7. Багрій М.М., Левченко Л.О., Тихенко О.М., Колумбет В.П., Рєзнік Д.В. Розроблення та дослідження властивостей текстильного матеріалу від впливу електромагнітних полів. Вісник національного університету водного господарства та природокористування. 2019. Вип. 1(85). С. 237-244
8. Glyva V.A., Podoltsev A.D., Bolibrukh B.V., Radionov A.V. A Thin Electromagnetic Shield of a Composite Structure Made On the Basis of a Magnetic Fluid. Tekhnichna elektrodynamika. 2018. № 4. P.14−18. https://doi.org/10.15407/techned2018.04.014.
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. Рябов Ю.Г., Гуров И.Б. Экранирование встроенных трансформаторных подстанций. Технологии ЭМС. 2014. № 3. С. 21–28.
11. Y. Wang, C. Ma, S. Xie, Z. Wu, Z. Ji. An ultra-wideband electromagnetic shielding concrete based on multi-scale conductive fillers. 2025. Materials Letters, Volume 399, 15 November 2025, 139080. https://doi.org/10.1016/j.matlet.2025.139080
12. Панова О.В., Тихенко О.М., Ніколаєв К.Д., Ходаковський О.В., Сапельнікова О.Ю. Дослідження захисних властивостей металевих електромагнітних екранів та визначення умов їх максимальної ефективності. Системи управління, навігації та зв'язку. 2019. Вип. 5(57). С. 103−107.
13. Глива В.А., Панова О.В., Тихенко О.М., Левченко Л.О., Колумбет В.П. Дослідження амплітудно-частотних залежностей захисних властивостей магнітних екранів на основі аморфних сплавів. Системи управління, навігації та зв'язку. 2019. Вип. 6(58). С. 102−107.
14. Левченко О.Г., Левчук В.К., Тимошенко О.Н. Экранирующие материалы и средства индивидуальной защиты сварщика от магнитных полей. Автоматическая сварка. 2011. № 3. С.49–55.
15. Екрануючий комплект: пат. 90892 Україна: МПК G12B 17/00. № 201400841; заявл. 30.01.2014; опубл. 10.06.2014, Бюл. № 11. 4 с.
16. Глива В. А. Моніторинг та нормалізація фізичних факторів виробничого середовища при експлуатації автоматизованих систем: дис. …д-ра техн. наук: 05.26.01. Київ, 2012. 320 с.

Received (Надійшла) 15.08.2025
Accepted for publication (Прийнята до друку) 29.10.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Бурдейна Наталія Борисівна – доктор технічних наук, доцент, професор кафедри фізики, Київський національний університет будівництва та архітектури, Київ, Україна;
Nataliia Burdeina – Doctor of Technical Sciences, Associate Professor, Professor of the Department of Physics, Kyiv National University of Construction and Architecture, Kyiv, Ukraine;
e-mail: burdeina.nb@knuba.edu.ua, ORCID Author ID: https://orcid.org/0000-0002-2812-1387;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57220047954.

Підлісний Ярослав Анатолійович – аспірант кафедри технологій захисту навколишнього середовища та охорони праці, Київський національний університет будівництва і архітектури, Київ, Україна;
Yaroslav Pidlisnyi – PhD student at the Department of Environmental Protection Technologies and Labour Protection, Kyiv National University of Construction and Architecture, Kyiv, Ukraine;
e-mail: Pidlisnyi97@gmail.com; ORCID Author ID: https://orcid.org/0009-0008-4906-3164.

Means of normalising the levels of magnetic fields of local sources

Nataliia Burdeina, Yaroslav Pidlisnyi

Abstract. The protective properties of soft magnetic metal alloys for shielding ultra-low frequency magnetic fields have been investigated. The dependence of the shielding coefficients of industrial frequency magnetic fields by the most common electrical steels on the induction of the shielded magnetic field has been established. It has been determined that this dependence is non-monotonic. The presence of the maximum shielding coefficient is due to the achievement of the saturation induction of the material. The values of the shielding coefficients depend on the specific losses in steel, which are reference values. It has been shown that for the manufacture of shielding structures, it is advisable to use low-grade electrical steels, which are least suitable for their main application – the manufacture of cores and magnetic conductors. A two-layer structure is proposed for shielding high-intensity magnetic fields, for example, in built-in transformers. A two-layer structure made of electrical steel with a gap between the layers has higher industrial frequency magnetic field shielding coefficients than several layers of the same steel. This reduces the weight of the shielding structure and increases its efficiency. Filling the gap between the layers reduces equipment noise, which is a regulatory parameter. The dependence of the magnetic field shielding coefficients of amorphous cobalt alloys on the induction of an external magnetic field has been studied. This dependence is non-monotonic, which must be taken into account in the design of shielding structures. Amorphous ferromagnetic alloys are also effective in shielding low-frequency magnetic fields. The advantages of such alloys are high magnetic permeability and low thickness. Thermomagnetic treatment of amorphous alloys significantly increases their protective properties. The feasibility of developing a composite material with a filler of finely dispersed amorphous soft magnetic alloy is shown.

Keywords: magnetic shield, ultra-low frequency, amorphous soft magnetic alloy.

УДК 629.5.015.5-6:331.45 **doi: 10.26906/SUNZ.2025.4.160-163**

В. А. Глива, М. С. Кашлев

Київський національний університет будівництва і архітектури, Київ, Україна

ДОСЛІДЖЕННЯ ЕФЕКТИВНОСТІ ТОНКИХ ШУМОЗАХИСНИХ ЕКРАНІВ В УМОВАХ ОБМЕЖЕНИХ ПРОСТОРІВ

Анотація. Досліджено можливості застосування тонких шумозахисних екранів для зниження акустичного впливу на працюючих і населення в умовах обмежених просторів. Такі умови притаманні будівельному шуму при виконанні робіт серед щільної забудови. Наведено найбільш зручний математичний апарат для прогнозування ефективності акустичного екрана. Виміряні амплітудно-частотні характеристики індексів зниження шуму у залежності від площі акустичного екрана. Встановлено прийнятну збіжність розрахункових і експериментальних даних. Встановлено, що у низькочастотній частині звукового спектра ефективність тонких екранів недостатня. Досліджено вплив отворів у захисній конструкції на індекси зниження шуму. Встановлено, що на середніх та високих частотах отвори стають критичними для захисту від шуму. При цьому на низьких частотах вони практично не впливають на ефективність екрана. Виміряно індекси зниження шуму у октавних смугах частот і їх зміни з відстанню від екрана до приймача звуку. Встановлено, що на низьких частотах рівні звуку можуть зростати (від'ємні індекси) внаслідок інтерференційних явищ. Запропоновано в умовах обмежених просторів надавати екранам резонансних властивостей. За резонансну частоту доцільно обрати частоту звуку з максимальною амплітудою. Для підвищення ефективності екрана в області середніх та високих частот доцільне виготовлення двошарових панелей і заповнення проміжку між шарами шумопоглинальним пористим матеріалом.

Ключові слова: будівельний шум, акустичний екран, індекс зниження шуму, обмежений простір.

Вступ

Тонкі шумозахисні екрани є традиційними засобами зниження акустичного впливу на працюючих і населення. Такі конструкції мають різну форму і габаритні розміри у залежності від умов застосування. Це довгі екрани уздовж транспортних потоків, екрани, які огороджують окремі робочі місця та звукоізолюючі конструкції навколо обладнання з високими рівнями генерованого шуму. Перевагою тонких шумозахисних екранів є відносно низька вартість, простота вироблення та зручність у монтажі. Тому поширеною практикою є застосування мобільних захисних екранів. Але висока ефективність таких конструкцій забезпечується за умови наближеності екрана до джерел звуку та достатньо великих відстаней до приймачів звуку. За інших умов екрануючі конструкції можуть мати низьку ефективність. Наприклад, при виконанні будівельних робіт на території житлової забудови тонкі шумозахисні екрани можуть бути неефективними. Тому доцільно визначити межі ефективності тонких шумозахисних екранів у залежності від їх розмірів, відстаней до захищеної зони тощо. Доцільно визначити можливість зміни індексів зниження шуму внаслідок інтерференційних явищ при відбиті звукових хвиль та дифракцій на кромках екрана. Це дозволить обрати найбільш раціональну форму і площу екрана як з функціональних так і економічних міркувань.

Стан питання

Рівні шуму, що впливають на населення регламентуються національними санітарними нормами [1]. Ці норми достатньо жорсткі і відповідають за гранично допустимими значеннями нормативам Євросоюзу. Це вимагає досліджень і розроблення засобів захисту людей від впливу техногенного шуму, зокрема будівельного. Дослідження [2, 3] присвячені саме цій проблематиці. Але вони стосуються впливу на людей будівельного шуму. Робота [4] розглядає моделі поширення будівельного шуму, що використовується у проектних рішеннях. Такий підхід прийнятний для нової забудови, але виконання робіт на території існуючої забудови стикається з низкою проблем. Головною є неможливість забезпечення нормативних значень шуму існуючої будівельної техніки у оточуючих будівлях. Моніторинг рівнів такого шуму свідчить про його ненормативні значення у побутовому середовищі [5]. Найбільш ефективним засобом зниження акустичного навантаження на населення є застосування звукоізолюючих матеріалів і конструкцій [6, 7]. Але, враховуючи тимчасову потребу у зниженні рівнів шуму конкретного об'єкту будівництва, застосування таких звукопоглинальних матеріалів є проблематичним. Найбільш простими й економічно вигідними є тонкі акустичні екрани, які широко використовуються у виробничих умовах та для зниження шуму транспортний потоків. Ефективність таких екранів досліджено у роботі [8]. Зокрема, надано математичний апарат для оцінки ефективності акустичних екранів. Наведені дослідження стосуються достатньо великих відстаней. При цьому показано, що внаслідок дифракційних явищ на кромках екранів та інтерференції прямих та відбитих хвиль ефективність екрануючої конструкції може знижуватись до незадовільних значень. Найбільш ефективними є поглинаючі та резонансні конструкції [9, 10]. Вони дозволяють отримати високі індекси зниження низькочастотного звуку та інфразвуку. Проте, ці конструкції достатньо складні і мають високу вартість. їх недоліком є необхідність налаштування на певні частоти у залежності від амплітудно частотних характеристик шум конкретних джерел. Тому доцільно дослідити межі ефективності тонких шумозахисних екранів в умовах, коли відстані від джерел шуму невеликі, а площі захисних конструкцій обмежені.

Мета роботи – дослідити ефективності шумозахисних екранів у залежності від умов їх застосування.

Викладення основного матеріалу

Необхідність застосування для зниження рівнів шуму у обмежених просторах або розповсюдження – наближеності до житлової забудови обумовлене низьким згасанням шуму у повітрі. У загальному випадку зниження рівнів звукового тиску з відстанню визначається співвідношенням:

- для сферичних акустичних хвиль

$$\Delta L(r) = 20\lg r,$$

- для циліндричних акустичних хвиль:

$$\Delta L(r) = 10\lg r.$$

При цьому спостерігається значна залежність згасання від частоти звукової хвилі.

У середньому значення рівня звукового тиску знижується на 6 дБ за подвоєння відстані для сферичних акустичних хвиль та на 3 дБ для циліндричних хвиль. Це обумовлене ефектом дивергенції звуку. У реальних умовах навіть для відстаней 100–500 м відчутний ефект згасання має місце для складових звуку 2000 Гц і вище.

Тобто, для більш конкретних розрахунків:

$$\Delta L(f) = \alpha(f) r,$$

де α – коефіцієнт поглинання звуку у повітрі,
f – частота звукової хвилі.

Загальновизнаною практикою є визначення α за методикою міжнародної організації цивільної авіації (ICAO). У цій методиці враховується залежність коефіцієнта згасання від відносної вологості та температури повітря.

А це не є суттєвим для будівельних майданчиків, де житлова забудова перебуває на відносно невеликій відстані, а низькочастотна складова будівельного шуму істотна.

Зазвичай акустичні екрани мають прямокутну форму. Для таких конструкцій індекси зниження шуму можна визначити зі співвідношення:

$$\Delta L = -10\lg\left(\tau + \sum_{i=1}^{4} K_{\partial i} + \sum_{n=1}^{N} K_{вп}\right) + 10\lg\left(1 + \sum_{n=1}^{N} K_{вп}\right),$$

де τ – коефіцієнт звукопровідності матеріалу екрана; $K_{\partial i}$ – коефіцієнт дифракції і-тої кромки екрана; $K_{вп}$ – коефіцієнт відбиття n-ї поверхні; N – кількість відбиваючих поверхонь.

Коефіцієнт звукопровідності пов'язаний зі звукоізоляцією R співвідношенням:

$$\tau = 10^{-0,3R}.$$

Цей параметр для більшості стандартних матеріалів є довідковим.

Коефіцієнт дифракції розраховується як:

$$K_{\partial i} = \frac{\gamma_i (10 - 3\varphi\alpha) h}{f \gamma_0 (r_i + l_i - r_0)},$$

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, $h = 1$ для відстані від поверхні до нижньої кромки екрана більшої за 1,5 м.

Для точного розрахунку прогнозованої ефективності тонких екранів можна використати співвідношення, наведені у [8].

Було проведено експериментальні дослідження залежності ефективності акустичного екрана від його площі та порівняно результати з розрахунковими даними. Випробування здійснювалися за стандартною процедурою згідно ISO 140 у ревербераційній камері. Результати дослідження наведено на рис. 1.

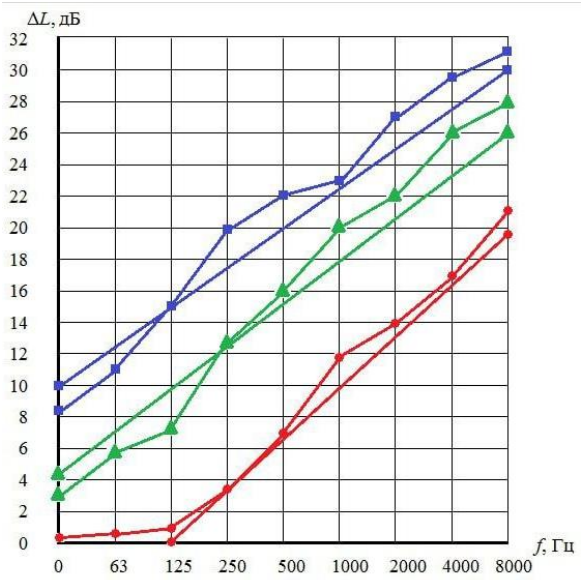


Рис.1. Залежність індексів зниження шуму від площі тонкого шумозахисного екрана:
— – 12 м², — – 6 м², — – 3м²,
ламана крива – експеримент, пряма – розрахунок

Усі екрани мали прямокутну форму однакових пропорцій довжини й ширини. Як видно з рис. 1, ефективність екранування очікувано підвищується зі збільшенням площі екрана. В цілому експериментальні дані збігаються з розрахунковими. Розбіжності можна пояснити похибкою вимірювальної апаратури та неврахованими акустичними ефектами.

У багатьох випадках шумозахисні конструкції мають технологічні опори. Найбільш це стосується виробів подвійного призначення – парканів, огороджувальних конструкцій тощо. Було досліджено ефективність тонкої захисної конструкції суцільної структури та кількома отворами (рис. 2).

Як видно з рис. 2, наявність отворів знижує ефективність захисту. При цьому це зниження зростає з підвищенням частот звукових хвиль. У практичній роботі слід враховувати, що суцільні тонкі екрани мають низьку ефективність у низькочастотній області звукового спектра. В той же час, як показано у [10], можна розраховувати параметри отворів, які забезпечують підвищення звукопоглинання низькочастотної складової шуму.

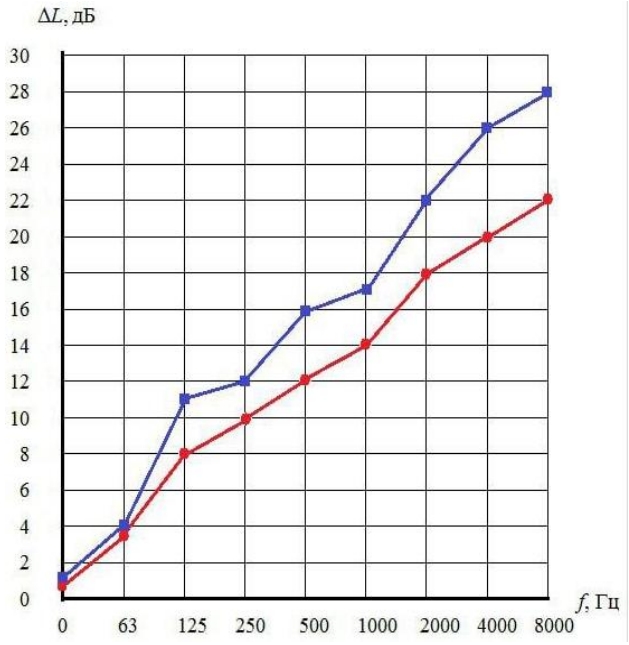


Рис. 2. Залежність індексів зниження шуму захисною конструкцією від наявності отворів:
— без отвору, — з отвором

Як зазначалося, в умовах обмежених просторів критичною є відстань від шумозахисного екрана до приймача звуку. Було визначено індекси зниження шуму у октавних смугах частот у залежності від відстані приймача до тонкого екрана (табл. 1).

Отримані дані свідчать, що у низькочастотній області звукового спектра у захищеній зоні відбувається підсилення шуму. Це обумовлене інтерференційними явищами, що може мати негативні наслідки для людей. Натурні вимірювання свідчать, що дані наведені у таблиці 1 не є характерними й не можуть розглядатися як узагальнені. У залежності від габаритних розмірів екрана частоти підсилення звуку або його зниження змінюються. Загальною тенденцією є збільшення ефективності екрана та зниження ефекту підвищення шуму у певних смугах. Найбільше значення для прояви інтерференції мають імпедансні характеристики відбивальних поверхонь. За сталих розмірів екранів, виготовлених з одного матеріалу на різних підстилаючих поверхнях ефект відбиття складає 5–7 дБ у низькочастотній області. Тобто, при визначені розмірів й позиціювання захисної конструкції необхідно врахувати наявність і показники усіх відбиваючих поверхонь.

Таблиця 1 – **Залежність індексу зниження шуму від відстані до тонкого екрана**

***l*, м**	**Δ*L*, дБ**							
	31	63	125	250	500	1000	2000	4000
10	10	5	-8	-5	0	10	8	8
20	8	0	-7	-5	2	8	10	12
30	0	0	-7	-2	5	8	10	12
40	0	-2	-8	-2	4	7	5	10
50	-2	-4	-15	0	2	6	5	8
60	-2	-5	-15	-2	2	4	5	7

У практичній роботі проєктування екранів під конкретний будівельний майданчик недоцільно. Це пояснюється тим, що відновлювальні роботи та реконструкція ушкоджених будівель тимчасові. Тому для даного виду діяльності доцільно створити більш-меш універсальний захисний екран прийнятної ефективності у низькочастотній області звукового спектра та високої ефективності для середніх і високих частот.

Відомо, що низькочастотний шум є значною складовою будівельного шуму. Тому для отримання прийнятних еквівалентних значень зниження шуму будівельної техніки захисні конструкції повинні бути комбінованими.

Усі захисні екрани мають кінцеві розміри, тому уникнути дифракційних явищ на кромках екранів практично неможливо. Необхідно максимально знизити проходження звуку крізь матеріал екрана, у тому числі й за рахунок резонансних явищ. Найбільш раціональним є обирання масогабаритних параметрів екранів із забезпеченням максимального поглинання звуку частоти максимальної амплітуди. При цьому один з боків конструкції або проміжок між двома поверхнями доцільно заповнювати шумопоглинальним матеріалом, наприклад, базальтовими волокнами. Це забезпечить прийнятні індекси зниження шуму у низькочастотній області звукового спектра.

Висновки

1. Визначено найбільш прийнятний у практичній діяльності математичний апарат для прогнозування захисних властивостей тонких акустичних екранів. Встановлено прийнятну збіжність розрахункових даних з експериментом. Отримано залежності індексів зниження шуму від площі акустичного екрана.

2. Проведено порівняння індексів зниження шуму суцільним екраном та екраном з отворами. Встановлено, що на середніх та високих частотах наявність отворів суттєво знижує ефективність захисного екрана. При цьому ефективність захисту у низькочастотній області звукового спектра незадовільна.

3. Показано, що на низьких частотах тонкий екран може підсилювати шум внаслідок інтерференційних явищ (від'ємні індекси зниження шуму). Для уникнення такого ефекту доцільно проєктувати захисні конструкції, налаштовані на резонансну частоту, якою є частота звуку найбільшої амплітуди.

СПИСОК ЛІТЕРАТУРИ

1. ДСН допустимих рівнів шуму в приміщеннях житлових та громадських будинків і на території житлової забудови. Наказ Міністерства охорони здоров'я України 22 лютого 2019 року № 463
2. Mostafa Mir, Farnad Nasirzadeh, SangHyun Lee, Densil Cabrera, Anthony Mills. Construction noise management: A systematic review and directions for future research, Applied Acoustics, V. 197, 2022, 108936, https://doi.org/10.1016/j.apacoust.2022.108936
3. Seulbi Lee, Sungjoo Hwang, Meesung Lee, Sungchan Lee. The impact of different types and levels of construction noise on physiological responses: Focusing on standardization and habituation, Sustainable Cities and Society, V. 112, 2024, 105644, https://doi.org/10.1016/j.scs.2024.105644
4. Jantien Stoter, Ravi Peters, Tom Commandeur, Balázs Dukai, Kavisha Kumar, Hugo Ledoux. Automated reconstruction of 3D input data for noise simulation. Computers, Environment and Urban Systems, V 80, 2020, 101424, https://doi.org/10.1016/j.compenvurbsys.2019.101424
5. Bhan Lam, Woon-Seng Gan, DongYuan Shi, Masaharu Nishimura, Stephen Elliott. Ten questions concerning active noise control in the built environment. Building and Environment, V.200, 2021, 107928, https://doi.org/10.1016/j.buildenv.2021.107928
6. 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, Ber- lin (2018)
7. 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
8. Шевченко Ю. С. Розробка моделей оцінки та підвищення ефективності зниження шуму транспортних потоків: дис. к.т.н., 21.06.01, Київ, 2016, 236 с.
9. 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
10. Glyva, V., Gusev, V., Biruk, Y., & Kashlev, M. (2024). Засади зниження рівнів низькочастотного звуку та інфразвуку у виробничих та побутових умовах. Системи управління, навігації та зв'язку. Збірник наукових праць, 1(75), 170-173. https://doi.org/https://doi.org/10.26906/SUNZ.2024.1.170

Received (Надійшла) 25.07.2025
Accepted for publication (Прийнята до друку) 22.10.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Глива Валентин Анатолійович – доктор технічних наук, професор, завідувач кафедри фізики, Київський національний університет будівництва та архітектури, Київ, Україна;
Valentyn 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.

Research on the effectiveness of thin noise barriers in confined spaces

Valentyn Glyva, Mykhailo Kashlev

Abstract. The possibilities of using thin noise barriers to reduce acoustic impact on workers and the population in confined spaces have been investigated. Such conditions are typical for construction noise when working in densely built-up areas. The most convenient mathematical apparatus for predicting the effectiveness of an acoustic screen is presented. The measured amplitude-frequency characteristics of noise reduction indices depending on the area of the acoustic screen are given. An acceptable convergence of calculated and experimental data has been established. It has been established that thin screens are insufficiently effective in the low-frequency part of the sound spectrum. The influence of openings in the protective structure on noise reduction indices has been investigated. It has been established that at medium and high frequencies, openings become critical for noise protection. At the same time, at low frequencies, they have practically no effect on the effectiveness of the screen. Noise reduction indices in octave frequency bands and their changes with distance from the screen to the sound receiver were measured. It was found that at low frequencies, sound levels can increase (negative indices) due to interference phenomena. It is proposed to give screens resonant properties in confined spaces. It is advisable to choose the sound frequency with the maximum amplitude as the resonant frequency. To increase the effectiveness of the screen in the mid and high frequency ranges, it is advisable to manufacture two-layer panels and fill the gap between the layers with noise-absorbing porous material.

Keywords: construction noise, acoustic screen, noise reduction index, limited space.

УДК 629.5.015.5-6:331.45 **doi: 10.26906/SUNZ.2025.4.164-166**

Г. Ю. Краснянський, Я. І. Бірук

Київський національний університет будівництва і архітектури, Київ, Україна

ОПТИМІЗАЦІЯ РОЗРАХУНКОВОГО АПАРАТУ ДЛЯ ПРОЕКТУВАННЯ ПОРИСТИХ ЗВУКОПОГЛИНАЮЧИХ МАТЕРІАЛІВ

Анотація. Головним недоліком існуючих моделей поширення звуку крізь звукопоглинальні пористі матеріали є складність їх практичного застосування. Усі розрахунки наведені в уявній формі, що збільшує обсяги обчислень й ускладнює процес автоматизації проєктування звукопоглинальних матеріалів потрібної ефективності. Тому доцільно перевести усі розрахунки у дійсну форму, що спростить процес автоматизації проєктування звукопоглинальних матеріалів і конструкцій. Показано, що використання емпіричних та напівфеноменологічних моделей для попереднього оцінювання звукопоглинальних властивостей пористих матеріалів утруднено через їхню залежність від численних параметрів, які неможливо виміряти безпосередньо. Запропоновано оптимізацію розрахункового апарату на основі моделі Джонсона-Шампу-Алларда-Лафаржа (JCAL). Показано, що модель JCAL може бути зведена до двох неакустичних параметрів – пористості і опору повітряному потоку, які достатньо точно визначаються експериментально, при врахуванні їх кореляції з чотирма параметрами, що залишилися. Пористість матеріалу легко визначається зважуванням тестових зразків, а опір повітряному потоку можна виміряти у імпедансній трубі стандартної конструкції або на лабораторній установці для визначення опору продування потоком повітря з використанням витрати повітря та давача тиску. Зменшення кількості параметрів моделі дозволяє прискорити прогнозування звукопоглинаючих характеристик та роблять модель застосовнішою до стратегій оптимізації та автоматизованого проектування захисних пористих матеріалів. Створення прикладного програмного забезпечення для автоматизації проєктування звукоізоляції з великою швидкістю обчислень дозволить оптимізувати обирання потрібних параметрів матеріалів методом перебору варіантів. Запропоновані зміни значно знижують обчислювальну складність, підвищуючи придатність розрахункового апарату для додатків реального часу та великомасштабного моделювання.

Ключові слова: поглинання звуку, пористі матеріали, акустичні моделі, розрахунковий апарат.

Вступ

У відповідності до міжнародних та національних стандартів і санітарних норм рівні шуму звукового діапазону жорстко регламентується [1]. У сучасних умовах задача зниження рівня звукового навантаження всередині будівель та споруд набуває особливої актуальності. Одним із шляхів її вирішення є використання пористих звукопоглинаючих матеріалів. Процес проектування таких матеріалів може бути суттєво спрощений за рахунок попереднього оцінювання їх поглинальних властивостей розрахунковими методами. Зазвичай використовувані для цих цілей емпіричні та напівфеноменологічні моделі непрактичні через їхню залежність від численних параметрів, які неможливо виміряти безпосередньо. Тому доцільним є вдосконалення математичного апарату для розрахункового проектування пористих звукопоглинаючих матеріалів.

Огляд літературних джерел. Напівфеноменологічні акустичні моделі для пористих матеріалів розроблялися з середини 20 століття. Вони поєднують емпіричні дані з теоретичними основами для опису взаємодії звуку з матеріалом. Ці моделі використовують набір із чотирьох-восьми неакустичних параметрів для апроксимації коефіцієнта поглинання звуку [2, 3]. Серед напівфеноменологічних моделей виділяється модель Джонсона-Шампу-Алларда-Лафаржа (JCAL) внаслідок її надійності та точності в описі поширення звуку в пористих середовищах у широкому діапазоні частот [4]. Основне обмеження моделі JCAL полягає в труднощах, пов'язаних з точним виміром та оцінкою неакустичних параметрів, від яких вона залежить, оскільки лише деякі з цих параметрів піддаються безпосередньому виміру [5]. Для прогнозування акустичних властивостей пористих матеріалів простіше використовувати емпіричні моделі, такі як моделі Делані-Безлі та інші, оскільки для них потрібний лише один неакустичний параметр – опір повітряному потоку [6, 7]. Проте точність прогнозування емпіричних моделей нестабільна. Ця нестабільність виникає через те, що ці моделі ґрунтуються на методах регресії з використанням великих масивів експериментальних даних для конкретних матеріалів. Таким чином, напівфеноменологічні та емпіричні моделі не є практичними рішеннями для оцінки акустичних властивостей пористих матеріалів. Отже, доцільним є вдосконалення цих моделей для подолання зазначених обмежень.

Метою дослідження є проведення оптимізації розрахункового апарату для проектування пористих звукопоглинаючих матеріалів.

Викладення основного матеріалу

Для оцінки звукопоглинаючих характеристик пористих матеріалів, як базову, будемо використовувати модель JCAL. Коефіцієнт поглинання звуку α визначається виразом:

$$\alpha = 1 - |R|^2 = 1 - \left|\left(Z_s - \rho_0 c_0\right)/\left(Z_s + \rho_0 c_0\right)\right|^2, \quad (1)$$

де R – коефіцієнт відбиття звуку; ρ_0 – густина повітря; c_0 – швидкість звуку в повітрі;
поверхневий акустичний імпеданс:

$$Z_s = Z_c \coth(kl); \quad (2)$$

характеристичний імпеданс:

$$Z_c = \sqrt{\rho(\omega)K(\omega)}; \quad (3)$$

комплексне хвильове число:

$$k = \omega\sqrt{\rho(\omega)/\ K(\omega)}, \quad (4)$$

де l – товщина пористого шару; ω – циклічна частота звукової хвилі; $\rho(\omega)$ – комплексна густина; $K(\omega)$ – комплексний динамічний модуль об'ємної пружності пористого матеріалу.

Відповідно до моделі JCAL:

$$\rho(\omega)=\frac{\alpha_\infty\rho_0}{\phi}\left[1-i\frac{\phi\eta}{k_0\rho_0\alpha_\infty\omega}\sqrt{1+i\frac{4k_0^2\rho_0\alpha_\infty^2\omega}{\eta\Lambda^2\phi^2}}\right], \quad (5)$$

$$K(\omega)=\frac{\gamma P_0/\phi}{\gamma-\dfrac{\gamma-1}{1-i\dfrac{\phi\eta}{k_0'\rho_0 P_r\omega}\sqrt{1+i\dfrac{4k_0'^2P_r\rho_0\omega}{\eta\Lambda'^2\phi^2}}}}. \quad (6)$$

Як видно із (5), (6) дані величини залежать від шести неакустичних параметрів, а саме – відкритої пористості ϕ, високочастотної межі звивистості α_∞, в'язкісної Λ та теплової Λ' характеристичних довжин та в'язкісної $k_0=\eta/\sigma$, пов'язаної з опором повітряному потоку σ, і теплової k_0' проникностей. Інші характеристики повітря, що входять в (5), (6), – η (динамічна в'язкість), γ (постійна адіабати), P_0 (атмосферний тиск), P_r (число Прандтля) – табличні величини, i – уявна одиниця. Основне обмеження моделі полягає у визначенні цих параметрів, лише деякі з яких можна виміряти безпосередньо. Крім того, комплексний вигляд співвідношень ускладнює їх практичне використання.

Підвищення практичності моделі за збереження точності може бути досягнуто при врахуванні кореляції пористості і опору повітряному потоку, які визначаються експериментально з мінімальною похибкою, з чотирма неакустичними параметрами, що залишилися. Для оцінки критичних параметрів моделі через ϕ і σ використаємо співвідношення [5]:

$$\alpha_\infty=1-11\cdot\ln\phi;\quad \Lambda'=2\Lambda;$$
$$\Lambda=\sqrt{\frac{8\eta(1-11\cdot\ln\phi)}{\sigma\phi}};\quad k_0'=\frac{\phi^2}{1-\phi}\cdot10^{-10}. \quad (7)$$

Тоді вирази (5) і (6) є такими:

$$\rho(\omega)=a\left(1-i\frac{\sqrt{1+ib}}{2b}\right); \quad (8)$$

$$K(\omega)=\frac{d}{1-f(1-im\sqrt{1+in})^{-1}}, \quad (9)$$

де $a=\dfrac{\rho_0(1-11\cdot\ln\phi)}{\phi};\ b=\dfrac{n\omega}{2\sigma};\ d=\dfrac{P_0}{\phi};\ f=1-\dfrac{1}{\gamma};$

$$m=\frac{(1-\phi)\eta}{\phi\rho_0\,\mathrm{Pr}\,\omega}\cdot10^{10};\ n=\frac{\phi^2\sigma\cdot10^{-10}}{8m(1-\phi)\eta(1-11\cdot\ln\phi)}. \quad (10)$$

Представимо (1) у вигляді:

$$\alpha=\frac{4t\,\mathrm{Re}(Z_s)}{|Z_s|^2+2t\,\mathrm{Re}(Z_s)+t^2}, \quad (11)$$

де $t=\rho_0c_0$; $\mathrm{Re}(Z_s)$ – дійсна частина Z_s.

Оскільки аналітичний вираз для α, який можна отримати із (11) з використанням (2) – (10), вкрай громіздкий, запишемо кінцеву формулу через проміжні величини. Введемо позначення:

$$P=\sqrt{1+ib};\ \ Q=\sqrt{1+in}. \quad (12)$$

Для дійсної і уявної (Im) частин P і Q маємо:

$$\mathrm{Re}(P)=\sqrt{\frac{\sqrt{1+b^2}+1}{2}};\ \ \mathrm{Im}(P)=\sqrt{\frac{\sqrt{1+b^2}-1}{2}}; \quad (13)$$

$$\mathrm{Re}(Q)=\sqrt{\frac{\sqrt{1+n^2}+1}{2}};\ \ \mathrm{Im}(Q)=\sqrt{\frac{\sqrt{1+n^2}-1}{2}}. \quad (14)$$

Тоді $$\rho(\omega)=a\left(\left(1+\frac{\mathrm{Im}(P)}{2b}\right)-i\frac{\mathrm{Re}(P)}{2b}\right); \quad (15)$$

$$K(\omega)=\frac{d(1+m(\mathrm{Im}(Q)-i\,\mathrm{Re}(Q)))}{1-f+m(\mathrm{Im}(Q)-i\,\mathrm{Re}(Q))}. \quad (16)$$

Далі, обчислюємо $\sqrt{\rho(\omega)K(\omega)}$, $\sqrt{\dfrac{\rho(\omega)}{K(\omega)}}$ і, використовуємо співвідношення:

$$\coth(x+iy)=\frac{\sinh(2x)-i\sin(2y)}{\cosh(2x)-\cos(2y)}. \quad (17)$$

Розраховуємо:

$$\coth\left(\omega l\sqrt{\rho(\omega)/K(\omega)}\right). \quad (18)$$

Після знаходження Z_s за (2) – (4), виділяємо $\mathrm{Re}(Z_s)$ і обчислюємо $|Z|^2=(\mathrm{Re}(Z))^2+(\mathrm{Im}(Z))^2$, які необхідні для розрахунку α.

Повна послідовна підстановка всіх формул призводить до багатокомпонентних виразів для $\mathrm{Re}(Z_s)$ і $|Z_s|^2$ через вкладені радикали, гіперболічні функції та операції з комплексними числами. На практиці для аналітичних обчислень доцільним є використання математичних пакетів, наприклад, MatLab, Mathematica або Maple. Для чисельних розрахунків оптимальним є використання Python з бібліотеками NumPy і SciPy для роботи з комплексними числами і гіперболічними функціями.

Послідовність обчислень:

1. Розраховують P і Q.

2. Знаходять $\rho(\omega)$, $K(\omega)$.

3. Обчислюють Z_s.

4. Знаходять α.

Це дозволить достатньо швидко отримати шуканий результат й обрати найбільш сприйнятливе співвідношення вихідних даних для проєктування звукопоглинального матеріалу потрібної ефективності.

Висновки

1. Недоліком моделі Джонсона-Шампу-Аллард-Лафаржа (JCAL) є залежність коефіцієнта звукопоглинання від шести неакустичних параметрів, які

не можуть бути визначені у лабораторних умовах прямими вимірюваннями, що ускладнює її практичне застосування.

2. Розроблений розрахунковий апарат оцінювання звукопоглинальних властивостей пористих матеріалів на основі моделі JCAL зводить розрахунки до двох параметрів – пористості й опору повітряному потоку, які можна виміряти достатньо точно із врахуванням їх кореляції з чотирма неакустичними параметрами, які залишаються.

3. Зменшення кількості параметрів моделі, що визначаються експериментально, дозволяє прискорити прогнозування звукопоглинаючих характеристик та роблять модель застосовнішою до стратегій оптимізації та автоматизованого проектування захисних матеріалів.

Список літератури

1. Directive 2003/10/EC – noise. Of 6 February 2003 on the minimum health and safety requirements regarding the exposure of workers to the risks arising from physical agents (noise) (Seventeenth individual Directive within the meaning of Article 16(1) of Directive 89/391/EEC). European Agency for Safety and Health at Work. Latest update: 19/03/2021. URL: https://osha.europa.eu/en/legislation/directives/82
2. Taban E., Tajpoor A., Faridan M. et al. Acoustic absorption characterization and prediction of natural coir fibers. Acoustics Australia. 2019. 47. P. 67–77. https://doi.org/10.1007/s40857-019-00151-8
3. Attenborough K., Bashir I., Taherzadeh S. Outdoor ground impedance models. The Journal of the Acoustical Society of America. 2011. 129 (5). P. 2806-2819. https://doi.org/10.1121/1.3569740
4. Niskanen M., Groby J.-P., Duclos A., Dazel O., Le Roux J. C., Poulain N., Huttunen T., L€ahivaara T. Deterministic and statistical characterization of rigid frame porous materials from impedance tube measurements. The Journal of the Acoustical Society of America. 2017. 142(4). P. 2407–2418. https://doi.org/10.1121/1.5008742
5. Yang Tao, M. Eser, Xiong Xiaoman, Groby J.-P., Schmid J. M., Maeder M., Chang Yu-Hao, Marburg S. The Journal of the Acoustical Society of America. 2025. 157. P. 3497–3507. https://doi.org/10.1121/10.0036644
6. Pelegrinis T., Horoshenkov K. V., Burnett A. An application of Kozeny–Carman flow resistivity model to predict the acoustical properties of polyester fibre. Applied acoustics. 2016. 101. P. 1–4. https://doi.org/10.1016/j.apacoust.2015.07.019
7. Burdeina N., Glyva V., Levchenko L., Krasnianskyi G., Biruk Y., Zozulya S., Zozulya L., Kashlev M., Grzelakowski T. Innovative approaches to designing sound insulation in historic buildings during reconstruction. International Journal of Conservation Science. 2025. Vol. 16, Special Issue. P. 373-382. https://doi.org/10. 36868/IJCS.2025.si.01

Received (Надійшла) 15.08.2025
Accepted for publication (Прийнята до друку) 08.10.2025

Відомості про авторів / About the Authors

Краснянський Григорій Юхимович – кандидат фізико-математичних наук, доцент, доцент кафедри фізики, Київський національний університет будівництва та архітектури, Київ, Україна;
Grygorii Krasnianskyi – Candidate of Physical and Mathematical Sciences, Associate Professor, Associate Professor of the Department of Physics, Kyiv National University of Construction and Architecture, Kyiv, Ukraine;
e-mail: krasnianskyi.giu@knuba.edu.ua, ORCID Author ID: https://orcid.org/0000-0002-2421-1270;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=6507131193.

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.

Optimization of the calculation apparatus for designing porous sound-absorbing materials

Grygorii Krasnianskyi, Yana Biruk

Abstract. The main drawback of existing models of sound propagation through sound-absorbing porous materials is the complexity of their practical application. All calculations are presented in an imaginary form, which increases the volume of calculations and complicates the process of automating the design of sound-absorbing materials of the required efficiency. Therefore, it is advisable to convert all calculations into a real form, which will simplify the process of automating the design of sound-absorbing materials and structures. It has been shown that the use of empirical and semi-phenomenological models for preliminary assessment of the sound-absorbing properties of porous materials is difficult due to their dependence on numerous parameters that cannot be measured directly. It is proposed to optimise the calculation apparatus based on the Johnson-Champ-Allard-LaFarge (JCAL) model. It is shown that the JCAL model can be reduced to two non-acoustic parameters – porosity and airflow resistance, which can be determined quite accurately experimentally, taking into account their correlation with the four remaining parameters. The porosity of the material is easily determined by weighing test samples, and the air flow resistance can be measured in a standard impedance tube or on a laboratory setup for determining air flow resistance using air flow and a pressure sensor. Reducing the number of model parameters speeds up the prediction of sound absorption characteristics and makes the model more applicable to optimisation strategies and automated design of protective porous materials. The creation of application software for the automation of sound insulation design with high calculation speed will allow the optimisation of the selection of the required material parameters by trial and error. The proposed changes significantly reduce the computational complexity, increasing the suitability of the calculation apparatus for real-time applications and large-scale modelling.

Keywords: sound absorption, porous materials, acoustic models, calculation apparatus.

УДК 537.87:669.162.12 (045) **doi: 10.26906/SUNZ.2025.4.167-170**

К. Д. Ніколаєв[1,2], А. С. Білик[1,3], К. М. Сапожников[1]

[1] Науково-дослідний інститут воєнної розвідки, Київ, Україна
[2] Міжрегіональна академія управління персоналом, Київ, Україна
[3]Київський національний університет будівництва і архітектури, Київ, Україна

ЗАСАДИ РОЗРОБЛЕННЯ ЕЛЕКТРОМАГНІТНИХ ЕКРАНІВ З МАЛИМИ МАСОГАБАРИТНИМИ ПАРАМЕТРАМИ

Анотація. Проаналізовано можливості розроблення матеріалу малої товщини для екранування електромагнітних випромінювань ультрависоких і вищих частот. Показано, що покриття завтовшки меншої за чверть довжини падаючої електромагнітної хвилі – 400–500 мкм не забезпечують малих коефіцієнтів відбиття при нанесенні на металеву поверхню. Зниження коефіцієнтів відбиття до 0,7–0,6 відбувається випадковим чином на різних частотах. Це пов'язане з резонансними явищами у покритті. Частотна смуга мінімального відбиття вузька й не має практичного значення. Показано, що забезпечити прийнятний баланс магнітних та електрофізичних властивостей тонкого покриття для мінімізації коефіцієнтів відбиття складно. Це обумовлено залежністю показників кінцевого продукту не тільки від магнітних та діелектричних властивостей компонентів та їх концентрацій, а й від морфології частинок, їх розподілу у тілі матриці, режиму полімеризації. Крім того магнітні й електрофізичні властивості мають амплітудно-частотну залежність. Обґрунтовано, що для забезпечення малих коефіцієнтів відбиття електромагнітних хвиль й високого поглинання електромагнітної енергії доцільно застосовувати багатошарову структуру. Шари повинні мати різні магнітні й електрофізичні властивості. При цьому найменші концентрації наповнювача – у верхньому шарі. Ці концентрації повинні мати значення на межі перколяційного ефекту – різкого підвищення електропровідності композиції. Ці концентрації можна оціночно визначити із застосуванням теорії протікання електродинаміки суцільних середовищ. За можливості між шарами доцільно передбачити діелектричний прошарок. Наявність границь розділу підвищує загальну ефективність захисного покриття.

Ключові слова: електромагнітне випромінювання, електромагнітний екран, коефіцієнт відбиття.

Вступ

Зниження рівнів електромагнітних випромінювань відповідними матеріалами і конструкціями поступово стає багатовекторною задачею. Це обумовлене як загальним підвищенням електромагнітного фону, у навколишньому середовищі, так і насиченістю побутових та виробничих приміщень електронною технікою, яка генерує електромагніті випромінювання широкого частотного діапазону. При цьому спостерігається стійка тенденція до підвищення робочих частот усіх типів бездротового зв'язку, радіолокаційного обладнання тощо. Така ситуація вимагає наявності ефективних матеріалів для зниження рівнів електромагнітних випромінювань у будь-яких умовах. При цьому для випромінювань ультрависоких і вищих частот критичним є коефіцієнт відбиття, який впливає на перерозподіл потоків випромінювання у просторі. Складність розробки відповідних матеріалів полягає у тому, що значна частина таких досліджень має закритий характер через застосування у військовій сфері – забезпечення електромагнітної сумісності електронного обладнання та електромагнітного камуфляжу. Технічно та економічно доцільно забезпечувати захист окремих об'єктів та технічних засобів. Такі об'єкти та пристрої зазвичай мають складну конфігурацію поверхні, тому масогабаритні параметри засобів захисту є критичними. Зазвичай дослідження у цьому напрямі стосуються конкретної вузької задачі. Тому не сформований системний підхід до розроблення й впровадження електромагнітних екранів потрібної ефективності у залежності від умов їх використання. Це обумовлює актуальність визначення базових засад проєктування та виготовлення електромагнітних екранів малих масогабаритних параметрів.

Огляд досліджень і розробок

Металеві й металізовані поверхні мають високі загальні коефіцієнти екранування електромагнітних випромінювань [1, 2]. Але для частот випромінювання дуже високих, ультрависоких і вищих – така ефективність досягається, в основному, за рахунок відбиття електромагнітних хвиль. Уникнення такого явища можливо за рахунок чвертьхвильового захисту, виходячи з фундаментальних фізичних принципів [3]. Такі конструкції передбачають наявність діелектричного зазору між шарами конструкції, кратного чверті довжини падаючої електромагнітної хвилі. Це збільшує габаритні розміри конструкцій і робить її практично монохромною (розрахованою на екранування випромінювання однієї частоти). Для вирішення задач високої ефективності екранування за рахунок поглинання електромагнітної енергії розроблено багато композиційних матеріалів [4, 5]. У частини цих матеріалів співвідношення коефіцієнтів поглинання та відбиття електромагнітних хвиль прийнятне, але товщина матеріалів 5 мм і більше. Крім того більшість таких матеріалів низькотехнологічні у практичному використанні й схильні до деградації під впливом фізико-хімічних чинників. Найбільш зручними для використання є рідкі композиції, які можна наносити на будь-які поверхні. У роботі [6] запропоновано використання стандартних геополімерних фарб (матриця) й магнетиту (наповнювач). Але за малих концентрацій наповнювача недостатня загальна ефективність екранування з низьким коефіцієнтом відбиття. За великих концентрацій, 40 % за масою і більше, різко зростає коефіцієнт відбиття, що у більшості випадків є небажаним. Більш ефективні й збалансовані рідкі суміші, описані у [7, 8]. Але вони не враховують зміни коефіцієнтів відбиття зі

зміною електрофізичних характеристик компонентів композиції. Крім того, ці матеріали мають складні технології виготовлення, наприклад перетворення полівінілбутералю, і високі вартості. Найбільш раціональним шляхом підвищення загальної ефективності захисного матеріалу разом з малим коефіцієнтом відбиття електромагнітних хвиль є створення у товщі матеріалу градієнту електрофізичних властивостей. У роботі [9] це реалізовано послідовним нанесенням шарів з різною концентрацією наповнювача. Дана робота повністю експериментальна й не дає можливість раціоналізувати показники кожного шару й знизити загальну товщину захисного покриття. Тому доцільно проаналізувати показники складу й ефективності усіх доступних матеріалів і визначити загальні засади проєктування та виготовлення тонких матеріалів для екранування електромагнітних випромінювань.

Викладення основного матеріалу

Мінімізація товщини електромагнітного екрана у загальному випадку є задачею оптимізації кількох показників захисного матеріалу або конструкції. Така оптимізація виконується експериментально або теоретично у залежності від умов і цілей застосування екранування. Головними показниками екрана є загальний коефіцієнт екранування, внесок у нього екранування за рахунок відбиття електромагнітних хвиль та поглинання електромагнітної енергії. Усі ці показники залежать від магнітних та електрофізичних характеристик матеріалу. При цьому поглинання визначається показниками усього матеріалу, а відбиття – властивостями поверхневого шару – поверхневим імпедансом.

На сьогодні багато уваги приділяється захисту від електромагнітних випромінювань ультрависоких та вищих частот. Більшість матеріалів мають достатньо високі коефіцієнти відбиття електромагнітних хвиль малої довжини, а у конструкційних сплавів він має значення до 0,9. Це призводить до перерозподілу випромінювання у просторі, підвищуючи ступені впливу випромінювання на людей, погіршуючи умови експлуатації електронного обладнання та підвищуючи електромагнітну помітність рухомих об'єктів.

Найбільш надійним способом мінімізації коефіцієнта відбиття є створення покриття, товщина якого дорівнює чверті довжини падаючої електромагнітної хвилі. Хвилі, які відбиваються від зовнішньої та внутрішньої поверхонь матеріалів взаємно гасяться внаслідок інтерференційних явищ. Але, наприклад, для електромагнітного випромінювання частотою 10 ГГц (λ = 30 мм) захисний шар повинен мати товщину 7,5 мм. Це не завжди прийнятно. Товщину покриття можна зменшити за певного співвідношення магнітних та діелектричних характеристик шару:

$$d = c \Big/ \left(4 f \sqrt{\varepsilon\mu}\right),$$

де d – товщина покриття, c – швидкість розповсюдження електромагнітних хвиль, f – частота електромагнітного випромінювання, ε, μ – відносні діелектрична та магнітна проникності матеріалу. Наприклад, якщо забезпечити ε = 6, μ = 1,2, то товщина захисного шару може складати 3 мм. Очевидно, що потрібні співвідношення діелектричних та магнітних властивостей можна отримати тільки у композиційних матеріалах. Але їх проєктування складне через низку чинників. Визначення прогнозованої діелектричної проникності із застосуванням співвідношення Максвелла-Гарнета дає значні похибки й розбіжності з експериментом:

$$\frac{\varepsilon - \varepsilon_d}{\varepsilon + 2\varepsilon_d} = \upsilon_m \frac{\varepsilon_m - \varepsilon_d}{\varepsilon_m + 2\varepsilon_d},$$

де ε_d, ε_m – діелектричні проникності матриці та наповнювача, υ_m – об'ємна доля наповнювача у діелектричні матриці. Це ж стосується формули Оделевського для визначення магнітної проникності композиції.

$$\varepsilon = \varepsilon_d \left[1 + \frac{\upsilon_m \left(\varepsilon_m - \varepsilon_d\right)}{\left(1 - \upsilon_m / \upsilon_k\right) F \left(\varepsilon_m - \varepsilon_d\right) + \varepsilon_d}\right],$$

де ε_d, ε_m – діелектричні проникності матриці та наповнювача, υ_m – об'ємний вміст наповнювача, υ_k – критичний об'ємний вміст наповнювача, за якого екрануючі частинки контактують між собою, F – коефіцієнт деполяризації.

Навіть введення у формулу показника деполяризації F (врахування морфології частинок наповнювача у діелектричній матриці) не вирішує проблеми. Крім того, значний вплив на електрофізичні та магнітні властивості композиції має рівномірність розподілу частинок у матриці. Це вимагає застосування при виготовленні вихідної суміші швидкісних змішувачів – дисольверів. Підвищити точність розрахунків можна застосуванням співвідношень електродинаміки суцільних середовищ [10]. Але це вимагає великих обсягів обчислень й наявності експериментальних даних щодо питомої провідності (питомого опору) композиції з відомим об'ємний вмістом екрануючого матеріалу у діелектричній матриці. Наявність діелектричних та магнітних властивостей у композиції з відчутною електропровідністю матеріалу забезпечують ферити на основі заліза та нікелю. Було досліджено можливості мінімізації коефіцієнтів відбиття електромагнітних хвиль покриттям з товщинами, меншими за чверть довжини падаючої хвилі. Покриття товщиною 400–500 мкм наносилося на металеву поверхню й досліджувався коефіцієнт відбиття (рис. 1). Як видно з рис. 1, різке зниження коефіцієнтів відбиття має випадковий характер (у залежності від концентрацій наповнювача) й пов'язане з резонансними явищами. Смуги малого коефіцієнта відбиття надзвичайно вузькі на значеннях, які мають практичне значення.

Це значення, нижчі за 0,7. Усі джерела електромагнітних випромінювань, які формально працюють на фіксованій частоті, мають дещо відмінні робочі частоти. Тому у більшості випадків зниження коефіцієнтів відбиття за рахунок резонансних явищ не ефективне. Спроби підібрати потрібне співвідношення компонентів для мінімізації коефіцієнтів відбиття зазвичай виявляються невдалими. Відомо, що ефективність відбиття залежить від поверхневого імпедансу матеріалу. Його наближення до імпедансу середовища розповсюдження електромагнітних хвиль 377 Ом. У загальному випадку коефіцієнт відбиття визначається значенням $\sqrt{\mu/\varepsilon}$, де μ, ε – абсолютні магнітна та діелектрична проникності матеріалу.

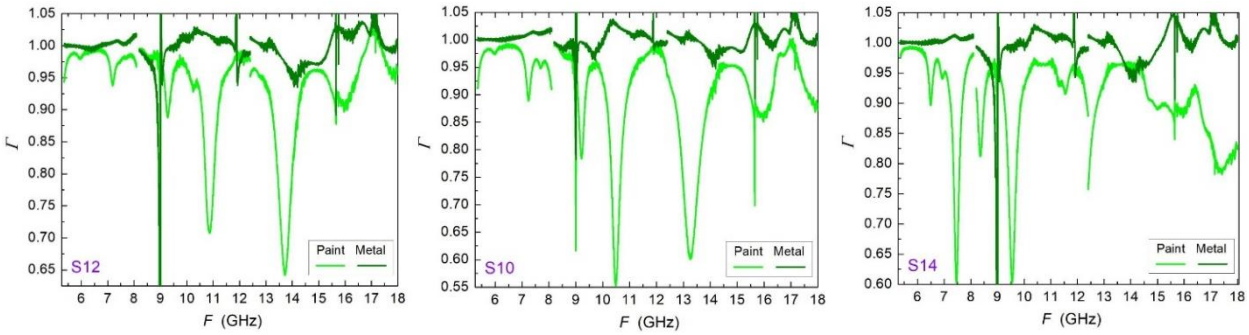


Рис. 1. Залежність коефіцієнтів відбиття від довжини падаючої електромагнітної хвилі: ─ зразок без покриття, ─ зразок з покриттям

Практично він розраховується порівнянням поверхневого імпедансу матеріалу з показником повітря: $K_e = (Z_m - Z_n)/(Z_m + Z_n)$, де Z_n, Z_m – імпеданси повітря та матеріалу, а $Z_m = \sqrt{2f\mu/\sigma}$, де σ – питома електропровідність матеріалу.

Фактично питома електропровідність визначає ступінь діелектричності матеріалу.

Для отримання прийнятного коефіцієнта відбиття необхідно наблизити значення діелектричних та магнітних проникностей. На практиці, навіть за різних вмістах магнітних та діелектричних домішок вони сильно відрізняються. На рис. 2 наведено виміряні співвідношення цих показників.

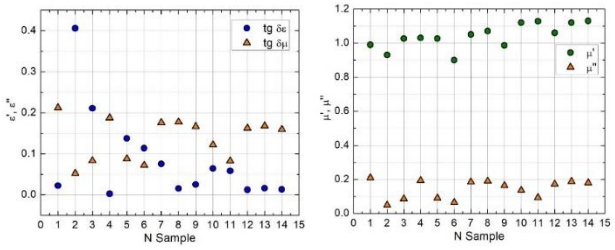


Рис. 2. Значення діелектричної та магнітної проникностей композиційних матеріалів з різним вмістом магнітного наповнювача

При цьому $Z_m = \sqrt{(\mu' - j\mu'')/(\varepsilon' - j\varepsilon'')}$, де μ′, ε′ – дійсні складові магнітної та діелектричної проникності, μ″, ε″ – уявні складові.

Зі зміною частоти електромагнітного випромінювання змінюється і його магнітні та електричні властивості. При цьому при зниження коефіцієнта відбиття електромагнітних хвиль знижується й коефіцієнт поглинання електромагнітної енергії. За умови нанесення захисного покриття на неметалеву поверхню це є небажаним. Тому найбільш ефективним покриттям є покриття з градієнтом магнітних та електричних властивостей по глибині захисного шару. Реалізувати це для тонкого композиційного матеріалу практично неможливо. Тому для досягнення прийнятного ефекту доцільно робити покриття принаймні двошаровим. Для мінімізації коефіцієнта відбиття верхній шар покриття повинен мати концентрацію провідної субстанції, яка може мати й магнітні властивості, на межі перколяційного ефекту – різкого зростання електропровідності з певної концентрації наповнювача у діелектричній матриці. Цей ефект настає на певній частоті електромагнітного випромінювання. Концентрація наповнювача у наступному шарі збільшується, що повинне відповідати нижчій частоті випромінювання. Такий підхід збільшує частотну смугу ефективності електромагнітного екрана. Слід враховувати, що поріг протікання електричного струму у кожному окремому випадку слід визначати експериментально. Цей показник залежить не тільки від електрофізичних показників компонентів і матеріалу в цілому, а й від дисперсності наповнювача, форми частинок та характеру їх розподілу у діелектрику. Орієнтовні концентрації наповнювача за об'ємом можливо оцінити з використанням співвідношень теорії протікання [10], але лабораторні дослідження є обов'язковими. За можливості (певних максимальних товщин матеріалів) між захисними шарами доцільно передбачити узгоджувальний діелектричний шар. Наявність границь розділу підвищує загальну ефективність захисного шару.

Висновки

1. Встановлено, що захисні покриття товщиною, меншою за чверть довжини падаючої електромагнітної хвилі при нанесенні на металеві поверхні можуть знижувати коефіцієнти відбиття електромагнітних хвиль до 0,7–0,6 випадковим чином, що пов'язане з резонансними явищами. Частотні смуги зниження коефіцієнтів відбиття вузькі й непередбачувані. Значного зниження коефіцієнтів відбиття за різних концентрацій провідної та магнітної субстанції у композиції не спостерігається.

2. Визначено, що досягти балансу співвідношень магнітних та електрофізичних показників кінцевого продукту розрахунковими та експериментальними методами складно через залежність цих параметрів не тільки від значень окремих компонентів композицій, а й морфології магнітних та провідних частинок, їх розподілів у тілі діелектричної матриці, режиму полімеризації тощо.

3. Для зниження коефіцієнтів відбиття електромагнітних хвиль й підвищення поглинання електромагнітної енергії запропоновано застосування багатошарового тонкого покриття. Концентрація магнітної та провідної субстанції у зовнішньому шарі найнижча й зростає у внутрішніх шарах. Умовою ефективності покриття є концентрація провідних частинок у кожному шарі на межі перколяційного ефекту. Оцінка необхідної концентрації наповнювача можлива із застосуванням теорії протікання, виходячи зі співвідношень електродинаміки суцільних середовищ.

СПИСОК ЛІТЕРАТУРИ

1. Панова О.В., Тихенко О.М., Ніколаєв К.Д., Ходаковський О.В., Сапельнікова О.Ю. Дослідження захисних властивостей металевих електромагнітних екранів та визначення умов їх максимальної ефективності. Системи управління, навігації та зв'язку. 2019. Вип. 5(57). С. 103−107.
2. Глива В.А., Панова О.В., Тихенко О.М., Левченко Л.О., Колумбет В.П. Дослідження амплітудно-частотних залежностей захисних властивостей магнітних екранів на основі аморфних сплавів. Системи управління, навігації та зв'язку. 2019. Вип. 6(58). С. 102−107.
3. Glyva V., Levchenko L., Panova O., Tykhenko O., Radomska M. The composite facing material for electromagnetic felds 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
4. Tudose I.V., Mouratis K., Ionescu O.N., Romanitan C., Pachiu C., Popescu M., Khomenko V., Butenko O., Chernysh O., Kenanakis G., Barsukov V.Z., Suchea M.P., Koudoumas E. Novel Water-Based Paints for Composite Materials Used in Electromagnetic Shielding Applications. Nanomaterials. 2022, 12(3). P. 487. https://doi.org/10.3390/nano12030487
5. Alina Ruxandra Caramitu, Ioana Ion, Adriana Mariana Bors, Violeta Tsakiris, Jana Pintea, Ana-Maria Daniela Caramitu. Preparation and Spectroscopic Characterization of Some Hybrid Composites with Electromagnetic Shielding Properties Exposed to Different Degradation Factors. MATERIALE PLASTICE. 2023. 59. 82-94 https://doi.org/10.37358/MP.22.4.5627
6. 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
7. 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
8. Butenko O., Boychuk V., Savchenko B., Kotsyubynsky V., Khomenko V., Barsukov V. Pure ultrafine magnetite from carbon steel wastes. Materials Today: Proceedings. 2019. V. 6, pp. 270–278
9. Бурдейна Н.Б., Бірук Я.І., Ніколаєв К.Д. (2023). Розроблення матеріалів багатошарової структури градієнтного типу на основі рідких композицій для екранування електромагнітних полів. Екологічна безпека та природокористування. 45 (1). С. 68–75. https://doi.org/10.32347/2411-4049.2023.1.68-75.
10. G. Krasnianskyi, V. Glyva, N. Burdeina, Y. Biruk, L. Levchenko, O. Tykhenko. 2024. Methodology For Designing Facing Building Materials with Electromagnetic Radiation Shielding Functions. INTERNATIONAL JOURNAL OF CONSERVATION SCIENCE. Volume 15, Special Issue 1, 2024: 53-62. DOI: 10.36868/IJCS.2024.SI.05

Received (Надійшла) 25.07.2025
Accepted for publication (Прийнята до друку) 08.10.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Ніколаєв Кирило Дмитрович – доктор наук з державного управління, кандидат сільськогосподарських наук, професор кафедри публічного адміністрування, Міжрегіональна академія управління персоналом, старший науковий співробітник Науково-дослідного інституту розвідки, Київ, Україна;
Kyrylo Nikolaiev – Doctor of Science in Public Administration, PhD (Agrarian Sciences), Professor of the Department of Public Administration, Interregional Academy of Personnel Management, Senior Researcher at the Defence Intelligence Research Institute, Kyiv, Ukraine;
e-mail: nikolaevkirill@gmail.com ORCID ID: https://orcid.org/0000-0003-0404-6113,
Scopus ID https://www.scopus.com/authid/detail.uri?authorId=42762116000#.

Білик Артем Сергійович – кандидат технічних наук, доцент, Київський національний університет будівництва та архітектури, Начальник науково-дослідної лабораторії Науково-дослідного інституту воєнної розвідки, Київ, Україна;
Artem Bilyk – Candidate of Technical Sciences, PhD, Associate Professor, Kyiv National University of construction and architecture, Head of the Scientific Research Laboratory in the Defence Intelligence Research Institute, Kyiv, Ukraine;
e-mail: artem.bilyk@gmail.com ORCID ID: https://orcid.org/0000-0002-9219-920X.

Сапожников Костянтин Миколайович – начальник відділу, Науково-дослідний інститут воєнної розвідки, Київ, Україна;
Kostiantyn Sapozhnykov – Head of Department, Defence Intelligence Research Institute;
e-mail: sapozhnykov_kos@meta.ua, ORCID ID: https://orcid.org/0000-0003-0259-3690.

Principles of development of electromagnetic screens with small dimensions and weight

Kyrylo Nikolaiev, Artem Bilyk, Kostiantyn Sapozhnykov

Abstract. The possibilities of developing thin-film materials for shielding ultra-high and higher frequency electromagnetic radiation are analysed. It is shown that coatings with a thickness less than a quarter of the wavelength of the incident electromagnetic wave (400–500 μm) do not provide low reflection coefficients when applied to a metal surface. The reduction of reflection coefficients to 0.7–0.6 occurs randomly at different frequencies. This is due to resonance phenomena in the coating. The frequency band of minimum reflection is narrow and has no practical significance. It has been shown that it is difficult to achieve an acceptable balance of magnetic and electrophysical properties of a thin coating to minimise reflection coefficients. This is due to the dependence of the final product's performance not only on the magnetic and dielectric properties of the components and their concentrations, but also on the morphology of the particles, their distribution in the matrix body, and the polymerisation mode. In addition, magnetic and electrophysical properties have an amplitude-frequency dependence. It has been proven that in order to ensure low electromagnetic wave reflection coefficients and high electromagnetic energy absorption, it is advisable to use a multilayer structure. The layers must have different magnetic and electrophysical properties. At the same time, the lowest filler concentrations are in the upper layer. These concentrations should be at the limit of the percolation effect – a sharp increase in the electrical conductivity of the composition. These concentrations can be estimated using the theory of electrodynamics of continuous media. If possible, it is advisable to provide a dielectric layer between the layers. The presence of boundaries increases the overall effectiveness of the protective coating.

Keywords: electromagnetic radiation, electromagnetic screen, reflection coefficient.

UDC 621.396.67 **doi: 10.26906/SUNZ.2025.4.171-175**

Oleh Berdnykov[1], Serhii Mazor[1], Tetiana Khranovska[1], Pavlo Dimitrov[2]

[1] Institute of Special Communications and Information Protection of the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine
[2] Defence Intelligence Research Institute, Kyiv, Ukraine

SUBSTITUTION OF RESEARCH ON ANTENNA SYSTEMS IN SHORTWAVE AND ULTRASHORTWAVE RANGES

Abstract. To improve the short-wave antenna based on low-positioned vibrators, an analysis of current trends in the development of new and modernization of existing antenna systems for short-wave radio communications was conducted. The possibility of studying the use of low-positioned horizontal vibrators is due to the fact that it is possible to create rather complex antenna arrays with large gain coefficients from available materials at ionospheric wave antenna sites and not only for them. Approaches to improving antenna systems depending on the type of radio wave propagation, physical parameters of antenna-feeder devices and their directional properties were analyzed. In this regard, an approach to antenna synthesis and the use of this approach in further work to improve the antenna system and obtain the necessary directional characteristics for organizing stable radio communications were substantiated. It is proposed to develop a method of constructive synthesis based on the methods of induced electromotive forces and mirror reflection. At the same time, a set of scientific tasks has been defined to achieve new results, classical mathematical methods have been selected, with the help of which it will be possible to calculate the electromagnetic field created by an improved antenna system. To achieve the set goal, the main directions have been defined, such as: improving the mathematical model of the antenna system based on low-lying radiators; the next step is to improve the mathematical method for calculating the electromagnetic field taking into account the radial components of the electric and magnetic fields; and, as a result, to obtain a convenient tool, to develop a methodology for the constructive synthesis of an improved short-wave antenna system from low-lying radiators with the ability to change the direction of the antenna radiation pattern in the vertical and horizontal planes. The article has formed an objective function that determines the main parameters that need to be achieved. Controlled and uncontrolled variables have been determined, on which the objective function of the specified process directly depends. In conclusion, an approach to calculating the efficiency indicator for evaluating the obtained parameter value from their previous values is presented.

Keywords: mobile radio communication, short-wave antenna, low-lying vibrators, induced electromotive force method, mirror reflection method, antenna systems, gain.

Introduction

In conditions of maneuverable fast-paced modern combat with a sharp change in the situation and in the absence of a continuous line of contact of troops, reliable high-quality radio communication is a guarantee of stable and flexible command of troops.

The main trends in the development of military radio communication means of the tactical command link are a reduction in their weight and dimensions, an increase in the duration of autonomous operation, as well as an improvement in the main technical characteristics of such means, due to the improvement of systems and individual elements of these means. On the one hand, these trends meet the requirements for military radio stations, and on the other hand, the possibility of improving their characteristics due to the rapid growth of technologies in modern technology. Therefore, the creation of modern means of communication and the modernization of existing radio stations requires the development of new and if possible, improved standard antenna systems [1].

Statement of the problem

The search for the optimal variant of the antenna type, its parameters, design, technology, element base, metrological support, etc., which best meet the requirements, forms the basis of the design (synthesis) of promising antenna systems. Thus, in the theory of antenna technology, a number of scientific tasks arise that are related to the need to further improve the general theory of antennas, methods for their calculation, finding new ways to build antenna systems and solving no less important design and technological problems [2–4].

Design and construction of antennas according to the specified requirements, that is, the synthesis of antenna systems, is the main task of any process of developing such systems. The process of performing this task can be conditionally divided into two parts - the development of theoretical principles (mathematical methods) and practical implementation (design and technological implementation) of antennas (Fig. 1).

The theoretical (physical-mathematical) theory of antenna system synthesis usually consists of two separate tasks, namely: solving the external task – determining currents (electromagnetic fields) according to given directivity characteristics; solving the internal task – determining the elements of the antenna design, excitation device, distribution systems and others [5].

These two tasks are considered independently of each other in most cases. For the construction of antenna systems, it is advisable to use constructive

synthesis of antennas. According to the views of well-known experts in the field of antenna technology, constructive synthesis is a process by which, based on the specified requirements for electrical characteristics, a design solution for an antenna with the appropriate technical characteristics is found [2, 6]. Tactical control link communications equipment in service has a standard set of antenna-feeder devices with limited technical characteristics. At the same time, these equipment cannot operate without changing the height of the antenna devices, which complicates the formation of the desired antenna beam pattern and does not provide control over it when the situation changes. Today, there is no method for forming the necessary beam pattern taking into account the antenna orientation and its height of suspension.

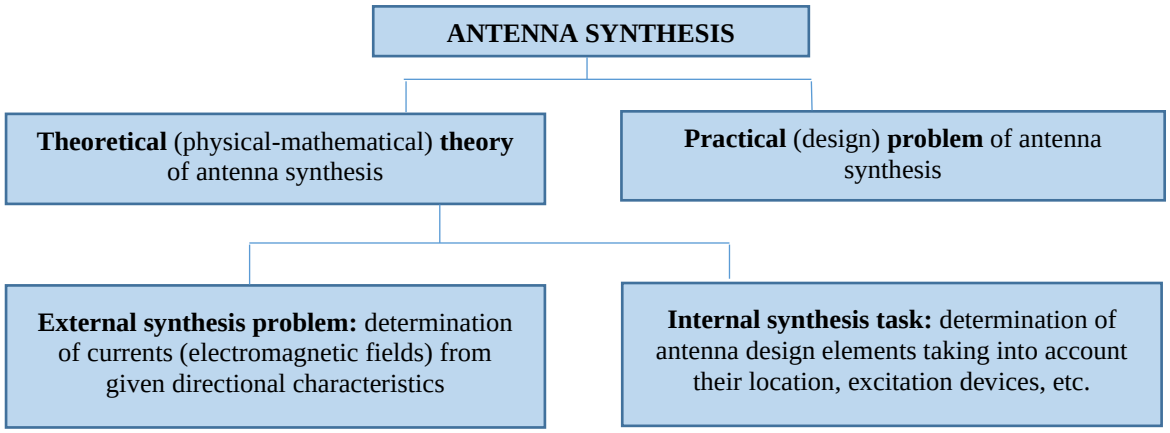


Fig. 1. The principle of antenna construction

Presentation of the main material

To achieve operational and economic benefits of modernizing mobile radio communications, it is advisable to pay attention not only to the development of new antenna systems, but also to the improvement of existing ones, which have a limited number of standard antenna-feeder devices, by improving the technical characteristics of the antenna to obtain the required width of the directivity diagram, as well as taking into account the direction of information transmission and the influence of the earth's surface on the propagation of radio waves. Currently, stationary speakers made from improvised materials are used for the shortwave (HF) and ultrashortwave (VHF) ranges. For example, for two parallel lines, one of which is shown in Fig. 2, with 10 vibrators in each, it is possible to obtain a gain of up to 12 dB. The line is a vibrator (conductors from the antenna cable) that are suspended above the ground on wooden poles (Fig. 2).

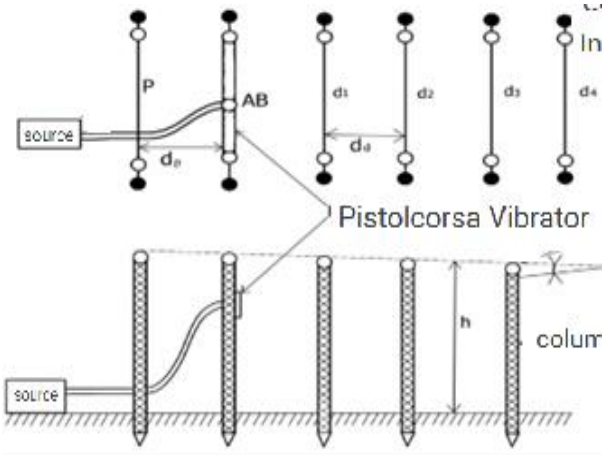


Fig. 2. Antenna system made of horizontal vibrators

Such antenna system can be installed because radio equipment can be moved approximately once every 2-3 days, and in territorial defense once a month (from the experience of practical application of communication systems).

Such AS can be used for both automobile communication stations and portable ones, the input/output impedance of which is 50 Ohms [3, 4].

Analysis of the design features of antenna systems of mobile radio communication equipment shows that for vibrator-type antennas with a width of the main lobe of the DS Θ=80o, a radio communication range of up to 50 km is provided. Under the conditions of reducing the width of the main lobe of the DS to 50o, the radio communication range increases to 70 km, with unchanged transmitter power.

Clarifying the essence of the processes of propagation of electromagnetic waves directly near the planet Earth and in its surrounding space plays an important role in organizing communication [3, 4].

Radio wave propagation is a spatial process that encompasses the propagation of electromagnetic waves of the radio range in the atmosphere, outer space and the Earth's interior, while electromagnetic waves can propagate along rectilinear trajectories, skirting the convex surface of the Earth, reflecting from the ionosphere, etc.

The methods of electromagnetic waves propagation significantly depend on the wavelength λ, the illumination of the atmosphere by the Sun, the location of radio paths relative to the Earth's surface.

A significant role in the process of electromagnetic waves propagation relative to the Earth's surface is played by a part of space that has the shape of an ellipsoid of rotation, at the foci of which the transmitter and receiver are located, while radio waves can be weakened or amplified.

In addition, since the Earth has a heterogeneous surface, when assessing the impact of the Earth's surface on the propagation of electromagnetic waves, it is advisable to take into account such physical effects and phenomena as the properties of the Earth's surface (type of soil, the presence of water, forest resources, urban developments, transport arteries, power lines, etc.); properties of natural and artificial objects and the boundaries of the distribution of boundaries between them and the multi-beam formation of the final signal; multi-beam formation of the weakened power of radio waves due to their absorption by rain, snow, dust; reflection of radio waves from rain, snow, dust, flocks of birds; curvature of the propagation paths of radio waves due to the heterogeneity of the layers of the atmosphere [3, 4].

Fig. 3 shows a typical diagram of a short-wave radio line taking into account reflection from the F2 ionosphere layer. The figure also shows that in short-wave radio communication, space and ground waves are necessarily present, and there is also a ring zone of silence around the receiver, when the ground wave is no longer possible, and the space wave is not yet possible. This is due to the significant attenuation of the ground wave along the earth's surface and the angles of incidence and reflection of the space radio wave from the ionosphere layers [7, 8].

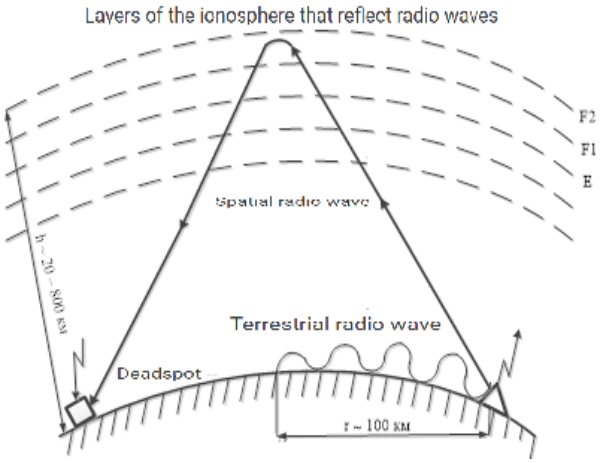


Fig. 3. Typical shortwave radio line diagram

Thus, an urgent scientific task arises regarding the constructive synthesis of an improved AC for short-wave radio communication and the development of an appropriate methodology that takes into account the design features of the AC, the direction of information transmission and the influence of the earth's surface on the propagation of radio waves, which will allow for the rapid construction of a short-wave antenna based on low-lying vibrators.

Therefore, the purpose of the research is to develop a methodology for the constructive synthesis of an improved AC for a mobile radio communication device, which will allow the head of the communication station to calculate a short-wave AC based on low-lying vibrators, in accordance with the specified technical characteristics, and the service personnel to quickly deploy the antenna.

To achieve the research goal, it is necessary to:

– improve the mathematical model of the antenna system based on low-lying emitters;

– improve the mathematical method for calculating the electromagnetic field taking into account the radial components of the electric and magnetic fields;

– develop a method for the constructive synthesis of an improved short-wave antenna system from low-lying emitters with the ability to change the direction of the antenna directivity diagram in the vertical and horizontal planes.

Existing programs for calculating the directivity characteristics of vibrator antennas are developed based on the theory of radiation of a symmetrical vibrator in the far zone, which has significant inaccuracies in the calculation of the components of the electromagnetic field.

Therefore, it is proposed to take into account the radial component of the electric field, which directly affects closely located wire segments.

To calculate the mutual influence of coupled vibrators, it is advisable to use the induced electromotive force method, which allows to find the induced and self-resistances of vibrators, as well as the amplitudes and phases of currents in passive vibrators.

To take into account the influence of real radiators and their mirror reflections on the antenna directivity characteristics in the induced induced electromotive force method, it is advisable to use the mirror reflection method, which allows to partially take into account the influence of real soil parameters on the antenna characteristics.

The essence of the mirror reflection method is that when determining the electromagnetic field generated by a vibrator located above a perfectly conducting surface, secondary currents are excluded from consideration by introducing a probable vibrator, which is a mirror image of the existing vibrator [1, 6].

To achieve the goal of the research, it is necessary to solve the following set of scientific tasks:

1. Analysis of the design features of existing antenna systems of radio communication facilities.

2. Analysis of mathematical methods for calculating the design, construction, modeling and modernization or improvement of antenna systems of radio communication facilities.

3. Development of a physical model of an antenna array of low-lying radiators based on the method of induced electromotive forces and the method of mirror reflection to a system of wire radiators of a special shape, taking into account their common relative location and the influence of the soil.

4. Improving the calculation method for creating a system of emitters in order to increase the energy of the radio line in the desired direction.

5. Developing technical proposals for the use of an antenna system from standard antenna-feeder devices of mobile radio communication equipment and a grounding system, taking into account soil parameters. Using the specified antenna system, form the necessary directivity diagrams in the vertical and horizontal planes, which can be quickly controlled.

6. Developing recommendations for the communications chief of the unit (station chief) for calculating the deployment of the antenna array based on the initial data: frequency, elements of the antenna system, the probable location of the correspondent and approximate soil parameters.

When using modernized vertically located antenna systems (this applies to VHF antennas), it is necessary to achieve the following characteristics:

1) The radiation pattern should change so that the width of the main lobe in the vertical plane is $\theta_0 \le 50^0$. For standard antennas, this parameter is 80º or more.

2) It is assumed that the antenna gain (G_a) will increase to 6 dB. Currently, this parameter is up to 3 dB.

3) Based on the width of the directional pattern in the vertical (θ_0) and horizontal (φ_0) planes and the antenna gain, the electric field E strength should increase accordingly, which at the reception point improves the quality of communication, or while maintaining the same quality of communication, the communication range will increase [9].

Thus, the objective function of this process will be:

$$R(x) = \max R\left(x^*\right); \quad x^* \in \Delta,$$

where $R(x)$ – the main parameter of the objective function (communication range); x – parameters on which the objective function depends (controlled and uncontrolled variables); x^* – parameter values at the maximum value of the objective function; Δ – permissible range of parameter changes.

Limitations and assumptions:

– limitations on increasing the transmitter power and receiver sensitivity, which must remain unchanged;

– limitations on deployment time (the deployment time of the AS is equal to or less than the current one);

– the qualification of the performer corresponds to a certain position (station chief);

– other limitations may arise based on the direct operational task when developing a methodology for the station chief.

The controlled variables include: the number of antenna array elements (n), the location of the antenna array elements (the distance between the elements is r), the dimensions of these elements (the length of the element is l; the diameter of the element is d), the height of the antenna rise (h), the operating frequency (f), the wavelength (λ), and the antenna gain (G_a).

The controlled variables determine the main parameters of the objective function. The communication range is chosen as the main objective function.

The increase in the communication range is directly proportional to the antenna gain, i.e. $R = \sqrt{G_a}$.

The narrowing of the DS width is approximately 2 times.

Accordingly, the electric field strength must increase, which follows from the equation:

$$E = \frac{\sqrt{30 P_{пер} G_a}}{R}.$$

Uncontrolled variables: climatic conditions, such as season of the year, weather (heat, rain, etc.), soil parameters, maximum transmitter power, receiver sensitivity, type of signal encoding, etc.

The efficiency indicator has the form:

$$\eta = \left[\frac{\left(R_{нов} - R_{ст}\right)}{R_{ст}}\right] \cdot 100\%,$$

where $R_{нов}$ – the new received parameter value, $R_{ст}$ – the previous value of the parameter.

For example, when increasing the communication range from 50 to 70 km, we get:

$$\eta = \left[\frac{70-50}{50}\right] \cdot 100\% = 40\%.$$

Conclusions

The practical implementation of the research results should be an improved antenna system based on low-lying vibrators with the ability to change the direction of information transmission depending on the location of the correspondents. This will provide the technical possibility of improving the efficiency of performing HF and VHF communication and ground wave communication, and will also provide guaranteed radio communication at the required set distance between correspondents.

By modernizing the antenna system without changing the transmitter power and receiver sensitivity, it is expected to increase the communication range by 20–40% [1].

REFERENCES

1. Pozdniak V., Makarov S., Vysotskyy O., Pavlichenko A., Lopatin A., Chekunova O. Method of automated calculation of electrical characteristics and parameters of nonsymetric vibrator antennas. *Scientific Works of Kharkiv National Air Force University*. 2022. No. 4 (74). P. 64-71. DOI: https://doi.org/10.30748/zhups.2022.74.09
2. Ільницький Л. Я., Савченко О. Я., Сібрук Л. В. Пристрої надвисоких частот та антени: навч. посіб.: К: НАУ, 2013. 188 с. URL: https://kafelec.nau.edu.ua/Materialu/Ultrahigh-frequency_and_antenna_devices_training_text.pdf
3. Wang, Y., Gao, H., Tian, Y., Lu, T. and Zhang, X. "Novel multi-mode shortwave broadcast transmitting antenna array", *Scientific Reports*, vol. 12, art. no. 11094, Jun. 2022. DOI: https://doi.org/10.1038/s41598-022-15336-x, URL: https://www.nature.com/articles/s41598-022-15336-x
4. Author Basu (VU2NSB), "Why Might Antenna Height Matter More Than Gain?," Amateur Radio Tech Journal, 2020. Available at: https://vu2nsb.com/why-might-antenna-height-matter-more-than-gain
5. Harold Melton (KV5R), "NVIS Antennas Gain vs. Height Analysis," KV5R.com, 2019. Available at: https://kv5r.com/ham-radio/nvis-antennas/nvis-page-3

6. AGU Publications, “Influence of Ground Conductivity and Permittivity on HF Antenna Patterns”, *Earth and Space Science*, vol. 9, 2022. DOI: https://doi.org/10.1029/2021RS007343.
7. Шолудько В. Г., Єсаулов М. Ю., Вакуленко О. В., Гурський Т. Г., Фомін М. М. .Організація військового зв’язку: навч. пос. К.: Вид. дім «СКІФ», 2023. 280 с. URL: https://shron1.chtyvo.org.ua/Sholudko_Vasyl/Orhanizatsiia_viiskovoho_zviazku.pdf
8. Ільїнов М. Д., Гурський Т. Г., Борисов І. В., Гриценок К. М. Лінії радіозв’язку та антенні пристрої. Навчальний посібник. К.: ВІТІ, 2018. 250 с. URL: https://sprotyvg7.com.ua/wp-content/uploads/2023/05/%D0%B0%D0%BD%D1%82%D0%B5%D0%BD%D0%B8%D1%96%D0%BB%D1%96%D0%BD%D1%96%D1%97.pdf
9. Манойлов В.П., Мартинчук П.П. Методи розрахунку та вимірювання параметрів і характеристик антен НВЧ. Житомир: ПП "Рута", 205 с. URL: https://library.ztu.edu.ua/doccard.php/137573

Received (Надійшла) 13.08.2025
Accepted for publication (Прийнята до друку) 05.11.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Бердников Олег Михайлович – кандидат технічних наук, доцент, доцент спеціальної кафедри № 3 Інституту спеціального зв’язку та захисту інформації Національного технічного університету України “Київський політехнічний інститут імені Ігоря Сікорського”, Київ, Україна;
Oleh Berdnykov – Candidate of Technical Sciences, Associate Professor, Associate Professor of a Special Department № 3 Institute of Special Communications and Information Protection of the National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Kyiv, Ukraine;
e-mail: 2507bom@gmail.com; ORCID Author ID: https://orcid.org/0009-0000-2537-8796.

Мазор Сергій Юрійович – кандидат технічних наук, доцент спеціальної кафедри № 3 Інституту спеціального зв’язку та захисту інформації Національного технічного університету України “Київський політехнічний інститут імені Ігоря Сікорського”, Київ, Україна;
Serhii Mazor – Candidate of Technical Sciences, Associate Professor of a Special Department № 3 Institute of Special Communications and Information Protection of the National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Kyiv, Ukraine,
e-mail: smazor@gmail.com; ORCID Author ID: https://orcid.org/0009-0003-2103-4883;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57204147596.

Храновська Тетяна Василівна – старший викладач спеціальної кафедри № 3 Інституту спеціального зв’язку та захисту інформації Національного технічного університету України “Київський політехнічний інститут імені Ігоря Сікорського”, Київ, Україна;
Tetiana Khranovska – senior lecturer of a Special Department № 3 Institute of Special Communications and Information Protection of the National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Kyiv, Ukraine;
e-mail: tanyakhranovskaya@gmail.com; ORCID Author ID: https://orcid.org/0009-0007-7873-4919.

Дімітров Павло Євстафійович – Науково-дослідний інститут воєнної розвідки, Київ, Україна;
Pavlo Dimitrov – Defence Intelligence Research Institute, Kyiv, Ukraine;
e-mail: strenia@ukr.net; ORCID Author ID: https://orcid.org/0009-0007-6184-8223.

Обґрунтування досліджень антенних систем в короткохвильових та ультракороткохвильових діапазонах

О. М. Бердников, С. Ю. Мазор, Т. В. Храновська, П. Є. Дімітров

Анотація. Для удосконалення короткохвильової антени на основі низькорозташованих вібраторів проведено аналіз сучасних тенденцій у розробці нових та модернізації існуючих антенних систем для короткохвильового радіозв’язку. Можливість дослідження застосування низькорозташованих горизонтальних вібраторів пов’язана з тим, що є можливість на стоянках для антен іоносферних хвиль і не тільки для них створювати з підручних матеріалів досить складні антенні решітки з великими коефіцієнтами підсилення. Проаналізовано підходи до удосконалення антенних систем у залежності від виду розповсюдження радіохвилі, фізичних параметрів антено-фідерних пристроїв та їх спрямованих властивостей. У зв’язку з цим обґрунтовано підхід щодо синтезу антен та використання даного підходу у подальшій роботі для удосконалення антенної системи й отримання потрібних характеристик спрямованості для організації стійкого радіозв’язку. Запропоновано розробити методику конструктивного синтезу на основі методів наведених електрорушійних сил та дзеркального відображення. Для досягнення поставленої мети визначено основні напрямки, такі як: удосконалення математичної моделі антенної системи на основі низькорозташованих випромінювачів; наступний крок – це удосконалення математичного методу розрахунку електромагнітного поля з врахуванням радіальних складових електричного та магнітного полів; та, як наслідок, для отримання зручного інструменту розробити методику конструктивного синтезу удосконаленої короткохвильової антенної системи із низькорозташованих випромінювачів з можливістю зміни напряму діаграми спрямованості антени в вертикальній та горизонтальній площинах. У статті сформовано цільова функція, що визначає основні параметри, які необхідно досягти. Визначені керовані та некеровані змінні, від яких напряму залежить цільова функція зазначеного процесу. Як підсумок, наведено підхід до розрахунку показника ефективності для оцінки отриманого значення параметрів від їх попередніх значень.

Ключові слова: мобільний засіб радіозв’язку, короткохвильова антена, низько розташовані вібратори, метод наведених електрорушійних сил, метод дзеркального відображення, антенні системи, коефіцієнт підсилення.

UDC 621.39.83 **doi: 10.26906/SUNZ.2025.4.176-179**

Mykola Bikchentayev, Bohdan Boriak

National University "Yuri Kondratyuk Poltava Polytechnic", Poltava, Ukraine

DESIGN AND IMPLEMENTATION OF A SOFTWARE-DEFINED SPECTRUM ANALYZER BASED ON PLUTO-SDR

Abstract. This article studies how a Software-Defined Radio (SDR) can work as a real-time spectrum analyzer and compares two spectrum analyzer program designs for PlutoSDR. **The aim of the article**: to demonstrate how an SDR can be used as a spectrum analyzer that makes fast sweeps and clearly displays data over wide frequency spans. We compare two approaches to the program design: (1) a step-by-step sweep and (2) a fast sweep with fewer tuner steps and RBW/VBW used during signal processing. **The results obtained**: the second program has shorter sweep times and more stable power level estimates across the band. It also makes burst signals easier to see and supports continuous monitoring like an entry-level Real-Time Spectrum Analyzer (RTSA). **Conclusions**: an SDR configured with proper RBW/VBW and efficient rendering can provide good-enough real-time spectrum analysis for educational and engineering tasks.

Keywords: software-defined radio, SDR, spectrum analyzer, signal processing.

Introduction

Wireless signals are now an essential part of daily life: phones, Wi-Fi, Bluetooth, IoT devices, sensors in factories, RF trackers all share the same air [1]. Since many devices communicate at once, they can interfere with each other and cause slow data transfer or dropouts. Some bands are licensed and tightly managed [2]. Others (like 2.4 and 5.6 GHz) are unlicensed and represent free, public portion of the radio spectrum that anyone can use, so devices must use low power, short bursts, and polite spectrum access (i.e., a technique called Listen Before Talk, LBT) [3]. Designers try to keep energy inside the right channel (limit “adjacent-channel” leakage) and follow basic rules on power and emissions, while still providing good speed and battery life.

Real-world RF environment changes quickly—often in milliseconds. Signals can hop in frequency, start and stop in short bursts, or collide with strong nearby transmitters. Classic sweep tools (i.e., swept-tuned spectrum analyzers that tune across the band and measure one narrow slice at a time) often miss these brief events, so modern practice observes the spectrum as it changes over time. Even simple checks, such as confirming that a device keeps its energy inside the assigned channel, need an instrument that can see fast behavior, not only slow averages. This need leads directly to a different kind of tool [1].

A Real-Time Spectrum Analyzer (RTSA/RSA) is a test instrument that continuously watches the radio spectrum. It can catch very short events, trigger when a chosen condition occurs (for example, a burst above a limit or a mask violation), record the signal around that moment, and let us study it by frequency and by time. In practice, an RSA helps engineers find and explain interference, out-of-band emissions, spurious tones, and timing issues. Because it captures and stores the exact data segment, teams can repeat the analysis, consult the requirements, and issue fixes—useful both for specialists and enthusiasts. For this purpose different analyzer architectures exist [1].

RSA Architecture Types

Swept Spectrum Analyzer (SA). An SA, also called superheterodyne spectrum analyzer, downconverts the input and sweeps a narrow filter (resolution bandwidth, RBW) across the chosen frequency span, measuring one small slice at a time (Fig. 1, a). This approach gives strong dynamic range and accurate results for steady, continuous signals. However, because it looks at only one slice at any instant, rapid changes in a signal may be missed if they happen between slices [1].

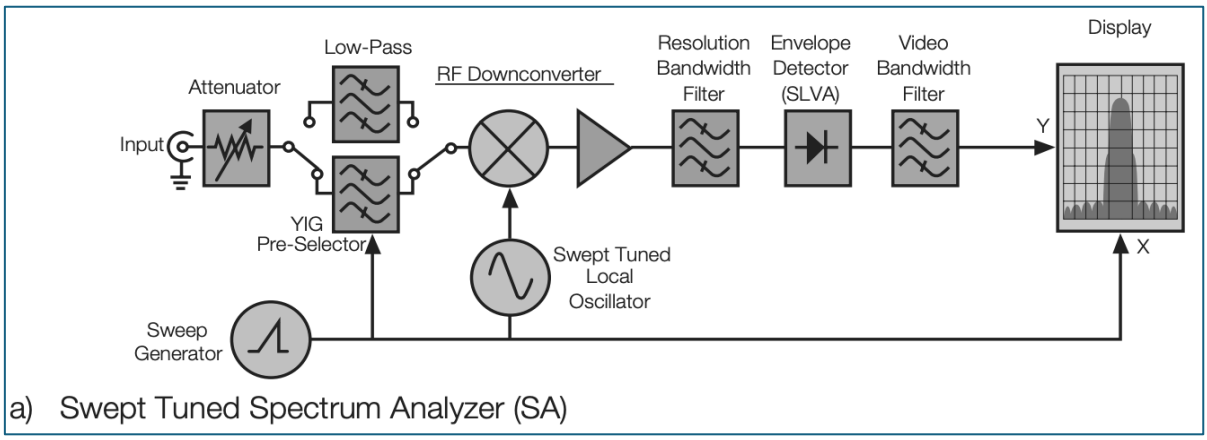


Fig. 1. Block diagram of SA analyser

Vector Signal Analyzer (VSA). A VSA downconverts a signal to IF/baseband, digitizes it, and uses digital downconversion, filtering, and an FFT to produce spectrum/time views and modulation metrics (e.g., EVM, channel power).

Because it analyzes stored IQ blocks, it is excellent for known, repeatable waveforms (Fig. 2, b). Its limits are that it is blind between captures (so rare or very short events can be missed).

Thus, a VSA is powerful tool, but it is not designed for continuous, gap-free watching of the spectrum [1].

Real-Time Spectrum Analyzer (RTSA/RSA). "Real-time" originated in early digital system simulations and means the simulation processes events as fast as the real system does (Fig. 2, c). An RTSA continuously watches a chosen slice of the radio band. The signal is sampled fast enough to satisfy Nyquist criteria to catch very short events like brief bursts, hops, or unwanted tones [1].

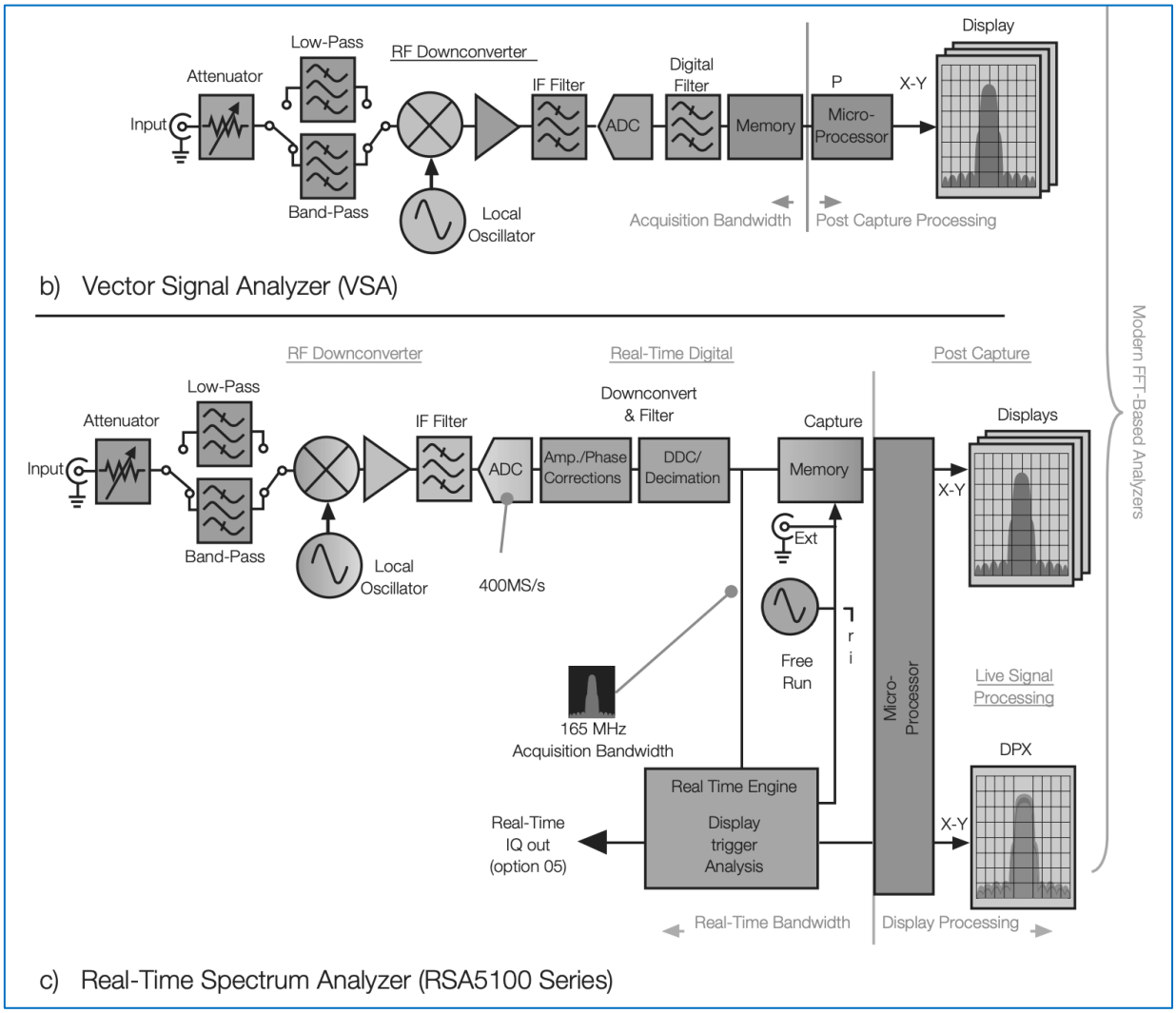


Fig 2. Block diagram of VSA and RTSA analyzers

SDR as RSA

Software-Defined Radio (SDR) can be programmed to act as an RSA. Due to the nature of SDR, this approach offers flexible, affordable solution that can be adapted to a given project. The viewing range and detection rules can be set in software. We can also record data for replay and sharing, and as well as plug it into tools or tests—all without buying a new RSA device each time a change is needed.

Compared with small swept analyzers (e.g. TinySA), an SDR-based setup is better at catching short, changing signals and avoids the blind spots that appear when reading the signal one slice at a time as all these issues can be addressed in the programming. Although professional benchtop RSAs deliver the highest calibrated accuracy and widest coverage, they are usually much more expensive; an SDR provides a "good enough" real-time visibility at a fraction of the cost. The main trade-off is setup effort: a reliable radio front end and a capable computer are needed, plus a bit of tuning, to make sure everything runs smoothly. We will use PlutoSDR as a basis for our RSA. The program that will monitor the spectrum will be based off the existing ADALM Pluto SDR Spectrum Analyzer publicly available on GitHub under the MIT license [4] (Fig. 3). This program uses Python as a programming language and pyadi-iio library to communicate with PlutoSDR. Below you can see a block diagram of this program.

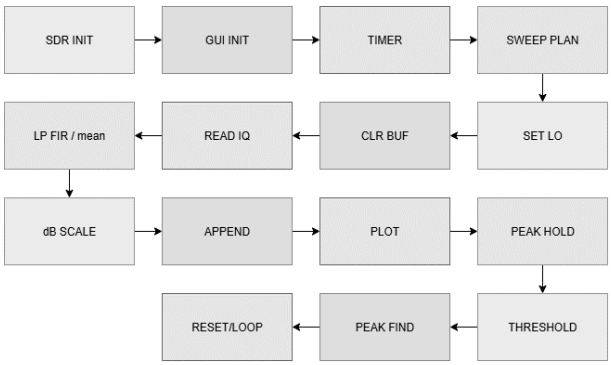


Fig 3. Block diagram of ADALM-Pluto-Spectrum-Analyzer program

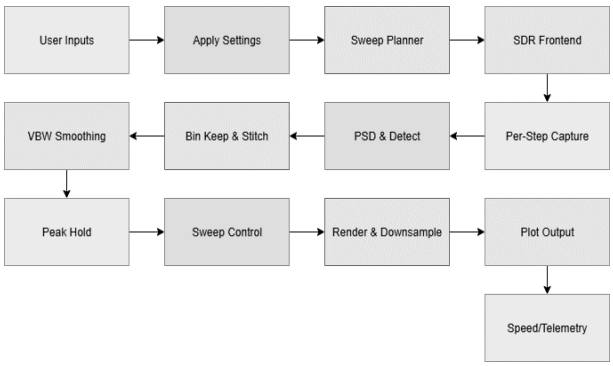


Fig 4. Block diagram of a "fast spectrum analyzer" program

The program begins with SDR INIT and GUI INIT: it connects to PlutoSDR, sets sample rate, gain (AGC or manual), and buffer sizes, then builds the plot, controls, markers, and status labels. A TIMER (fires every 10–30 Hz) breaks work into small chunks so the UI stays smooth. Each timer tick follows the SWEEP PLAN: SET LO tunes to the next center frequency, CLR BUF drops stale samples, and READ IQ captures a fresh block of IQ data. LP FIR / mean applies a filter to reduce jitter. dB SCALE converts to log units, APPEND stores frequency–level pairs, and PLOT draws these pairs across the span.

PEAK HOLD keeps the highest level for each frequency. THRESHOLD compares the current (or peak-hold) amplitude to a user-set threshold. If exceeded, trigger a visual cue (and optionally a log/notification). PEAK FIND runs at sweep end to list local maxima. RESET/LOOP wraps to the start of the frequency plan and continues the program. This ensures continuous, real-time monitoring: the loop repeats with the same parameters unless changed by the user, refreshing traces and peak-hold indefinitely.

Although this program offers helpful overlays, markers, threshold alerts, live point-by-point updates, and a UI with lots of inputs, it has limits on wide spans. It changes the tuner at every step, clears buffers, and draws each new point on the screen, so sweeping is slow, and very short bursts can be missed between steps. Its detector uses simple smoothing rather than standard RBW/VBW, so levels and narrow peaks may be less reliable. The edges between chunks of processed data can add small errors (stitching artifacts), and heavy screen updates can make the app feel busy and laggy when there are many points. Thus, a sweep from 100 MHz to 5.8 GHz may take up to 2 minutes.

For this reason, we need a program that runs the radio at a higher sample rate to reduce LO hops, sets the FFT size from a chosen RBW for clear and predictable resolution, and applies VBW after detection on linear power to smooth the trace in a proper, comparable way. It should draw once per sweep with fewer points, so the UI stays smooth. Block diagram of the new program can be seen in Fig. 4.

The sweep starts with User Inputs and Apply Settings: a user chooses start/stop, RBW (detail), VBW (smoothing), gain, and view mode (clean, with less points displayed on the chart, or granular).

The program validates these, computesFFT size from RBW, clears old peaks/buffers, and initializes timers. Next, Sweep Planner splits the span into overlapping steps for processing and builds one global frequency axis. In SDR Frontend, PlutoSDR is configured (sample rate, RF bandwidth, gain) and will tune to each step central frequency and stream IQ data.

Per step, Per-Step Capture tunes to the center and reads a fixed IQ block. PSD & Detect converts it with an FFT to power-versus-frequency and scales to dB. Bin Keep & Stitch takes only the middle “good” bins and places them into the global trace, so all steps join into one spectrum. After detection, VBW Smoothing applies a gentle time average at each frequency to reduce flicker. Peak Hold stores the highest seen level per bin to catch short bursts. Sweep Control advances steps, and at the end records sweep time for the speed readout.

For display, Render & Downsample keeps the UI fast by lightly averaging (clean view) and drawing fewer points while preserving shape. Plot Output shows live dB versus frequency, with an optional dashed peak trace. Speed/Telemetry reports last sweep time and sweeps per second. In contrast, with the first program, this one is built for speed and stable results on wide spans: it has a faster sample rate, needs fewer tuner steps, sets FFT size from RBW (for proper frequency detail), and applies VBW after detection to smooth the trace in a standard way. It renders once per sweep and limits plotted points, so the UI stays smooth, and the levels are more trustworthy across the whole band. This makes this program work faster over big frequency ranges, gives us faster sweeps, cleaner displays, and more consistent measurements. In this case a sweep from 100 MHz to 5.8 GHz takes around 400 ms.

Conclusions

An SDR working as a real-time spectrum analyzer can watch one band continuously, trigger on quick events, and save IQ data for later study. These features support faster interference hunting, clearer compliance checks, and an open toolchain that we can improve over time. Two spectrum analyzer program designs for PlutoSDR were compared. The first design is simple and shows each point on the screen, but it becomes slow for wide spans and may miss narrow or short signals. The second design reduces tuner steps, sets FFT size from the chosen RBW, applies VBW after detection, and draws once per sweep. As a result, it gives faster scans, a cleaner display,

and more stable levels across the band. The performance of the program is limited only by the chosen RBW/VBW, memory needs for long captures, and SDR computational resources, nevertheless, the proposed method is practical for study and engineering work. Future work may include adding a better level calibration, custom triggers, and report export to bring the system closer to professional instruments while keeping the benefits of open, software-based tool.

The screenshot of a program can be seen on Fig. 5.

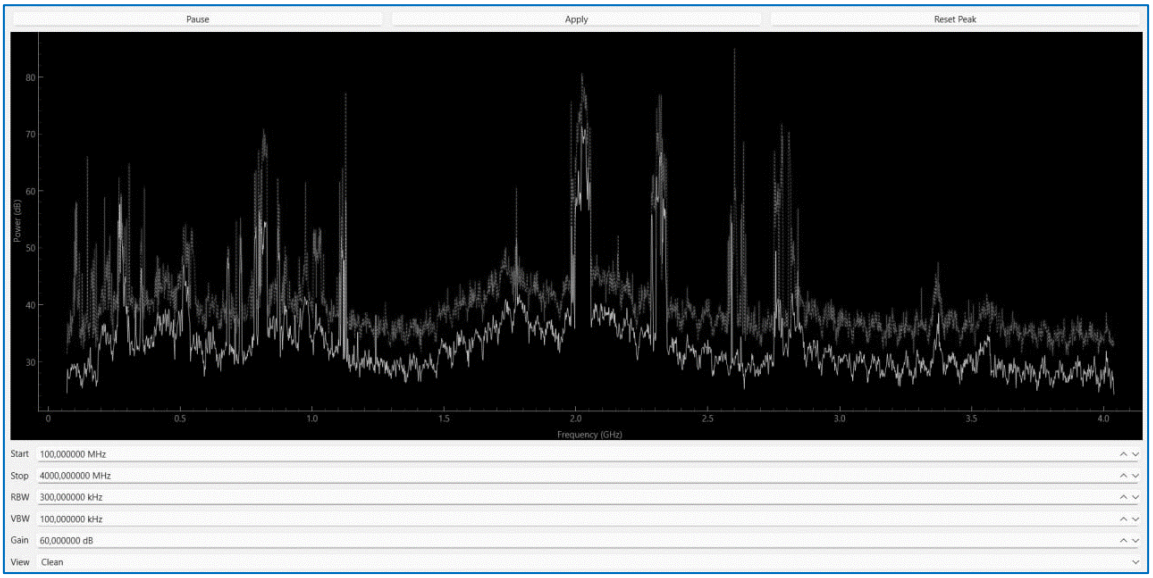


Fig 5. "Fast spectrum analyzer" program UI

References

1. Tektronix. Spectrum Analyzer How-To Guide: What They Are, What They Measure, & How to Use Them. Tektronix. URL: https://www.tek.com/en/documents/primer/what-spectrum-analyzer-and-why-do-you-need-one
2. Schirn A. IEEE 1932.1-2024: Licensed and Unlicensed Wireless Networks. The ANSI Blog. 20.05.2025 URL: https://blog.ansi.org/ansi/ieee-1932-1-2024-licensed-and-unlicensed-spectrum/
3. Wirepas. How wireless mesh and WLAN can live happily together. Wirepas Blog. 11.06.2019 URL: https://wirepas.com/blog/how-wireless-mesh-and-wlan-can-live-happily-together/
4. Fromconcepttocircuit. ADALM-Pluto-Spectrum-Analyzer: Real-time spectrum analyzer for ADALM Pluto SDR with GUI, peak hold, and markers. Репозиторій GitHub. URL: https://github.com/fromconcepttocircuit/ADALM-Pluto-Spectrum-Analyzer

Received (Надійшла) 20.08.2025
Accepted for publication (Прийнята до друку) 29.10.2025

Відомості про авторів / About the Authors

Бікчентаєв Микола Олексійович – аспірант, Національний університет «Полтавська політехніка ім. 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 автоматики, 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 Author ID: http://orcid.org/0000-0002-8114-7930;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=58983614200.

Дослідження використання Pluto-SDR як аналізатора спектру

М. О. Бікчентаєв, Б. Р. Боряк

Анотація. У статті досліджується використання програмно-керованого радіо (SDR) у ролі аналізатора спектру та порівнюються два підходи до розробки програми аналізатора для PlutoSDR. **Мета статті** — показати, як SDR можна використати у ролі аналізатора спектру, що здатен швидко здійснювати обробку широкого радіочастотного діапазону та наочно відображати отримані дані. Ми порівнюємо два підходи до проектування: (1) покроковий аналіз та (2) швидкий аналіз який здійснює менша переналаштуавань частоти та застосує RBW/VBW під час обробки сигналу. **Результати**: другий варіант має вищу швидкість роботи та більш стабільні оцінки рівня потужності по заданому діапазону. Він також краще фіксує короткочасні сигнали і при цьому безперервно здійснює обробку сигналу у вибраному діапазоні, подібно до RTSA початкового рівня, на кшталт TinySA. **Висновок**: SDR із коректно налаштованими RBW/VBW та оптимізованим процесом відображення результатів може забезпечити достатньо якісний аналіз спектру в реальному часі.

Ключові слова: програмно-кероване радіо, SDR, аналізатор спектру, обробка сигналів.

УДК 621.396 **doi: 10.26906/SUNZ.2025.4.180-185**

О. В. Жила[1], В. В. Кошарський[2]

[1] Харківський національний університет радіоелектроніки, Харків, Україна
[2] Національний аерокосмічний університет "Харківський авіаційний інститут", Харків, Україна

РОЗРОБКА І АНАЛІЗ ЧИСЕЛЬНОЇ МОДЕЛІ ОБЧИСЛЕННЯ ЯСКРАВІСНОЇ ТЕМПЕРАТУРИ АТМОСФЕРИ В СЕРЕДОВИЩІ MAPLE НА ОСНОВІ РЕКОМЕНДАЦІЙ ITU-R

Анотація. Стаття присвячена розробці та аналізу імітаційної моделі для обчислення низхідної яскравісної температури, реалізованої в середовищі Maple на основі рекомендацій ITU-R. **Метою статті** є створення методології та імітаційної моделі для чисельного моделювання яскравісної температури, а також оцінка ефективності ітеративного та рекурсивного методів з урахуванням їх продуктивності та практичного застосування**. Завдання дослідження** включають розробку алгоритмів обчислення яскравісної температури, реалізацію моделі в Maple, аналіз залежності яскравісної температури від кутів місця та частот, порівняння ітеративного й рекурсивного підходів за ключовими параметрами та оцінку точності моделі відносно еталонних даних. **Отриманий результат** демонструє коректність моделі, підтверджену близькістю обчислених значень яскравісної температури до реальних даних. Ітеративний метод виявився швидшим, тоді як рекурсивний підхід краще ілюструє фізичну модель. Чисельне моелювання відображає залежність яскравісної температури від кутів місця, що узгоджується з фізичними принципами. **Галузь застосування:** охоплює метеорологію (аналіз атмосферного випромінювання), телекомунікації (оцінка затухання сигналів 5G/6G), дистанційне зондування Землі та освіту (демонстрація RTE і алгоритмів). Перспективи включають інтеграцію розсіювання та даних реального часу.

Ключові слова: яскравісна температура, імітаційна модель, радіометрія, мікрохвильове випромінювання.

Вступ

Аналіз останніх досліджень і публікацій. Яскравісна температура є важливим параметром у радіофізиці та метеорології і використовується для описання випромінювання атмосфери Землі і космосу у різних діапазонах частот. Сучасні методи її обчислення спираються на розв'язання радіаційного трансферного рівняння, яке враховує поглинання, випромінювання та розсіювання електромагнітних хвиль у атмосфері. Існує багато моделей, що описують яскравісну температуру, кожна з яких має унікальні особливості. Серед таких моделей вирізняється Radiative Transfer for TOVS (RTTOV) – швидка модель прямого переносу випромінювання, що розроблена Європейським центром прогнозів погоди ECMWF. Ця модель інтегрує профілі температури, вологості та хмарності для аналізу даних супутникових сенсорів, а також використовується в асиміляції даних метеорологічних прогнозів. Також ця модель використовується для інтерпретації вимірювань мікрохвильових радіометрів завдяки своїй оптимізації для роботи в реальному часі [1–3]. Іншою цікавою розробкою є модель Atmospheric Radiative Transfer Simulator (ARTS), яка враховує розсіювання на дрібних частинках, таких як дощ, сніг та лід, що особливо важливо для частот вище 50 Ггц. Ця модель базується на стохастичних та детермінованих методах і використовується для аналізу даних пасивних мікрохвильових сенсорів, наприклад, GPM Microwave Imager, в дослідженні хмар та опадів [4,5]. Нещодавні дослідження, такі як роботи із використанням нейронних мереж, пропонують заміну традиційних чисельних моделей на перенавчені алгоритми, які здатні передбачати яскравісну температуру на базі основних даних про атмосферні профілі. Такі підходи прискорюють обробку даних у реальному часі, що має важливе значення для супутникового зв'язку [6].

Особливий інтерес представляють моделі, засновані на рекомендаціях Міжнародного союзу електрозв'язку (ITU), зокрема, стандарті ITU-R P.676-13 [7]. Ця модель враховує вклад кисню, водяної пари та рідкої води, спираючись на емпіричні залежності і спектральні лінії, що робить її популярною у супутниковій метеорології та системах зв'язку для оцінки затухання сигналів. Висока точність при стандартних умовах і регулярні оновлення стандарту забезпечують її практичну значимість.

Яскравісна температура, алгоритм знаходження якої детально описано у [7], відіграє важливу роль у радіолокації та радіометрії як ефективна температура випромінювання, що реєструється приймальними системами з урахуванням атмосферного поглинання. Цей параметр критичний для інтерпретації даних у мікрохвильовому діапазоні, де атмосфера значною мірою модулює сигнал. Вона також є еталоном для калібрування радіолокаційних та радіометричних систем, забезпечуючи точність вимірювань відображень від земної поверхні або космічних об'єктів.

Сучасні дослідження у [8] описують калібрування даних із застосуванням яскравісної температура для сенсорів типу VIIRS та NOAA-21. У поєднанні з моделями атмосферного поглинання яскравісна температура дозволяє розділяти вклади атмосфери та цільових об'єктів, що важливо для дистанційного зондування Землі, дослідження планет та астрономічних спостережень. Наприклад, стаття [9] демонструє її застосування для аналізу земних поверхонь з використанням даних SMAP.

У телекомунікаціях яскравісна температура використовується для оцінки завад та затухання сигналів, особливо у високошвидкісних системах 5G та 6G у різних діапазонах частот. Аналіз її впливу на затухання у террагерцовому діапазоні для 6G представл-

ено у роботі [10], а дослідження [11] об'єднує експериментальні та теоретичні дані про вплив атмосфери на 6G-системи. В роботі [12] акцент зроблено на оптимізації каналів з навколишніми поверхнями.

Постановка проблеми. Дослідження присвячене створенню імітаційної моделі розрахунку яскравісної температури у середовищі Maple на базі стандартів [7, 13, 14]. Запропонований підхід поєднує ітеративний та рекурсивний методи чисельного розв'язання RTE, забезпечуючи гнучкість та адаптивність до різних атмосферних профілів. Модель включає дискретизацію атмосфери з кроком 0.5 км до висоти 100 км, інтеграцію функції поглинання та траєкторії радіохвиль, а також порівняння продуктивності двох підходів. Це не лише відповідає міжнародним стандартам, але і представляє зручний інструмент для освітніх та дослідницьких цілей, зберігаючи баланс між точністю, швидкістю та наочністю коду.

Розроблена у програмі Maple імітаційна модель має значні переваги. Використання даного пакету прикладних програм, відомого своїми символьними та чисельними обчисленнями, забезпечує легку модифікацію алгоритмів та адаптацію до різних умов або частотним діапазонам, включаюсь можливість інтеграції з іншими стандартами завдяки модульній структурі. Платформа також підтримує інтерактивність та візуалізацію, що робить модель цінним ресурсом для науковців. Порівняння двох підходів – ітеративного та рекурсивного – дозволило визначити оптимальний підхід для розрахунку яскравісної температури. Висока точність, що підтвердилась відповідністю стандартам, робить модель надійним інструментом для наукових та практичних задач.

Мета статті. Основною метою дослідження було створення методології та імітаційної моделі в середовищі Maple, що переводить теоретичні положення різних рекомендацій по обчисленню яскравісної температури у єдину практичну форму для чисельних експериментів. Це мотивовано також представленням інструмента для ілюстрації фізичних основ та алгоритмічних розрахунків яскравісної температури. Проведене дослідження можна вважати основою для подальших покращень, таких як додавання модулів інтеграції з даними реального часу, що відкриває широкі перспективи для більш широкого класу задач. Отримані результати можуть знайти застосування у дистанційному зондуванні, телекомунікаціях (наприклад, у аналізі даних мікрохвильових радіометрів), а також у моделюванні впливу атмосфери на радіолокаційні системи та навігаційні прибори космічних апаратів.

Методологія

Яскравісна температура мікрохвильового випромінювання – це фундаментальна концепція у радіофізиці, метеорології та астрономії, яка використовується для описання інтенсивності випромінювання в термінах температури еквівалентного чорного тіла та може бути обчислена по формулі [7]:

$$T_B(f_{GHz}, T_j) = 0,048 \cdot f_{GHz} \times \times \left(1 \Big/ \left(\exp\left(0,048\, f_{GHz}/T_j\right) - 1\right)\right),\ (K), \quad (1)$$

де T_J – фізична температура j-го шару. Яскравісна температура $T_B(f_{GHz}, T_j)$ добре апроксимується значенням T_J на частоті $f_{GHz} < 0,42$.

Тобто, яскравісна температура T_B визначається як температура ідеального тіла, що випромінює електромагнітна хвиля з тією ж інтенсивністю, що і середовище при певній частоті. Для обчислення яскравісної температури T_B випромінювання потрібно враховувати низхідне випромінювання атмосферних газів (кисню та водяної пари), а також вплив затухання і геометрії траси [7]. Згідно рекомендаціям [7] для моделювання низхідної яскравісної температури необхідно розглядати модель, де атмосфера поділяється на певну кількість шарів, а вибір кількості цих шарів відповідає за точність значення яскравісної температури. Цей вибір залежить від багатьох факторів, включаючи мету дослідження, характеристики атмосфери, частотний діапазон радіометра, обчислювальні ресурси тощо. Як правило, перший шар атмосфери розглядається на поверхні Землі, а останній – це верхній шар атмосфери. Для обчислення сумарної яскравісної температури потрібно розглянути суму значень яскравісної температури випромінювання в кожному шарі, помноживши отриману суму на втрати при переході до нового шару с урахуванням точки спостереження. У цьому випадку розсіюванням можна знехтувати, припустивши, що атмосфера знаходиться у стані локальної термодинамічної рівноваги.

При відомих атмосферних профілях фізичної температури, атмосферного та парціального тиску [12] яскравісна температура *низхідного* мікрохвильового випромінювання обчислюється за формулою, що є розв'язком рівняння переносу випромінювання для низхідного мікрохвильового випромінювання за набором дискретних шарів [7]:

$$T_{downwelling} = T_B(f_{GHz}, 2,73) \cdot 10^{-\left(\sum_{j=1}^{k} a_j \gamma_j \Big/ 10\right)} + \\ + \sum_{j=1}^{k} T_B(f_{GHz}, T_j) \cdot \left(10^{a_j \gamma_j / 10} - 1\right) \cdot 10^{-\left(\sum_{i=1}^{j} a_i \gamma_i \Big/ 10\right)},\ (K), \quad (2)$$

де $T_{downwelling}$ – яскравісна температура низхідного випромінювання (в Кельвінах); $T_B(f_{GHz}, T)$ – яскравісна температура для частоти f_{CHz} і температури T; f_{CHz} – частота в гігагерцах; T_j – температура j-го шару атмосфери; a_j – коефіцієнт поглинання для j-го шару атмосфери; T_J – оптична товщина j-го шару атмосфери; k – загальн а кількість шарів атмосфери.

Важливим етапом обчислення яскравісної температури є розрахунок погонного затухання в атмосферних газах [7] на основі моделювання поглинання киснем і водяною парою з використанням даних про спектральні лінії. При цьому, необхідно врахувати додаткові коефіцієнти, що відповідають за нерезонансний спектр поглинання киснем на низьких частотах, поглинання азоту на частотах від 100 ГГц та додаткове поглинання у смузі неперервного поглинання водяною парою.

Формула для обчислення погонного затухання в атмосферних газах має вигляд [7]:

$$\gamma = \gamma_O + \gamma_W = 0{,}1820 \cdot f \times$$
$$\times\left(\sum_i {}_{кис} S_i F_i + N_D^{"}(f) + \sum_i {}_{вод.пар} S_i F_i \cdot \right), (дБ/км), \quad (3)$$

де $\gamma_O -$ погонне затухання, обумовлене сухим повітрям (киснем); $\gamma_W -$ погонне затухання, обумовлене водяною парою; $S_i -$ спектральні лінії, $F_i -$ коефіцієнт форми спектральних ліній, $N_D^{"}(f) -$ неперервний спектр для сухого повітря, обумовлений поглинанням азоту та дебаєвським спектром.

Окрім погонного затухання потрібно враховувати траєкторію поширення випромінювання, тобто, обчислювати довжину ділянки траси через кожен шар. Довжина ділянки траси через *i*-й шар товщиною δ_i обчислюється за формулою [7]:

$$a_i = -r_i \cos\beta_i + \sqrt{r_i^2 \cos^2\beta_i + 2r_i\delta_i + \delta_i^2}, \text{ (км)}, \quad (4)$$

де $r_i -$ довжина радіус-вектора із центра Землі до нижньої межі *i*-го шару; $r_{i+1} = r_i + \delta_i$, $r_1 -$ довжина радіус-вектора Землі до нижньої межі нижнього шару, як правило – це середній радіус Землі (6371 км); $\beta_1 -$ місцевий зенітний кут при поверхні Землі або близько біля неї; $\beta_1 = 90 - \varphi$, де $\varphi -$ видимий кут місця; $\beta_{i+1} = \arcsin\left(n_1 r_1 \sin\beta / n_i r_{i1}\right)$, $n -$ індекс рефракції із [13].

В рекомендації [7] також зазначено, для обчислення яскравісної температури зручно застосовувати рекурсію. Це пов'язано з природою радіаційного трансферного рівняння, оскільки воно виражається як рекурентне співвідношення, у якому внесок кожного шару атмосфери у значення яскравісної температури залежить від попереднього шару. Окрім того, рекурсія робить модель більш інтуїтивною для розуміння, адже вона чітко демонструє, як випромінювання накопичується і явно відображає фізичність рівняння переносу.

Проте, слід зазначити, що використання рекурсії не є обов'язковим, хоча і мотивовано бажанням відобразити природню рекурентність рівняння переносу. Альтернативний підхід із застосування циклів (ітеративний) також має сенс і може бути більш оптимальним для розрахунку яскравісної температури.

Алгоритми обчислення яскравісної температури та їх порівняння

Даний розділ присвячений детальному опису алгоритмів обчислення низхідної яскравісної температури, реалізованих у середовищі Maple. Розглянуто два підходи: ітеративний та рекурсивний, а також проведено їх порівняння за ключовими характеристиками.

Ітеративний алгоритм базується на циклічному обчисленні яскравісної температури для послідовних щарів атмосфери. Процес починається з ініціалізації параметрів: висота $h = 0$, яскравісна температура $T_B = 0$, індекс шару $i = 0$ та загальна кількість шарів $N = 100$. У межах циклу виконуються такі кроки:

1. Обчислення коефіцієнту поглинання γ_j для поточного шару.
2. Визначення довжини шляху.
3. Оновлення яскравісної температури за формулою [7]^

$$T_{B,downwelling,last} = \\ = T_{B,downwelling} \cdot L_j + (1 - L_j) \cdot T_B, \quad (5)$$

де $L_j = 10^{-a_j \gamma_j / 10}$.

4. Збільшення індексу $i = i + 1$ та висоти $h = h + \delta h$. Цикл повторюється, доки $i < N$. По завершенні повертається остаточне значення яскравісної температури.

Цей підхід забезпечує лінійну обробку даних і не потребує додаткових ресурсів пам'яті.

На рис. 1 у вигляді блок-схеми зображено алгоритм виконання ітеративного методу обчислення яскравісної температури.

Рекурсивний алгоритм реалізовано через окрему рекурсивну функцію, яка викликає сама себе для обробки наступного шару і показано на рис. 2 у вигляді блок-схеми. Ініціалізація параметрів виконується аналогічно ітеративному алгоритму, а далі рекурсивний алгоритм включає:

1. Перевірку базового випадку: якщо $i \geq N$, повертається поточне значення яскравісної температури.
2. Обчислення коефіцієнта поглинання та довжини шляху
3. Оновлення яскравісної температури.
4. Виклик рекурсивної функції з новими параметрами.

Цей метод відображає рекурентну природу рівняння переносу, де кожен шар залежить від попереднього.

Порівняння ітеративного та рекурсивного підходів показало, що ітеративний метод демонструє вищу швидкість (середній час виконання 781,7 мс), що приблизно на 6% швидше, ніж рекурсивний (830 мс) завдяки відсутності накладних витрат на виклики функцій. Також перший підхід використовує фіксований об'єм пам'яті, тоді як другий накопичує стекові кадри, що може призвести до переповнення при великій кількості шарів.

Проте, рекурсивний метод краще ілюструє рекурентру структуру базового рівняння переносу, що корисно для освітніх цілей, тоді як ітеративний простіший для модифікації та відлагодження. Ітеративний підхід легше адаптувати до паралельних обчислень, а рекурсивний вимагає переробки для інтеграції складних умов. Тому, рекурсивний підхід більше відповідає умовам демонстрації фізичних принципів обчислення яскравісної температури, а ітеративний краще підходить для масових обчислень, зокрема, обробки даних у реальному часі.

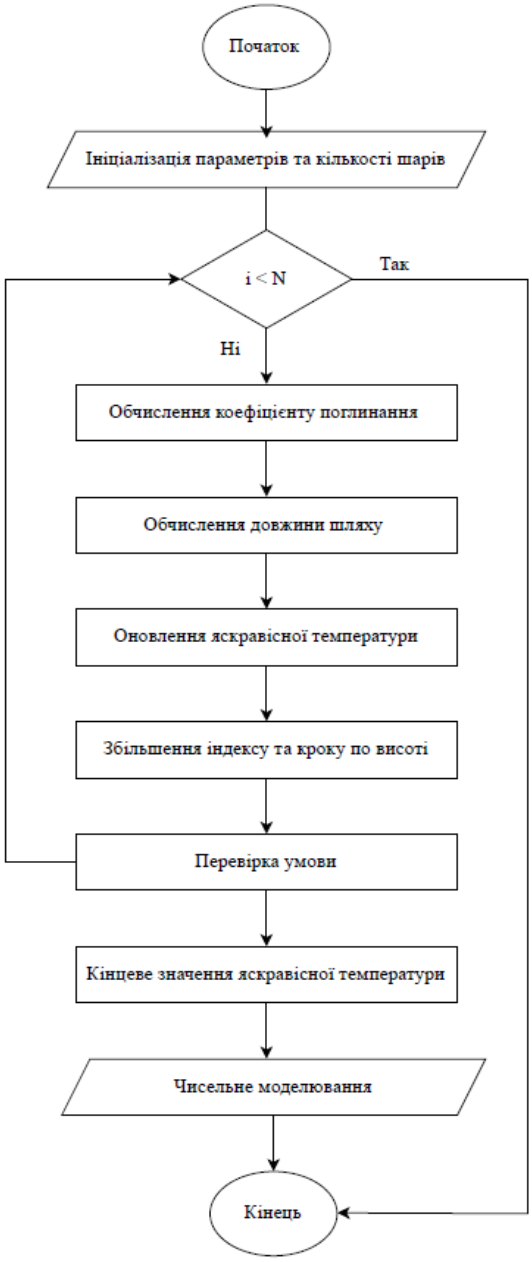


Рис. 1. Ітеративний алгоритм обчислення T_B

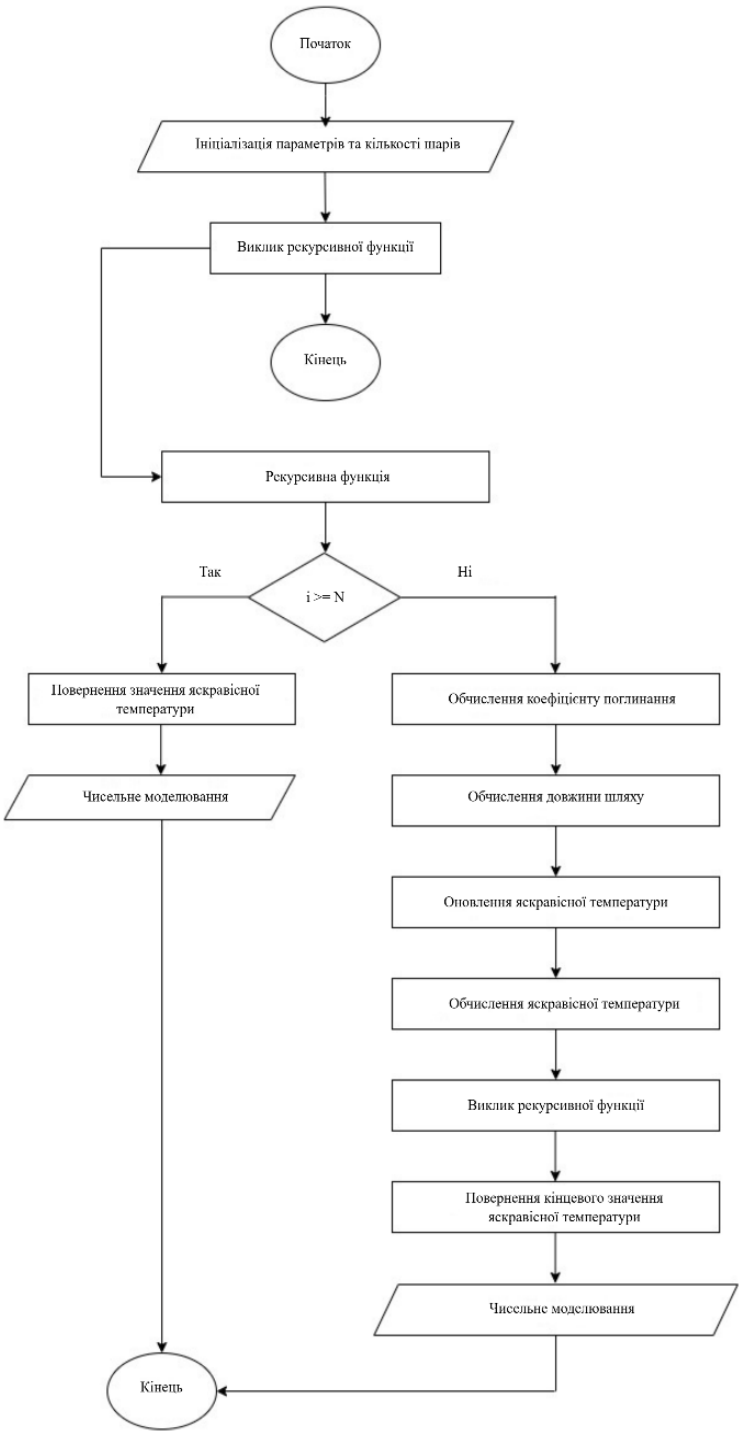


Рис. 2. Рекурсивний алгоритм обчислення T_B

Обидва алгоритми коректно реалізують обчислення яскравісної температури відповідно до рекомендацій, демонструючи близькість результатів до еталонних даних. Вибір між ними залежить від цілей.

Чисельне моделювання

У цьому розділі представлено аналіз чисельного моделювання низхідної яскравісної температури у мікрохвильовому діапазоні, виконаних за допомогою розробленої імітаційної моделі в середовищі Maple.

На рис. 3 показано розподіл яскравісної температури залежно від кутів місця для різних частот.

Графік на рис. 3 відображає залежність яскравісної температури від кутів місця у діапазоні від 5° до 90° для кількох частот у мікроволновому діапазоні, зокрема 11 ГГц, 20 ГГц, 67 ГГц і 90 ГГц. Модель базується на віщезазначених рекомендаціях і використовує дискретизацію атмосфери з кроком 0.5 км до висоти 100 км.

На графіку представлено кілька кривих, кожна з яких відповідає певній частоті, із чіткими позначками осей: горизонтальна вісь відображає кут місця (у градусах), а вертикальна — значення яскравісної температури (у кельвінах). Кожна крива, ймовірно, має різний колір або стиль лінії для

розрізнення частот, із відповідними легендами. Графік демонструє, що значення яскравісної температури, отримані моделлю, близькі до реальних даних, що підтверджує коректність реалізації. Маленькі відхилення можуть бути пов'язані з відсутністю у моделі врахування розсіювання на гідрометеорах (дощ, сніг), яке стає значним при частотах > 50 ГГц, або з використанням еталонних профілів атмосфери [14], які можуть не повністю відображати локальні умови.

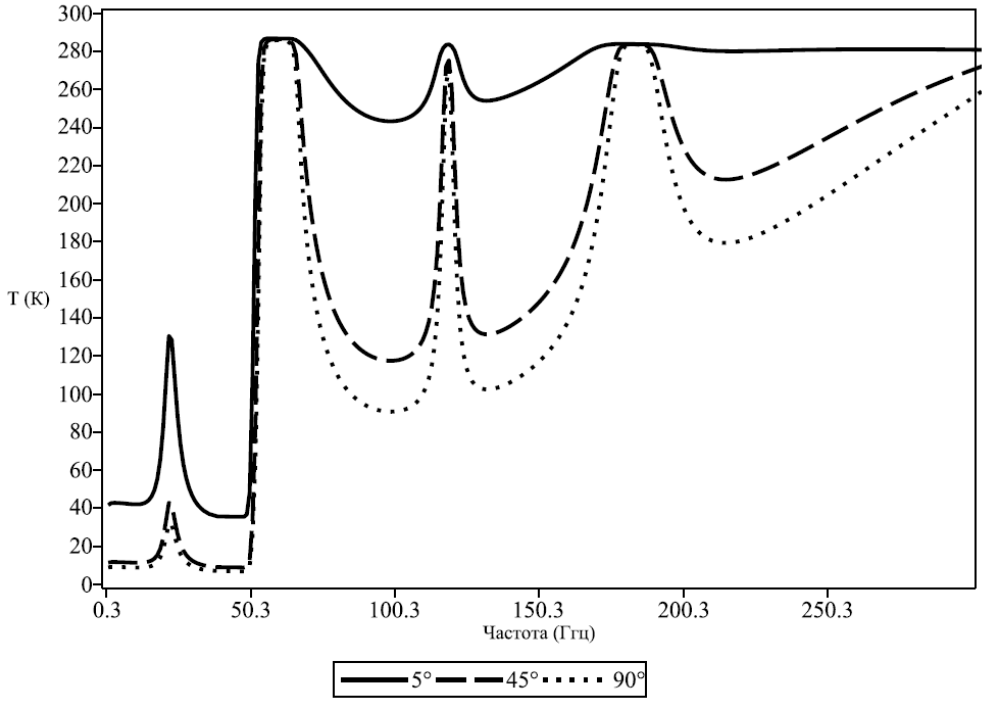


Рис. 3. Яскравісна температура низхідного мікрохвильового випромінювання

Графік на рис. 3 є наочним підтвердженням працездатності імітаційної моделі в Maple.

Зростання яскравісної температури зі зменшенням кута місця та підвищенням частоти узгоджується з фізичними принципами, а близькість до реальних значень свідчить про високу точність моделі в стандартних умовах. Для покращення можна додати модуль розсіювання та інтегрувати локальні атмосферні профілі.

Висновки

Дане дослідження представило детальну розробку та аналіз імітаційної моделі для обчислення низхідної яскравісної температури у мікрохвильовому діапазоні, реалізованої в середовищі Maple на основі рекомендацій. Модель, що базується на радіаційному трансферному рівнянні, продемонструвала високу точність і практичну цінність для наукових та освітніх цілей.

Розроблено та порівняно два алгоритмічні підходи: ітеративний і рекурсивний. Ітеративний метод виявився ефективнішим за швидкістю, що робить його придатним для масових обчислень, тоді як рекурсивний підхід вирізняється наочністю і краще відображає рекурентну природу трансферного рівняння, що є цінним для навчальних цілей. Обидва методи показали близькість результатів до еталонних даних, підтверджуючи коректність реалізації. Близькість графіків до реальних значень підкреслює надійність моделі, хоча невеликі відхилення вказують на необхідність урахування розсіювання та локальних атмосферних профілів.

Модель має значний потенціал для застосування в метеорології, телекомунікаціях (особливо для 5G/6G) та дистанційному зондуванні Землі, а також як освітній інструмент для демонстрації фізичних і алгоритмічних аспектів рекурентних рівнянь. Перспективи подальшого розвитку включають інтеграцію модулів розсіювання, використання даних реального часу та оптимізацію рекурсивного методу (наприклад, хвостовою рекурсією). Дослідження закладає міцну основу для майбутніх удосконалень і практичного використання в сучасних наукових та інженерних задачах.

СПИСОК ЛІТЕРАТУРИ

1. Bruna Barbosa Silveira, Emma Catherine Turner, and Jérôme Vidot, Global evaluation of fast radiative transfer model coefficients for early meteorological satellite sensors, EGU24, vol. 17, issue 4, 2024, 1279-1296, https://doi.org/10.5194/amt-17-1279-2024
2. Yongbo Zhou, Tianrui Cao, and Lijian Zhu, Optimizing cloud optical parameterizations in Radiative Transfer for TOVS (RTTOV v12.3) for data assimilation of satellite visible reflectance data: an assessment using observed and synthetic images, EGU25, vol. 18, issue 14, 2024, 3267-3285, https://doi.org/10.5194/amt-18-3267-2025

3. Vargas Jiménez, F. and De los Reyes, J. C., Automatic Optical Depth Parametrization in Radiative Transfer Model RTTOV v13 via LASSO-Induced Sparsity for Satellite Data Assimilation, EGUsphere [preprint], 2025. https://doi.org/10.5194/egusphere-2025-950
4. Buehler, S. A., R. Larsson, O. Lemke, S. Pfreundschuh, M. Brath, I. Adams, S. Fox, F. E. Roemer, P. Czarnecki, and P. Eriksson ,The Atmospheric Radiative Transfer Simulator ARTS, Version 2.6 — Deep Python Integration, J. Quant. Spectrosc. Radiat. Transfer, 2025, 341, 109443, https://doi.org/10.1016/j.jqsrt.2025.109443.
5. Buehler, S. A., J. Mendrok, P. Eriksson, A. Perrin, R. Larsson, and O. Lemke, ARTS, the atmospheric radiative transfer simulator — version 2.2, the planetary toolbox edition, Geosci. Model Dev., 2018, 11(4), 1537–1556, https://doi.org/10.5194/gmd-11-1537-2018.
6. Shixiong Wang, Wei Dai, Geoffrey Ye Li, Distributionally Robust Receive Combining, Electrical Engineering and Systems Science, 2025, 1-16, https://doi.org/10.48550/arXiv.2401.12345
7. ITU, “Attenuation by atmospheric gases,” Recommendation ITU-R P.676-13, Aug. 2022. www.itu.int. https://www.itu.int/rec/R-REC-P.676-13-202208-I/en. (accessed Jul. 3, 2025).
8. N. Charaf, J. Haase, A. Kulisch, C. Von Elm and D. Göhringer, "RTASS: a RunTime Adaptable and Scalable System for Network-on-Chip-Based Architectures," 2023 26th Euromicro Conference on Digital System Design (DSD), Golem, Albania, 2023, pp. 585-592, https://doi.org/10.1109/DSD60849.2023.00086.
9. Nandy A, Phinn S, Grinham A, Albert S. Developing a Semi-Automated Near-Coastal, Water Quality-Retrieval Process from Global Multi-Spectral Data: South-Eastern Australia. Remote Sensing. 2024; 16(13):2389. https://doi.org/10.3390/rs16132389
10. S. Yan, X. Shi, X. Liu, Q. Chen and L. Lin, "Mechanism Analysis of Instability in Grid-Connected PV Systems With Volt-Var Control," 2024 IEEE 10th International Power Electronics and Motion Control Conference (IPEMC2024-ECCE Asia), Chengdu, China, 2024, pp. 2824-2828, https://doi.org/10.1109/IPEMC-ECCEAsia60879.2024.10567812.
11. F. Li, L. Pang and T. Dai, "CPG Motion Controller Based on Van der Pol Nonlinear Oscillator for a Quadruped Robot," 2023 5th International Conference on Robotics, Intelligent Control and Artificial Intelligence (RICAI), Hangzhou, China, 2023, pp. 236-239, https://doi.org/10.1109/RICAI60863.2023.10489012.
12. E Silva, J.D.S.; Ribeiro, J.A.P.; Adanvo, V.F.; Mafra, S.B.; Mendes, L.L.; Li, Y.; de Souza, R.A.A. A Survey on the Impact of Intelligent Surfaces in the Terahertz Communication Channel Models. Sensors **2024**, 24, 33. https://doi.org/10.3390/s24010033
13. ITU, “Sky-wave field-strength prediction method for the broadcasting service in the frequency range 150 to 1600 kHz,” Recommendation ITU-R P.435, Jen. 2017. www.itu.int. https://www.itu.int/rec/R-REC-P.435/en. (accessed Jul. 30, 2025).
14. ITU, “Reference atmospheres,” Recommendation ITU-R P.835, Feb. 2025. www.itu.int. https://www.itu.int/rec/R-REC-P.835/en. (accessed Jul. 30, 2025).

Received (Надійшла) 21.08.2025
Accepted for publication (Прийнята до друку) 29.10.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Жила Ольга Володимирівна – кандидат фізико-математичних наук, доцент кафедри вищої математики, Харківський національний університет радіоелектроніки, Харків, Україна;
Olha Zhyla – candidate of Physical and Mathematical Sciences, Associate Professor of the Department of Higher Mathematics, National University of Radio Electronics, Kharkiv, Ukraine;
e-mail: olha.kuryzheva@nure.ua; ORCID Author ID: https://orcid.org/0000-0002-6888-8953;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=56784340900.

Кошарський Володимир Віталійович – доктор філософії, доцент кафедри аерокосмічних радіоелектронних систем, Національний аерокосмічний університет ім. М. Є. Жуковського "Харківський авіаційний інститут", Харків, Україна;
Volodymyr Kosharskyi – PhD in Telecommunications and Radioengineering, Associate Professor at the Aerospace Radio-Electronic Systems Department, National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine.
e-mail: v.kosharsky@khai.edu; ORCID Author ID: https://orcid.org/0000-0002-8569-2047;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57218710916 .

Development and analysis of a numerical model for calculating the brightness temperature of the atmosphere in the Maple environment based on ITU-R recommendations

Olha Zhyla, Volodymyr Kosharskyi

Abstract. The article is devoted to the development and analysis of a simulation model for calculating the downward luminous temperature, implemented in the Maple environment based on ITU-R recommendations. The purpose of the article is to create a methodology and simulation model for numerical modelling of luminance temperature, as well as to evaluate the effectiveness of iterative and recursive methods, taking into account their performance and practical application. The research tasks include the development of algorithms for calculating luminance temperature, the implementation of the model in Maple, the analysis of the dependence of luminance temperature on elevation angles and frequencies, the comparison of iterative and recursive approaches in terms of key parameters, and the evaluation of the model's accuracy relative to reference data. The results demonstrate the correctness of the model, confirmed by the proximity of the calculated luminous temperature values to the actual data. The iterative method proved to be faster, while the recursive approach better illustrates the physical model. Numerical modelling reflects the dependence of brightness temperature on elevation angles, which is consistent with physical principles. Field of application: covers meteorology (analysis of atmospheric radiation), telecommunications (assessment of 5G/6G signal attenuation), remote sensing of the Earth, and education (demonstration of RTE and algorithms). Prospects include the integration of scattering and real-time data.

Keywords: brightness temperature, simulation model, radiometry, microwave radiation.

UDC 004.738.5:004.7: 621.39 **doi: 10.26906/SUNZ.2025.4.186-189**

Stanislav Myhal

National University "Yuri Kondratyuk Poltava Polytechnic", Poltava, Ukraine

MATHEMATICAL MODEL OF A DATA FLOW MANAGEMENT SYSTEM IN A CLUSTER-BASED MULTICONTROLLER SDN

Abstract. The problem of load balancing among controllers and ensuring network state consistency is one of the key challenges in multicontroller Software-Defined Networks (SDN). To address this issue, it is essential to develop adequate mathematical models that accurately describe the processes of data flow management. The subject of this study is the data flow control system within a clustered multicontroller SDN environment. **The objective** of the research is to develop a mathematical model that accounts for the dynamic clustering of SDN controllers when receiving data from network switches, enabling balanced load distribution across controllers and facilitating rapid reallocation of links between controllers and switches**. The following results** have been obtained. A three-tier hierarchical clustering approach for control and data transmission/reception devices within the system has been proposed. The relationships between controllers and OpenFlow switches within the developed clustered architecture have been formalized. Computational analysis of the control element loads has been performed. **Conclusion.** The proposed mathematical model enables efficient load balancing among controllers and ensures dynamic redistribution of workloads in the event of controller failure.

Ключові слова: telecommunication network, communication network, controller, airborne network, OpenFlow switch, 5G standard, SDN.

Introduction

Modern computer networks are becoming increasingly complex due to the growing volume of data transmission, the diversity of services, and the demand for flexible resource management. One of the most effective approaches to organizing network infrastructure is Software-Defined Networking (SDN). The core concept of SDN lies in the separation of the control plane from the data forwarding plane, enabling centralized network management through software-based tools. This paradigm enhances the flexibility, scalability, and adaptability of network systems, while also simplifying the implementation of security policies and traffic optimization mechanisms.

However, as SDN networks expand in scale, there arises a need for distributed control systems, where control functions are no longer concentrated in a single node but are distributed among multiple controllers. Consequently, cluster-based multicontroller SDN architectures have gained significant popularity. In such architectures, multiple controllers operate collaboratively to ensure load balancing, fault tolerance, and reduced latency. Each controller is responsible for a specific segment of the network, yet must continuously interact with others to maintain network state consistency.

At the same time, data flow management in cluster-based multicontroller SDNs presents several challenges related to coordination among controllers, synchronization of routing tables, and maintenance of network stability. Additional difficulties arise from traffic heterogeneity, inter-controller communication delays, and the need for dynamic resource redistribution. Without an efficient flow management mechanism, these systems may experience performance degradation, excessive load concentration on individual controllers, or even instability, ultimately leading to network service deterioration.

To address these challenges, it is essential to develop adequate mathematical models that describe the processes of data flow management in cluster-based multicontroller SDNs. Such models make it possible to formalize controller interactions, determine equilibrium or steady states of the system, and evaluate the impact of network parameters on operational efficiency. Mathematical modeling thus provides a foundation for analytical studies, control algorithm optimization, network behavior prediction, and overall system stability improvement.

Therefore, the development of a mathematical model of data flow management in a cluster-based multicontroller SDN represents a crucial step toward enhancing the mechanisms of distributed control, improving network reliability, and ensuring the efficient performance of modern software-defined networks.

1. Review of Current Research

Recent scientific studies have devoted considerable attention to the development of Software-Defined Networking (SDN), which is recognized as a promising approach for building scalable and flexible network systems. The fundamental principles of the SDN concept and the OpenFlow architecture were established in early works, where the separation of the control plane from the data forwarding plane was first proposed, enabling centralized network management [1].

Subsequent research has focused on analyzing the architecture, advantages, and challenges of SDN. In particular, studies have examined issues of security, controller performance, scalability, and the efficiency of inter-controller communication [2]. It has been demonstrated that a purely centralized control model poses certain limitations, leading to the development of distributed and multicontroller architectures [3].

One of the major challenges in multicontroller SDNs is load balancing among controllers and network state synchronization. A number of studies have proposed synchronization mechanisms between controllers and assessed their impact on request processing delays and system stability [4]. Other works have focused on dynamic controller assignment algorithms that allocate controllers to network nodes based on distance and communication latency, thereby improving the efficiency of data flow distribution [5].

An important research direction in SDN development is controller clustering, which enhances fault tolerance and ensures network scalability. Studies in this area have proposed hierarchical clustering models and described mechanisms for inter-cluster coordination [6]. Analytical models have also been developed for evaluating controller performance and identifying critical load points [7].

A significant contribution has been made in the field of mathematical modeling of data flow management processes. The application of queueing theory has enabled the formalization of relationships among traffic volume, controller performance, and request processing delays. Other researchers have proposed dynamic interaction models of controllers represented by systems of differential equations, allowing for the analysis of transient and steady states of the network. Furthermore, adaptive mathematical models have been developed to predict network traffic behavior and optimize flow distribution [8].

Thus, the literature review demonstrates that the mathematical representation of data flow control processes in cluster-based multicontroller SDNs remains a highly relevant research topic. The development of such models is essential for further optimization of controller architectures, enhancement of system stability, and improvement of network resource management efficiency.

Research Objective: To develop a mathematical model that accounts for dynamic clustering of SDN controllers during data acquisition from network switches, enabling balanced load distribution across controllers and real-time reallocation of links between controllers and switches.

2. Hierarchical Clustering for Multicontroller SDN Networks

The proposed hierarchical clustering algorithm operates in a dynamic mode, adapting to changes in network conditions in real time. The input parameters for the algorithm are the network traffic characteristics and data flow requests. The set of control devices is divided into clusters of SDN controllers, with each cluster assigned a main SDN controller responsible for coordination. Fig. 1 illustrates the hierarchical structure of a clustered SDN network.

The SDN network consists of a centralized alpha-controller (Cα) and several clusters of regular SDN controllers, with an equal number of controllers in each cluster. Each cluster has a main node, referred to as a beta-controller (Cβ). The beta-controller is responsible for cluster configuration and load balancing among the controllers within that cluster.

In Fig. 1, the following designations are used: β – beta-controller; *c* –SDN controller, a member of the cluster; *OFS* – OpenFlow Switchboard.

The primary objective of the developed clustering algorithm is to balance the load among controllers within the set of control devices and to mitigate issues related to controller failures. This approach reduces the overall operational cost of the SDN network while enhancing its availability, reliability, and fault tolerance.

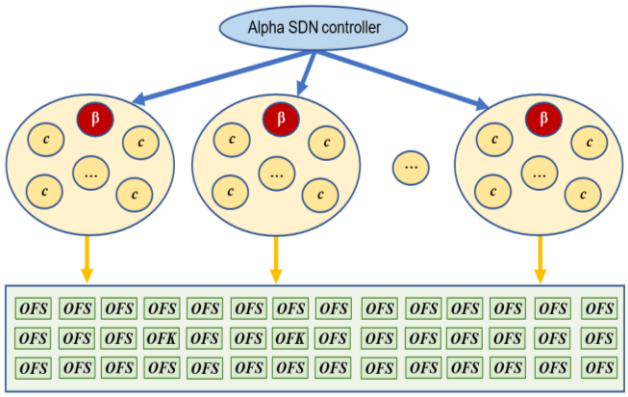


Fig. 1. General structure of a clustered multicontroller SDN network

The clustering process can be divided into three main phases: 1) controller configuration; 2) switch connection establishment; 3) formation of the steady state.

During the cluster formation phase, the alpha-controller initializes the SDN clusters by selecting beta-controllers and the member controllers for each cluster. The controller with the lowest expected load is selected as the beta-controller, assuming the role of coordinator for its respective cluster. All clusters are formed homogeneously, meaning that each contains the same number of SDN controllers. Once the clusters are established, the controller configuration phase concludes, and the switch connection phase begins.

Each cluster is then assigned a group of OpenFlow switches. During the dynamic allocation of switches among the member controllers, the system aims to achieve near-optimal delay and cost parameters.

At this stage, each SDN controller assumes a specific role – either as a regular cluster member or as a beta-controller – and establishes optimal connections with the assigned OpenFlow switches.

The process then proceeds to the steady-state phase. Each cluster member controller manages its connected OpenFlow switches and maintains the flow tables generated by those switches. Each beta-controller supervises both its assigned OpenFlow switches and the member controllers of the cluster, effectively serving as the head node of that cluster.

The beta-controller continuously monitors the traffic among the nodes within its cluster and reports to the SDN alpha-controller. When the beta-controller detects load imbalance among cluster members — for instance, when certain controllers are overloaded beyond a specified percentage of their maximum capacity, while others are underloaded — it initiates a load-balancing process among the cluster controllers through the following actions: 1) transferring the role of head node to a member controller with the lowest current load, which becomes the new beta-controller; 2) reassigning the OpenFlow switch connections among the cluster's SDN controllers. The reconnection process consists of the same three phases as the initial clustering procedure. However, unlike the overall clustering process, which is controlled by the alpha-controller, the reconnection process is managed locally by the beta-controller within

its cluster. These dynamic processes are executed automatically whenever network imbalance is detected, continuing until the load among clusters is equalized. The alpha-controller collects reports from all beta-controllers and monitors inter-cluster traffic. If a significant imbalance is detected between clusters, it triggers a new round of the clustering procedure, forming updated clusters using the previously defined phases. This dynamic hierarchical approach ensures continuous load balancing among distributed SDN controllers throughout the entire network operation period, enabling optimal utilization of computational resources and maintaining high network performance and stability.

2. Development of a Mathematical Model for the Functioning of a Multicontroller SDN

The set of deployed SDN controllers that constitute the collection of control devices can be represented as a vector *C*, defined as follows:

$$C = \{C_1, C_2, \ldots, C_i, \ldots C_N\}, \quad (1)$$

where C_i denotes an SDN controller with index *i*, and *N* is the total number of deployed SDN controllers.

The controllers responsible for managing clustering are also included in the set *C*, that is:

$$C_\alpha \in C; \quad C_\beta \in C. \quad (2)$$

The total number of clusters created is denoted as *M*, and this value may vary as the system state changes. The set of beta-controllers constitutes the subset C_β, which is defined as follows:

$$\begin{gathered} C\beta = \{C_{\beta 1}, C_{\beta 2}, \ldots, C_{\beta j}, \ldots, C_{\beta M}\}, \\ \operatorname{card} C\beta = M; \quad C\beta \subset C; \\ C_{\beta j} \in C \quad \forall j \in \overline{1, M}. \end{gathered} \quad (3)$$

Each formed cluster has an associated set of member controllers, the length of which is *L*. The set of member controllers for each *j*-th cluster can be expressed as:

$$\begin{gathered} Cj = \{C_{j1}, C_{j2}, \ldots, C_{j\ell}, \ldots, C_{jL}\}, \\ \operatorname{card} Cj = L \quad \forall j \in \overline{1, J}; \quad Cj \subset C; \\ C_{j\ell} \in C \quad \forall \ell \in \overline{1, L}. \end{gathered} \quad (4)$$

Let the set of SDN control devices manage *K* OpenFlow switches deployed in the network, distributed among the controller clusters (as shown in Fig. 1).

Each OpenFlow switch is connected to a specific SDN controller, with the allocation determined by the controller placement algorithm. The set of deployed OpenFlow switches can be represented as:

$$S = \{S_1, S_2, \ldots, S_k, \ldots S_K\}. \quad (5)$$

Each ℓ - th controller of the *j*-th SDN cluster has a specific set of connected OpenFlow switches, which can be defined by the following set:

$$\begin{gathered} Sj\ell = \{S_{j\ell 1}, S_{j\ell 2}, \ldots, S_{j\ell r}, \ldots S_{j\ell R_{j\ell}}\}, \\ \operatorname{card} Sj\ell = R_{j\ell}; \quad Sj\ell \subset S; \\ S_{j\ell r} \in S \quad \forall \ell \in 1, L; \quad \forall j \in 1, J, \end{gathered} \quad (6)$$

where $S_{j\ell r}$ is the *r*-th OpenFlow switch currently connected to the ℓ-th controller of the *j*-th SDN controller cluster.

Connections between OpenFlow switches and SDN controllers are represented by a Boolean switching matrix *W*, where rows correspond to SDN controllers and columns correspond to OpenFlow switches. The matrix *W* therefore reflects the total number of switches connected to each SDN controller. If the row corresponding to a specific controller consists entirely of zeros, that controller acts as an alpha-controller, which communicates exclusively with beta-controllers. The beta-controllers, in turn, manage the cluster controllers and have the lowest number of ones in the matrix *W* among all elements of their respective clusters.

Next, consider an example fragment of an SDN network that includes five SDN controllers and six OpenFlow switches, distributed across two clusters:

$$C = \{C_1, C_2, C_3, C_4, C_5\}; \; C_\alpha = C_1; \; C_{\beta 1} = C_2;$$

$$C_{\beta 2} = C_4; \quad C_{11} = C_3 \in C1; \quad C_{21} = C_5 \in C2;$$

$$\operatorname{card} C1 = \operatorname{card} C2 = 2; \; S = \{S_1, S_2, S_3, S_4, S_5, S_6\}.$$

For this fragment of the SDN network, the Boolean switching matrix *T* may be represented as follows:

$$W = \begin{array}{c} \\ C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \end{array} \begin{array}{c} \begin{array}{cccccc} S_1 & S_2 & S_3 & S_4 & S_5 & S_6 \end{array} \\ \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \end{array}. \quad (7)$$

One of the ways to evaluate the performance of a controller is by measuring its response time, which is mainly affected by queueing delay. The controllers can be modeled using a multi-server *M*/*M*/*s* queueing model, where each controller is assumed to have *s* cores. The transmitted packets arrive at the controller with a defined rate corresponding to a Poisson process, forming a single queue on the selected controller. The average response time T_i of controller Ci is the sum of the waiting time in the queue and the processing time, and it can be calculated using the Erlang formula as a function of the data packet arrival rate λ_i and the service rate μ:

$$T_i(\lambda) = P(s, \lambda_i/\mu)/(s \cdot \mu - \lambda_i) + 1/\mu, \quad (8)$$

where $P(s, \lambda_i/\mu)$ is the probability that all servers in the system are occupied and any incoming packet will be placed in the corresponding queue. This probability can be calculated as follows:

$$\begin{aligned} P(s, \lambda_i/\mu) &= \frac{\left((s \cdot \rho)^s / s!\right) \cdot (1-\rho)^{-1}}{\sum_{k=0}^{s-1} \left((s \cdot \rho)^k / k!\right) + \left((s \cdot \rho)^s / s!\right) \cdot \frac{1}{1-\rho}} = \\ &= 1 \Big/ \left(1 + \frac{1}{1-\rho} \cdot \frac{s!}{(s \cdot \rho)^s} \cdot \sum_{k=0}^{s-1} \left((s \cdot \rho)^k / k!\right)\right), \end{aligned} \quad (9)$$

where ρ is the server utilization factor, which indicates the stability of the system and is calculated as follows:

$$\rho = \lambda_i / (s \cdot \mu). \quad (10)$$

The system has a stable distribution only if the value of ρ is less than one. When the number of incoming requests in the queue exceeds the number of server cores, the controller operates at maximum throughput. The arrival rate of requests to a controller can be calculated as the sum of the average arrival rates of requests from all switches connected to that controller, i.e.

$$\lambda_i = \sum_{k_s} \lambda_{k_s} \quad (11)$$

where the summation is performed over all switches currently connected to the considered controller C_i. The average load on controller C_i can be calculated as the average number of requests that are queued for processing and those being processed. Using (8), the average load on controller C_i can be calculated as follows:

$$L_i(\lambda_i) = s \cdot \rho + (\rho/(1-\rho)) \cdot P(s, \lambda_i/\mu). \quad (12)$$

Thus, the developed mathematical model, represented by (1–6) and (8–12), allows for balancing the load among controllers and dynamically redistributing traffic in case of failure of any controller that belongs to the set *C*.

Conclusions

The paper presents a mathematical model for the operation of a clustered multicontroller SDN. The proposed approach involves performing three-level hierarchical clustering of the system's control and data transmission devices. The relationships between controllers and OpenFlow switches in the implementation of this clustered structure have been formalized. Calculations of the load distribution on the control elements have been carried out.

The use of the proposed mathematical model makes it possible to balance the load among controllers and quickly redistribute it in the event of a controller failure. The direction of further research is the development of an algorithm for distributing OpenFlow switches among SDN network controllers.

REFERENCE

1. McKeown, N. et al. (2008), "OpenFlow: Enabling Innovation in Campus Networks", *ACM SIGCOMM Computer Comm. Review*, vol. 38, is. 2, pp. 69–74, doi: https://doi.org/10.1145/1355734.1355746
2. Kreutz, D. et al. (2015), "Software-Defined Networking: A Comprehensive Survey", *Proceedings of the IEEE*, vol. 13, is. 1, pp. 14–76, doi: https://doi.org/10.1109/JPROC.2014.2371999
3. *Tootoonchian, A. and Ganjali, Y. (2010),* "HyperFlow: A Distributed Control Plane for OpenFlow," *INM/WREN*, available at: https://www.usenix.org/legacy/events/inmwren10/tech/full_papers/Tootoonchian.pdf
4. Yeganeh, S. H., Tootoonchian, A. and Ganjali, Y. (2013), "On Scalability of Software-Defined Networking", *IEEE Commun. Magazine*, available at: https://www.cs.toronto.edu/~soheil/papers/sdnscalability-ieeemag.pdf
5. Hu, Y., Wendong, W., Gong, X.; Que; X. and Shiduan C. (2012), "Reliability-Aware Controller Placement for Software-Defined Networks" *IEEE Communications Letters*, Gent, available at: https://ieeexplore.ieee.org/document/6573050/authors
6. Suarez-Varela, J. and Barlet-Ros, P. (2017), "Towards a NetFlow Implementation for OpenFlow Software-Defined Networks", *Proc. of the 29th Int. Teletraffic Congress Itc 2017*, 1, pp. 187–195, 8064355, doi: https://doi.org/10.23919/ITC.2017.8064355
7. Shekhawat, V.S., Kulshrestha, R., Yadav, P., Singh, A. and Firdous, F. (2025), "Modeling and performance evaluation of OpenFlow switches using a MAP/PH/1/n queueing model", *Computer Networks*, 266, doi: https://doi.org/10.1016/j.comnet.2025.111338
8. Zhang, J., Huang, X., Li, J., Sun, Q. and Lu, J. (2022), "A Dynamic Flow Table Management Method Based on Real-time Traffic Monitoring", IEEE International Conference on High Performance Switching and Routing Hpsr, 2022-June, pp. 212–217, doi: https://doi.org/10.1109/HPSR54439.2022.9831366

Received (Надійшла) 03.09.2025
Accepted for publication (Прийнята до друку) 05.11.2025

ВІДОМОСТІ ПРО АВТОРІВ/ ABOUT THE AUTHORS

Мигаль Станіслав Вікторович – аспірант кафедри автоматики, електроніки та телекомунікацій, Національний університет «Полтавська політехніка імені Юрія Кондратюка», Полтава, Україна;
Stanislav Myhal – PhD student of the Department of Automation, Electronics and Telecommunications, National University "Yuri Kondratyuk Poltava Polytechnic", Poltava, Ukraine;
e-mail: stas.migal1998@gmail.com; ORCID Author ID: https://orcid.org/0009-0007-4675-7358.

Математична модель системи управління потоками даних кластерної мультиконтролерної SDN

С. В. Мигаль

Анотація. Проблема балансування навантаження між контролерами та узгодження станів мережі є однією з ключових у мультиконтролерних SDN. Для подолання цієї проблеми необхідно розробляти адекватні математичні моделі, які описують процеси управління потоками даних. **Предметом** дослідження є система управління потоками даних в середовищі кластерної мультиконтролерної SDN. **Метою** дослідження є розробка математичної моделі, яка враховує динамічну кластеризацію контролерів SDN при отриманні даних з комутаторів мережі, та дозволяє збалансувати навантаження на контролери та проводити оперативний перерозподіл зв'язків між контролерами та комутаторами.. **Отримані наступні результати**. Розглянутий підхід до проведення трирівневої ієрархічної кластеризації пристроїв управління і прийому/передачі даних системи. Формалізовані зв'язки між контролерами та OpenFlow-комутаторами при реалізації даної кластерної структури. Проведені розрахунки навантаження на елементи управління. **Висновок**. Використання запропонованої математичної моделі дозволяє збалансувати навантаження між контролерами та оперативно перерозподілити навантаження при відмові якогось із контролерів.

Ключові слова: телекомунікаційна мережа, мережа зв'язку, котролер, літаюча мережа, OpenFlow-комутатор, стандарт 5G, стандарт 5G, SDN.

УДК 621.391.83 **doi: 10.26906/SUNZ.2025.4.190-194**

О. В. Михайліченко

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

МАТЕМАТИЧНА МОДЕЛЬ РАДІУСА ЗВ'ЯЗКУ МОБІЛЬНОГО НАЗЕМНОГО РЕТРАНСЛЯТОРА З УРАХУВАННЯМ ДИНАМІЧНОЇ ОРІЄНТАЦІЇ АНТЕНИ

Анотація. Підвищення вимог до ефективності мобільних ретрансляторів у гібридних мережах типу FANET і LoRa зумовлює потребу у створенні спрощених моделей оцінювання зв'язності. Об'єктом дослідження є процес поширення радіосигналу між мобільним ретранслятором і множиною вузлів у динамічному середовищі. Метою статті є розроблення спрощеної математичної моделі радіуса зв'язку, що враховує швидкість руху, орієнтацію антени та втрати сигналу на перешкодах. Запропонована модель забезпечує скорочення часу обчислень на 10–15 % при похибці не більше 15 % порівняно зі складною моделлю. Вона придатна для швидкого моделювання великих FANET/LoRa мереж, забезпечує баланс між точністю та швидкодією та може бути використана для оптимізації розташування мобільних ретрансляторів у телекомунікаційних системах.

Ключові слова: радіус зв'язку, мобільний ретранслятор, орієнтація антени, лог-дистанційна модель, гібридні мережі, FANET, LoRa, моделювання, швидкодія, телекомунікаційні системи.

Вступ

Постановка проблеми. У сучасних гібридних телекомунікаційних системах, де одночасно використовуються безпілотні літальні апарати (UAV) і наземні роботизовані платформи (UGV), стабільність радіозв'язку визначається не лише енергетичними характеристиками передавача, а й динамікою орієнтації антени під час руху.

Більшість існуючих моделей розрахунку дальності зв'язку враховують лише потужність передавання, коефіцієнт загасання середовища та випадкову тіньову складову, але не описують вплив зміни положення антени у просторі. Такі моделі є перевантаженими, оскільки включають десятки параметрів – кут нахилу, азимут, статистику федінгу, фазові зсуви тощо – що робить їх непрактичними для симуляцій великих мереж із сотнями вузлів.

У цьому контексті наземні ретранслятори (UGV-MR), які працюють як проміжні ланки зв'язку між UAV та базовими станціями, потребують спрощеної моделі, що одночасно враховує вплив швидкості, типу антени й перешкод у середовищі.

Аналіз останніх досліджень і публікацій. У більшості робіт UAV і UGV розглядаються окремо, без порівняння їх енергетичних характеристик у єдиному середовищі. Як показано у дослідженні Цзен та ін. [1], енергетична ефективність безпілотних літальних апаратів визначається насамперед геометрією траєкторії польоту, тоді як основні втрати енергії пов'язані з підтриманням висоти та стабілізацією. Це, своєю чергою, обмежує можливість застосування різноспрямованих антен і скорочує тривалість автономної роботи системи.

На відміну від цього, наземні платформи UGV не витрачають ресурсів на компенсацію підйомної сили й здатні працювати безперервно протягом тривалого часу.

Практичні результати, отримані Міллер та ін. [2], підтверджують, що використання енергоорієнтованого планування місій для UGV дозволяє істотно зменшити витрати енергії та частково перенести комунікаційні функції з UAV на наземні вузли. Такий підхід формує багаторівневу архітектуру управління, у якій UAV виконують роль систем збору даних, а UGV виступають стаціонарними або мобільними ретрансляторами.

Узагальнений аналіз Чжан та ін. [3] показує, що більшість відомих стратегій підвищення енергоефективності орієнтовані на оптимізацію мережевого рівня, тоді як антенно-фізичні обмеження, що безпосередньо впливають на якість зв'язку, залишаються поза увагою, а методи представлені Діао [4] та Омоніва [5], демонструють, що впровадження алгоритмів глибокого навчання дозволяє забезпечити гнучкий компроміс між якістю каналу та споживанням енергії, доводячи, що розумний розподіл функцій між платформами підвищує тривалість місій і зберігає стабільність зв'язку навіть у складних середовищах.

У систематичному огляді Мунасінгхе та ін. [6] зазначено, що більшість існуючих моделей нехтують точним урахуванням орієнтації антени. Автори наголошують на необхідності включення геометричних характеристик антен у математичні моделі радіоканалів.

Робота Юань та ін. [7] доводить, що у мережах FANET орієнтація антени та напрям руху вузлів мають визначальний вплив на стійкість зв'язку та пропускну здатність, що підтверджують експериментальні дані Азеведо та ін. [8].

У дослідженнях Джі [9] та Ван [10] підкреслюється, що надмірно складні та статистично перевантажені моделі не масштабуються для великих мереж і непридатні для оперативного аналізу, що актуалізує необхідність спрощених аналітичних підходів.

Метою роботи є розроблення спрощеної математичної моделі радіуса зв'язку мобільного наземного ретранслятора, яка враховує вплив швидкості руху, динамічної орієнтації антени та втрат на перешкодах, але не перевантажена надлишковими параметрами.

Запропонована модель повинна забезпечити можливість швидкого моделювання великих мереж у середовищі Python, зберігаючи при цьому фізичну достовірність і узгодженість із результатами класичних моделей поширення радіосигналу.

Основна частина

У класичних роботах для визначення радіуса дії ретранслятора використовується повна лог-нормальна модель загасання з урахуванням тіньового федінгу, багатопроменевого розсіяння, втрат на поляризації та зміни орієнтації антени.

Загальний вираз має вигляд:

$$P_{rad} = P_p + K_p(\theta,\varphi) + K_r - VS(r,f,n) - \\ - VS\text{fad} - VS_{\text{obs}} - VS_{\text{pol}}, \quad (1)$$

де P_p – потужність передавача,

$K_p(\theta,\varphi)$ – коефіцієнт підсилення передавальної антени залежно від її орієнтації (кутів нахилу θ та повороту φ);

K_t – підсилення антени приймача;

$VS(r,f,n)$ – втрати сигналу на відстані r;

VS_{fad} – випадкові втрати федінгу;

VS_{obs} – додаткові втрати від перешкод;

VS_{pol} – втрати через невідповідність поляризацій.

Таке рівняння є точним, проте при моделюванні великих мереж воно вимагає задання десятків параметрів і реалізації стохастичних розподілів, що значно ускладнює обчислення. Для систем, де важлива швидкість оцінювання зв'язності, цей підхід вважається перевантаженим.

Тому вводиться агрегування: дані групуються (1) у кілька «узагальнених» параметрів зі зрозумілою фізичною інтерпретацією, зберігаючи головні частки впливу (відстань, смуга/шум, тип антени, швидкість і перешкоди). Основою будь-якої моделі радіозв'язку є рівняння балансу потужності, яке описує, скільки енергії передавача реально досягає приймача після врахування всіх підсилень і втрат на шляху сигналу. У логарифмічній (дБ) шкалі ця залежність має такий вигляд:

$$\text{SNR}(r) = P_p + K_p + K_r - VS(r) - N, \quad (2),$$

де P_p – потужність передавача, дБм;

K_p, K_r – антенні підсилення, дБі;

$VS(r)$ – втрати шляхом, дБ;

N – шумова потужність на вході приймача, дБм.

На граничній відстані $r = R$, а SNR дорівнює пороговій чутливості SNR_{th}:

$$\text{SNR}_{\text{th}} = P_p + K_p + K_r - VS(R) - N. \quad (3)$$

Під час руху наземного ретранслятора орієнтація антени неідеальна, тому детальна геометрія замінюється лінійним штрафом.

Замість явної залежності $K_p(\theta,\varphi,t)$ (кут/час/джиттер) вводимо:

$$K_p^{\text{eq}} \triangleq K_0 - kv, \quad (4)$$

де K_0 – номінальне підсилення передавальної антени (коли платформа статична або добре стабілізована), дБ;

v – швидкість руху ретранслятора, м/с;

k – лінійний штраф у дБ на кожен м/с (емпіричний коефіцієнт, 0,2–0,5 дБ/(м/с)).

Для малого/середнього діапазону швидкостей (0-20 м/с) кутова похибка наведення зростає приблизно пропорційно швидкості (через обмежену смугу/потужність стабілізатора).

У вузькоспрямованих антенах втрата підсилення поблизу осі – квадратична по куту; у середньому це можна замінити ефективним лінійним штрафом по швидкості, що добре калібрується в польових умовах.

Сума складних ефектів (стіни/листя, частково – тіньову логнормальну варіацію) згортається у дискретний або агрегований тип:

$$VS_{\text{obs}} \in \{0,\ 10,\ 20\}\ \text{дБ}, \quad (5),$$

де внутрішніми значенням виступають «немає перешкод/помірні перешкоди/сильні перешкоди».

Шум у приймачі:

$$N = -174 + 10\log_{10} B + NF, \quad (6)$$

де – 174дБм/Гц є спектральною щільністю теплового шуму при 290 К; додавання $10\log_{10} B$ переводить у дБм на смугу B; NF – запас шуму.

У вільному просторі втрати з відстанню описує формула Фріса, де сигнал поширюється не лише у вільному просторі, а й через відбиття, дифракцію та розсіювання.

У реальних середовище умовах (місто / забудова / рослинність) посилює залежність від відстані, тому замість квадрату відстані використовують лог-дистанційну модель з показником загасання n.

Втрата на опорній відстані d_0(зазвичай 1 м) є:

$$VS(d_0) = 32{,}44 + 20\log_{10}(f) + 20\log_{10}(d_0), \quad (7)$$

А лог-дистанційна модель на довільній відстані r буде такою:

$$VS(r) = VS(d_0) + 10n\log_{10}\left(\frac{r}{d_0}\right) + VS_{\text{obs}}, \quad (8)$$

Після процесу підстановки та скорочень фінальна формула має вигляд:

$$R = d_0 \times 10^{\frac{P_p+K_0+K_r-kv-N-VS_{obs}-VS(d_0)-SNR_{th}}{10\times n}}, \quad (9)$$

Завдяки логарифмічній структурі рівняння розрахунок зв'язку для кожного вузла виконується за сталий час $O(1)$, що робить модель масштабованою для великих мереж (MANET, FANET, LoRa). Крім того, всі параметри мають чітку фізичну інтерпретацію і легко калібруються за експериментальними даними, що спрощує практичне застосування у симуляціях і реальних сценаріях.

Для визначення ефективності формули, змодельовано мережу за допомогою Python, розміром 1000 на 1000 м, що містить 40 вузлів (UAV) та 12 прямокутних перешкод із втратами 10–20 дБ (рис. 1).

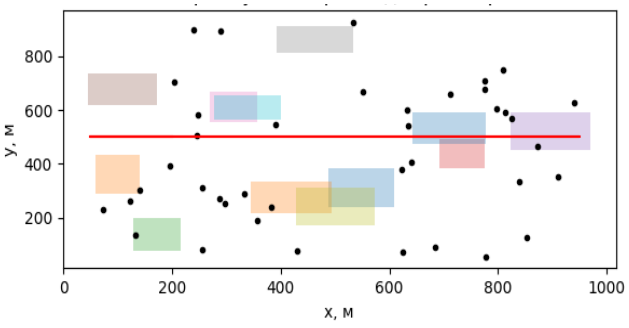


Рис. 1. Карта вузлів, перешкод і траєкторії

Ретранслятор (UGV) рухається по центральній горизонталі зі швидкістю 10 м/с. Рис. 2 демонструє зміну частки покритих вузлів у часі відносно цієї швидкості для LoRa 915 МГц і Wi-Fi 2.4 ГГц. Пунктирні лінії (класична модель) лежать вище суцільних (спрощена модель), оскільки не враховують втрати орієнтації антени та тіньові області. Спостерігається періодичний характер кривих, пов'язаний із циклічним рухом ретранслятора. Для LoRa середній рівень покриття перевищує 0,8 навіть у присутності перешкод, тоді як для 802.11 g він коливається біля 0,4.

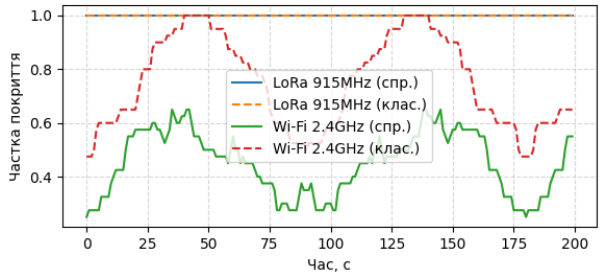


Рис. 2. Динаміка покриття у часі

Для обох технологій спостерігається плавне зниження покриття зі збільшенням швидкості, що пояснюється втратами через погіршення стабілізації антени (рис. 3). Класична модель демонструє завищені значення, не враховуючи орієнтаційні втрати, тоді як спрощена дає реалістичнішу картину поведінки зв'язку під час руху.

На рис. 4 показано порівняння граничного радіуса зв'язку $R(v)$ у спрощеній і класичній моделях. Пунктирні лінії (класична модель) відображають ідеальні умови, у яких дальність не змінюється з v, тоді як суцільні криві враховують зменшення підсилення антени з швидкістю. Це підтверджує адекватність спрощеної моделі для сценаріїв з мобільними платформами.

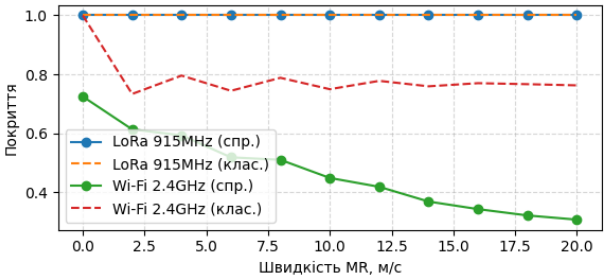


Рис 3. Середнє покриття в залежності від швидкості

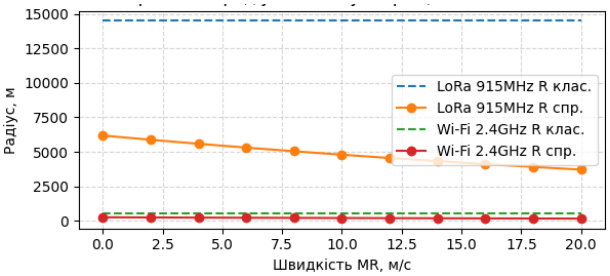


Рис. 4. Радіус зв'язку при різних швидкостях

На рис. 5, зображено порівняння є радіуси зв'язку для трьох типів антен при швидкості 10 м/с та втраті 10 дБ. Вузьконаправлена антена забезпечує найбільший радіус, всеспрямована найменшу дальність. Спостерігається нелінійна залежність між підсиленням антени й приростом дальності, що зумовлено впливом показника загасання n.

Таким чином, для низькошвидкісних мобільних платформ доцільно застосовувати антени із середнім підсиленням, які забезпечують компроміс між стабільністю і дальністю.

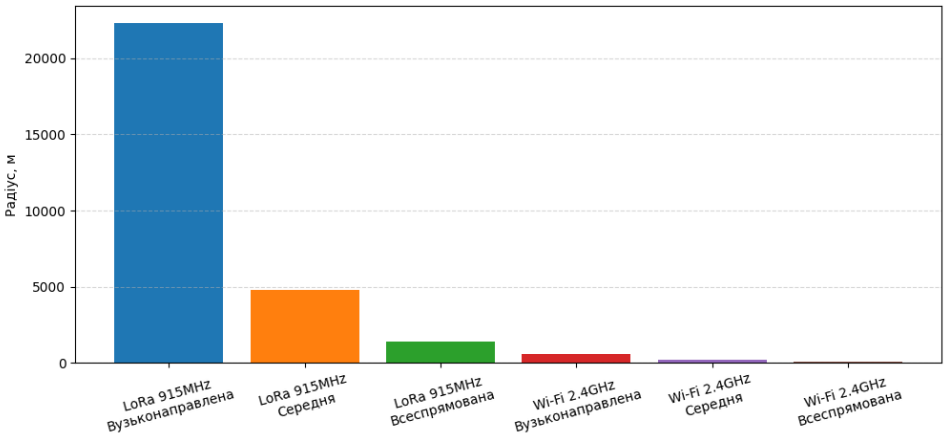


Рис. 5. Вплив типу антени на радіус дії

Для підтвердження збільшеної швидкодії, проведене моделювання складної моделі, що додатково враховує: кутову залежність антени, тіньовість, маломасштабний федінг та стохастичну варіацію втрат на перешкодах. Спрощена модель замінює ці ефекти одним агрегованим параметром штрафу орієнтації сталим значенням, що дає схожу динаміку зв'язності, але вимагає суттєво менше обчислень. Таким чином, відбувається заміна кількох стохастичних змінних одним детермінованим параметром — це спрощує обчислення та полегшує математичний аналіз мережі. Порівняльне моделювання двох підходів виконано для однакових стартових умов: однакова карта вузлів (40–320), швидкість руху ретранслятора $v = 10$ м/с, та два типи технологій – LoRa 915 МГц і 802.11g 2.4 ГГц.

У випадку LoRa 915 МГц, час моделювання для 320 вузлів склав 2,10 с у спрощеній моделі проти

2,36 с у складній, що відповідає прискоренню ≈ 1,13. Для 802.11g 2.4 ГГц різниця аналогічна — 2.15 с проти 2,38 с, або ≈ 1,12×. При цьому середня похибка між моделями для метрики покриття становить MAE = 0,334 (LoRa) і 0,238 (802.11g), що в межах 10–15 % від середнього значення, а отже не впливає на якісні закономірності результатів.

Усі досліджені моделі демонструють лінійну залежність часу обчислення від кількості вузлів, тобто складність обчислень масштабується як $O(N)$.

Висновки

Представлена спрощена математична модель розрахунку радіуса дії мобільного наземного ретранслятора з урахуванням динамічної орієнтації антени, типу середовища та швидкості руху відповідно до результатів симуляцій продемонструвала, що зі зростанням швидкості руху ретранслятора радіус дії зменшується майже лінійно через втрату ефективного підсилення антени. Для LoRa цей ефект проявляється помірно (зменшення радіусу приблизно на 25 %), тоді як для Wi-Fi — істотно (до 40 %). Порівняння зі «класичною» моделлю показало, що ігнорування орієнтаційних втрат і фізичних перешкод призводить до переоцінки покриття на 10–20 %. Результати порівняння зі складною моделлю підтвердили, що спрощена модель є більш ефективною для практичного моделювання мереж з великою кількістю вузлів. Вона забезпечує прийнятну точність (до 15 %) при зменшенні часу обчислень на ≈ 10–15 %, зберігаючи фізичну узгодженість формули балансу потужності.

Отримані залежності підтверджують, що логарифмічна форма моделі з мінімальною кількістю параметрів є оптимальною для практичного моделювання зв’язності у гібридних мобільних мережах. Запропонований підхід може бути використаний для адаптивного планування маршрутів мобільних ретрансляторів, вибору типу антен і прогнозування стійкості каналів у сценаріях змішаних наземно-повітряних систем зв’язку.

Список літератури

1. Zeng, Y., & Zhang, R. (2017). Energy-Efficient UAV Communication With Trajectory Optimization. IEEE Transactions on Wireless Communications, 16(6), 3747–3760. https://doi.org/10.1109/twc.2017.2688328
2. Miller, N., Goulet, N., & Ayalew, B. (2023). Energy-aware mission planning for unmanned ground vehicle fleets. Proc. of the Ground Vehicle Systems Engineering and Technology Symp. U.S. Army TARDEC. https://doi.org/10.4271/2024-01-4071
3. Zhang, Y., Zhao, R., Mishra, D., & Ng, D. W. K. (2024). A comprehensive review of energy-efficient techniques for uav-assisted industrial wireless networks. Energies, 17(18), 4737. https://doi.org/10.3390/en17184737
4. Diao, Y., Hu, Y., & Fu, J. (2024). Deep reinforcement learning-based UAV control for optimized energy efficiency and throughput in uav-assisted communication. У 2024 43rd chinese control conference (CCC) (pp. 2524–2530). IEEE. https://doi.org/10.23919/ccc63176.2024.10661522
5. Omoniwa, B., Galkin, B., & Dusparic, I. (2022). Optimising energy efficiency in uav-assisted networks using deep reinforcement learning. IEEE Wireless Communications Letters, 1. https://doi.org/10.1109/lwc.2022.3167568
6. Munasinghe, I., Perera, A., & Deo, R. C. (2024). A comprehensive review of UAV-UGV collaboration: Advancements and challenges. Journal of Sensor and Actuator Networks, 13(6), 81. https://doi.org/10.3390/jsan13060081
7. Yuan, Y., Ren, G., Cai, X., & Li, X. (2024). An adaptive 3D neighbor discovery and tracking algorithm in battlefield flying ad hoc networks with directional antennas. Sensors, 24(17), 5655. https://doi.org/10.3390/s24175655
8. Azevedo, J. A., & Santos, F. E. (2022). Performance evaluation of directional antennas in zigbee networks under NLOS propagation conditions. Electronics, 11(13), 2032. https://doi.org/10.3390/electronics11132032
9. Ji, A., & Wu, J. (2022). Joint deployment and power optimization for UAV relay in post-disaster situations. Wireless Communications and Mobile Computing, 2022, Article ID 9560806. https://doi.org/10.1155/2022/9560806
10. Wan, F., Yaseen, M. B., Riaz, M. B., Shafiq, A., Thakur, A., & Rahman, M. O. (2024). Advancements and challenges in uav-based communication networks: A comprehensive scholarly analysis. Results in Engineering, 103271. https://doi.org/10.1016/j.rineng.2024.103271

Received (Надійшла) 18.08.2025
Accepted for publication (Прийнята до друку) 22.10.2025

Відомості про авторів/ About the Authors

Михайліченко Олексій Валерійович – аспірант, Національний університет «Полтавська політехніка імені Юрія Кондратюка», Полтава, Україна;
Oleksii Mykhailichenko – PhD student, National University "Yuri Kondratyuk Poltava Polytechnic", Poltava, Ukraine;
e-mail: aleksejmikhajlichenko@gmail.com; ORCID Author ID: https://orcid.org/0009-0009-3512-0030.

Mathematical model of the communication radius of a mobile ground-based retransmitter taking into account the dynamic orientation of the antenna

Oleksii Mykhailichenko

Abstract. Increased requirements for the efficiency of mobile repeaters in hybrid networks such as FANET and LoRa necessitate the creation of simplified models for assessing connectivity. The object of the study is the process of radio signal propagation between a mobile repeater and a set of nodes in a dynamic environment. The purpose of the article is to develop a simplified mathematical model of the communication radius that takes into account the speed of movement, antenna orientation, and signal loss due to obstacles. The proposed model reduces computation time by 10–15% with an error of no more than 15% compared to the complex model. It is suitable for rapid modelling of large FANET/LoRa networks, provides a balance between accuracy and speed, and can be used to optimise the location of mobile repeaters in telecommunications systems.

Keywords: communication radius, mobile repeater, antenna orientation, log-distance model, hybrid networks, FANET, LoRa, modelling, performance, telecommunication systems.

УДК 004.9 **doi: 10.26906/SUNZ.2025.4.194-197**

С. С. Пироженко[1], С. С. Даценко[2]

[1] Харківський національний університет радіоелектроніки, Харків, Україна
[2] Національний технічний університет «Харківський політехнічний інститут», Харків, Україна

ТЕХНОЛОГІЯ СЕМАНТИЧНОГО ПЕРЕТВОРЕННЯ ІНФОРМАЦІЙНИХ ПОВІДОМЛЕНЬ У СЕРЕДОВИЩІ ПРОМИСЛОВОГО ІНТЕРНЕТУ РЕЧЕЙ

Анотація. Актуальність дослідження полягає у розробленні та впровадженні технології семантичного перетворення інформаційних повідомлень, яка підвищить ефективність обробки даних і сприятиме розвитку інтелектуальних систем управління промисловими процесами. **Мета дослідження:** розроблення технології семантичного перетворення інформаційних повідомлень у середовищі промислового Інтернету речей, яка забезпечує узгоджене представлення, інтерпретацію та обмін даними між гетерогенними пристроями і системами. **Результати.** У статті запропонована архітектура гетерогенного шлюзу промислового Інтернету речей. Доведена можливість реалізації технології семантичного шлюзу, який вирішує проблеми взаємодії різних прикладних технологій на рівні метаданих. У гетерогенному середовищі такий шлюз відповідає за перетворення протоколів промислового Інтернету речей. Проведений аналіз протоколів передачі даних, що використовуються в мережах IIoT з метою проведення семантичного перетворення. **Висновок.** Запропонований підхід дозволяє підвищити рівень інтероперабельності, зменшити інформаційні втрати під час трансляції повідомлень та створити основи для побудови інтелектуальних сервісів обробки даних у промислових кіберфізичних системах.

Ключові слова: промисловий Інтернет речей, семантичне перетворення, гетерогенний шлюз, протокол IIoT, інтероперабельність.

Вступ

Сучасний етап розвитку інформаційних технологій характеризується інтенсивною інтеграцією кіберфізичних систем, сенсорних мереж і хмарних платформ у межах концепції промислового Інтернету речей (IIoT, Industrial Internet of Things). У цьому середовищі ключову роль відіграє ефективний обмін даними між гетерогенними пристроями, що мають різні протоколи, формати та семантику інформаційних повідомлень. Забезпечення узгодженого тлумачення даних є критичним для підвищення надійності, гнучкості та адаптивності виробничих процесів.

Однією з головних проблем у IIoT є семантична неоднорідність інформації, яка ускладнює автоматизовану взаємодію між системами управління, аналітичними платформами та сенсорними пристроями. Традиційні методи обміну повідомленнями не гарантують збереження змістового контексту, що призводить до втрати смислової узгодженості або помилок при обробці даних.

Технологія семантичного перетворення інформаційних повідомлень орієнтована на вирішення цієї проблеми шляхом формування спільного онтологічного простору, який дозволяє інтерпретувати зміст даних незалежно від їхнього джерела або формату представлення. Такі підходи забезпечують інтелектуальну інтеграцію інформаційних потоків, автоматичну адаптацію даних до вимог користувача та контексту застосування, а також створюють основу для побудови когнітивних виробничих систем.

Отже, актуальність дослідження полягає у розробленні та впровадженні технології семантичного перетворення інформаційних повідомлень, яка забезпечить інтероперабельність компонентів IIoT, підвищить ефективність обробки даних і сприятиме розвитку інтелектуальних систем управління промисловими процесами.

Метою статті є розроблення технології семантичного перетворення інформаційних повідомлень у середовищі промислового Інтернету речей, яка забезпечує узгоджене представлення, інтерпретацію та обмін даними між гетерогенними пристроями і системами.

Запропонований підхід спрямований на підвищення рівня інтероперабельності, зменшення інформаційних втрат під час трансляції повідомлень та створення основи для побудови інтелектуальних сервісів обробки даних у промислових кіберфізичних системах.

1. Аналіз літературних джерел

Проблематика семантичної інтеграції даних у промисловому Інтернеті речей (IIoT) активно досліджується в останні роки, що пов’язано з необхідністю забезпечення узгодженого обміну інформацією між численними пристроями, системами управління та аналітичними платформами.

У роботах [1, 2] розглядаються підходи до формування семантичних моделей даних на основі онтологій, які дозволяють описувати структуру, взаємозв’язки та контекст інформаційних повідомлень. Зокрема, використання стандартів **OWL (Web Ontology Language)** і **RDF (Resource Description Framework)** забезпечує уніфіковане представлення знань і підтримку машинної інтерпретації.

У праці [3] підкреслюється роль семантичних шлюзів (semantic gateways) у забезпеченні взаємодії між гетерогенними пристроями та платформами IIoT. Такі шлюзи виконують функції трансляції протоколів, нормалізації форматів даних і контекстної інтерпретації повідомлень, що надходять із сенсорних мереж. Автори наголошують, що ефективна робота семантичних шлюзів потребує застосування методів онтологічного узгодження (ontology alignment) та семантичного мапінгу (semantic mapping).

Дослідження [4, 5] фокусуються на інтеграції семантичних технологій із хмарними та периферійними обчисленнями. Така комбінація дає змогу реалізувати розподілену обробку даних у реальному часі

з урахуванням контексту та пріоритетів задач. Застосування **Semantic Web Services** і **Linked Data** забезпечує автоматизовану інтерпретацію запитів та оптимізацію маршрутів передачі інформації.

Окрему увагу приділено питанням безпеки та довіри в процесі семантичного обміну повідомленнями [6]. Дослідники пропонують використовувати механізми цифрових підписів, шифрування метаданих і контроль доступу на основі семантичних політик, що підвищує надійність IIoT-комунікацій.

Таким чином, проведений аналіз показує, що основні напрями розвитку технологій семантичного перетворення в середовищі IIoT пов'язані з онтологічним моделюванням, автоматичним узгодженням даних, побудовою розподілених семантичних сервісів і забезпеченням кібербезпеки. Проте залишається відкритим питання створення комплексної технології, яка б поєднувала ці аспекти для ефективної адаптації повідомлень у динамічних промислових середовищах.

2. Структура гетерогенного шлюзу промислового Інтернету речей

Гетерогенний шлюз Інтернету речей – пристрій, що забезпечує взаємодію різних технологій Інтернету речей як між собою, так і з мережею зв'язку загального користування (МЗЗК) на всіх рівнях моделі OSI [7]. На рис. 1 зображено архітектуру гетерогенного шлюзу промислового Інтернету речей.

Гранична база даних	Граничний аналіз	…	Панель керування	…	*Послуги в екосфері IIoT*
Емуляція	Віртуальна машина	…	Контейнер	…	*Віртуальна екосфера*
IPv4/ IPv6	6LoWPAN	…	LoRaWAN	…	*Мережевий рівень*
IEEE 802.3	IEEE 802.15.1	…	IEEE 802.15.4	…	*Фізичний та канальний рівні*

Рис. 1. Архітектура гетерогенного шлюзу промислового Інтернету речей

Дана система забезпечує взаємодію різних технологій фізичного, канального та мережевого рівня та за допомогою спеціального програмного забезпечення (ПЗ), що функціонує у віртуальному оточенні, дозволяє перетворення даних між використовуваними технологіями на прикладному рівні. Таким чином, дана система складається з наступних компонентів:

• фізичні інтерфейси, що підтримують різні технології Інтернету;

• операційна система, системні драйвера, ПЗ, що дозволяють підтримувати функціонування протоколів фізичного, канального, мережевого та транспортного рівнів на гетерогенному шлюзі;

• віртуальне оточення (емулятори, контейнери, віртуальні машини та ін.), що дозволяє додавати нові можливості та послуги для гетерогенного шлюзу;

• ПЗ, що функціонує у віртуальному оточенні та дозволяє гетерогенним шлюзам розширювати існуючу функціональність. Наприклад, це ПЗ дозволяє розширити можливості гетерогенного шлюзу за рахунок підтримки нових протоколів, послуг та технологій.

Гетерогенний шлюз складається з основної частини, що дозволяє підтримувати різні технології Інтернету речей, системи віртуалізації, і прикладної частини, що функціонує у віртуальному оточенні і дозволяє впроваджувати без зміни основного робочого оточення гетерогенного шлюзу нові технології та послуги. Однією з таких технологій є семантичний шлюз промислового Інтернете речей.

3. Завдання семантичного перетворення повідомлень для гетерогенного шлюзу IIoT

Семантичний шлюз Інтернету речей – програмне забезпечення, що дозволяє здійснювати перетворення прикладних протоколів ІВ між собою та забезпечує загальний адресний простір для всіх пристроїв, що взаємодіють із цим ПЗ. Таким чином, семантичний шлюз вирішує проблеми взаємодії різних прикладних технологій на рівні метаданих. Проблема перетворення даних між різними протоколами на семантичному рівні є актуальною для промислового Інтернету речей. При використанні різних промислових рішень виникає проблема стикування технологій як на фізичному, канальному, мережевому, транспортному рівнях, так і прикладному рівні. Для вирішення цієї проблеми пропонується використовувати гетерогенний шлюз. Проблема взаємодії протоколів лише на рівні метаданих вирішується використанням спеціального ПЗ, що і буде промисловим семантичним шлюзом. Дане ПЗ виділяє ключову інформацію від кожного з пакетів даних, прикладний рівень яких заснований на одному з промислових протоколів, і перетворює на загальний проміжний формат Industrial Internet of Things Conversion Format (IIoTCF). У разі потреби використовуються функції ПЗ для перетворення отриманих даних у протоколи, що підтримуються, і подальшого їх відправлення в пункт призначення. На рис. 2 зображено структуру для перетворення протоколів IIoT. Це ПЗ складається з таких основних функцій:

• «Отримання пакету» – відповідає за прийом та фільтрацію пакетів, заснованих на одному з протоколів IIoT, що підтримуються. Дане ПЗ аналізує всі пакети, що приходять на активний мережевий інтерфейс і приймає ті пакети, формат яких заснований на одному з підтримуваних протоколів;

• «Черга 1..N» – черга пакетів, що надходять у програмне забезпечення, як для обробки, так і для відправлення;

Рис. 2. Структура семантичного шлюзу IIoT

• «В IICF» – перетворення пакетів, що надходять, у загальний формат IICF;

• "База даних IICF" – відповідає за зберігання пакетів. Може являти собою систему зберігання даних, що вбудовується в програмне забезпечення, обмежений буфер, програмний інтерфейс для взаємодії із зовнішніми системами зберігання даних та ін. Може зберігати інформацію про протокол, який підтримує пункт призначення, прописаний у конкретному пакеті;

• «IICF» – відповідає за перетворення даних із формату IICF у необхідний формат згідно з протоколом;

• «Надіслати пакет» – відповідає за відправку сформованого пакета до потрібного пункту призначення.

Як прикладне рішення для семантичного шлюзу промислового Інтернету речей пропонується використовувати гетерогенний шлюз IIoT, шлюз, що встановлюється в рамках одного рішення IIoT і відповідає за перетворення протоколів промислового Інтернету речей між собою.

На рис. 3 зображено приклад програмно-апаратного комплексу (ПАК), що поєднує різні підмережі, які функціонують на базі різних протоколів промислового Інтернету речей.

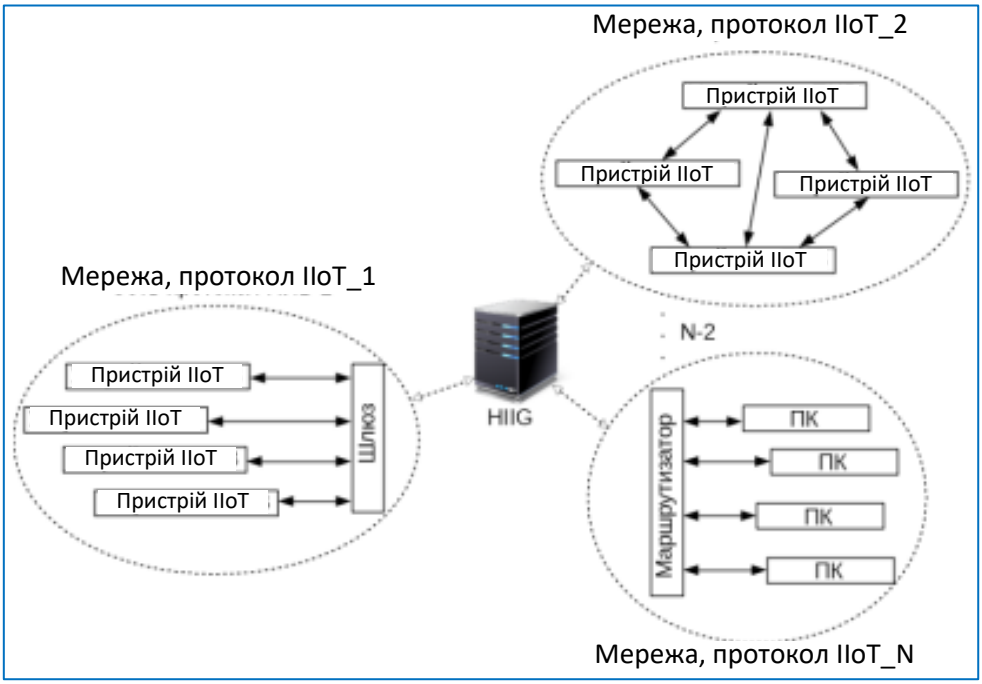


Рис. 3. Структура програмно-апаратного комплексу, що містить гетерогенний шлюз IIoT

4. Аналіз протоколів передачі даних, що використовуються в мережах IIoT

Для розробки структури системи, що відповідає за перетворення та формування пакетів на базі

Відомих форматів даних різних протоколів IB у форматі IICF, необхідно досліджувати формати даних різних промислових протоколів і виділити загальні поля даних. Для цього дослідження було обрано такі найпоширеніші протоколи: CoAP; MQTT; XMPP; ModBus RTU; ModBus TCP; OPC UA; HTTP.

Основні поля в протоколі CoAP є такими: версія CoAP (2 біти); тип повідомлення (2 біти); кількість додаткових опцій (4 біти); код – метод запиту даних або код запиту (8 біт), наприклад GET (1), POST (2), PUT (3), DELETE (4); ідентифікатор повідомлення (16 біт); опції та корисні дані повідомлення.

Основні поля в протоколі MQTT є такими: тип повідомлення (4 біти); прапор перезапиту повідомлення - DUP (1 біт); рівень QoS (2 біти); прапор збереження повідомлення (1 біт); довжина опціональної інформації (8 біт); довжина заголовка та даних.

Формат метаданих для протоколу XMPP такий: версії заголовка XML; тип повідомлення; ідентифікатор сесії; версія XML; ідентифікатори повідомлення, відправника, отримувача; мова повідомлення; адреса потоку; текст повідомлення.

Формат метаданих для протоколу ModBus RTU має такі основні поля: стартовий прапор; адресу ModBus (8 біт); функції ModBus (8 біт); поле даних.

Формат метаданих для протоколу HTTP орієнтується на такі основні поля: тип запиту; URI-домен, що запитується; версія HTTP-запиту; URI-адреса призначення; опції з'єднання; опції контролю кешу; дані про користувача клієнта; дані запиту; кодування даних; мова даних; дані про сервер; час.

Висновки

У статті запропоновано технологію семантичного перетворення інформаційних повідомлень у середовищі промислового Інтернету речей, яка забезпечує узгоджене представлення, інтерпретацію та обмін даними між гетерогенними пристроями і системами. Розроблена архітектура гетерогенного шлюзу промислового Інтернету речей. Доведена можливість реалізації технології семантичного шлюзу, який вирішує проблеми взаємодії різних прикладних технологій на рівні метаданих. У гетерогенному середовищі такий шлюз відповідає за перетворення протоколів промислового Інтернету речей. Проведений аналіз протоколів передачі даних, що використовуються в мережах IIoT з метою проведення семантичного перетворення.

Запропонований підхід дозволяє підвищити рівень інтероперабельності, зменшити інформаційні втрати під час трансляції повідомлень та створити основи для побудови інтелектуальних сервісів обробки даних у промислових кіберфізичних системах.

Список літератури

1. Yang, X., Huang, J., Ao, F. and Yin, J. (2023), “Ontology-based Semantic Data Model for Command and Control”, 2023 9th Int. Conf. on Big Data and Inf. Analytics Bigdia Proc., pp. 330–335, doi: https://doi.org/10.1109/BigDIA60676.2023.10429478
2. Chalapathi, G.S.S., Chamola, V., Vaish, A. and Buyya, R. (2022), “Industrial internet of things (Iiot) applications of edge and fog computing: A review and future directions”, *Advances in Information Security*, vol. 83, pp. 293–325, doi: https://doi.org/10.1007/978-3-030-57328-7_12
3. Zuev, A., Karaman, D. and Olshevskiy, A. (2023), “Wireless sensor synchronization method for monitoring short-term events”, *Advanced Information Systems*, vol. 7, no. 4, pp. 33–40, doi: https://doi.org/10.20998/2522-9052.2023.4.04
4. Gramoli, V. (2020), “From blockchain consensus back to Byzantine consensus”, *Future Generation Computer Systems*, vol. 107, pp.760–769, doi: https://doi.org/10.1016/j.future.2017.09.023
5. Kovalenko, A. and Kuchuk, H. (2022), “Methods to Manage Data in Self-healing Systems”, *Studies in Systems, Decision and Control*, vol. 425, pp. 113–171, doi: https://doi.org/10.1007/978-3-030-96546-4_3
6. Kuchuk, H. and Malokhvii, E. (2024), “Integration of IOT with Cloud, Fog, and Edge Computing: A Review”, *Advanced Information Systems*, vol. *8*(2), pp. 65–78, doi: https://doi.org/10.20998/2522-9052.2024.2.08
7. Decker, C. and Wattenhofer, R. (2014), “Bitcoin transaction malleability and MtGox”, *European symposium on research in computer security*, pp. 313–326, doi: https://doi.org/10.1007/978-3-319-11212-1_18

Received (Надійшла) 25.07.2025
Accepted for publication (Прийнята до друку) 29.10.2025

Відомості про авторів / About the Authors

Пироженко Сергій Станіславович – аспірант кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Serhii Pyrozhenko – PhD candidate of Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;
e-mail: serhii.pyrozhenko@nure.ua; ORCID Author ID: http://orcid.org/0009-0006-4209-2144;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=59675302400&origin=resultslist.

Даценко Сергій Сергійович – аспірант кафедри комп’ютерної інженерії та програмування, Національний технічний університет «Харківський політехнічний інститут», Харків, Україна;
Serhii Datsenko – PhD candidate of Department of Electronic Computers, National Technical University «Kharkiv Polytechnic Institute», Kharkiv, Ukraine;
e-mail: Serhii.Datsenko@gmail.com; ORCID Author ID: http://orcid.org/0009-0004-4372-1913.

Technology of semantic transformation of information messages in the Industrial Internet of Things environment

Serhii Pyrozhenko, Serhii Datsenko

Abstract. The relevance of the research is the development and implementation of the technology of semantic transformation of information messages, which will increase the efficiency of data processing and promote the development of intelligent industrial process control systems. **The purpose of the research**: the development of the technology of semantic transformation of information messages in the industrial Internet of Things environment, which provides a consistent representation, interpretation and exchange of data between heterogeneous devices and systems. **Results.** The article proposes the architecture of a heterogeneous gateway for the industrial Internet of Things. The possibility of implementing the technology of a semantic gateway, which solves the problems of interaction of various applied technologies at the metadata level, is proven. In a heterogeneous environment, such a gateway is responsible for the transformation of industrial Internet of Things protocols. An analysis of data transmission protocols used in IIoT networks for the purpose of semantic transformation is carried out. **Conclusion.** The proposed approach allows to increase the level of interoperability, reduce information losses during message transmission, and create a basis for building intelligent data processing services in industrial cyber-physical systems.

Keywords: industrial Internet of Things, semantic transformation, heterogeneous gateway, IIoT protocol, interoperability.

УДК 621.396.67 **doi: 10.26906/SUNZ.2025.4.198-202**

В. М. Почерняєв[1], М. С. Магомедова[2], Н. М. Сивкова[1], О. С. Ястреба[3]

[1] Національна академія Служби безпеки України, Київ, Україна
[2] Київський фаховий коледж зв'язку, Київ, Україна
[3] Національний університет «Полтавська політехніка імені Юрія Кондратюка», Полтава, Україна

ДАТЧИК ПОТУЖНОСТІ НВЧ НА ЧАСТКОВО ЗАПОВНЕНИХ ДІЕЛЕКТРИКОМ ПРЯМОКУТНИХ ХВИЛЕВОДАХ

Анотація. У системах автоматичної стабілізації потужності функції вимірювального елемента виконує датчик потужності. Тракти приймачів модулів НВЧ для низки мобільних радіотехнічних систем пропонується реалізовувати на частково заповнених діелектриком хвилеводах. Вихідна потужність в діапазоні НВЧ є одним з найбільш важливих параметрів джерел сигналів і підсилювачів НВЧ, передавальних трактів НВЧ і радіотехнічних систем різного призначення в цілому. Від рівня вихідної потужності НВЧ залежить можливість використання джерела сигналів НВЧ у радіолокаційних станціях, наземних телекомунікаційних станціях НВЧ, супутникових системах передачі. При проектуванні радіотехнічної системи рівень потужності джерела сигналів НВЧ визначає склад радіоелектронної бази передавального тракту НВЧ, компонування антенно-фідерного тракту та установки електроживлення системи. В роботі приводяться схеми передавального тракту НВЧ радіотехнічної системи НВЧ та конструкція датчика потужності НВЧ на частково заповненому діелектриком хвилеводі. Оскільки існують радіотехнічні системи НВЧ, в тому числі цифрові тропосферні станції, які працюють у декількох піддіапазонах частот, то необов'язково весь діапазон апроксимулювати ортогональними поліномами, що мають рівнохвильовий характер. В роботі запропоновано, що для таких радіотехнічних систем НВЧ доцільніше використовувати ортогональні поліноми Лежандра та ортогональні поліноми Гегенбауера. Розрахункові дані показані у вигляді графіків, де порівнюються коефіцієнти відображення в нормованому частотному діапазоні, що отримані за поліномами четвертого та п'ятого порядків. В висновках роботи вказано, що такі датчики потужності НВЧ можуть бути застосовані в вузловій цифровій тропосферній станції, високошвидкісній цифровій тропосферній станції та у комбінованих цифрових тропосферних станціях.

Ключові слова: датчик потужності надвисоких частот, цифрові радіотехнічні системи НВЧ, ступінчатий перехід на частково заповненому діелектриком хвилеводі, поліноми Лежандра, поліноми Гегенбауера.

Вступ

Вихідна потужність в діапазоні НВЧ є одним з найбільш важливих параметрів джерел сигналів і підсилювачів НВЧ, передавальних трактів НВЧ і радіотехнічних систем різного призначення в цілому. Від рівня вихідної потужності НВЧ залежить можливість використання джерела сигналів НВЧ. У радіолокаційних станціях, наземних телекомунікаційних станціях НВЧ, супутникових системах передачі потужність передавача НВЧ визначає дальність дії системи. При проектуванні радіотехнічної системи рівень потужності джерела сигналів НВЧ визначає склад радіоелектронної бази передавального тракту НВЧ, компонування антенно-фідерного тракту та установки електроживлення системи.

У процесі дослідно-конструкторської розробки, серійного випуску виробу, контролю параметрів під час експлуатації цього виробу необхідний контроль над рівнем потужності НВЧ. Як правило, такий контроль за прохідної потужності НВЧ покладається на датчик потужності НВЧ (ДП).

У сучасних радіотехнічних системах застосовуються фазовані антенні решітки (ФАР), що складаються з великої кількості приймально-передавальних модулів НВЧ. Для забезпечення необхідних діаграм спрямованості, коефіцієнта посилення, мінімуму бічних пелюсток ФАР формується певний розподіл амплітуди сигналів приймально-передавальних модулів НВЧ. При цьому важливо забезпечити симетричне відносно центру ФАР амплітудне розподілення. Для забезпечення необхідних характеристик ФАР у різних умовах експлуатації мобільних радіотехнічних систем необхідно зберегти обране значення вихідної потужності кожного приймально-передаваль-ного модуля НВЧ з високою точністю. Це завдання вирішується шляхом застосування систем автоматичної стабілізації потужності (АСП). У системах АСП функції вимірювального елемента виконує ДП НВЧ. Тракти приймачів модулів НВЧ для низки мобільних радіотехнічних систем пропонується реалізовувати на частково заповнених діелектриком хвилеводах (ЧЗДХ).

Аналіз літератури. Слід зазначити, що ДП НВЧ використовуються не тільки в телекомунікаційних та радіолокаційних системах [1–3], але в таких галузях науки та техніки в яких використовується НВЧ обладнання, як вимірювальна техніка [4–6], криптографія [7], дослідження атмосфери та космічного простору [8], дослідження природних ресурсів [8–10], біофізика та медицина [11–15].

Метою роботи є створення та дослідження датчика потужності НВЧ на ЧЗДХ.

Основна частина

Враховуючи, що в роботах [16–19] побудовані та досліджені на ЧЗДХ система автоматичного регулювання потужності передавача (АРПП), пристрій управління, регулювання, обмеження потужності НВЧ, то є необхідним створення ДП на ЧЗДХ. Крім того, багатокаскадні передавальні тракти НВЧ мобільних цифрових тропосферних станцій мають у своєму складі декілька ДП (рис. 1).

На рис. 1 показаний двокаскадний передавальний тракт НВЧ радіотехнічної системи НВЧ, де ПП1, ПП2 – підсилювач потужності; НВ – направлений відгалужувач; ДП – датчики потужності;

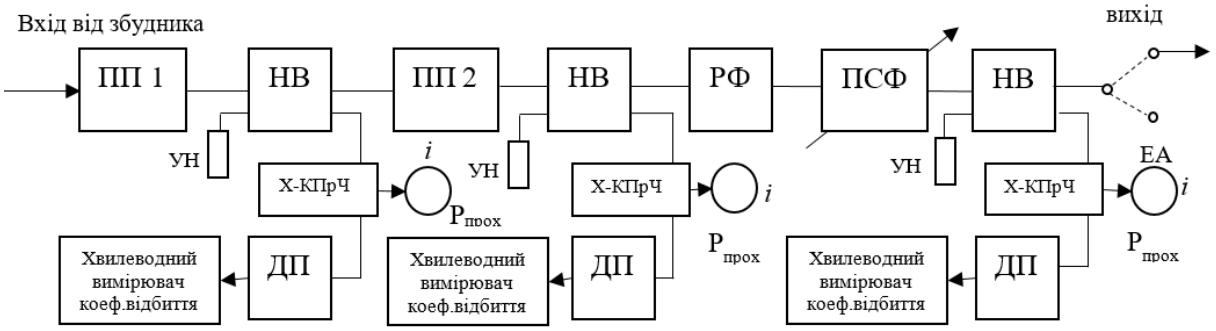


Рис. 1. Структура передавального тракту НВЧ радіотехнічної системи НВЧ

«*i*» – індикатор прохідної потужності $P_{прох}$; РФ – режекторний фільтр; ПСФ – смуговий фільтр, що перебудовується; ЕА – еквівалент антени; УН – узгоджувальне навантаження; Х-КПЧ – хвилеводно-коаксіальний перетворювач частоти; ХВ – хвилеводний вимірюавач коефіцієнта відбиття $|\Gamma_н|$. Відмітимо, що в якості Х-КПЧ та ХВ використовуються стандартні вимірювачі, які випускаються виробниками вимірювальної техніки. Контроль за вихідною потужністю на кожному каскаді та всього тракту НВЧ в цілому полягається на ДП. Цим забезпечується також контроль системи АРПП [16], так як дана система включає пристрій управління потужності [17], пристрій регулювання потужності [18] та обмежувач потужності [19]. Дані пристрої реалізовані на ЧЗДХ.

На рис. 2 показано конструкцію ДП: 1 – відкрита нелінійна структура (ВНС) [20]; 2 – діелектрична пластина; 3 – металевий прямокутний хвилевід.

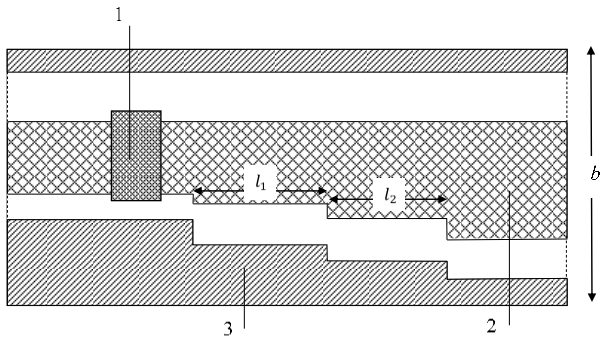


Рис. 2. Конструкція датчика потужності НВЧ

В основі розрахунку падаючої потужності, що проходить по тракту передачі НВЧ покладено співвідношення:

$$P_{пад} = P_{прох}/(1 - |\Gamma_н|^2),$$

де $P_{прох}$ – потужність, що надходить від НВ (потужність, що відгалужується з передаючого тракту НВЧ); $\Gamma_н$ – коефіцієнт відбиття від навантаження (конструкції ДМ).

Коефіцієнт відображення $\Gamma_н$ визначається за такими формулами:

$$\Gamma_н = -\frac{y}{2+y},$$

$$y = g_{ВНС} + jb,$$

$$jb = jb_{ВНС} + jb_{перех},$$

де $g_{ВНС}, jb_{ВНС}$ – активна та реактивна провідність ВНС, визначені в роботах [17–19]; $jb_{перех}$ – реактивна провідність ступінчатого переходу.

З даних виразів знаходимо $|\Gamma_н|$:

$$|\Gamma_н| = \frac{\sqrt{[g_{ВНС} + (2 + g_{ВНС}) + jb^2]^2 + 4b^2}}{(2 + g_{ВНС})^2 + b^2}.$$

Особливість роботи деяких радіотехнічних систем у тому, що частотний інтервал між робочими піддіапазонами не використовується. Звідси випливає, що в неробочому інтервалі частот досягнення заданого допуску на неузгодження не обов'язково, а достатньо забезпечити необхідний допуск на неузгодженість у робочих піддіапазонах частот. Так, частотні характеристики, які апроксимуються поліномами Чебишева першого роду є рівнохвильовими і мають коливальний характер у всій робочій смузі частот. Але чебишевські переходи, окрім задоволення вимог по заданому допуску на неузгодженість на граничних частотах робочої смуги, повинні задовольняти цю вимогу всередині інтервалу ще $(n - 1)$ - разів, де n – порядок полінома, тобто допуск на неузгодженість повинен досягатися у робочій смузі частот $(n - 1)$ - разів. Це накладає жорсткі вимоги на точність виготовлення чебишевських ступінчатих переходів і є серйозним недоліком при їх серійному виробництві.

З вищого викладеного випливає, що функція робочого згасання ступінчатих переходів має бути описана іншими ортогональними поліномами.

Враховуючи ускладнення конструкції ступінчатих переходів та особливості робочої смуги частот, наприклад, цифрових тропосферних станцій, доцільно мати частотну характеристику коливального, але нерівнохвильового характеру. Вибір апроксимуючого полінома ґрунтується на порівнянні властивостей поліномів Чебишева першого роду з сферичними поліномами Лежандра, ультрасферичними поліномами Гегенбауера, поліномами Чебишева другого роду, які ортогональні на інтервалі (-1;+1). Ці поліноми являються рішенням лінійного диференціального рівняння другого порядку виду $(1 - x^2)y'' - (2m + 1)xy' + n(2m + n)y = 0$. При $m = 0$, маємо поліноми Чебишева першого роду $T_n(x)$, при $m = 1$ – поліноми Чебишева другого роду $U_n(x)$, при $m = 1/2$ – сферичні поліноми Лежандра $P_n(x)$ або ультрасферичні поліноми Гегенбауера виду $C_n^{1/2}(x)$.

Порівняльний аналіз ортогональних поліномів на інтервалі (-1;+1) показує:

– поліноми Чебишева першого роду $T_n(x)$ ортогональні з вагою $(1-x^2)^{-1/2}$ і має норму $\pi/2(n\neq 0), \pi(n=0)$;

– поліноми Чебишева другого роду $U_n(x)$ ортогональні з вагою $(1-x^2)^{1/2}$ і має норму $\pi/2$;

– сферичні поліноми Лежандра $P_n(x)$ ортогональні з вагою 1 і мають норму $2/(2n+1)$;

– ультрасферичні поліноми Гегенбауера виду $C_n^{1/2}(x)$, як часний випадок ультрасферичних поліномів, що співпадає з сферичними поліномами Лежандра, ортогональні з вагою 1 і мають норму $\Gamma(n+1)/n!\,(n+1/2)$.

Ультрасферичні поліноми Гегенбауера виду $C_n^{1/2}(x)$ з вагою $(1-x^2)^{\alpha-1/2}$ і мають норму $[\pi 2^{1-2\alpha}\Gamma(n+2\alpha)]/[n!\,(n+\alpha)[\Gamma(\alpha)]^2]$.

В даному випадку при синтезі ступінчатих переходів будемо використовувати сферичні поліноми Лежандра $P_n(x)$ або ультрасферичні поліноми Гегенбауера виду $C_n^{(\alpha)}(x)$. Наведемо явний вигляд виразів для сферичних поліномів поліноми Лежандра для $n=1\ldots5$:

$$\begin{aligned} P_1(x) &= x,\\ P_2(x) &= (3x^2-1)/2,\\ P_3(x) &= (5x^3-3x)/2,\\ P_4(x) &= (35x^4-30x^2+3)/8,\\ P_5(x) &= (63x^5-70x^3+15x)/8. \end{aligned} \quad (1)$$

Частотні характеристики ступінчатих переходів можна апроксимувати різним числом членів рядів ортогональних поліномів залежно від їх вагової функції та норми і заданої точності, або виділити ділянки інтервалу, в яких при заданому допуску на неузгодженість можна використовувати поліном з меншим ступенів.

Функція робочого згасання L ступінчатого переходу має вигляд:

$$L = 1 + h^2 Q_n^2(x), \quad (2)$$

де $h = |\Gamma_н|_{max}/\sqrt{1-|\Gamma_н|^2_{max}}$ – амплітудний множник; $|\Gamma_н|_{max}$ – найбільше значення модуля коефіцієнта відображення у робочій смузі частот (допуск на неузгодженість), Q_n – поліном n-порядку, n – число ступенів узгоджувального переходу, $x = cos\Theta/S$, $\Theta = \beta l$, β – фазова постійна, l – довжина ступеня, $S = sin[\pi(\beta_В-\beta_Н)/2(\beta_В+\beta_Н)]$ – масштабний множник; $\beta_В, \beta_Н$ – фазові постійні на граничних частотах робочої смуги.

Аналіз (1) та (2) при $Q_n(x) = P_n(x)$ показує, що лежандровський ступеневий перехід при коливальному характері частотної характеристики в робочій смузі має $|\Gamma_{11}| < |\Gamma_{11}|_{max}$ в (n-1) - точці всередині інтервалу. Це свідчить про те, що при серійному виробництві лежандрівським переходом можуть бути менш жорсткі вимоги щодо точності виготовлення, ніж до чебишевських переходів.

Точно повинно дотримуватися вимог допуску на неузгодження лише на граничних частотах робочої смуги, що завжди можливо.

Оскільки ступінчаті переходи на частково заповнених діелектриком прямокутних хвилеводах (ЧЗДПХ) конструктивно складніше, ніж переходи на порожнистих прямокутних хвилеводах, недоліки переходів чебишевського типу проявляється сильніше. Тому розраховуємо ступінчаті переходи на ЧЗДПХ лежандрівського типу.

Досліджується ступінчатий перехід на ЧЗДПХ зі змінним розміром вузької стінки (стрибок поперечного перерізу хвилеводу в Е – площині). Такий перехід може бути використаний для узгодження підсилювального НВЧ каскаду із НВЧ трактом на ЧЗДПХ передавального пристрою цифрової тропосферної станції сантиметрового діапазону хвиль.

Методика розрахунку конструкції ступінчатого переходу заснована на методі невизначених коефіцієнтів та методі власних функцій, застосованого до розрахунку реактивних провідностей плоскопоперечних стиків двох ЧЗДПХ. Представляючи ступінчатий перехід у вигляді взаємного, реактивного, антиметричного чотириполюсника, на підставі маємо:

$$\begin{aligned} 2hQ_n(x) &= \left(\sqrt{R}-\frac{1}{\sqrt{R}}\right)cos^2\beta l -\\ &-(\rho_1\sqrt{R}/\rho_2-\rho_2/\rho_1\sqrt{R})sin^2\beta l, \end{aligned} \quad (3)$$

де $\rho_1 = b_1/b_0$, $\rho_2 = b_2/b_0$, $R = b/b_0$, b_1, b_2 – вузькі стінки відрізка хвилеводів, що утворюють першу та другу сходинки. Представивши в (3) замість $Q_n(x)$ величину $P_2(x)$ з (1), отримаємо наступні співвідношення для амплітудного множника h та масштабного множника S:

$$\begin{aligned} h &= \rho_1\sqrt{R}/\rho_2-\rho_2/\rho_1\sqrt{R},\\ S &= 1/\sqrt{[1+(R-1)/h\sqrt{R}]/3}. \end{aligned} \quad (4)$$

Отримані для лежандрівського переходу вирази (4) відрізняються від аналогічних виразів для Чебишевського переходу числовими множниками. Далі знаходимо:

$$b_1 = b_0\sqrt{h\sqrt{R}/2+\sqrt{h^2R/4+R}},\ b_2 = bb_0/b_1. \quad (5)$$

Тепер визначимо довжини сходинок:

$$l_1 = (\pi-\widetilde{b_1})/2\beta,\ l_2 = (\pi-\widetilde{b_2})/2\beta, \quad (6)$$

де $\widetilde{b_1}$,$\widetilde{b_2}$ – нормовані реактивні провідності плоскопоперечних стиків ЧЗДПХ.

Розрахунок двоступінчатого переходу ілюструється прикладом узгодження ЧЗДПХ із поперечним перерізом $a\times b = 40\times 20$мм^2, $\varepsilon_r = 2{,}4$; $c/a = 0{,}2$; $d/b = 0{,}8$ (c, d - поперечні розміри діелектричної пластини з відносною діелектричною проникністю ε_r) з ЧЗДПХ з поперечним перерізом

$$a\times b_0 = 40\times 8\ \text{мм}^2,$$

діелектрична пластина якого має такі самі відносні параметри. Допуск на неузгодженість $|\Gamma_н|_{max}$=0,09 (КСХ≤ 1,2) повинен дотримуватися у 10% робочої смуги. На рис. 3 показані залежності модуля коефіцієнта відбиття $|\Gamma_н|^2$ у нормованому діапазоні довжин хвиль, яку відповідають робочим частотним піддіапазоном 4,4 ... 4,6 ГГц та та 4,7 ... 4,9 ГГц радіотехнічної системи НВЧ.

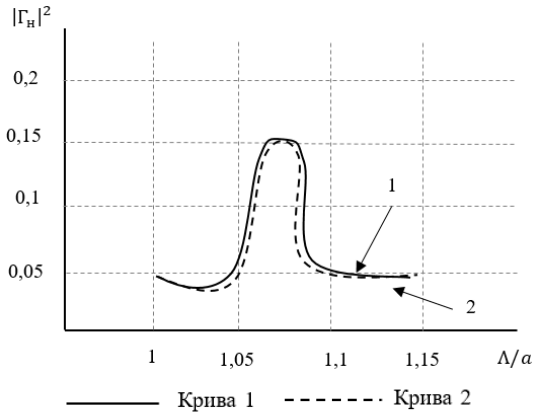


Рис. 3. Залежності модуля коефіцієнта відбиття від нормованого частотного діапазону радіотехнічної системи НВЧ

Крива 1 побудована для випадку, коли в якості апроксимуючих поліномів є поліном Лежандра $P_4(x)$ та ультрасферичний поліном Гегенбауера виду $C_4^{(1/2)}(x)$, крива 2 – коли в якості апроксимуючих поліномів є поліном Лежандра $P_5(x)$ та ультрасферичний поліном Гегенбауера виду $C_5^{(1/2)}(x)$. Ортогональні поліноми $P_4(x)$ та $C_4^{(1/2)}(x)$,як $P_5(x)$ та $C_5^{(1/2)}(x)$ точно збігаються, а апроксимація поліномами степенів $n = 4$ та $n = 5$ практично дають однаковий результат. Така точність апроксимації дозволяє не збільшувати ступінь полінома.

Висновки

Як і вряду передавальних трактів НВЧ радіотехнічних систем НВЧ, так і в системах АСП, необхідні датчики прохідної потужності НВЧ. Цього вимагає також і фазована антенна решітка. Однак, ДП вимагають узгодження з основними трактами, що зручно робити за допомогою ступінчастих переходів. Але є особливість у робочих діапазонах частот радіотехнічних систем. Наприклад, цифрові тропосферні станції мають декілька робочих піддіапазонів частоти, інтервали між якими не використовуються. У цьому випадку, як показало дослідження, доцільно в якості апроксимуючого поліному застосовувати не чебишевські, а сферичні поліноми Лежандра або ультрасферичні поліноми Гегенбауера. Для мобільної вузлової цифрової тропосферної станції та високошвидкісної цифрової тропосферної станції, мобільної комбінованої цифрової тропосферно-радіорелейної станції та комбінованої цифрової тропосферно-космічної станції, мобільної комбінованої цифрової тропосферно-іоносферної станції розроблені ДП, які вбудовуються в передавальний тракт НВЧ зі стандартними вимірювальними пристроями [21, 22].

СПИСОК ЛІТЕРАТУРИ

1. AK Jha, A. Lamecki, M. Mrozowski, M. Bozzi, A Highly Sensitive Planar Microwave Sensor for Detecting Direction and Angle of Rotation // IEEE Transactions on Microwave Theory and Techniques, 2020. Vol. 68(4). P.1598-1609 http://doi.10.1109/TMTT.2019.2957369
2. Muñoz-Enano J., Vélez P., Gil M., Martín F. Planar Microwave Resonant Sensors: A Review and Recent Developments // Applied Sciences. 2020, Vol.10(7). P.2615. https://doi.org/10.3390/app10072615
3. J.-H.Deng, H.Xiong, Q.Yang, M.Suo, J.-Y. Xie, H.-Q. Zhang Metasurface-Based Microwave Power Detector for Polarization Angle Detection // IEEE Sensors Journal, 2023. Vol. 23(19). pp. 22459-22465. https://doi.org/10.1109/JSEN.2023.3306462
4. Dazhen Gu NIST-Traceable Microwave Power Measurement in a Waveguide Calorimeter With Correlated Uncertainties // IEEE Transactions on Instrumentation and Measurement, 2019. Vol. 68, №. 6. P. 2280-2287. https://ieeexplore.ieee.org/document/8599075
5. Zhihao Chen, Yu Fu, Hiroshi Kawarada, Yuehang Xu. Microwave diamond devices technology: Field-effect transistors and modeling // National Natural Science Foundation of China, 26 August 2020. Vol.34, №1. https://doi.org/10.1002/jnm.2800
6. E.Nazemosadat, S.García, I.Gasulla Heterogeneous multicore fiber-based microwave frequency measurement // Opt. Express, 2022. Vol. 30(15). P.26886-26895. http://doi.10.1364/OE.463152
7. Shijie Song, Suen Xin Chew, Linh Nguyen, Xiaoke Yi High-resolution microwave frequency measurement based on dynamic frequency-to-power mapping // Optics Express, 2021. №29, P.42553-42568 https://opg.optica.org/oe/fulltext.cfm?uri=oe-29-26-42553&id=465714
8. Alahnomi R.A., Zakaria Z., Yussof Z.M., Althuwayb A.A., Alhegazi A., Alsariera H., Rahman N.A. Review of Recent Microwave Planar Resonator-Based Sensors: Techniques of Complex Permittivity Extraction, Applications, Open Challenges and Future Research Directions // Sensors, 2021 Vol. 21. P.2267. https://doi.org/10.3390/s21072267
9. Li Dai, Xue Zhao, Jiuchuan Guo, Shilun Feng, Yusheng Fu, Yuejun Kang, Jinhong Guo Microfluidics-based microwave sensor // Sensors and Actuators A: Physical, 2020 Vol. 309. P. 111910 https://doi.org/10.1016/j.sna.2020.111910
10. Kazemi N., Schofield K., Musilek P. A High-Resolution Reflective Microwave Planar Sensor for Sensing of Vanadium Electrolyte // Sensors, 2021. Vol.21. P.3759. https://doi.org/10.3390/s21113759
11. Kandwal et al., Surface Plasmonic Feature Microwave Sensor With Highly Confined Fields for Aqueous-Glucose and Blood-Glucose Measurements // IEEE Transactions on Instrumentation and Measurement, 2021. Vol. 70, P.1-9. https://doi.org/10.1109/TIM.2020.3017038
12. M.C. Cebedio, L.A. Rabioglio, I.E. Gelosi, R.A. Ribas, A.J. Uriz, J.C. Moreira Analysis and Design of a Microwave Coplanar Sensor for Non-Invasive Blood Glucose Measurements // IEEE Sensors Journal, 2020. Vol. 20(18). P.10572-10581 https://doi.org/10.1109/JSEN.2020.2993182
13. S. Kiani, P. Rezaei, M. Fakhr Dual-Frequency Microwave Resonant Sensor to Detect Noninvasive Glucose-Level Changes Through the Fingertip // IEEE Transactions on Instrumentation and Measurement, 2021. Vol.70. P.1-8. https://doi.org/10.1109/TIM.2021.3052011
14. G. Govind, M. J. Akhtar Metamaterial-Inspired Microwave Microfluidic Sensor for Glucose Monitoring in Aqueous Solutions // IEEE Sensors Journal, 2019. Vol.19(24), P.11900-11907. https://doi.org/10.1109/JSEN.2019.2938853

15. C. G. Juan. Study of Q_u-Based Resonant Microwave Sensors and Design of 3-D-Printed Devices Dedicated to Glucose Monitoring // IEEE Transactions on Instrumentation and Measurement, 2021. Vol.70, pp.1-16 https://doi.org/10.1109/TIM.2021.3122525
16. Почерняєв В.М., Повхліб В.С., Сивкова Н.М. Система автоматичного регулювання потужності передавача НВЧ для комбінованих мобільних цифрових тропосферно-радіорелейних станцій // Вісник НТУУ «КПІ», 2021. №84. с.40-47. http://radap.kpi.ua/radiotechnique/article/view/1692
17. Почерняєв В.М., Сивкова Н.М. Пристрій управління потужністю НВЧ на частково заповненому діелектриком прямокутному хвилеводі // Інфокомунікаційні та комп'ютерні технології, 2021. №1(01). с.81-89. https://visn-icct.uu.edu.ua/index.php/icct/article/view/28/7
18. Почерняєв В.М., Сивкова Н.М., Магомедова М.С. Пристрій регулювання потужністю НВЧ на частково заповнених діелектриком прямокутних хвилеводах // Інфокомунікаційні та комп'ютерні технології, 2021. №2(02). с.161-171. https://visn-icct.uu.edu.ua/index.php/icct/article/view/49/36
19. Почерняєв В.М., Сивкова Н.М., Магомедова М.С. Обмежувач потужності НВЧ на частково заповнених діелектриком прямокутних хвилеводах // Інфокомунікаційні та комп'ютерні технології, 2022. №1(03). С.90-101. https://visn-icct.uu.edu.ua/index.php/icct/article/view/69/55
20. Почерняєв В.М., Сивкова Н.М. Зовнішні параметри з'єднання прямокутного хвилевода, частково заповненого лінійним діелектриком з прямокутним хвилеводом, частково заповненим нелінійним діелектриком» // Вісник Університету «Україна», 2020. №1(24). с.100-105. https://visn-it.uu.edu.ua/old_site/article.php?full=vysnyk-606f1968be142
21. Почерняєв В.М., Сивкова Н. М., Магомедова М.С. Мобільна вузлова цифрова тропосферна станція // Системи озброєння і військова техніка. 2024, №4(76). С.6-15. https://journal-hnups.com.ua/index.php/soivt/article/view/1542/1408
22. Почерняєв В.М., Магомедова М.С., Сивкова Н. М. Мобільні цифрові тропосферні станції з комбінуванням просторово-рознесених сигналів // Системи управління, навігації та зв'язку, 2024. №3. С. 211-215. https://journals.nupp.edu.ua/sunz/article/view/3475/2896

Received (Надійшла) 27.07.2025
Accepted for publication (Прийнята до друку) 15.10.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Почерняєв Віталій Миколайович – доктор технічних наук, професор, професор кафедри, Національна академія Служби безпеки України, Київ, Україна;
Vitaly Pochernyaev – Doctor of Technical Sciences, Professor, Professor of the Department, National Academy of the Security Service of Ukraine, Kyiv, Ukraine;
e-mail: vpochernyaev@gmail.com; ORCID Author ID: https://orcid.org/0000-0001-7130-8668;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=6603517180.

Магомедова Марія Сергіївна – доктор філософії, викладач вищої категорії, Київський фаховий коледж зв'язку, Україна;
Mariia Mahomedova – PhD, Higher Category Teacher, Kyiv Professional College of Communications, Kyiv, Ukraine;
e-mail: kkz.praktika@ukr.net; ORCID Author ID: https://orcid.org/0000-0003-1936-5555.

Сивкова Наталія Максимівна – доктор філософії, доцент кафедри, Національна академія Служби безпеки України, Київ, Україна;
Natalia Syvkova – PhD, Associate Professor of the Department, National Academy of the Security Service of Ukraine, Kyiv, Ukraine;
e-mail: natsivonat@gmail.com; ORCID Author ID: https://orcid.org/0000-0002-4934-4109;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57223023373.

Ястреба Олег Сергійович – аспірант кафедри автоматики, електроніки та телекомунікацій, Національний університет «Полтавська політехніка імені Юрія Кондратюка», Полтава, Україна;
Oleh Yastreba – postgraduate student of the Department of Automation, Electronics and Telecommunications, National University "Yuri Kondratyuk Poltava Polytechnic", Poltava, Ukraine;
e-mail: ol.yastr17@gmail.com; ORCID Author ID: https://orcid.org/0000-0003-3911-4109401.

Power microwave sensor on partially filled rectangular waveguides by dielectric

Vitaly Pochernyaev, Mariia Mahomedova, Natalia Syvkova, Oleh Yastreba

Abstract. In automatic power stabilization systems the functions of the measuring element are performed by a power sensor. It is proposed to implement the microwave module receiver paths for a number of mobile radio engineering systems on partially filled waveguides by dielectric. The output power in the microwave range is one of the most important parameters of microwave signal sources and amplifiers, microwave transmitting paths and radioengineering systems of various purposes in general. The ability to use a microwave signal source depends on the level of microwave output power. In radar stations, ground-based microwave telecommunication stations and satellite transmission systems the microwave transmitter power determines the range of the system. The article presents the schemes of the microwave transmission path of a digital troposcatter station and the design of a microwave power sensor on a waveguide partially filled by dielectric. Since there are microwave radiosystems, including digital troposcatter stations, operating in several frequency subranges, it is not necessary to approximate the entire range with orthogonal polynomials of equalwave nature. The article suggests that for such microwave radio systems it is more appropriate to use Legendre orthogonal polynomials and Gegenbauer orthogonal polynomials. The calculated data are shown in the form of graphs, where the mapping coefficients obtained from fourth- and fifthorder polynomials are compared. The conclusions of the article indicate that such microwave power sensors can be used in a nodal digital troposcatter station, a highspeed digital troposcatter station and in combined digital troposcatter stations.

Keywords: power microwave sensor, digital microwave radiosystems, step transition on a partially filled waveguide by dielectric, Legendre polynomials, Gegenbauer polynomials.

UDC 621.396.96 **doi: 10.26906/SUNZ.2025.4.203-206**

Vitalii Rudenko

National University "Yuri Kondratyuk Poltava Polytechnic", Poltava, Ukraine

DISTANCE MEASUREMENT USING TDOA METHOD BASED ON LORA PROTOCOL

Abstract. The article focuses on the implementation of distance measurement between objects using the Time Difference of Arrival (TDoA) method, specifically through the Round-Trip Time of Flight (RTToF) technique, based on the LoRa protocol and utilized by the SX1280 transceiver. **The aim of the article** is to detail the temporal characteristics of packet exchange and formalize the calculation of distance. The key parameters significantly affecting measurement accuracy are identified as Bandwidth (BW) and Spreading Factor (SF). The research includes a comprehensive analysis of error sources, such as Reference Oscillator Error (timing offset) and Multipath Propagation, and provides formulas for their quantification. A multi-stage correction methodology is proposed, incorporating Frequency Error Correction (based on FEI), LNA Compensation (based on RSSI), Statistical Filtering (using the median), Polynomial Curve Correction, and Environment-Specific Correction. **The results obtained** from field measurements indicate that with this comprehensive correction approach, a ranging precision of approximately 1 meter can be achieved in controlled line-of-sight conditions. Furthermore, practical outdoor applications demonstrate mean errors below 10 meters even at distances exceeding 1 km. **Conclusions**: The implementation and comprehensive correction of RTToF measurements using the LoRa-based SX1280 module allow for the achievement of high accuracy while maintaining an acceptable balance between precision, time-on-air (energy consumption), and communication range, which is crucial for object detection and positioning systems.

Keywords: TDoA, LoRa, SX1280, RTToF, ToF, Ranging.

TDoA

Measuring the distance between objects is a key task in target detection systems, such as radars. To detect an object, a radar sends powerful radio pulses that reflect off the object, with part of the pulse energy returning to the radar. The propagation speed of the pulses is close to the speed of light, which is a known value. By receiving the pulse return time, the radar determines the distance to the object.

The TDoA method operates on a similar principle [2]. The master device sends a ranging request to a slave device and activates a timer. When the slave device receives the ranging request, it immediately returns a response to the master device. The master device stops the timer when the response arrives and stores it in a result register. The measurement result is the Time of Flight (ToF), from which we derive the distance between objects using the formula:

$$d = (T_f/2) \cdot c, \quad (1)$$

where d is the distance between objects; $T_f/2$ is the response reception time divided by 2, because the signal travels double the distance; c is the speed of light.

The slave takes time T_r to process, form, and send the response, which must be subtracted (Fig. 1):

$$d = ((T_f - T_r)/2) \cdot c, \quad (2)$$

Thus, the TDoA method involves message exchange between master and slave. During the exchange, the master activates a timer that records the time when it receives the response from the slave. From this time, the distance between objects is calculated.

This method is implemented on radio modules such as the SX1280 from Semtech, where it operates based on LoRa modulation at a 2.4 GHz frequency [1]. The SX1280 performs one round of exchange, and based on the response reception time, the master calculates the result, a number that is relative to the distance between objects.

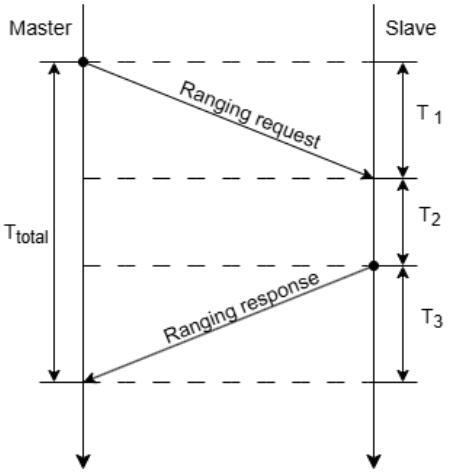


Fig. 1. TDoA timing diagram

There is also the ATA8352 radio module from Microchip [2]. This chip operates in the 6.2 GHz–8.3 GHz range and uses Impulse-Radio Ultra-Wideband for distance determination. Unlike the SX1280, the ATA8352 radio module supports Double-Sided Two-Way-Ranging (DS-TWR). DS-TWR application exchanges a sequence of data telegrams between the nodes and captures the timestamps of these data telegrams at the transmitter and receiver nodes to measure the distance between them.

Implementation of RTToF Method on SX1280 Radio Module and Data Processing Algorithm

The SX1280 module performs distance measurement using its integrated Ranging Engine [3]. Here, TDoA is implemented as Round Trip Time of Flight (RTToF). This is a simple form of TDoA where the measurement session consists of one round of ranging packet exchange. Each exchange is a complete measurement operation. However, for accurate measurements, it is advisable to conduct several exchanges to obtain a more precise result.

In RTToF, there are clearly defined roles: master and slave [3]. Before the actual measurement, roles must be explicitly assigned between radio modules. Also, distance measurement is a directional operation. The slave receives a logical address. The Master performs measurement with a specific slave in its coverage area, embedding the slave's address in the ranging packet.

The ranging packet structure is similar to a regular LoRa packet but with some differences [3] (Fig. 2). The preamble is the same as in a regular LoRa packet. The header includes the slave address and checksum. The payload carries no information. In this form, the Master sends this packet to the slave. The slave receives it and sends back a response that has no preamble or header, only payload. The timer on the master counts until the response from the slave arrives.

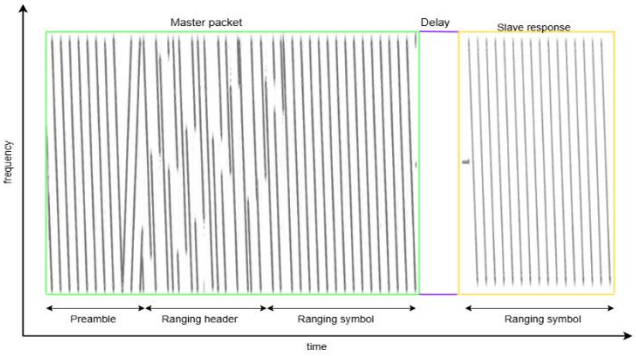


Fig. 2. Ranging packet spectrogram

The total packet time depends on such factors as: bandwidth, spreading factor, preamble length, number of ranging symbols, and the time to switch slave from RX to TX mode. In simplified form, the packet time $T_{ranging}$ is the time of one symbol T_s multiplied by their total number $N_{ranging\_symbol\_exchange}$ [1]:

$$T_{ranging} = T_s * N_{ranging\_symbols}. \quad (3)$$

The symbol time T_s depends on Spreading Factor SF and Bandwidth BW according to the equation [1]:

$$T_s = 2^{SF}/BW. \quad (4)$$

$N_{ranging\_symbol\_exchange}$ consists of the sum of master packet symbols $N_{symbols\_master}$, slave response symbols $N_{symbols\_slave}$, and transmission delay N_{symbol_delay} [1]:

$$N_{ranging\_symbol} = N_{symbols\_master} + \\ + N_{symbol\_delay} + N_{symbols\_slave};$$
$$N_{ranging\_symbol} = N_{symbols\_master} + \\ N_{symbol\_delay} + N_{symbols\_slave}. \quad (5)$$

As already mentioned, the packet from the master consists of the sum of preamble symbols $N_{preamble}$, header symbols N_{symbol_header}, and payload symbols $N_{ranging\_symbols}$ [1]:

$$N_{symbols\_master} = N_{preamble} + \\ + N_{symbol\_header} + N_{ranging\_symbols}. \quad (6)$$

The response from the slave consists only of payload $N_{ranging\_symbols}$ [1]:

$$N_{symbols\_slave} = N_{ranging\_symbols}. \quad (7)$$

N_{symbol_delay} is the deterministic symbol equivalent duration of the silence between the end of ranging request reception and the beginning of ranging response transmission, which is defined as a constant $N_{symbol\_delay} = 2$ [1]. The preamble $N_{preamble}$ consists of the sum of preamble symbols $N_{symbol\_preamble}$ and the constant 4.25, which corresponds to special end-of-preamble symbols [1]:

$$N_{preamble} = N_{symbol\_preamble} + 4.25. \quad (8)$$

The number of ranging header symbols is fixed at 16 symbols $N_{symbol\_header} = 16$ [1]. This gives us the general formula for measuring the exchange time $T_{ranging}$ [1]:

$$T_{ranging} = 2^{SF}/BW \times (N_{symbol\_preamble} + \\ + 2 \cdot N_{ranging\_symbols} + 22.25). \quad (9)$$

From which we can separately extract the master component $T_{ranging\_master}$ and slave component $T_{ranging\_slave}$ [1]:

$$T_{ranging\_master} = 2^{SF}/BW \times \\ (N_{preamble} + N_{ranging\_symbols} + 16); \quad (10)$$

$$T_{ranging\_slave} = (2^{SF}/BW) \cdot N_{ranging\_symbols}. \quad (11)$$

Ranging result Traw is stored in a specific register of the transceiver as raw data (a 24-bit two's complement number); thus, the result should be transformed into distance using the following formula [4]:

$$D = (T_{raw}/BW) \cdot 36621.09375. \quad (12)$$

It is evident that the key parameters are BW and SF. The SX1280 radio module for distance measurement supports BW of 400 kHz, 800 kHz, 1600 kHz, and SF from 5 to 10.

An individual ranging measurement on a single channel can be expected to have a ranging precision of approximately 1 m at high SF and high BW [5]. In the broader context of link design trade-offs, this implies that there is a compromise to be struck between accuracy, time-on-air (equivalently energy consumption), and the range of the link. Higher accuracy can be attained by increasing the number of measurements or increasing the SF and BW. Lower time on air can be achieved by reducing the SF and increasing BW. Longer range is possible by reducing the BW and increasing the SF. Recalling that SF and modem BW are the main variables available to the designer, the graphic illustrates this design compromise for a given LoRa modem setting (Fig. 3).

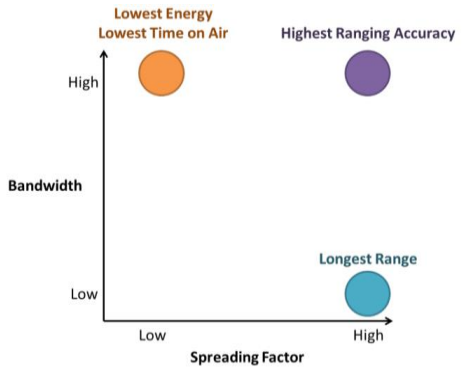


Fig. 3. Bandwidth and SF relation for a given application [5]

Result Correction Techniques and Measurement Accuracy

The accuracy of distance measurements using the SX1280 radio module is affected by several sources of error that must be considered and corrected to achieve optimal results.

We need to draw an important delineation between static sources of error and those that can change during the ranging operation. It is relatively simple to correct static bias, so it should be corrected through transceiver calibration., so it should be corrected by transceiver calibration. The remaining dynamic sources of error can be further divided into those internal to the radio and those due to the channel in which the ranging signals propagate. Assuming a perfect communication channel, we see only errors due to circuit level phenomenon: Reference oscillator drift and Analog group delay [4].

Reference oscillator drift is a frequency difference between master and slave oscillators. The timing measurements on the master and synchronization on the slave are performed using local crystal reference oscillators. Any timing offset between the crystals of the master and slave results in distance measurement error. Since the same reference oscillator is used to derive both the ranging timer and the RF carrier frequency, the measured frequency error between the transmitter and receiver can reliably indicate the timing (and thus distance) offset between devices.

For a worst-case offset of ±40 ppm and maximum response delay of 42.9 ms (at BW = 400 kHz, SF = 10), the theoretical maximum ranging error is approximately ±257 m. This error reduces to ±2 m at minimum settings [5]. The distance offset is not linear, especially at higher spreading factors.

Because the RF signal and the timing reference for the ranging synchronization operation are the same crystal reference oscillator, a simple frequency error measurement of the frequency error between master and slave, using the LoRa Frequency Error Indicator (FEI), can be used to accurately evaluate the timing (and equivalently, distance) error.

Analogue Group Delay Error occurs in a LNA. The SX1280 employs Automatic Gain Control (AGC) to adapt the LNA gain to the received signal strength. The delay through the LNA varies as a function of amplifier gain. Measurements have shown that variation in attenuation at a fixed distance can result in over 8 m of measurement error [4].

In non-line-of-sight (nLoS) conditions, signals propagate via multiple reflected paths of different lengths, causing frequency selectivity and distance overestimation. The multipath resolution capability, expressed as the difference in length approximately [8]:

$$d_r \approx \mathrm{c}/BW. \qquad (13)$$

For BW = 1600 kHz, this yields a measured distance of approximately 185 m, while true distance is 150 m. Field measurements have confirmed that multipath propagation typically causes ranging distance overestimation of 30 – 40 m in nLoS scenarios [8].

Due to this error several correction techniques were introduced [5–8].

Semtech applies correction based on FEI. Before ranging operations, the master gets the FEI factor from the slave node by a separate communication packet. After converting the raw ranging result to meters, a correction is applied based on the FEI obtained during the communication phase. The corrected result is calculated as:

$$d_{result} = d_{raw} - k_{FEI} \cdot FEI/1000, \qquad (14)$$

where k_{FEI} is a gradient value stored as a function of SF and BW. It is a mapped collection provided by Semtech. For example, at SF9 and BW = 1600 kHz, k_{FEI}= -0.424[4]. Also, Semtech provides maps with LNA correction values. Same as k_{FEI} it a gradient value, but it is a function RSSI, SF and BW. It is applied as:

$$d_{result} = d_{raw} + d_{rssi\ error}, \qquad (15)$$

where $d_{rssi\ error}$ is a corrective value got from RSSI map.

After applying corrections, the median value is calculated from multiple frequency-hopped ranging exchanges. The median has been found empirically to provide better immunity to outliers than the arithmetic mean, which is crucial for avoiding inaccurate channel measurements. Semtech performs 40 – 80 frequency-hopped exchanges across the 2.4 GHz ISM band using the Bluetooth Low Energy channel plan.

Even in line-of-sight conditions, measurements show range underestimation at short distances (< 10 m), overestimation at medium distances (10-80 m), and underestimation beyond 80 m. A polynomial correction is applied to linearize results:

$$d_{\text{corrected}} = \sum_{\mathrm{k=0}}^{\mathrm{n}} a_k \cdot d_{\text{median}}^{n-k}, \qquad (16)$$

where the coefficients k are determined through field measurements and curve fitting. Different polynomial orders and coefficients are used for each SF-BW combination. For outdoor ranging up to 100 m, a 7th-order polynomial has been used successfully [4].

Stuart Robinson introduced an additional linear correction for practical applications. It can be derived from measurements in the specific operating environment [10]. Field tests have shown that applying a simple correction formula:

$$d_{corrected} = d_{raw} \cdot k_{adjust}. \qquad (17)$$

To calibrate the distance measurement system, the SX1280 was used in its ranging mode to measure a known, long-distance baseline. A suitable location was identified where the master unit maintained a clear line of sight across the entire path. This known distance, precisely 4.405 km as verified by a 1:25,000 Ordnance Survey map, was compared against the average measured distance of 4.424 km reported by the device. The adjustment factor for the distance measurement system was determined to be $k_{adjust} = 0.99571$ [8].

Ranging accuracy measurements conducted between 10 m and 200 m demonstrated that the SX1280 can achieve accuracy comparable to laser range finders when averaging 2000 RTToF measurements across frequency-hopped channels. The accuracy improves with increased BW: measurements at 1600 kHz showed

significantly tighter clustering around ground truth compared to 400 kHz configurations.

A critical finding is the relationship between measurement quantity and accuracy. Analysis at SF 9 with 1600 kHz BW revealed that 80 frequency-hopped ranging exchanges are sufficient to achieve approximately 1 m absolute accuracy at 170 m distance. Additional exchanges beyond this point provide diminishing returns [5].

Comparative studies between the SX1280 and coherent multi-channel ranging implementations show that for stationary line-of-sight scenarios, the SX1280 achieves mean ranging error of 75 m with standard deviation of 69 m at distances up to 500 m [6].

For distances exceeding 1 km, field measurements in rural environments demonstrate mean distance errors of 46.4 m with standard deviation of 83.6 m over an 8.2 km test route. The mean error increases with distance, reaching a maximum relative error of 9.6% at 950 m in areas with significant shadowing [6]. At extreme distances of 2 km in line-of-sight conditions, the SX1280 maintained a ranging precision of +2/-5 m using SF10 at 1600 kHz [7].

With this comprehensive correction approach, ranging precision of approximately 1 m can be achieved in controlled line-of-sight conditions, while practical outdoor applications can achieve mean errors below 10 m even at distances exceeding 1 km.

References

1. Semtech. "SX1280 Datasheet." Available at: https://semtech.my.salesforce.com/sfc/p/#E0000000JelG/a/3n000000l9OZ/Kw7ZeYZuAZW3Q4A3R_IUjhYCQEJxkuLrUgl_GNNhuUo
2. Microchip. "ATA8352 Datasheet," Rev. A, Feb. 2021. Available: https://www.microchip.com/content/dam/mchp/documents/RFA/ProductDocuments/DataSheets/ATA8352_Datasheet_RevA_FEB2021_70005450A.pdf
3. Semtech. An Introduction to Ranging with the SX1280 Transceiver, App. Note [AN1200.29]. Available: https://semtech.my.salesforce.com/sfc/p/#E0000000JelG/a/44000000MDiH/OF02Lve2RzM6pUw9gNgSJXbDNaQJ_NtQ555rLzY3UvY
4. Semtech. Design of the SX1280 Ranging Protocol and Result Processing, App. Note [AN1200.50]. p. 16 Available at: https://semtech.my.salesforce.com/sfc/p/#E0000000JelG/a/2R000000UypY/5mprGH6TIzeLnfosUgj1xK5ftoqDpoCnRk_dzY_jAx4
5. Semtech. How to Perform Ranging Tests with the SX1280 Development Kit, App. Note [AN1200.31]. Available at: https://semtech.my.salesforce.com/sfc/p/#E0000000JelG/a/44000000MDcY/ZsmAVCVenZkc0lUrr3RuxWSfdFxY2Tjmsk4N9DAhBo
6. Outdoor Ranging and Positioning based on LoRa Modulation / P. Muller et al. *2021 International Conference on Localization and GNSS (ICL-GNSS)*, Tampere, Finland, 1–3 June 2021. 2021. URL: https://doi.org/10.1109/icl-gnss51451.2021.9452277 (date of access: 05.10.2025).
7. Robinson S. Semtech SX1280 2.4ghz LoRa Ranging Transceivers. *GitHub*. URL: https://github.com/StuartsProjects/SX1280_Testing
8. Robinson S. Semtech SX1280 2.4ghz LoRa Ranging Transceivers. *GitHub*. URL: https://github.com/StuartsProjects/SX12XX-LoRa/tree/master/examples/SX128x_examples/Ranging

Received (Надійшла) 11.08.2025
Accepted for publication (Прийнята до друку) 29.10.2025

Відомості про авторів / About the Authors

Руденко Віталій Віталійович – аспірант, Національний університет «Полтавська політехніка імені Юрія Кондратюка», Полтава, Україна;
Vitalii Rudenko – PhD student, National University "Yuri Kondratyuk Poltava Polytechnic", Poltava, Ukraine;
e-mail: vitaliirudenko40@gmail.com; ORCID Author ID: https://orcid.org/0009-0006-9094-222X.

Вимірювання відстані методом різниці прибуття (TDoA) на базі протоколу LoRa

В. В. Руденко

Анотація. Стаття присвячена реалізації вимірювання відстані між об'єктами з використанням методу різниці часу прибуття (Time Difference of Arrival, TDoA), зокрема через техніку часу двостороннього поширення сигналу (Round-Trip Time of Flight, RTToF), на основі протоколу LoRa та із застосуванням трансивера SX1280. **Мета статті** полягає в детальному описі часових характеристик обміну пакетами та формалізації розрахунку відстані. Ключовими параметрами, що суттєво впливають на точність вимірювання, визначено ширину смуги (Bandwidth, BW) та коефіцієнт розширення спектра (Spreading Factor, SF). Дослідження включає всебічний аналіз джерел похибок, таких як похибка опорного генератора (зсув синхронізації) та багатопроменеве поширення (Multipath Propagation), і надає формули для їх кількісної оцінки. Запропоновано багатоетапну методологію корекції, що включає корекцію частотної похибки (на основі FEI), компенсацію LNA (на основі RSSI), статистичну фільтрацію (за допомогою медіани), поліноміальну корекцію кривої та корекцію, специфічну для навколишнього середовища. **Результати** польових вимірювань показують, що за допомогою цього комплексного підходу до корекції можна досягти точності визначення відстані приблизно 1 метр в контрольованих умовах прямої видимості. Крім того, практичні зовнішні застосування демонструють середні похибки менше 10 метрів навіть на відстанях, що перевищують 1 км. **Висновки**: Реалізація та комплексна корекція вимірювань RTToF з використанням LoRa-модуля SX1280 дозволяють досягти високої точності, зберігаючи при цьому прийнятний баланс між точністю, часом перебування в ефірі (енергоспоживанням) та дальністю зв'язку, що є критично важливим для систем виявлення та позиціонування об'єктів.

Ключові слова: TDoA, LoRa, SX1280, RTToF, ToF, Ranging.

UDC 004.056.53:004.415:004.89:621.39 **doi: 10.26906/SUNZ.2025.4.207-213**

Ievgen Samborskyi[1], Heorhii Krykhovetskyi[2]

[1] State University "Kyiv Aviation Institute", Kyiv, Ukraine
[2] Defence Intelligence Research Institute, Kyiv, Ukraine

SYNTHESIS OF THE DIGITAL TWIN OF THE LOGICAL-DYNAMIC INFORMATION AND EVENTS MANAGEMENT SYSTEM FOR THE SECURITY OF COMPUTER SYSTEMS OF THE MOBILE CELLULAR INFORMATION AND COMMUNICATION NETWORK

Abstract. The article focuses on the aspect of information security and notes that currently the modern mobile information and communication cellular network is one of the most vulnerable and at the same time important objects of the critical information infrastructure of the state. It serves a wide range of users who make decisions for the organization of public administration, and also provides digital communication to a number of other important systems from the population to departmental structures. That is why this network acts as a priority object in the context of organizing effective management of its information security events. To organize the reliable functioning of this important object, a new approach to the synthesis of a digital twin of the information and security event management system of computer systems of the cellular mobile information and communication network is proposed. The proposed synthesis is based on a logical-dynamic approach to modeling security events in modern computer systems, attack scenarios and mechanisms for responding to these information security incidents by forming appropriate effective control influences. The architecture of the digital twin, the algorithm for its synthesis are considered, and possible approaches for implementing the integration of this virtual object with such platforms as Wazuh, Streamlit, Neo4j, AWS IoT are proposed. Verification and testing are carried out using the example of a DDoS scenario, and the results of the synthesis algorithm implementation are presented. The effectiveness of the model in detecting threats and adapting to intensive changes in the security environment of the computer system of the mobile digital network is shown.

Keywords: digital twin, management system, information security, security event, synthesis, logical-dynamic model, security event management, mobile network, integration, SIEM, Wazuh.

Statement of the problem

In the current conditions of information confrontation and hybrid aggression, digital cellular communication currently plays a key role in the functioning of critical information infrastructure facilities of the state, including energy, the financial system, logistics, state registers, and especially corporate secure communications. Ensuring the security of computer systems, which are the core of the mobile information and communication network system, has become extremely relevant due to the increase in the number and complexity of threats to the information security of these controlling computing facilities.

A vivid example of these threats is a large-scale cyberattack on one of the largest mobile operators in our country - "Kyivstar" in December 2023. This attack led to a long and serious disruption of digital communications, disruptions in the work of banking, logistics, energy and administrative services. This incident, in addition to organizational problems in the management of this digital cellular structure, revealed both the technical and conceptual inability of traditional information and security event management systems, which are currently operated in the information and communication network system, to effectively and reliably resist complex multi-level information security incidents in real time.

Organizing effective information and security event management in the conditions of dynamic, heterogeneous and distributed mobile cellular digital networks requires the urgent implementation of significantly new approaches. Methods that allow synthesizing adaptive, attack-resistant solutions with a high level of efficiency in information and security event management deserve special attention. In this context, an important The concept of digital twins plays a role, which allows you to synthesize virtual models of PCM objects and display them in real time. As a rule, the concept of digital twins synthesis is based on modern technologies, namely: Industry 4.0 technologies. At the same time, it is imperative to take into account that mobile information and communication network system is one of the most vulnerable and at the same time critically important super-complex objects of the critical information infrastructure. It serves a very wide range of users and systems – from the population to state departmental structures. That is why information and communication network system acts as a priority object of information security event management. Therefore, an urgent scientific and applied task of synthesizing digital twins of logical-dynamic information management systems and security events of the information and communication network system arose and is emerging, which requires the development of new models, methods and algorithms to increase the level of security, interference and functional stability of these digital means.

Analysis of recent research and publications

In recent years, there has been a sharp increase in scientific interest in the concept of creating and improving digital twin technologies as an effective tool for improving the security of information and communication networks. In the field of information and security event management of a computer system, digital twins allow creating virtual copies of critical

information infrastructure objects capable of monitoring, analyzing states, and modeling the development of security events in real time. Let us analyze the latest research presented in a number of fundamental publications. It should be noted that the scientific work [1] is decisive for the synthesis of logical-dynamic models of the digital twin of the information and security event management system of a computer system which confirms the relevance of this approach for complex information systems. In articles [2–3], the authors substantiate the feasibility of using a logical-dynamic approach in computer system security tasks with a high degree of criticality, demonstrating its adaptability, scalability, and compliance with the dynamic nature of events in cyberspace.

Particular attention is paid to the problem of detecting attacks such as APT, MITM, DDoS in mobile networks, which have a high dynamic topology and limited depth of response from classical SIEM/SOAR platforms. In this context, a digital twin with a built-in logical-dynamic model is able not only to reproduce the architecture of the information and communication network, but also to model the consequences of security events, predict critical states and launch control actions. A systematic approach to the synthesis of such models is disclosed in the works [4–7]. They propose algorithmic mechanisms for distributing response efforts based on models of radio-electronic influence and logical-event schemes for organizing control processes.

The conducted analysis of scientific sources indicates the relevance and existing powerful potential of the logical-dynamic approach for the tasks of algorithmization of the digital twin of the information and security event management system in distributed critical environments – the information and communication network. Information and security event management systems of the computer system of the information and communication network with a digital twin, in which this approach is implemented, have significant advantages over classical SIEM/SOAR systems due to the possibility of adaptive response to security events, forecasting the development of information security incidents, modeling the mutual correlation of security events in information and communication networks as well as the operational and effective formation of security event management in information and communication networks. This determines the scientific novelty and practical value of further research on the synthesis of a digital twin with a logical-dynamic core in the computer system of mobile cellular information and communication networks.

The purpose of this work is to substantiate and develop an effective approach to the synthesis of digital twins of the logical-dynamic information and event management system of the computer system of the mobile cellular information and communication network as a critical object of the national information infrastructure. It is advisable to focus particular attention on the construction of a synergized architecture of the digital twin, the creation of an algorithm for its synthesis based on the logical-dynamic model, as well as the integration of the digital twin with modern platforms for monitoring and processing security events, such as Wazuh, Streamlit, Neo4j, AWS IoT, etc. Along with this, the goal is also to demonstrate the effectiveness of the proposed model in conditions of DDoS scenarios and highly dynamic processes in the computer system of the information and communication network. This will allow to ensure a high level of security and functional stability of the computer system, to predict the development of possible threats (especially “zero-day”), to automatically form response strategies and adaptively manage security events and a number of other risks in critical conditions. Therefore, the main goal of the article is to propose new and effective approaches to the synthesis of digital twins of logical-dynamic information and event management systems that provide reliable protection of the computer system from modern existing threats, as well as from zero-day threats.

Presentation of the main material

To develop fundamentally new and effective methods for synthesizing a digital twin of logical-dynamic information management systems and security events that provide reliable protection of the computer system of the information and communication network from a wide range of existing threats to their security, and, especially, from zero-day threats we take into account that this network is the most vulnerable among all state objects of critical information infrastructure.

Among the key types of possible information threats that are generators of computer system security events we should especially note the following:

– traffic interception (man-in-the-middle);

– jamming (suppression) of digital communication radio channels;

– attacks related to malicious software (malware), including zero-day;

– internal threats to the computer system;

– APT attacks with a phased impact on the infrastructure of the information and communication network.

The specified security events are characterized by high dynamics, ambiguity and interdependence which makes it impossible to effectively process them by traditional means. In this regard, the task of formalizing the process of managing the security events of the information and communication network as a logical-dynamic system arises in order to integrate it into the structure of the information and security event management system of the digital twin. It should be noted that the mobile cellular information and communication network has a number of specific characteristics that significantly affect the requirements for the synthesis of the digital twin. Let us consider them in more detail:

– dynamic topology: network nodes (base stations, routers, switches, etc.) constantly and dynamically change their location and load depending on the structural geography of a significant number of users.

– heterogeneity of devices: different types of terminals, protocols, standards (4G, 5G, LTE).

– high traffic density: especially in highly urbanized areas which significantly complicates monitoring and rapid response to computer system security events.

– criticality of services: servicing emergency and special services, medical facilities, energy and logistics facilities.

The above features require the digital twin of the information and event management system for the security of the computer system of the information and communication network to have the ability to adaptively model, quickly respond to changes in parameters, and integrate with the information and event management systems for the security of the computer system available in the network.

To detail the wide range of requirements for the digital twin, we propose a formalized model:

$$DT = (F, T, A, O), \quad (1)$$

where F – functional requirements; T – technical requirements; A – analytical requirements; O – organizational requirements.

Each of these elements of the proposed synthesized model (1) is detailed in the form of subsets, namely:

$$F = (f_1, f_2, ..., f_n);\ T = (t_1, t_2, ..., t_n);$$
$$A = (a_1, a_2, ..., a_n);\ O = (o_1, o_2, ..., o_n).$$

The implementation of effective security event management processes for a computer system of a mobile cellular information and communication network requires that the digital twin of the security information and event management system meet a number of requirements covering functional, technical, analytical and organizational aspects. Its architecture must be flexible, scalable, secure and fully comply with international security standards. Let us define these requirements, which will become the basis for building an effective logical-dynamic security information and event management system using a digital twin. At the same time, this virtual object must provide the implementation of the necessary functions, namely:

– monitoring of current and previous security events;

– reconstruction of security event chains and their destructive effects on the processes of functioning of the computer system of the information and communication network;

– forecasting the state of the information and communication network in response to all possible incidents of information security of the computer system;

– generation of management actions for prompt and effective counteraction to threats to the computer system of the information and communication network.

At the same time, the requirements for the digital twin of the information and event management system for the security of a computer system of a mobile information and communication network include the following:

– the ability to integrate with telemetry, logs, NetFlow, syslog;

– display of the current state of the topology of the information and communication network and all its component segments;

– formalization of system behavior in the form of a logical-dynamic model;

– prediction of the consequences of security events and the prompt formation of management responses to incidents;

– compatibility with real security platforms for information and communication networks such as SIEM/SOAR.

The above indicates that the synthesis of digital twins of the information and event management system of a computer system of an information and communication network with a logical-dynamic core is a complex but critically necessary scientific task that includes architecture design, state formalization, transition modeling, and integration with information and management platforms of digital networks.

We synthesize the functional architecture of the digital twin of the information and security event management system of the computer system of the information and communication network. In the process of synthesis, we take into account that this algorithmic and software tool is a virtual analogue of the physical system, which provides real-time monitoring, analysis, forecasting and management of security events. The architecture of such a twin should be based on a logical-dynamic approach, which provides modeling of discrete states of the system and their transitions under the influence of external and internal security events.

The synthesized structure of this virtual analogue of the physical system - the information and security event management system of the computer system of the information and communication network consists of the following synergized modular subsystems, namely:

– information and communication network structure module – a virtual model of the topology of a mobile cellular information and communication network. It includes nodes, channels, base stations, switches, routers. It is implemented through Neo4j – a graph database and is constantly updated in real time based on information from the API *(Application Programming Interface)* and telemetry about the network status.

– data module (data aggregation level module) – aggregates data from information and communication network telemetry, log files, network security events, SNMP queries, NetFlow, syslog. Provides normalization, time synchronization and data processing. Works with Kafka, Fluentd or OpenTelemetry.

– logical-dynamic modeling module (*logical-dynamic core*) – implements a logical-dynamic model in the form of a tuple $LDM = (S, E, M, T)$, where S – a set of system states – an information and communication network; E – multiple security events; M – computer system status monitoring functions, T – rules for transitions between states of a computer system. This module allows you to formalize the behavior of this computing and control system and identify its critical states, taking into account (1).

– visualization and interpretation module – an interactive interface for displaying the current state of the system, forecasting results, and recommendations for prompt response through the formation and implementation of control actions. Implemented using Streamlit, Grafana, Kibana.

– SIEM/SOAR integration interface – exchange of events, notifications and control commands with systems such as Wazuh, Splunk, TheHive, OpenCTI. Provides two-way exchange with SOC analytics.

– response orchestrator – an event response automation subsystem that supports playbook scenarios implemented in the form of logical trees. These logical trees are activated by a digital twin depending on the threat vector model for the computer system of the information and communication network.

The interaction between the components is implemented via a data bus with support for the publish–subscribe model. This allows you to dynamically update the model, adapt it to changes in the states of the information and communication network and provide a cyclic mechanism in security event processing algorithms, namely: *"Security event → Analysis → Forecast → Management decision → Response → Feedback"*.

The synthesized architecture of the digital twin is a system for managing information and security events of a computer system of an information and communication network, built on the basis of a logical-dynamic approach, presented in Fig. 1.

The developed algorithm for synthesizing a digital twin consists of a number of synergized sequential stages. Each of them involves the gradual formalization, modeling and integration of a logical-dynamic approach in the synthesized architecture (Fig. 1) of the digital model. The main goal of the proposed synthesis is to build a digital twin capable of independent analysis of security events, predicting their consequences, and, especially, optimizing decision-making support in the security environment of a mobile cellular information and communication network. The synthesis algorithm includes the following step-by-step stages:

Stage 1. Identification of the modeling object:

– identification of the components of the information and communication network as an object of critical information infrastructure;

– identification of key control points of security events, information flows and typical information security incidents.

Stage 2. Formalization of events and states:

– construction of a set of discrete states (for example: "normal", "threat detected", "response activated");

– definition of a set of events (incidents, anomalies, intrusions, etc.);

– construction of monitoring functions that determine under what conditions a certain security event is recorded;

– formalization of rules for transitions between states under the influence of security events (in the form of logical rules or graphs).

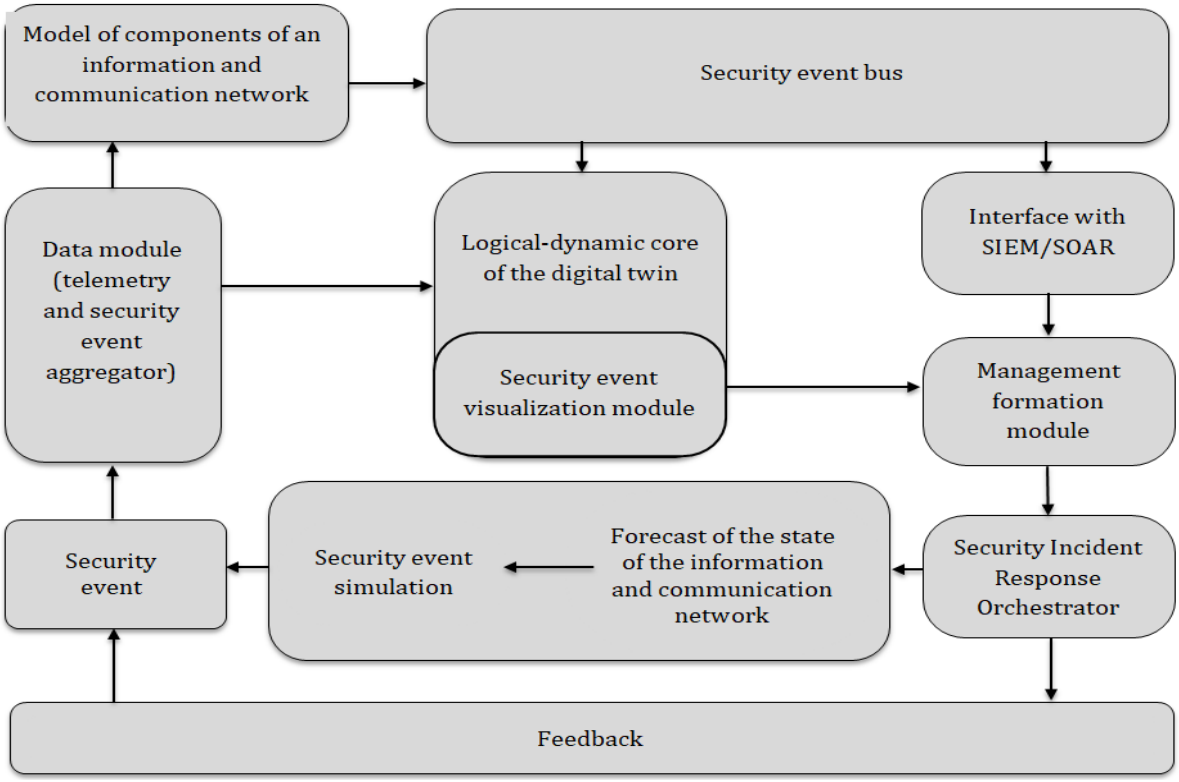


Fig. 1. Architecture of the digital twin of the information and event management system of the security control computer system of the information and communication network, built on a logical-dynamic approach

Stage 3. Construction of a logical-dynamic model:

– combining sets into a tuple $LD = (S, E, M, T)$;

– checking the model for conflicts, completeness, and the absence of cycles without a solution;

– modeling typical scenarios of information security incidents and checking the model's response to these events.

Stage 4. Integration with the digital infrastructure of the information and communication network:

– implementation of communication with telemetry sources (log files, network events, SNMP);

– construction of adapters for visualization, response orchestration, interaction with SOC.

Stage 5. Automation and training of the digital twin:

– implementation of mechanisms for adapting the model to new security events, and especially zero-day;

– application of machine learning methods to detect new event scenarios;

– training of transition rules based on statistical data on security events (formation of a reflexive model).

Analysis of the synthesized digital twin algorithm shows that its implementation in real systems for managing information and security events of a computer system of an information and communication network will allow us to gradually move from a description of an object to a functional model with predictive, reactive and adaptive capabilities for organizing security event management, which can be flexibly integrated into modern complex protection systems for these critically important digital networks.

Let us verify and validate the synthesized digital twin. It should be noted that after building this virtual digital tool for the information and security event management system of a computer system of an information and communication network, its verification and validation are critically important stages. These processes allow us to assess the correct functioning of the model, its compliance with the expected characteristics as well as the ability to detect and respond to all possible security events in a mobile network. Let us take into account that verification involves an internal comprehensive check of the logic of the model and its components, namely:

– checking the correctness of the definition of sets of states *S*, events *E* and transition rules *T*;

– the absence of logical conflicts, ambiguous transitions or looping of models;

– the conformity of the model to the given architecture and structure of the digital twin.

Validation is aimed at checking the model in real or close to real conditions of the functioning of a real information and communication network.

It involves:

– testing the behavior of the digital twin based on attack scenarios (for example, DoS, APT, MITM);

– comparing the model's response with reference response scenarios in a real SIEM – information and communication network system;

– determining the accuracy, completeness and timeliness of incident detection.

The validation scenario algorithm involves the following sequence of actions:

1. An "abnormal communication channel overload" event is generated.
2. The model records a change in metrics through a telemetry aggregator.
3. The logical-dynamic model enters the "threat detected" state.
4. The orchestrator's response scenario is activated to block traffic.
5. The result is saved and displayed in the visualization interface.

The verification performed ensures the structural consistency of the digital twin, and the validation results indicate its effectiveness in the context of compensating for the destructive consequences for the computer system of the information and communication network in the event of real threats to its security. Synergization of verification and validation allows to increase the confidence in the model and ensure its practical suitability for the implementation of integration processes in the real environment - modern mobile cellular information and communication networks (Table 1).

Table 1 – **Integration of the digital twin with the real information and communication network environment**

Component / platform	Task / role	Integration features	Example of implementation
Wazuh (SIEM)	Monitoring security events and logs of information and communication network nodes	Using agents, triggers, REST API; connecting to Wazuh manager	Detection of DoS, MITM, incidents in base stations of the information and communication network
Streamlit	Interactive visualization of model states	Web interface; display of states, graphs, what-if modeling	Real-time model state transition graph
AWS IoT Core	Telemetry collection, connecting devices to the information and communication network	MQTT, Lambda functions, Amazon Timestream for analyzing and storing security events	Channel congestion analysis, time series storage
Message brokers (Kafka, MQTT)	Real-time transmission of events to a digital twin	Providing publish–subscribe logic, buffering security events	Delivery of telemetric information from information and communication network nodes to the LDM module
Response mechanisms (SOAR/Playbook)	Initiating action in response to a real threat	Support for two-way exchange and implementation of security incident response scenarios	Automatic blocking of traffic in the event of a DoS attack

To confirm the operability of the synthesized digital twin of the information and event management system of the computer system of the information and communication network, an experimental model was implemented in a virtual environment. A conditional DDoS attack on a computer system, which is the control node of the mobile information and communication network was chosen as a test scenario.

The following implementation tool environments were selected [8–10]:

Streamlit – for building a digital twin visualization interface;

Python – implementation of the logical-dynamic kernel and analysis algorithms;

Neo4j – graph database for modeling the structure of the PCM and states;

MQTT – for modeling telemetry streams;

Wazuh – as a source of real logs and security events. During the experiment, the digital twin recorded a suspicious load coming from a certain network segment.

An event “traffic excess anomaly” was generated, which initiated the transition to the “threat detected” state.

The model activated a response scenario: temporary traffic isolation, redirection of logs to SIEM. During testing, the following results of the digital twin’s operation were recorded (Table 2).

Table 2 – **Digital twin testing results during a *DDoS*-scenario**

Parameters	Parameter values	Efficiency assessment	Comments of the experiment
Average threat detection time	up to 3 sec.	High speed	Within target threshold (<5 sec.)
Number of false positives	1 of 20 events	False positives – to 5%	Acceptable level for SIEM class
CPU load	45–60 %	Average load	Peak – when processing complex scenarios
Updating rules during an attack	3 new rules (with automatic update)	Adaptability confirmed	Saved to a graph database – an array of security events

Conclusions

As a result of the research, one of the possible approaches to building digital twins of mobile information and communication networks was implemented, the basis of which is the logical-dynamic modeling of security event management processes.

The synthesized digital event architecture provides structured synergization of information and communication network nodes, security events, computer system states and response scenarios to information security incidents.

The proposed model synthesis algorithm allows for the sequential implementation of key stages of building a digital representation of the system, including the formalization of security events, the construction of logical transitions and integration into real environments of the information and communication network.

Particular attention is paid to practical aspects of integration with security tools, namely: *Wazuh*, *AWS IoT*, *Streamlit*, *Neo4j*.

Testing in a virtual environment of the information and communication network showed the ability of the synthesized digital twin to:

– rapid identification of threats (less than 3 seconds);

– minimizing false positives (≈5%);

– dynamic updating of reaction logic;

– flexible visualization of the current state and “what-if” scenarios.

The results obtained indicate the effectiveness of the logical-dynamic approach to managing security events in a computer system of critical information infrastructure and the feasibility of using a digital twin to increase the level of situational awareness, adaptability and reactivity in the management of critical information and communication infrastructures.

REFERENCES

1. Samborskyi E. I., Peleshok E. V. Synthesis of Logical-Dynamic Information Management Systems and Security Events of Computer Structures. Control, Navigation and Communication Systems. – 2025. – № 2 (72). – P. 185–194. DOI: https://doi.org/10.26906/SUNZ.2025.2.185-194
2. Pavlenko P. M., Samborskyi Ye. I. Upravlinnia informatsiieiu i podiiamy bezpeky kompiuternykh system iz vykorystanniam lohiko-dynamichnykh modelei. Information Technology and Security. 2025. T. 13, № 1 (24). 43–54. DOI: https://doi.org/10.20535/2411-1031.2025.13.1.328764 [in Ukrainian].
3. Sholokhov S. M., Pavlenko P. M., Nikolaienko B. A., Samborsky I. I., Samborsky E. I. 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. DOI: https://doi.org/10.15588/1607-3274-2023-4-2
4. Cherdantseva Y., Burnap P., Blyth A. et al. A review of cyber security risk assessment methods for SCADA systems. Computers & Security. 2016. Vol. 56. P. 1–27. DOI: https://doi.org/10.1016/j.cose.2015.09.009
5. Radanliev P., De Roure D., Nurse J. et al. Digital twins: Concepts and use cases in cyber security risk assessment. Journal of Cyber Security Technology. 2022. Vol. 6(3). P. 147–174. DOI: https://doi.org/10.1080/23742917.2021.1982822

6. Vasyliev V. V., Kovalenko O. S. Intelektualni systemy vyiavlennia zahroz dlia kiberzakhystu krytychnoi infrastruktury. Kiberbezpeka: osvita, nauka, tekhnika. 2023. № 3. S. 42–49. DOI: https://doi.org/10.28925/2663-4023.2023.3.4249 [in Ukrainian].
7. Gamil A. et al. A framework for real-time threat detection and mitigation using digital twins in IoT networks. IEEE Internet of Things Journal. 2021. Vol. 8(12). P. 9740–9752. DOI: https://doi.org/10.1109/JIOT.2020.3046026
8. Wazuh. The Open-Source Security Platform. Documentation. URL: https://documentation.wazuh.com
9. AWS IoT Developer Guide. URL: https://docs.aws.amazon.com/iot
10. Neo4j Graph Data Platform. URL: https://neo4j.com.

Received (Надійшла) 13.08.2025
Accepted for publication (Прийнята до друку) 12.11.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Самборський Євген Іванович – аспірант кафедри організації авіаційних перевезень Державного університету "Київський авіаційний інститут", Київ, Україна;
Ievgen Samborskyi – Postgraduate student, Department of Air Transportation Organization, State University "Kyiv Aviation Institute", Kyiv, Ukraine;
e-mail: seinauedu@gmail.com; ORCID Author ID: https://orcid.org/0000-0003-4441-1947.

Криховецький Георгій Яремович – кандидат технічних наук, старший науковий співробітник, Науково-дослідний інститут воєнної розвідки, Київ, Україна;
Heorhii Krykhovetskyi – Candidate of Technical Sciences (PhD), Senior Researcher, Defence Intelligence Research Institute, Kyiv, Ukraine;
e-mail: kgeorg@ukr.net; ORCID: https://orcid.org/0009-0001-2981-7810;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=58697828100&origin=resultslist.

Синтез цифрового двійника логіко-динамічної системи управління інформацією та подіями безпеки комп'ютерних систем мобільної стільникової інформаційно–комунікаційної мережі

Є. І. Самборський, Г. Я. Криховецький

Анотація. У статті акцентовано особливу увагу на аспекті інформаційної безпеки та відмічено, що наразі сучасна мобільна інформаційно-комунікаційна стільникова мережа є одним із найбільш уразливих та водночас важливих об'єктів критичної інформаційної інфраструктури держави. Вона обслуговує широке коло користувачів, які приймають рішення для організації державного управління, а також забезпечує цифровим зв'язком низку інших важливих систем від населення до відомчих структур. Саме тому ця мережа і виступає як пріоритетний об'єкт у контексті організації ефективного управління подіями її інформаційної безпеки. Для організації надійного функціонування цього важливого об'єкта запропоновано новий підхід до синтезу цифрового двійника системи управління інформацією і подіями безпеки комп'ютерних систем стільникової мобільної інформаційно-комунікаційної мережі. В основу запропонованого синтезу покладено логіко-динамічний підхід до моделювання подій безпеки в сучасних комп'ютерних системах, сценаріїв атак та механізмів реагування на ці інциденти інформаційної безпеки за рахунок формування відповідних ефективних управляючих впливів. Розглянуто архітектуру цифрового двійника, алгоритм її синтезу, а також запропоновані можливі підходи для реалізації інтеграції цього віртуального об'єкта з такими платформами як Wazuh, Streamlit, Neo4j, AWS IoT. Проведено верифікацію та тестування на прикладі DDoS-сценарію, наведено результати реалізації алгоритму синтезу. Показано ефективність моделі у виявленні загроз та адаптації до інтенсивних змін безпекового середовища комп'ютерної системи мобільної цифрової мережі.

Ключові слова: цифровий двійник, система управління, інформаційна безпека, подія безпеки, синтез, логіко-динамічна модель, управління подіями безпеки, мобільна мережа, інтеграція, SIEM, Wazuh.

АЛФАВІТНИЙ ПОКАЖЧИК

Наукове видання

СИСТЕМИ УПРАВЛІННЯ, НАВІГАЦІЇ ТА ЗВ'ЯЗКУ

Збірник наукових праць

Випуск 4 (82)

Відповідальний за випуск *О. В. Шефер*

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

Підписано до друку 02.12.2025. Формат 60×84/8. Ум.-друк. арк. 26,75. Тираж 120 прим. Зам. 1202-25
Адреса редакції: Україна, 36011, м. Полтава, проспект Віталія Грицаєнка, 24, тел. (050) 302-20-71
Національний університет «Полтавська політехніка імені Юрія Кондратюка»

Віддруковано з готових оригінал-макетів у цифровій друкарні Impress
61002, м. Харків, вул. Пушкінська, 56, тел. **+ 38 (057) 714-52-11**
e-mail: **irina@impress.biz.ua**