Харківський національний університет радіоелектроніки
Державне підприємство
"Південний державний
проектно-конструкторський та науково-дослідний
інститут авіаційної промисловості"

Kharkiv National University of Radio Electronics
State Enterprise
"Southern National Design & Research Institute of Aerospace Industries"

Сучасний стан наукових досліджень та технологій в промисловості

Innovative technologies and scientific solutions for industries

Щоквартальний науковий журнал

№ 4 (34), 2025

Затверджений до друку
Науково-технічною Радою
Харківського національного університету радіоелектроніки
(Протокол № 11 від 10 грудня 2025 р.)

ЗАСНОВНИКИ
Харківський національний університет радіоелектроніки,
Державне підприємство "Південний державний проектно-конструкторський та науково-дослідний інститут авіаційної промисловості"

АДРЕСА РЕДАКЦІЇ:
Україна, 61166, м. Харків, проспект Науки, 14
Інформаційний сайт: http://itssi-journal.com
https://journals.uran.ua/itssi
E-mail редколегії: journal.itssi@gmail.com

Quarterly scientific journal

No. 4 (34), 2025

Approved for publication
by the Scientific and Technical Council
of the Kharkiv National University of Radio Electronics
(Protocol No. 11 d.d. 10 December 2025)

ESTABLISHERS
Kharkiv National University of Radio Electronics,
State Enterprise
"National Design & Research Institute
of Aerospace Industries"

EDITORIAL OFFICE ADDRESS:
14 Nauky Ave., Kharkiv, Ukraine, 61166
Information site: http://itssi-journal.com
https://journals.uran.ua/itssi
E-mail of the editorial board: journal.itssi@gmail.com

Ідентифікатор медіа R30-03878
Витяг з реєстру суб'єктів у сфері медіа – реєстрантів від 25.04.2024 № 1410

Журнал включено до Переліку наукових фахових видань України, в яких можуть публікуватися результати дисертаційних робіт на здобуття наукових ступенів доктора і кандидата наук наказом Міністерства освіти і науки України від 16.07.2018 №775 (додаток 7)

Харків – 2025

СУЧАСНИЙ СТАН НАУКОВИХ ДОСЛІДЖЕНЬ ТА ТЕХНОЛОГІЙ В ПРОМИСЛОВОСТІ

№ 4 (34), 2025

ЗМІСТ

Інформаційні технології

Інженерія та промислові технології

Електроніка, телекомунікаційні системи та комп'ютерні мережі

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

INNOVATIVE TECHNOLOGIES AND SCIENTIFIC SOLUTIONS FOR INDUSTRIES

No. 4 (34), 2025

CONTENTS

Information Technology

Engineering & Industry Technologies

Electronics, Telecommunication Systems & Computer Network

The author is responsible for the accuracy of the facts, quotations and other information

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

УДК 004.94
DOI: https://doi.org/10.30837/2522-9818.2025.4.005

М. Гринченко, Д. Куценко

МОДЕЛІ БІЗНЕС-ПРОЦЕСІВ МЕДИЧНОГО ЦЕНТРУ ДЛЯ ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ ПРИЙНЯТТЯ РІШЕНЬ

Предметом дослідження є процес цифровізації результатів лабораторних досліджень і формалізація бізнес-процесів медичного центру для створення інформаційної системи, орієнтованої на інтелектуальний аналіз даних. **Мета статті** – розроблення моделей, які формалізують бізнес-процеси медичного центру у взаємодії з пацієнтом, що дасть змогу визначити напрями діяльності, які потребують автоматизації для підвищення ефективності надання медичних послуг. У роботі необхідно виконати такі **завдання:** проаналізувати сучасний стан цифровізації медичної інформації та проблеми уніфікації результатів лабораторних досліджень; визначити особливості організації бізнес-процесів медичного центру та взаємодії його ключових учасників; сформувати концептуальну й функціональну модель надання послуг; побудувати модель бізнес-процесів, що відтворює структуровану організацію діяльності закладу у взаємодії з пацієнтом. **Методи дослідження**: SWOT- і PEST-аналіз медичного центру; аналіз внутрішньої та зовнішньої документації, наявної бази даних і алгоритму надання медичних послуг; методологія функціонального моделювання IDEF0; методологія моделювання бізнес-процесів організації BPMN. **Досягнуті результати:** проаналізовано сучасні дослідження та наявні рішення; подано результати SWOT- і PEST-аналізу медичного центру; створено концептуальну модель організації надання медичних послуг; побудовано функціональну модель з огляду на нормативно-правові, організаційні та клінічні аспекти; розроблено модель ключових бізнес-процесів на основі BPMN; визначено бізнес-процеси, що підлягають автоматизації за допомогою інтелектуальних методів аналізу даних, які формують контур майбутньої інформаційної системи. **Висновки**. Запропоновані моделі створюють науково-методологічне підґрунтя для автоматизації збирання та оброблення результатів лабораторних досліджень, їх уніфікації та інтеграції в єдиний інформаційний контур медичного центру. Цеьсприятиме підвищенню ефективності бізнес-процесів медичного центру у взаємодії з пацієнтом, а також скоротить час оброблення результатів медичної діагностики, знизить ризик помилок і забезпечить якісну підтримку надання медичних послуг.

Ключові слова: функціональна модель; бізнес-процес; автоматизація бізнес-процесів; інтелектуальний аналіз даних.

Вступ і постановка проблеми

Сучасна трансформація системи охорони здоров’я нерозривно пов’язана з упровадженням інформаційних технологій, що забезпечують підвищення якості, доступності та безпеки медичних послуг. Одним із ключових напрямів цифровізації є автоматизація процесів збирання, оброблення та аналізу результатів лабораторних досліджень, що становлять основу для визначення діагнозу, вибору лікувальної тактики й здійснення профілактичних заходів. Швидке зростання обсягів медичних даних, зокрема внаслідок використання зовнішніх лабораторій, і різнорідність форматів подання результатів актуалізує проблему їх уніфікації та інтеграції в єдиний інформаційний простір.

Важливою умовою успішного функціювання медичних центрів у цифрову епоху стає не лише технічна спроможність оброблення даних, а й організаційна готовність до перебудови бізнес-процесів і впровадження нових моделей взаємодії з пацієнтом. Інтеграція інтелектуальних методів аналізу даних у клінічні інформаційні системи відкриває можливості для автоматизованого виявлення критичних відхилень, ретроспективного аналізу динаміки показників і формування персоналізованих рекомендацій для лікарів.

Однак існує низка нерозв’язаних питань, пов’язаних із семантичною неоднорідністю інформації, відсутністю єдиних стандартів її подання та недостатньою інтеграцією в робочі процеси медичних центрів. У цих умовах виникає потреба в розробленні інформаційних систем, що основані на методах інтелектуального аналізу даних і здатні забезпечити автоматизоване оброблення результатів лабораторних досліджень. Водночас для створення таких систем необхідно попередньо вивчити й формально описати логіку функціювання медичного закладу, що стане підґрунтям для подальшої автоматизації, уніфікації та інтеграції даних у єдиний інформаційний контур охорони здоров’я.

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

Аналіз проблеми й наявних методів

Сучасні студії підтверджують, що цифровізація результатів лабораторних досліджень і впровадження електронних інформаційних систем стали ключовими чинниками підвищення ефективності медичної допомоги. Наукові роботи й практичні кейси свідчать, що використання електронних медичних записів не лише зменшує кількість помилок у клінічних процесах, але й створює умови для автоматизованого контролю показників, формування управлінської звітності та вдосконалення комунікації між різними рівнями системи охорони здоров'я [1]. Прикладом масштабного впровадження є система *eHealth* в Україні, яка вже охоплює понад 35 млн пацієнтів і надає доступ до електронних рецептів, направлень і лікарняних. Ця IT-система стала найбільшою у сфері охорони здоров'я в країні та підтвердила, що єдиний електронний простір підсилює соціальний захист і контроль якості медичної допомоги завдяки автоматизованій звітності та моніторингу [2]. Водночас міжнародні порівняльні дослідження юзабіліті показують, що навіть за наявності доступу пацієнтів до електронних записів користувацький досвід часто залишається "вузьким місцем", і для систем потрібні як кількісні, так і якісні метрики бенчмаркінгу зручності застосування [3, 4]. Політико-регуляторні порівняння між США та Великою Британією підтверджують: попри зростання прийняття EHR, ключем до безпечного масштабування залишаються стандарти інтероперабельності та належні правила безпеки / приватності інформації [5], а в Німеччині, де цифровізація підтримується на рівні телематичної інфраструктури (ePA, eRezept, KIM, TIM), збережено акцент на сумісності даних як обов'язковій передумові цифрової медицини [6]. Незважаючи на політичний прогрес, опитування лікарів університетських клінік у Німеччині демонструють низьку задоволеність юзабіліті EHR і запит на швидкодію, зменшення помилок і краще узгодження з реальними завданнями користувачів [4]. У Великій Британії великомасштабне дослідження в закладах невідкладної допомоги засвідчило, що жодна із широко використовуваних систем не досягла міжнародного порогу прийнятної зручності (SUS=68), що доводить критичну необхідність адаптувати цифрові інструменти до високонавантажених клінічних середовищ [7].

У педіатричній практиці Європи виявлено суттєву неоднорідність: рівень застосування EHR коливається від дуже низького до майже повного охоплення залежно від країни, але бракує базових педіатричних функцій – автоматичного відстеження росту та інтеграції з реєстрами імунізації, що вказує на потребу уніфікованих міжнародних стандартів для дитячої популяції [8]. У країнах, що розвиваються, локальні програми цифровізації підтверджують потенціал підвищення продуктивності, якості даних і координації догляду за умови належної організації процесів (зокрема навчання персоналу) і чіткого регламенту звітності. Водночас саме змішування ручних і електронних процедур і брак устаткування /каналів зв'язку часто нівелюють очікуваний ефект [9–11]. Показово, що у великих муніципальних мережах (на прикладі Мумбаї) "вузькими місцями" під час розгортання HMIS стають як технічні чинники (брак обладнання, низька пропускна здатність мережі), так і людські (спротив змінам, недостатній пост-тренінг), тож управління змінами потрібно планувати як невід'ємну частину технічного проєкту [11]. Загальні огляди HIS підтверджують, що ці системи здатні підвищувати ефективність, результативність і безпеку, але для стійкого ефекту необхідні стандартизовані процеси, релевантні модулі та поступова гармонізація звітності й контролю якості [12–14].

Дослідження критичних факторів успіху інформаційних систем ранжує надійність, зручність використання та "організаційну придатність" як найбільш впливові детермінанти результату, що безпосередньо пов'язує технологічні рішення з контекстом клінічної практики й структурою управління [15]. Технологічно обґрунтованим способом масштабування стають хмарні моделі: приватні хмари краще відповідають вимогам конфіденційності, однак потребують інтелектуального керування ресурсами для стабільного забезпечення якості сервісу під непередбачуваним навантаженням. Експериментальні підходи до адаптивного розподілу ресурсів у приватній хмарі демонструють ефективність у контексті HIS [16]. Паралельно розвиваються підходи до аналітики, що мінімізують переміщення медичної інформації: федеративне навчання дає змогу будувати прогностичні моделі поверх розподілених EHR-даних без передачі "сирих" записів, тобто знижуючи ризики приватності та регуляторні бар'єри під час збереження якості

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

аналітики [17, 18]. У підсумку навіть за різного рівня інфраструктури та управлінської зрілості узагальнений урок однаковий: цифровізація лабораторних і клінічних показників працює тоді, коли її впроваджують як реформу системи із стандартизованою семантикою та протоколами обміну [19], надійною інфраструктурою, орієнтацією на користувача й системним навчанням персоналу. У цьому різі вона приносить реальні клінічні, організаційні та суспільні вигоди [9, 12, 13].

У сучасних дослідженнях проблема вилучення клінічно значущих даних із неструктурованих джерел розв'язується за допомогою багатокомпонентних моделей, що інтегрують методи оптичного розпізнавання символів (OCR), оброблення природної мови (NLP) та алгоритми машинного навчання. Використання OCR на основі згорткових і рекурентних архітектур (CNN-LSTM, CNN-BiLSTM-CTC) забезпечує розпізнавання як друкованих, так і рукописних текстів у сканованих документах, а поєднання із методами аналізу макета сторінки дає змогу точно відновлювати структуру таблиць і форм, що є критичним для подальшої інтерпретації числових результатів лабораторних досліджень [20].

На наступному етапі ключову роль відіграють моделі NLP, орієнтовані на розпізнавання сутностей (*Named Entity Recognition*, NER) та вилучення відношень. Традиційні словникові й правило орієнтовані системи поступово витісняють глибокі контекстуальні моделі, зокрема трансформерні архітектури на кшталт BERT та його доменних адаптацій (BioBERT, ClinicalBERT, multilingual BERT). Вони демонструють високу ефективність у вилученні медичних понять, медикаментів, дозувань і результатів тестів, а також у їх нормалізації відповідно до міжнародних стандартів (SNOMED CT, LOINC, RxNorm) і структур медичної інформації (FHIR R4, OMOP CDM) [21, 25–27].

Окремий напрям пов'язаний з використанням великих мовних моделей (LLM), що допомагають реалізувати процеси екстракції даних у вигляді конвеєрів "кінець у кінець". Такі моделі інтегрують етапи попереднього оброблення (OCR, сегментація тексту, вбудовування), розпізнавання іменованих сутностей, класифікацію та семантичне структурування, забезпечуючи водночас значне підвищення повноти клінічно значущої інформації порівняно зі стандартними структурованими полями EHR. Прикладом є LLM-схема вилучення медикаментів, яка збільшила кількість виявлених онкологічних препаратів і пов'язаних з ними клінічних атрибутів на десятки відсотків, що безпосередньо підвищує якість формування планів лікування [21].

Для завдань, пов'язаних із великомасштабним архівним пошуком і ретроспективним аналізом, застосовуються системи *Retrieval-Augmented Generation* (RAG). Вони комбінують векторні вбудовування для семантичного пошуку релевантних фрагментів і генеративні моделі з метою їх подальшого структурування відповідно до задан их схем. Це дає змогу формувати узгоджені "карти пацієнтів", у яких результати численних аналізів інтегруються в єдину часову динаміку, що уможливлює автоматизовану стратифікацію ризику та оцінювання відповідності клінічним протоколам [28]. У разі, коли необхідне швидке й верифіковане формування списку пацієнтів для дослідження або оцінювання результатів терапії, ефективними виявилися класичні алгоритми машинного навчання, зокрема метод опорних векторів (SVM) [24]. На окрему увагу заслуговують мультимодальні підходи до аналізу "візуально багатих документів", які поєднують текстові, візуальні та просторові ознаки. Глибокі нейронні мережі з компонентами аналізу макета сторінки та злиттям ознак сприяють значному підвищенню точності вилучення числових показників, одиниць вимірювання та референсних значень, що є визначальним для коректного відтворення лабораторної динаміки й створення систем автоматизованого виявлення критичних тенденцій [29].

Отже, у сучасній практиці оброблення медичних аналізів формується ієрархія моделей: від базових OCR/HTR та макетних аналізаторів через контекстуальні NLP і трансформерні архітектури для вилучення сутностей та відношень до комплексних LLM- і RAG-підходів, здатних інтегрувати різнорідні показники в структуровані формати. Разом ці методи створюють підґрунтя для автоматизованого формування планів лікування, ретроспективного аналізу історій хвороби та побудови клінічних рекомендацій, що робить їх невід'ємним елементом сучасних систем підтримки медичних рішень [20–23].

Аналіз сучасних студій засвідчує, що значний прогрес у сфері оброблення результатів лабораторних досліджень пов'язаний із застосуванням методів оптичного розпізнавання символів, клінічного NLP та алгоритмів машинного навчання, зокрема великих мовних моделей, мультимодальних

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

нейронних мереж для роботи з візуально насиченими документами, а також федеративного навчання з метою збереження приватності [24–29]. Ці методи підтвердили свою ефективність у вилученні медичних показників із різнорідних текстових і графічних джерел, їх нормалізації до міжнародних стандартів (FHIR, OMOP, SNOMED CT, LOINC) та інтеграції в клінічні системи підтримки рішень, що створює основу для розвитку сучасної цифрової медицини.

Однак, попри наявні досягнення, у практичному застосуванні залишаються нерозв'язані деякі питання. Результати аналізів, отримані з різних внутрішніх і зовнішніх лабораторій, суттєво різняться за структурою, найменуваннями показників і системами одиниць, що створює бар'єри для автоматизованої семантичної уніфікації та інтеграції в єдине клінічне середовище. Сучасні моделі здебільшого зосереджуються на завданні екстракції даних, тоді як їх інтерпретація та використання для формування рекомендацій, побудови планів лікування чи оцінювання ретроспективної динаміки показників залишаються фрагментарно виконаними. Важливою прогалиною також є відсутність гнучких рішень для інтеграції даних із різних форматів – від CSV та HL7 до сканованих PDF-документів, що ускладнює повномасштабне впровадження автоматизації. Крім того, сучасні підходи обмежено підтримують інтерактивну візуалізацію часових рядів і прогнозних трендів, а концептуальні моделі інтеграції зазвичай розробляються для великих національних чи корпоративних систем, тоді як для окремих медичних центрів бракує цілісних архітектур, здатних забезпечити всі етапи – від збору до практичної інтерпретації показників.

Необхідність створення інформаційних технологій, здатних подолати семантичну неоднорідність лабораторних показників, забезпечити їх інтеграцію та інтелектуальний аналіз, є одною з умов зниження ризику помилок, скорочення часу оброблення результатів і підвищення ефективності клінічних рішень. Саме цим зумовлена актуальність цього дослідження, що закладає підґрунтя для створення інформаційної системи з метою автоматизації збору та уніфікації лабораторної інформації в медичних центрах.

Метою статті є розроблення моделей, які формалізують бізнес-процеси медичного центру у взаємодії з пацієнтом, що дасть змогу визначити напрями діяльності, які потребують автоматизації для підвищення ефективності надання медичних послуг. Запропоновані моделі стануть основою для створення інформаційної системи, що здатна забезпечити високу точність екстракції даних, відстеження динаміки лабораторних показників і формування рекомендацій лікарю. Це зі свого боку створить передумови для підвищення якості медичного обслуговування, скорочення часу оброблення результатів і підтримки прийняття обґрунтованих рішень.

Для досягнення поставленої мети необхідно виконати такі завдання:

– проаналізувати сучасний стан цифровізації медичних даних і визначити проблеми, пов'язані з уніфікацією результатів лабораторних досліджень і їх інтеграцією в інформаційні системи медичного центру;

– дослідити особливості організації бізнес-процесів у медичному центрі та взаємодії між основними учасниками (пацієнт, адміністративний персонал, лікарі, медичні працівники, зовнішні діагностичні центри);

– розробити концептуальну й функціональну моделі надання послуг у медичному центрі, які ідентифікують системні зв'язки між основними процесами діяльності медичного центру;

– створити модель бізнес-процесів організації діяльності медичного центру у взаємодії з пацієнтом із впровадженням BPMN-методології.

Використання концептуального моделювання для побудови інформаційної технології оброблення результатів лабораторних досліджень зумовлено необхідністю формалізованого підходу до автоматизації процесів збирання, уніфікації та аналізу медичних даних. У межах цієї роботи обґрунтовано вибір концептуального рівня розроблення, що дає змогу чітко визначити основні сутності ключових процесів діяльності медичного центру, потоки інформації та логіку взаємодії між підсистемами майбутньої інформаційної системи. Отже, моделі, що розробляються в цьому дослідженні, які зі свого боку основані на принципах формалізованого моделювання інформаційних систем, спрямовані на виконання завдань збирання та оброблення лабораторних результатів у діяльності медичного закладу на основі методів інтелектуального аналізу даних.

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

Розроблення моделей формалізації бізнес-процесів медичного центру

У сучасних умовах функціювання медичних установ актуальним завданням є формалізація та стандартизація процесів обслуговування пацієнтів з огляду на вимоги чинного законодавства, медичних протоколів і внутрішніх регламентів організації. Для цього доцільним є побудова концептуальних моделей, що забезпечують системні погляди про сукупність вхідних ресурсів, процедур, механізмів регулювання та вихідних результатів медичної діяльності.

У дослідженні об'єктом обрано медичний центр "Еввіва". Для виявлення особливостей його організаційної структури й бізнес-процесів організовано консультації та обговорення з керівництвом закладу, а також проаналізовано зовнішнє й внутрішнє середовище закладу, його документацію та алгоритм надання медичних послуг.

Для розроблення концептуальної моделі здійснено комплексне дослідження медичного центру, що передбачало PEST-, SWOT-аналіз та інтерв'ювання управлінського персоналу.

Результати PEST-аналізу подано в табл. 1. Фактори зовнішнього середовища створюють як можливості, так і виклики для "Еввіва". Ринок медицини в Україні швидко трансформується під впливом цифровізації, регуляторних змін і соціальних очікувань. Успішна адаптація до цих чинників вимагатиме від медичного центру гнучкості, впровадження інновацій та стратегічного планування, зокрема у сфері інтеграції новітніх технологій та зміцнення довіри пацієнтів.

Таблиця 1. *PEST-аналіз медичного центру*

Фактори	Опис
P – політичний (*Political*)	– регулювання діяльності приватних медичних закладів з боку МОЗ; – медична реформа й *eHealth*; – ліцензійні та акредитаційні вимоги; – законодавство щодо захисту персональної інформації; – воєнний стан у країні.
E – економічний (*Economic*)	– низькій рівень доходів населення; – зростання темпів інфляції; – цінова чутливість пацієнтів до платних послуг; – зростання попиту на якісну медицину.
S – соціальний (*Social*)	– зростання попиту на профілактику й ранню діагностику; – підвищення очікувань щодо якості обслуговування; – старіння населення; – зростання попиту на медичний супровід; – популярність цифрових сервісів.
T – технологічний (*Technological*)	– інтеграція лабораторних систем і електронних медичних карт; – упровадження штучного інтелекту для аналізу медичних даних; – автоматизація процесів (запис, супровід, документообіг).

У роботі досліджено медичний центр за допомогою методу SWOT-аналізу. Унаслідок застосування цього методу визначено внутрішні та зовнішні фактори, що впливають на діяльність медичного закладу. Вхідними даними була внутрішня інформація (фінансові звіти, опитування персоналу, статистика прийомів) та зовнішня (ринкові дослідження, аналіз конкурентів, соціальні показники, нормативно-правова база, відгуки пацієнтів). Результати SWOT-аналізу подано в табл. 2.

Аналіз демонструє, що медичний центр "Еввіва" має міцну внутрішню структуру й репутацію, орієнтовану на якісне обслуговування пацієнтів. Проте для зміцнення конкурентної позиції доцільно інтегрувати цифрові рішення. Особливо важливо зосередитися на мінімізації ризиків, пов'язаних із конкуренцією та нормативними змінами, а також активному реагуванні на технологічні можливості, що сприяють розвитку медичного бізнесу.

Унаслідок аналізу внутрішньої документації, наявної бази даних і технології надання медичних послуг сформовано концептуальну й функціональну моделі організації медичного обслуговування (рис. 1 та рис. 2), а також побудовано модель бізнес-процесів, що відтворює ключові процеси медичного центру (рис. 3).

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

На рис. 1 зображено концептуальну модель надання медичних послуг пацієнтам установи, що ґрунтується на ідеї подання цього процесу як інтегрованої системи, у межах якої здійснюється перетворення вхідних даних і регламентів у результативні медичні послуги. Центральним елементом є функціональний блок "Надання медичних послуг пацієнтам медичного центру", що акумулює нормативно-правові, організаційні, кадрові та інформаційні складники. У функціюванні системи надання медичних послуг вхідні дані цілеспрямовано перетворюються у вихідні результати за допомогою застосування відповідних правил, процедур і залучення персоналу установи.

Таблиця 2. *SWOT-аналіз медичного центру*

Категорія	Опис
S – переваги (*Strengths*)	– широкий спектр медичних послуг (терапія, діагностика тощо); – професійний медичний персонал; – високий рівень сервісу й організації супроводу пацієнтів; – наявність протоколів лікування та чітких процедур; – наявність власних онлайн-ресурсів.
W – недоліки (*Weaknesses*)	– обмежене географічне покриття (локальність діяльності); – потреба в більш глибокій цифровій трансформації; – висока залежність від персоналу в долученні даних; – відсутність інтеграції з єдиною електронною медкартою.
O – можливості (*Opportunities*)	– упровадження ІТ-рішень (застосування елементів штучного інтелекту, оновлення онлайн-кабінету); – масштабування мережі або запуск франчайзингу; – розширення партнерства із зовнішніми лабораторіями для обміну інформацією; – надання профілактичних послуг пацієнтам медичного центру; – вихід на корпоративний медичний ринок (страхові кампанії).
T – загрози (*Threats*)	– посилення конкуренції з боку великих приватних мереж; – зміни в державному регулюванні приватної медицини; – економічна нестабільність; – зменшення платоспроможності населення; – кіберзагрози й проблеми з безпекою медичної інформації; – загрози, пов’язані з воєнними діями.

Рис. 1. Концептуальна модель надання медичних послуг у медичному центрі

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

На початковому етапі вхідні дані містять запит пацієнта на потребу в медичній допомозі та необхідну документацію, яка забезпечує юридичне й процедурне оформлення процесу. Ці елементи поєднуються з регламентуючими факторами – чинним законодавством, протоколами лікування, правилами обслуговування й процедурами супроводу пацієнта. Водночас застосовуються правила оцінювання якості лікування, які визначають стандарти ефективності та безпечності медичних послуг.

Перетворення вхідних даних у вихідні здійснюється за допомогою персоналу медичного центру та супровідних департаментів (господарського, інформаційного, розвитку). Господарський департамент забезпечує ресурсну підтримку процесу, інформаційний – надає доступ до даних і комунікаційних каналів, департамент розвитку формує стратегічні інноваційні рішення, а медичний персонал виконує безпосередні діагностичні, лікувальні та консультативні дії. Ключовим інформаційним ресурсом є база даних, що акумулює медичну інформацію про пацієнта та забезпечує її використання в ухваленні клінічних рішень.

Унаслідок функціювання цієї системи формується вихідна інформація, що має практичну значущість для пацієнта. Зокрема це індивідуальні рекомендації щодо збереження й відновлення здоров'я та персоналізований план лікування, який визначає подальшу послідовність медичних заходів.

Отже, надання медичних послуг можна розглядати як кероване інформаційно-ресурсне перетворення, у якому вхідні дані (запити, документи, регламенти) під впливом правил, процедур та дій персоналу трансформуються у вихідні результати, що мають безпосереднє прикладне значення для пацієнта.

Ретельний аналіз організаційної структури, внутрішньої документації та алгоритм надання послуг медичного центру "Еввіва" дав змогу комплексно оцінити взаємодію між пацієнтами, адміністративним персоналом, лікарями та зовнішніми діагностичними центрами. На основі отриманих результатів розроблено моделі, які відтворюють логіку функціювання організації та забезпечують формалізоване подання процесів у нотації BPMN, а також функціональну модель надання медичних послуг. Таке моделювання створює основу для подальшої оптимізації управлінських процедур, інтеграції результатів лабораторних досліджень і побудови єдиного інформаційного контуру медичного центру [30].

Запропонована концептуальна модель дає змогу не лише описати логіку взаємодії елементів системи, а й забезпечує можливість формалізації процесу у вигляді діаграм IDEF0 наступного рівня чи BPMN. Це зі свого боку створює передумови для автоматизації надання медичних послуг для підвищення ефективності внутрішніх процесів і вдосконалення якості медобслуговування.

Наступним кроком було розроблено функціональну модель (рис. 2), що відтворює надання медичних послуг в установі та забезпечує системний підхід до взаємодії між пацієнтом і персоналом. Вона побудована на основі принципів нормативно-правового забезпечення, організаційної підтримки та управління якістю медичного обслуговування. Декомпозицію концептуальної моделі надання медичних послуг у медичному центрі продемонстровано з використанням методології функціонального моделювання IDEF0 [31].

На першому етапі здійснюється забезпечення життєдіяльності медичного центру, що передбачає ведення необхідної документації та дотримання правил надання послуг згідно з чинним законодавством. Цей етап формує організаційно-правові умови для функціювання закладу. Результатом виконання цього етапу є забезпечення надання якісних послуги відповідно до чинного законодавства, кваліфікований персонал, інформаційна та господарсько-технічна безпека.

Другий етап – запис пацієнта на прийом, який передбачає можливість обрання зручного часу й ознайомлення з правилами обслуговування в медичному центрі. Цю функцію виконує персонал адміністративних підрозділів, і вона спрямована на підвищення доступності послуг.

Третій етап – це зустріч пацієнта та його супровід до кабінету лікаря, що передбачає як інформаційну підтримку, так і оформлення супровідної документації. На цьому етапі відбувається інтеграція адміністративних і медичних процедур.

Ключовим є етап надання медичних послуг, який базується на протоколах лікування. Він передбачає консультацію лікаря, визначення діагностичних і терапевтичних заходів, а також формування листа призначень. У разі необхідності реалізується етап супроводу пацієнта на додаткові обстеження, до аптеки чи реєстратури, що забезпечує

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

розширення спектра послуг і комплексність медичного обслуговування.

Завершальний етап виконує функцію моніторингу якості наданих послуг і передбачає отримання зворотного зв'язку від пацієнта. На основі відгуків здійснюється коригування лікувальних планів, формуються нагадування про наступні візити та беруться до уваги індивідуальні побажання пацієнтів.

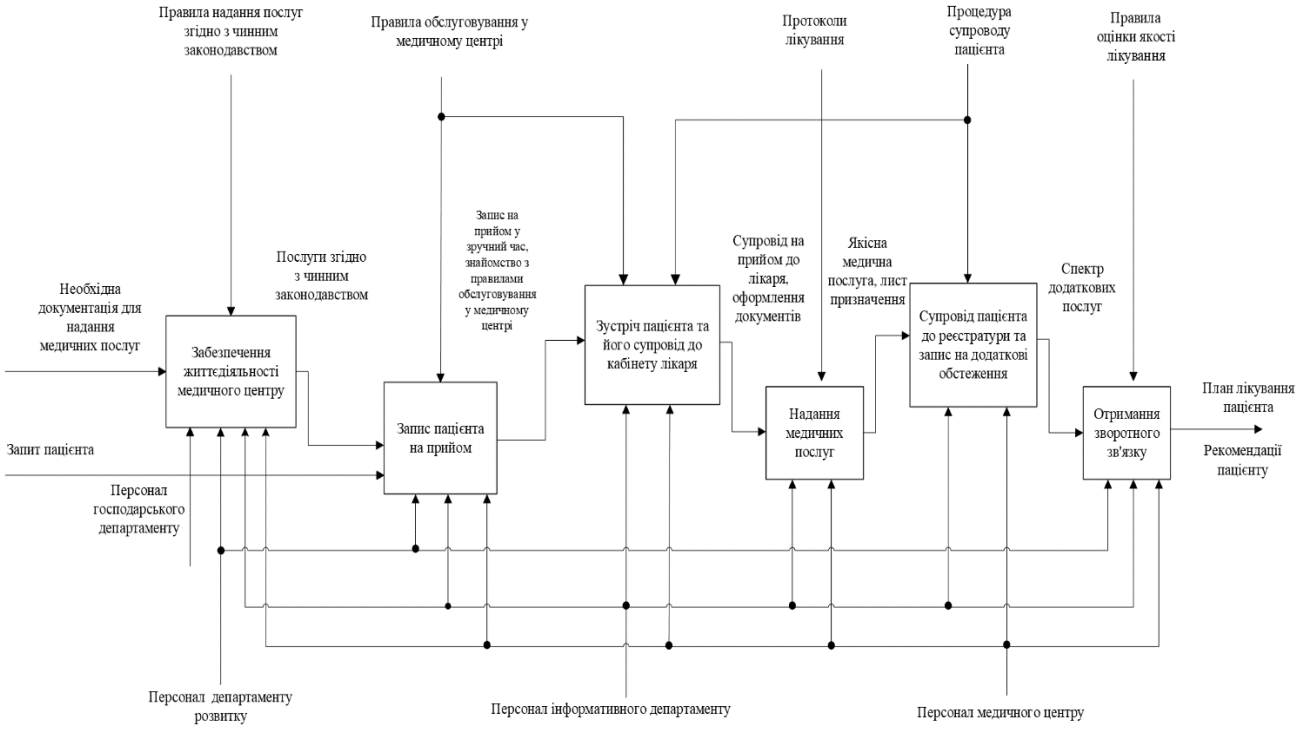


Рис. 2. Функціональна модель надання медичних послуг у медичному центрі

Отже, модель є інтегрованою системою організаційних і медичних процедур, спрямованою на підвищення ефективності функціювання медичного центру та покращення якості надання послуг. Її практичне застосування дає змогу оптимізувати процеси управління, забезпечити безперервність обслуговування та орієнтацію на потреби пацієнтів.

Запропонована модель бізнес-процесів (рис. 3) відтворює структуровану організацію діяльності медичного центру у взаємодії з пацієнтом.

Модель побудовано в нотації BPMN (*Business Process Model and Notation*) [32, 33], що забезпечує наочне подання логіки процесу, розподілу ролей та інформаційних потоків між учасниками.

Модель побудовано з огляду на багаторівневу взаємодію:

– **пацієнт** – ініціює процес, звертається до закладу, проходить етапи реєстрації, консультації, діагностики й лікування;

– **реєстратура** – відповідає за перевірку даних, оформлення медичної карти, запис на прийом та організаційний супровід пацієнта;

– **лікар** – консультує, призначає додаткові дослідження, аналізує результати та формує індивідуальний план лікування;

– **медичний працівник** – виконує лабораторні дослідження в межах закладу або формує запити до зовнішніх діагностичних центрів (ДЦ);

– **зовнішні діагностичні центри** – проводять дослідження в разі, якщо це неможливо виконати в медичному центрі;

– **база даних** – центральний вузол зберігання та передачі результатів досліджень і висновків.

Процес розпочинається зі звернення пацієнта до медичного центру (МЦ), після чого відбувається перевірка його персональної інформації та наявності медичної карти. За умови її відсутності створюється нова карта, що дає змогу інтегрувати пацієнта в систему медичного обслуговування. Наступним кроком є запис на прийом та супровід до кабінету лікаря. На етапі медичної консультації лікар збирає анамнез, формує первинний висновок і, якщо необхідно, призначає додаткові аналізи чи обстеження. Дослідження може здійснюватися як у межах

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

медичного центру, так і в зовнішніх діагностичних установах. У першому випадку медичний працівник безпосередньо виконує лабораторні дослідження, у другому – формується відповідний запит до зовнішніх установ. Усі результати інтегруються до централізованої бази даних, що забезпечує їх доступність для лікаря та гарантує цілісність інформаційного потоку. Подальший аналіз результатів дає змогу лікарю сформувати індивідуальний план лікування, який презентується пацієнтові у вигляді рекомендацій. Завершальний етап процесу передбачає отримання пацієнтом плану лікування, його оплату та закриття візиту. Отже, модель забезпечує повний цикл взаємодії від первинного звернення до фінального етапу лікування.

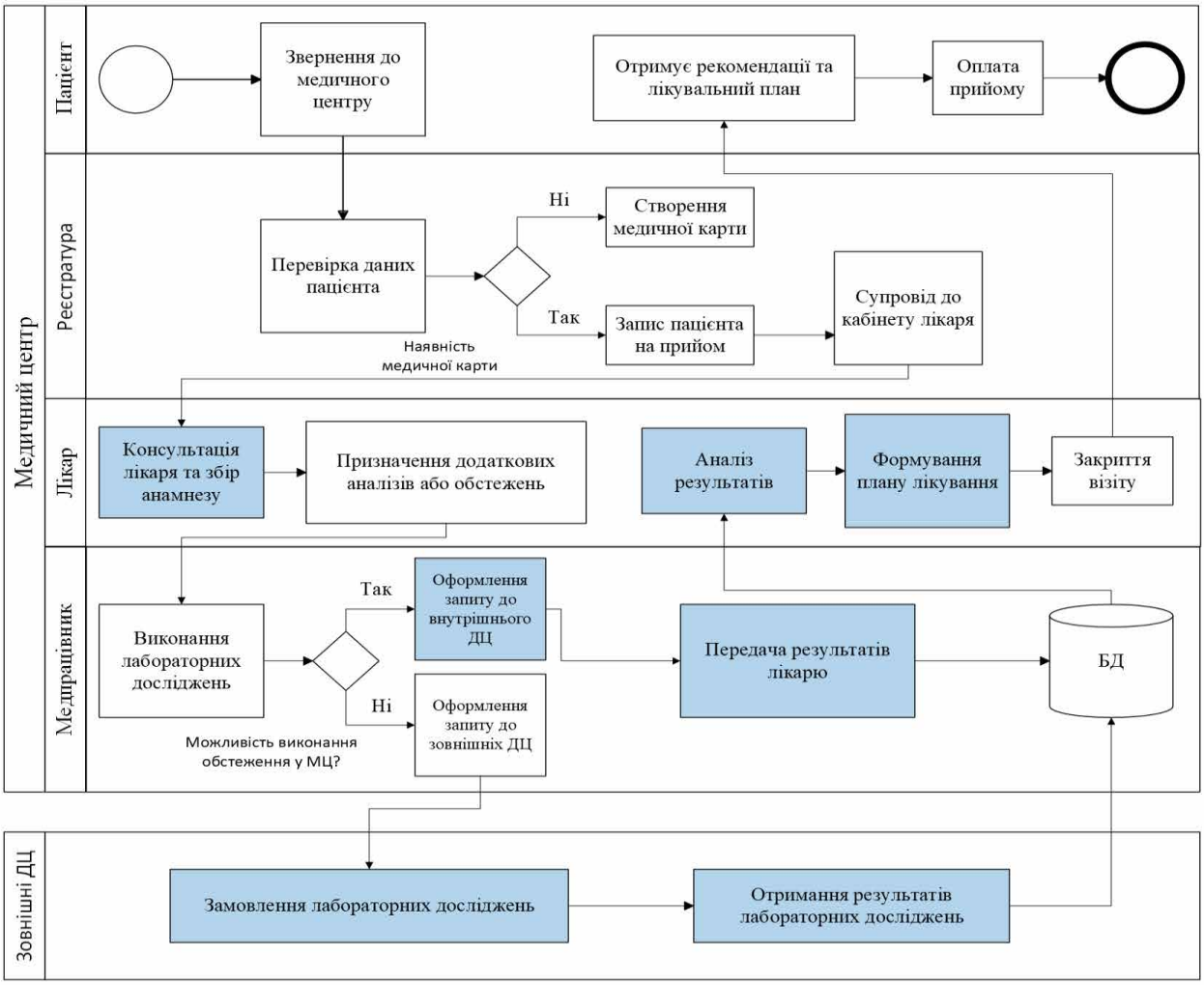


Рис. 3. Модель бізнес-процесів організації діяльності медичного центру у взаємодії з пацієнтом

Наукова цінність запропонованої моделі полягає в можливості її застосування для формалізації та оптимізації управлінських і медичних процесів, що відбуваються в медичних центрах. Використання BPMN допомагає не лише описати процеси на концептуальному рівні, а й забезпечує основу для їх подальшої автоматизації та інтеграції з інформаційними системами.

Важливо зауважити, що окремі бізнес-процеси, позначені на моделі синім кольором (рис. 3), визначено як такі, що підлягають автоматизації інформаційної системи, що сприятиме підвищенню ефективності як управлінських, так і клінічних процесів. Саме ці процеси формують основу цифровізації ключових етапів взаємодії між пацієнтом, лікарем, медичним персоналом та зовнішніми діагностичними центрами. Їх ідентифікація здійснена на основі аналізу сучасних технологій у медичному центрі "Евіва" та з огляду на актуальні запити закладу.

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

Результати досліджень та їх обговорення

За результатами проведеного аналізу побудовано концептуальну та функціональну моделі надання медичних послуг, а також модель бізнес-процесів, що описує організацію діяльності медичного центру. Запропоновані функціональна модель та модель бізнес-процесів дають змогу формалізувати основні етапи взаємодії між пацієнтом і персоналом, починаючи від первинного звернення до завершення лікування. Використання методології BPMN забезпечує структуроване відтворення логіки бізнес-процесів, ролей та інформаційних потоків, що створює основу для подальшої автоматизації та інтеграції з інформаційними системами. Така автоматизація мінімізує ризики дублювання або втрати інформації та забезпечує контроль на кожному етапі медичного обслуговування.

Функціональна модель репрезентує системний підхід до управління якістю медичних послуг, поєднуючи нормативно-правові, організаційні та клінічні аспекти. Вона описує послідовність процедур, що охоплюють як адміністративні, так і лікувальні процеси, водночас з отриманням зворотного зв'язку від пацієнта. Отже, модель підсилює пацієнтоорієнтованість організації, забезпечує комплексність медичного супроводу та створює механізми покращення надання послуг.

Окремі бізнес-процеси, ідентифіковані в межах моделі бізнес-процесів, визначені як пріоритетні для подальшої автоматизації. Їх інтеграція в інформаційну систему медичного центру стане основою для впровадження методів інтелектуального аналізу даних і автоматизованого розпізнавання результатів лабораторних досліджень. Це дасть змогу підвищити точність екстракції медичної інформації, забезпечити відстеження динаміки лабораторних показників і формувати рекомендації для лікаря, що сприятиме оптимізації клінічних рішень і підвищенню якості медобслуговування.

Висновки й перспективи подальшого розвитку

Огляд сучасних праць у сфері цифровізації медичної інформації та застосування методів інтелектуального аналізу продемонстрував, що основними проблемами залишаються семантична неоднорідність результатів лабораторних досліджень, обмежена інтеграція даних у клінічні бізнес-процеси та недостатня підтримка інструментів візуалізації для оцінювання ретроспективної динаміки показників.

Отже, проведений огляд і запропоновані моделі створюють науково-методологічне підґрунтя для автоматизації збирання та оброблення результатів лабораторних досліджень, їх уніфікації та інтеграції в єдиний інформаційний контур медичного центру. Це дасть змогу підвищити ефективність бізнес-процесів медичного закладу у взаємодії з пацієнтом, скоротити час опрацювання результатів діагностики, знизити ризик помилок і забезпечити якісну підтримку надання медичних послуг.

З метою реалізації методів інтелектуального аналізу та автоматизованого розпізнавання результатів клінічних досліджень заплановано провести порівняльний аналіз методів машинного навчання, оброблення природної мови (NLP), а також сучасних LLM- і RAG-підходів. Це допоможе визначити найбільш доцільні методи або їх комбінації для інтеграції в інформаційну систему, що забезпечить високу точність екстракції даних, формування динаміки лабораторних показників і автоматизоване надання рекомендацій лікарю. Отже, запропоновані моделі не лише відтворюють реальну логіку діяльності медичного центру, але й окреслює контур майбутньої інформаційної системи, спрямованої на практичне використання результатів методів інтелектуального аналізу медичних даних.

References

1. Pai, M., Ganiga, R., Pai, R., Sinha, R. (2021), "Standard electronic health record (EHR) framework for Indian healthcare system", *Health Services and Outcomes Research Methodology,* No. 21, P. 339–362. DOI: https://doi.org/10.1007/s10742-020-00238-0
2. Tyravska, Yu., Borysyuk, I., Vasylyuk-Zaitseva, S. (2023), "Ehrsystems in Ukraine: enhancing patient care and data management", *Prospects and innovations of science*, No. 14(32), P. 879–889. DOI: https://doi.org/10.52058/2786-4952-2023-14(32)-879-889

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

3. Kujala, S., Simola, S., Wang, B., Soone, H., Hagström, J., Bärkås, A., Hörhammer, I., Cajander, Å., Fagerlund, A.J., Kane, B., Kharko, A., Kristiansen, E., Moll, J., Rexphepi, H., Hägglund, M., Johansen, M.A. (2024), "Benchmarking usability of patient portals in Estonia, Finland, Norway and Sweden", *International Journal of Medical Informatics*, Vol. 181, 105302 p. DOI: https://doi.org/10.1016/j.ijmedinf.2023.105302
4. Stüer, T., Juhra, C. (2022), "Usability of Electronic Health Records in Germany – An Overview of Satisfaction of University Hospital Physicians", *Studies in Health Technology and Informatics*, No. 296, P. 90–97. DOI: https://doi.org/10.3233/SHTI220808
5. Wilson, K., Khansa, L. (2018), "Migrating to electronic health record systems: A comparative study between the United States and the United Kingdom", *Health Policy*, No. 122(11), P. 1232–1239. DOI: https://doi.org/10.1016/j.healthpol.2018.08.013
6. Stachwitz, P., Debatin, J.F. (2023), "Digitalisierung im Gesundheitswesen: heute und in Zukunft", *Bundesgesundheitsblatt*, No. 66, P. 105–113. DOI: https://doi.org/10.1007/s00103-022-03642-8
7. Bloom, B., Pott, J., Thomas, S., Gaunt, D., Hughes, T. (2021), "Usability of electronic health record systems in UK emergency departments: a national survey", *Emergency Medicine Journal,* No. 38(6), P. 410–415. DOI: https://doi.org/10.1136/emermed-2020-210401
8. Grossman, Z., del Torso, S., van Esso, D., Ehrich, J.H.H., Altorjai, P., Mazur, A., Wyder, C., Neves, A.M., Dornbusch, H.J., Jaeger Roman, E., Santucci, A., Hadjipanayis, A. (2016), "Use of electronic health records by child primary healthcare providers in Europe", *Child: Care, Health and Development*, No. 42(6), P. 928–933. DOI: https://doi.org/10.1111/cch.12374
9. Febrita, H., Martunis, M., Syahrizal, D., Abdat, M., Bakhtiar, B. (2021), "Analysis of hospital information management system using human organization fit model", *Indonesian Journal of Health Administration (Jurnal Administrasi Kesehatan Indonesia)*, No. 9(1), P. 23–32. DOI: https://doi.org/10.20473/jaki.v9i1.2021.23-32
10. Hertin, R.D., Al-Sanjary, O.I. (2018), "Performance of hospital information system in Malaysian public hospital: a review", *International Journal of Engineering & Technology*, No. 7(4.11), P. 24–28. DOI: https://doi.org/10.14419/ijet.v7i4.11.20682
11. Arora, L., Ikbal, F. (2023), "Experiences of implementing hospital management information system (HMIS) at a tertiary care hospital, India", *VILAKSHAN - XIMB Journal of Management*, No 20(1), P. 59–81. DOI: https://doi.org/10.1108/XJM-09-2020-0111
12. Darwis, M., Soraya, S., Nawangwulan, K., Ekawaty, D., Imran, A., Yusfik, Y. (2023), "Hospital management information system", *International Journal of Health Sciences*, No. 1(4), P. 485–492. DOI: https://doi.org/10.59585/ijhs.v1i4.174
13. Ashem, E.M., Shalaby, M.F.Y., Mabrouk, M.S., Marzouk, S.Y. (2023), "Integrative software design for hospital information systems", *Mansoura Engineering Journal*, No. 48(4), Article 7. DOI: https://doi.org/10.58491/2735-4202.3057
14. Narsale, S., Choramale, V., Kadam, P., Mishra, A. (2023), "Hospital management information system: an overview", *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, No. 11(5), P. 6948–6951. DOI: https://doi.org/10.22214/ijraset.2023.52562
15. Arpaci, I., Ghazisaeedi, M., Esmaeilzadeh, F., Barzegari, R., Barzegari, S. (2023), "Ranking the critical success factors for hospital information systems using a fuzzy analytical hierarchy process", *CIN: Computers, Informatics, Nursing*, No. 41(10), P. 765–770. DOI: https://doi.org/10.1097/CIN.0000000000001042
16. Gong, S., Zhu, X., Zhang, R., Zhao, H., Guo, C. (2023), "An intelligent resource management solution for hospital information system based on cloud computing platform", *IEEE Transactions on Reliability*, No. 72(1), P. 329–342. DOI: https://doi.org/10.1109/TR.2022.3161359
17. Antunes, R.S., da Costa, C.A., Küderle, A., Yari, I.A., Eskofier, B. (2022), "Federated learning for healthcare: systematic review and architecture proposal", *ACM Transactions on Intelligent Systems and Technology (TIST)*, No. 13(4), Article 54, P. 1–23. DOI: https://doi.org/10.1145/3501813
18. Woldemariam, M.T., Jimma, W. (2023), "Adoption of electronic health record systems to enhance the quality of healthcare in low-income countries: a systematic review", *BMJ Health & Care Informatics*, No. 30, Articles e100704. DOI: https://doi.org/10.1136/bmjhci-2022-100704
19. Almowalad, M., Alanazi, W., Alharthi, S., Alsahli, B., Aldhmaly, A., Alrashdi, R., Huntul, A. (2022), "Electronic Health Records (EHR): The Impact and Challenges of Implementing and Utilizing EHR Systems in Healthcare Organizations", *Journal of Survey in Fisheries Sciences*, No. 9(4), P. 96–98. DOI: https://doi.org/10.53555/sfs.v9i4.2545
20. Kostrinsky-Thomas, A., Hisama, F., Payne, T. (2021), "Searching the PDF haystack: automated knowledge discovery in scanned EHR documents", *Applied Clinical Informatics*, No. 12(2), P. 245–250. DOI: https://doi.org/10.1055/s-0041-1726103
21. Stuhlmiller, T., Rabe, A., Rapp, J., Manasco, P., Awawda, A., Kouser, H., Salamon, H., Chuyka, D., Mahoney, W., Wong, K., Kramer, G., Shapiro, M. (2025), "A scalable method for validated data extraction from electronic health records with large language models", *medRxiv.* DOI: https://doi.org/10.1101/2025.02.25.25322898

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

22. Wickramarathna, M., De Silva, K., Lekamalage, V., Senanayake, J., Perera, J., Ruggahakotuwa, L. (2022), "Oxygen: a distributed health care framework for patient health record management and pharmaceutical diagnosis", *Proceedings of the 2022 4th International Conference on Advancements in Computing (ICAC)*, P. 30–35. DOI: https://doi.org/10.1109/ICAC57685.2022.10025250
23. El Azzouzi, M., Bellafqira, R., Coatrieux, G., Cuggia, M., Bouzille, G. (2024), "Secure extraction of personal information from EHR by federated machine learning", *Studies in Health Technology and Informatics*, No. 316, P. 611–615. DOI: https://doi.org/10.3233/SHTI240488
24. Maarseveen, T., Maurits, M., Niemantsverdriet, E. et al. (2021), "Handwork vs machine: a comparison of rheumatoid arthritis patient populations as identified from EHR free-text by diagnosis extraction through machine-learning or traditional criteria-based chart review", *Arthritis Research & Therapy*, No. 23, 174 p. DOI: https://doi.org/10.1186/s13075-021-02553-4
25. Kundu, M. (2021), "Statistics and machine learning methods for EHR data – from data extraction to data analytics: by Hulin Wu et al." *Journal of Biopharmaceutical Statistics*, No. 31(4), P. 559–560. DOI: https://doi.org/10.1080/10543406.2021.1928833
26. Jasthi, V. (2021), "NLP-driven extraction of clinical insights from unstructured EHR data", *International Journal on Science and Technology (IJSAT)*, No. 12(3), P. 1–9. DOI: https://doi.org/10.71097/IJSAT.v12.i3.7551
27. Mukhopadhyay, A. (2024), "Electronic Health Records (EHR) Extraction Using Various NLP Models", *International Journal of Advanced Management, Science and Technology (IJAMST)*, No. 5(1), P. 1–4. DOI: https://doi.org/10.54105/ijamst.C3008.05011224
28. Joshi, R., Bubna, Y., Sahana, M., Shruthiba, A. (2024), "An approach to intelligent information extraction and utilization from diverse documents", *Proceedings of the 2024 8th International Conference on Computational System and Information Technology for Sustainable Solutions (CSITSS)*, P. 1–5. DOI: https://doi.org/10.1109/CSITSS64042.2024.10816908
29. Gbada, H., Kalti, K., Mahjoub, M.A. (2025), "Deep learning approaches for information extraction from visually rich documents: datasets, challenges and methods", *International Journal on Document Analysis and Recognition (IJDAR)*, No. 28, P. 121–142. DOI: https://doi.org/10.1007/s10032-024-00493-8
30. Kutsenko, D., Grinchenko, M. (2025), "Formulation of a conceptual model of information technology for intelligent patient data analysis", *International Scientific and Practical Conference "Information Systems and Innovative Technologies for Project and Program Management"* P. 38-40. available at: https://repository.kpi.kharkov.ua/handle/KhPI-Press/94105
31. "ISO/IEC/IEEE 31320-1:2012. Information technology. Modeling languages. Part 1. Syntax and semantics for IDEF0".
32. "BPMN for research", available at: https://github.com/camunda/bpmn-for-research (last accessed 23.08.202025)
33. Kopp, A., Orlovskyi, D. (2023), "Towards intelligent technology for error detection and quality evaluation of business process models", *In: IntelITSIS'2023: 4th International Workshop on Intelligent Information Technologies and Systems of Information Security*, P. 1–14. available at: https://ceur-ws.org/Vol-3373/keynote1.pdf

Received (Надійшла) 31.08.2025
Accepted for publication (Прийнята до друку) 30.11.2025
Publication date (Дата публікації) 28.12.2025

Відомості про авторів / About the Authors

Гринченко Марина Анатоліївна – кандидат технічних наук, доцент, Національний технічний університет "Харківський політехнічний інститут", завідувачка кафедри управління проєктами в інформаційних технологіях, Харків, Україна; e-mail: marinagrunchenko@gmail.com; ORCID ID: 0000-0002-8383-2675; Scopus ID: https://www.scopus.com/authid/detail.uri?authorId=57195064619

Куценко Дмитро Олександрович – Національний технічний університет "Харківський політехнічний інститут", аспірант кафедри управління проєктами в інформаційних технологіях, Харків, Україна; e-mail: cureghost17@gmail.com; ORCID ID: 0009-0005-6359-3143

Grinchenko Marina – PhD (Engineering Sciences), Associate Professor, National Technical University "Kharkiv Polytechnic Institute", Head of the Department of Project Management in Information Technologies, Kharkiv, Ukraine.

Kutsenko Dmytro – National Technical University "Kharkiv Polytechnic Institute", PhD Student at the Department of Project Management in Information Technologies, Kharkiv, Ukraine.

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

MODELS OF MEDICAL CENTER BUSINESS PROCESSES TO IMPROVE DECISION-MAKING EFFICIENCY

The **subject** of the article is the process of digitizing laboratory test results and formalizing the business processes of a medical center in order to develop an information system focused on intelligent data analysis. The **goal** of this work is to design models that formalize the business processes of a medical center in interaction with patients, enabling the identification of areas that require automation to improve the efficiency of medical services. The following **tasks** were solved: analyzing the current state of medical data digitalization and the issues related to the unification of laboratory test results; examining the organization of business processes in a medical center and the interaction of its key participants; developing a conceptual and functional model of service delivery; and constructing a business process model that reflects the structured organization of the institution's activities in interaction with patients. The following **methods** include SWOT and PEST analyses of the medical center, examination of internal and external documentation, analysis of the existing database and algorithms for providing medical services, the IDEF0 functional modeling methodology, and the BPMN business process modeling approach. The obtained **results**: an analysis of existing solutions and studies was conducted; the results of SWOT and PEST analyses of the medical center were presented; a conceptual model of medical service organization was created; a functional model was built considering regulatory, organizational, and clinical aspects; and a BPMN-based model of key business processes was developed. Business processes requiring automation using intelligent data analysis methods were identified, forming the framework of the future information system. **Conclusions:** the proposed models provide a scientific and methodological foundation for automating the collection and processing of laboratory test results, their standardization, and integration into a unified information framework of the medical center. This will enhance the efficiency of the medical center's business processes in interaction with patients, reduce the time required for processing diagnostic results, minimize the risk of errors, and ensure high-quality support for the provision of medical services.

Keywords: functional model; business process; business process automation; data analysis.

Бібліографічні описи / Bibliographic descriptions

Гринченко М. А., Куценко Д. О. Моделі бізнес-процесів медичного центру для підвищення ефективності прийняття рішень. *Сучасний стан наукових досліджень та технологій в промисловості.* 2025. № 4 (34). С. 5–17. DOI: https://doi.org/10.30837/2522-9818.2025.4.005

Grinchenko, M., Kutsenko, D. (2025), "Models of medical center business processes to improve decision-making efficiency", *Innovative Technologies and Scientific Solutions for Industries*, No. 4 (34), P. 5–17. DOI: https://doi.org/10.30837/2522-9818.2025.4.005

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

UDC 004.032.26:004.932.2:004.052 DOI: https://doi.org/10.30837/2522-9818.2025.4.018

S. DANYLENKO, K. SMELYAKOV

CONTENT-BASED IMAGE RETRIEVAL METHOD IN A MULTIDIMENSIONAL MODEL AT BIG DATA SCALE

The **subject** of the study is the method and algorithms for content-based image retrieval within the Multidimensional Cube (MDC) model. The **goal** is to develop a search method based on image descriptor vectors and an algorithm that implements this method in both sequential and parallel versions for MDC. The research **tasks** include: defining requirements for the search method; analyzing the MDC model structure and defining the approach to the search method; developing search methods and algorithms for scenarios where the model is stored in RAM or in a relational database; integrating parallel computing into the algorithm; analyzing alternative models based on multidimensional trees, graphs, hashing, inverted indexing, quantization and inverted multi-index structures; developing evaluation metrics and conducting experiments to compare the efficiency of the MDC-based method with alternative search models. **Methodology**: analytical and comparative methods for search algorithm evaluation, modeling, and experimental verification were applied. Thread-level parallelism and hardware optimization methods were used, along with comparative analysis of model efficiency (KD-tree, Locality-Sensitive Hashing, Hierarchical Navigable Small World, Inverted File with Flat Compression, Inverted Multi-Index). Statistical methods were employed to assess results using recall, search time, and model construction time metrics. Experiments were conducted with both web-sourced and synthetic image descriptors, as well as load testing to evaluate the model's throughput. **Results**: a new search method and the Wave-Search Algorithm were developed. Its parallel version achieves up to a 3x speedup. For top-10 and top-100 queries in a dataset of 1 million descriptors, MDC shows the best overall performance among the compared models based on the metrics and strong stability under load. **Conclusions**: the proposed search method and its implementation (Wave-Search Algorithm) efficiently utilize the MDC model's structure for search tasks, outperforms alternative search models in terms of effectiveness, demonstrates robustness under load, and has significant potential for further development, including the use of hardware acceleration.

Keywords: similarity search; big data; search algorithms; data structures; content-based image retrieval; parallel computing; high-performance computing; search efficiency; algorithm optimization.

Introduction

Information search is a fundamental need of humanity. Even in ancient times, when writing began to develop, there was a need to organize existing works and search for information contained in them. This led to the emergence of alphabetical indexes, catalogs, and indexes, which formed the basis of modern search systems [1].

Search engines have undergone a long period of development. Their implementation involves both classical approaches, such as Boolean, Vector Space, and Probabilistic models, and more modern ones, such as machine learning, deep learning, neural ranking models, contextual and semantic methods [2].

Search is used both in everyday tasks, such as renting accommodation, buying tickets, searching for goods, etc., and in professional fields such as medicine, education, business, e-commerce, and science. It greatly simplifies and speeds up the process of analyzing the necessary information [3].

A search engine is based on a model that uses a specific data structure. Each data structure has its own search algorithms, which are applied depending on the requirements. For example, for graphs, depending on the expected result, the classic Breadth First Search and Depth First Search algorithms can be applied. This makes search models universal and allows them to be widely used in various tasks [4].

With the growth of visual content, there is a need for effective image retrieval and search based on graphic information. Content-based image retrieval (CBIR) systems allow images to be found based on their visual characteristics, such as color, texture, and shape. Modern CBIR systems are actively used in medicine, e-commerce, and other industries. The implementation of such systems requires special data structures and search algorithms [5].

Despite significant progress in the development of CBIR systems, there are challenges related to scalability, efficiency, and search accuracy. In particular, there is the problem of effective representation of visual data in the context of Big Data and performing effective searches in these conditions. Therefore, the development of new models and search algorithms is an important task for modern research.

ISSN 2522-9818 (print)
Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34) ISSN 2524-2296 (online)

Analysis of recent studies and publications

Each CBIR model has its own search methods. They are determined by the data structures underlying the model and implemented by specific search algorithms. Basically, modern models are based on: multidimensional trees, hashing, graphs, and clustering. Special adapted vector databases can also be used. The chosen approach to indexing and searching directly affects the speed and accuracy of the CBIR system. Models use graphical information representation in the form of a descriptor, which often takes the form of a normalized vector [6], and compare them using special metrics [7].

Tree structures organize descriptors hierarchically, which allows you to effectively narrow down the search area. Searching in KD-Tree is performed by recursively comparing coordinates with node boundaries, which allows you to quickly filter out unnecessary branches. Quad-tree uses recursive division of space into quadrants, checking only those that contain the requested point. R-tree checks the minimum rectangles containing objects and recursively examines subtrees for range matching [8]. In B+ Tree, the search goes through internal nodes to leaf nodes, which optimizes range queries [9]. VP-Tree compares the distance to the reference point in each node, filtering out unnecessary areas. Ball-Tree compares the distance to "balls" (nodes) and recursively filters out irrelevant parts of the space [10]. Learned index uses machine learning models to predict the location of an element, significantly speeding up the search [11].

Hash methods organize the search for similar objects in different ways. Locality-Sensitive Hashing (LSH) performs a search by comparing hash codes: the more similar the objects, the more likely their hashes will match [12]. Spherical Hashing improves this process by using hyperspheres to divide the space for more accurate grouping of similar objects [13]. FlyHash speeds up the search thanks to sparse coding, which reduces the number of comparisons [14]. In Deep Hashing, the search is performed using binary codes that simultaneously take into account the content and semantics of images [15]. Adaptive multiscale hashing performs searches based on multi-level features, which allows for better differentiation between similar objects [16].

The most relevant graph-based nearest neighbor search methods are Hierarchical Navigable Small World (HNSW) [17] and NN-Descent [18]. HNSW demonstrates high accuracy and search speed thanks to its multi-level hierarchy and greedy navigation. Its architecture allows scaling to billion-scale datasets while maintaining high performance. It has numerous modifications: bh-DBSCAN uses bidirectional HNSW to reduce redundant queries of neighboring points [19], and HNSWQ improves entry points and adapts the algorithm to work with a graphics processor [20]. On the other hand, NN-Descent is an efficient approach to building search graphs – it does not require parametric fine-tuning and scales well. It is often used in systems in various fields, including bioinformatics [21, 22].

IVFFlat (Inverted File with Flat compression) combines vector space clustering (usually via k-means) with linear scanning only in selected clusters, reducing computational costs. Modifications to IVFFlat are proposed to improve search speed and flexibility. In particular, Fast IVF optimizes cluster selection to speed up search without significant loss of accuracy [23]. Semi-supervised IVF uses auxiliary labels to improve the distribution of vectors across the index, which increases the relevance of results [24]. Reconfigurable Inverted Index allows the index to be dynamically reconfigured, maintaining efficiency even when the number or composition of vectors changes [25].

Product Quantization (PQ) provides high compression and reduces the amount of stored data [26], and its use in Inverted Multi-Index (IMI) allows the application of multiple inverted indexes, which improves search speed and quality [27]. Approaches are proposed to improve efficiency: Hierarchical Hyperbolic Product Quantization uses hyperbolic geometry to preserve hierarchical semantic relationships between images, improving search accuracy in complex semantic structures [28]; Entropy-Optimized Deep Weighted Product Quantization provides balance in the representation of code words, improving search accuracy and coding efficiency [29]; Fuzzy Norm-Explicit Product Quantization proposes the use of second-type fuzzy sets to improve search completeness without increasing computational complexity [30].

Ready-made software products that provide efficient work with vectors, in particular vector databases or additions to relational databases, such as Pgvector [31], are also very popular.

Goals and objectives

A wide selection of search models allows you to use the one that is most effective for a specific task. Since MDC is a unique search model with its own advantages,

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

it requires its own search method that will utilize the features of its structure.

The purpose of this work is to develop a method for searching images in the MDC model based on image descriptor vectors and an algorithm that implements this method in sequential and parallel versions.

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

1) formulating requirements for the search method;

2) analyzing the MDC structure and determining the approach to the search method;

3) developing a search method and algorithms when placing the model in RAM and in a relational database;

4) implementing parallel computing in the search algorithm;

5) analyzing alternative models based on: multidimensional trees, graphs, hashing, reverse index, quantization, and reverse multi-index;

6) developing metrics and conducting experiments to compare the effectiveness of the MDC model with alternative search models.

The search method must meet the following requirements:

1) effectively use workstation resources;

2) ensure high search speed and quality;

3) adapt to variations in MDC placement in memory.

The research is comprehensive, and its result is to determine the effectiveness of the MDC model in comparison with alternative models in the context of image retrieval in Big Data repositories. It also determines the conditions under which the use of MDC is preferable to other models.

Search method in MDC

In MDC, the initial descriptor vectors are reduced by decreasing the initial dimension through pairwise aggregation of vector values. This results in an additional vector of small dimension. The range of possible values of vector elements is divided into intervals depending on the distribution of vector values within the range. After that, the values of the additional vector are replaced with the indices of the interval to which they belong. Thus, based on the values of the descriptor vector elements, a multidimensional cube (MDC) is formed, in which the dimension of the additional vector specifies the number of measurements, and when the measurements intersect, the aforementioned intervals form cells (clusters). The values of the index vector determine the location of the descriptor on the interval of each dimension and in space as a whole [32].

This structure determines the approach of the search method. The relationships between the values of the vectors of different descriptors are stored as indices in the index vector. For example, if for a given vector value the corresponding measurement interval index has a value of 5, then to find descriptors with slightly smaller and slightly larger values in this measurement, it is necessary to check the MDC cells that have an index equal to 4 and 6, respectively. This operation can be performed for all measurements, going through the cells around the specified one, visually imitating waves.

The algorithm of actions in this search method in MDC boils down to the following steps:

1) process the vector of the searched image descriptor and convert it to a vector of indices.

2) Use the index vector to determine the MDC cell and perform a full search for descriptors in this cell. Compare the descriptors with the searched one based on the selected metric and calculate the difference;

3) if the search was unsuccessful, perform 1 search wave, within which increase and decrease the index values by 1 for all measurements, form cell indexes from the values of all measurements, and perform a full search for descriptors in them;

4) if the search was unsuccessful, repeat step 3 until the search is successful, the maximum number of waves is reached, or all cells are checked;

5) sort the resulting list of descriptors in ascending order of the calculated difference value and return the required number of descriptors as the search result.

Performing a search wave means checking cells that are at a certain distance in terms of measurement index values relative to the initially defined cell for the descriptor being searched for.

The search is successful when the condition for its completion is met, for example, finding a certain number of the first most similar descriptors.

The number of descriptors in the search wave is determined by the formula:

$$wc_i = (j_i)^N \times rc - (j_{i-1})^N \times rc, \quad rc = \frac{dc}{k^N}, \qquad (1)$$

where wc is the number of descriptors on the search wave, i is the wave number, starting from 1, j is an ascending sequence of natural odd numbers: [3, 5, ...], N is the number of measurements, rc is the theoretical number of descriptors in one cell, dc is the number

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

of descriptors in the storage, k is the number of intervals in the measurements.

The parameters k and N are selected based on the requirements for the search system so that the value o f rc is close to the size of the search page [32].

The selected similarity metric does not affect the operation of MDC, but the quality of the search may depend on it. A wide range of metrics can be used: Manhattan, Euclidean, Quadratic Euclidean, etc. The metric that is best suited for working with a specific type of descriptor is selected. By default, the Manhattan distance is used, which is determined by the formula:

$$\mu = \sum_{i=1}^{N} \left| v_i^{(1)} - v_i^{(2)} \right|, \quad (2)$$

where μ is the difference between descriptors, i is the ordinal number of the vector element, N is the number of vector elements, $v^{(1)}, v^{(2)}$ are the vectors of the first and second descriptors.

The calculated resulting list is stored for a certain time in the search engine cache, because it usually contains more descriptors than were displayed to the user on the first page of results. This allows you to avoid performing further search waves until all the results found are displayed.

This search algorithm is called the Wave-Search Algorithm (WSA) because of the way it uses the MDC structure and is a hybrid of reverse multi-index search and exhaustive search.

An example of a search is shown in Figure 1. It shows an MDC with 2 dimensions and 10 equal intervals for each dimension

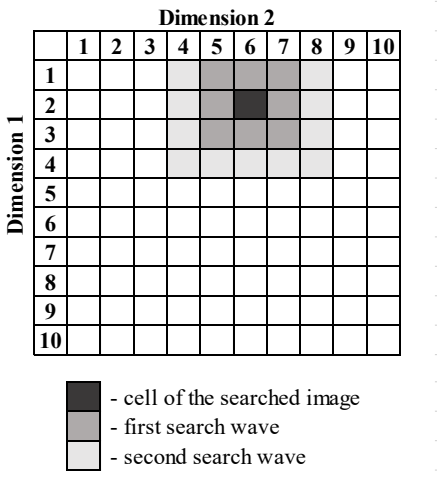


Fig. 1. Wave-Search visualization

Cell [2, 6] corresponds to the desired descriptor, and the search is initially performed in it. Next, one search wave is performed, and the indices are increased/decreased by 1 when possible. In the first wave, the following cells are checked: [1, 5], [1, 6], [1, 7], [2, 5], [2, 7], [3, 5], [3, 6], [3, 7]. In the second wave, the indexes are increased/decreased by 1 again, and so on.

Due to the peculiarities of the MDC structure, the values of the descriptor vector may be located on the border of cells, and therefore the descriptor relevant for the search may end up in a neighboring cell. And when searching only in the identified cell, it will be missed. To solve this, WSA always performs 1 search wave. Placing descriptors in cells is quite effective, and with the right parameters, each cell contains a small number of descriptors. This operation slightly increases the search time but improves the search quality.

The presented algorithm is a general implementation of the search method. For each of the MDC placements in memory, there are variations of the algorithm that technically refine some details of the implementation. There are also separate versions of algorithms that use parallel computing or can apply hardware acceleration.

Wave-Search algorithm for different ways of MDC placement in memory

When MDC is stored in RAM, the model has a "flat" structure – the index vector takes the form of a string with a separator and is then entered into a hash table, where the key is the cell index, for example "1-3-1-4", and the value for the key is a list of descriptors [32].

For such placement of MDC in memory, the following details of WSA implementation are specified:

– to form the indices of cells that fall into a specific search wave, numbers are extracted from the string representation of the cell and the specified index increase/decrease operations are performed on them. The calculated cell indexes are used to form their string representations, which are used to extract lists of descriptors from the hash table;

– the extracted lists are collected into a single resulting list. Next, for all descriptors found, the difference is calculated according to the selected metric, and the list is sorted according to its value.

When placing MDC in a relational database (DB), a "star" schema is used, in which there is a separate table for each dimension, where information about intervals is stored, a table for storing descriptors, and a linking table, in which the descriptor identifier is stored by the index vector [32].

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)
Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

In this implementation of WSA, the process of extracting descriptors from the repository is transferred to the DB. This is done using an SQL query, where cell indexes are specified as tuples. An example of such a query is shown in Figure 2.

```
SELECT d.prop_1, d.prop_2, d.prop_3, d.prop_4,
       d.file_id, d.original_vector
FROM descriptors d
    INNER JOIN links l ON d.id = l.descriptor_id
WHERE l.prop_1 IN (10, 9, 11) AND
      l.prop_2 IN (18, 17, 19) AND
      l.prop_3 IN (12, 11, 13) AND
      l.prop_4 IN (7, 6, 8)
```

Fig. 2. SQL query to perform a search

The descriptors found are returned as a list. Next, for all descriptors found, the difference is calculated according to the selected metric, and the list is sorted according to its value.

Parallel computations in the Wave-Search algorithm

From the technical details of the algorithm implementation, it follows that in the WSA implementation in RAM, it is possible to simultaneously perform the process of extracting descriptors from the hash table, calculating their similarity to the search term, and sorting the resulting list.

For parallel sorting, any of the currently known algorithms implemented in the programming language used to create MDC can be used. MDC has a basic implementation in the Java programming language. For implementation in Java, this is currently a parallel merge sort implementation.

In the DB implementation, extracting descriptors from different MDC cells is performed with a single query, which is more efficient in terms of working with the DB. Extracting descriptors from a hash table in RAM implementation is not a complex process, and running it in parallel mode is not expected to result in a significant performance gain; on the contrary, performance may decrease due to the cost of parallelization. Therefore, parallelism is not applied here.

Calculating the similarity of descriptors is a field for applying parallel computing. When it is necessary to compare a large number of descriptors or when comparing descriptors with large-dimensional vectors, parallel computing can significantly improve search performance.

In this work, to compare the list of found descriptors with the searched one, we suggest using parallelism at the thread level. The list of descriptors is divided into parts, and the comparison is performed for each part in a separate thread. The results from each thread are used to form the final list. This allows for the efficient use of workstation resources.

Using parallelism at the thread level to compare vectors is not very effective, as there are more specific approaches for this. In particular, hardware tools that are better suited for processing large amounts of similar data can be used, such as SIMD instructions of the central processing unit (CPU), such as SSE, AVX, AVX-512, etc., and graphics processing units (GPUs).

SIMD instructions efficiently utilize the computing power of a single CPU core by performing parallel operations on multiple data elements simultaneously. In contrast, a GPU has an architecture with a large number of simple cores and threads, allowing it to perform a large number of parallel computations. This can provide higher performance compared to CPUs, especially when processing large data arrays, although GPUs have higher overhead costs for initialization and data transfer.

To use SIMD in Java, you can use the Java Vector API, a low-level API for working with vector instructions [33]. There are also high-level libraries, such as ND4J, that can use SIMD or GPUs internally [34]. To perform calculations on GPUs in Java, JOCL (a wrapper over OpenCL) [35] is often used, as well as other libraries, such as JCuda, to access CUDA [36].

It is important to note that in order to process data on a GPU, it must first be transferred to video memory [37], which can become a bottleneck and significantly affect the overall performance of this approach.

This direction is extremely important in the context of MDC optimization, but it requires additional study and testing of a large number of software libraries and solutions. It also requires consideration of a large number of parameters and features of each processor, GPU, or driver version for them to work.

Therefore, this paper does not discuss in detail the improvement of search algorithm performance through the use of low-level hardware optimizations.

Experiment methodology

To verify the effectiveness of WSA, two experiments are conducted:

1) comparison of the effectiveness of sequential and parallel WSA variants;

ISSN 2522-9818 (print)
Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34) ISSN 2524-2296 (online)

2) comparison of the effectiveness of using MDC with sequential WSA with other CBIR search models.

The experiments test a real-life search scenario where the search engine finds a certain number of the most similar images (top-N), and the user determines which images are relevant to them.

In the experiments, WSA always performs a single search wave to ensure high search quality, as mentioned earlier.

All software implementations of the models presented in the experiments were developed using the Java 17 programming language and Spring Framework 6. PostgreSQL 15.2 is used to store MDC in the database. The results of the experiments were exported from the software solution in Excel table format.

All experiments were conducted on a MacBook Pro 2021: M1 Pro processor with ARM architecture, 10 cores with a frequency of up to 3.2 GHz, 16 GB of LPDDR5 SDRAM with a speed of up to 200 GB/s, 512 GB SSD, integrated GPU with 16 cores.

To conduct the experiment comparing the efficiency of the sequential and parallel versions of the algorithm, we used the MDC functionality, which allows generating descriptors and placing them in cells [32]. Artificial generation of descriptors allows creating vectors of different dimensions and the desired number of descriptors. This solves the problems of using real descriptors: the long process of searching for data sets to achieve the desired number of images and creating descriptors of different dimensions for them.

During the experiment, the time spent on the search is compared and the gain from the parallel algorithm is evaluated for different numbers of descriptors, vector dimensions, and number of threads.

The results provide an understanding of the conditions under which the use of a parallel algorithm provides advantages.

The second is an experiment in which the MDC model is compared with other search models for CBIR based on: multidimensional tree (KD-tree), hashing (LSH), graphs (HNSW), inverted index (IVFFlat), quantization and inverted multi-index (IMI), and vector database (pgvector). A comparison with exhaustive search is also made. For the experiment, basic models of each approach were selected because they demonstrate the general properties of a particular approach and are easier to implement or run than existing modifications. We create implementations of other models ourselves or use ready-made solutions if they are available and their configuration is simple.

This experiment is a large-scale test and comparison of the MDC model with other CBIR models in general, including testing the effectiveness of the presented search algorithm. It allows us to identify the strengths and weaknesses of the model and search algorithm compared to competitors and to predict the most favorable scenarios for using the model.

In all models, the metric for comparing descriptor vectors is Manhattan distance, unless otherwise specified.

To conduct this experiment, descriptors created for images from the following datasets are used:

1) 100,000 images from the COCO dataset [38] are used;

2) 750,000 images from Various tagged images [39] are used;

3) 150,000 images from Amazon bin images [40] are used.

This allows us to perform the experiment for 1 million descriptors. Of course, in real Big Data CBIR systems, the number of images and their corresponding descriptors can be significantly larger, but the results obtained allow us to compare the effectiveness of using different models in different conditions and scale the results as needed.

For this experiment, the main metrics are search time and recall. Recall determines the quality of the search. It is calculated using the following formula:

$$\mathrm{Re}\,call = \frac{TP}{TP + FN}, \tag{3}$$

where TP (true positive) is the number of correctly classified positive examples, and FN (false negative) is the number of positive examples that the model failed to detect.

Another key metric that is usually evaluated for CBIR models is precision, which is not evaluated due to the specifics of the experiment, which will be described below.

The first additional metric is the time it takes to build a search model. This parameter can be critically important if the search engine needs to ensure continuous operation and quickly adapt to updates in the storage content.

The second additional metric is labor intensity, which shows how many descriptors were compared during the search. It is determined when it is possible to obtain values from the model. It is provided for informational purposes.

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

Invariant Brightness Histogram descriptors were created for all images, as described in more detail in [32]. 100 images were randomly selected for the search. Two modifications were created for each image: a 180-degree rotation and a 2x reduction in scale. During the search, a positive result is finding a group of these three images among the top 10 and top 100 search results. For this reason, calculating the accuracy metric is inappropriate in this experiment, as the number of relevant results under the conditions of the experiment is small.

Each CBIR model has its own parameters that affect its structure and search efficiency. For each model, different combinations of parameters are experimentally tested to show the best result in terms of quality and search speed when using 1 million descriptors and the top 10 and top 100 search results. Testing the use of a larger number of found results is inappropriate due to the inefficiency of processing the obtained results. Determining the parameters of other models is not the subject of this work, so it will be briefly discussed in the next section. Detailed comparison results can be found on Google Drive [41]. The configurations of the models that showed the best results are compared with MDC.

Additionally, load testing is performed using the Postman utility [42]. 100 virtual users simultaneously searched for the top 100 most similar images for randomly selected descriptors from the repository for 1 minute, with a delay of 1-2 seconds between attempts. This allows us to determine the actual throughput and efficiency of the model in conditions close to real life, rather than simply testing them synthetically. During testing, the total number of requests made during this period, the throughput of the search engine, and the average response time are evaluated.

Parameters of alternative models participating in the experiment

For the multidimensional tree model, we implemented a KD-tree. The implementation is unbalanced, and the hierarchy is preserved as specified by the order in which elements are added to the tree. Each node contains one descriptor. The axis along which the space is divided at each level depends on the depth of that level, which is limited by the logarithm of the number of descriptors. An alternative branch is checked during the search if fewer than the required number of results are found or if the difference between the vectors is greater than the difference between the elements at the current level. Due to the last mentioned check, the selected vector similarity metric plays an important role. Therefore, for this model, the results for Manhattan and Quadratic Euclidean distances are given in the experiments. Other model parameters are not configured.

For hashing, we independently implemented the LSH model. Its main parameters are the number of tables t and the number of hash functions h. Each table has its own set of hash functions. The more hash functions and tables there are, the more accurate the results are and the smaller the search area is, but it takes more time to add elements to the table and search. It is important to determine a balanced value for these parameters. Combinations from the following parameter groups were tested: $t = (5,\ 10)$ and $h = (50,\ 75,\ 100)$. The following parameters were selected: $t = 10$, $h = 75$.

To test the vector database, we used an add-on to the PostgreSQL 15 database called Pgvector, which provides extensive capabilities for working with vectors within the database. In a simple usage scenario (exhaustive search), it does not require additional configuration and is ready to use immediately after installation in the DBMS.

For the reverse index, two IVFFlat implementations with k-means clustering are tested: one is implemented independently in RAM and uses the Manhattan metric during the search, the other is an implementation with Pgvector and uses the Euclidean metric during the search (Pgvector does not provide the ability to use the Manhattan metric for IVFFlat). For this type of model, the main parameter is the number of clusters. All descriptors are divided into clusters, and the search takes place only in one of the clusters or continues in the next cluster similar to the centroid. The following values were tested (10, 100, 100), and 100 was selected as the most effective.

For MDC, the main parameters are the number of measurements N and the number of intervals per measurement k. Their values change the number of cells in MDC and, accordingly, the number of descriptors in one cell. Combinations from the following parameter groups were predefined and tested: $N = 4$, $k = (18,\ 20)$, which provide about 10 descriptors in a cell. The parameters $N = 4$, $k = 20$ were selected.

The IMI model's own implementation is used to quantize the vector and the inverse multi-index. Its main parameters are the number of subspaces m and the number of clusters in each subspace k. These parameters change the number of combinations of

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

clusters from different subspaces, thereby changing the number of descriptors corresponding to each of them. Combinations from the following parameter groups were tested: $m = (2,\ 4,\ 8)$ and $k = (5,\ 10,\ 18,\ 20)$, and the parameters $m = 4$, $k = 10$ were selected.

The main parameters of HNSW are m – the maximum number of connections per layer, and efc – the size of the dynamic list of candidates for graph construction, which affect the quality of graph construction. Combinations from the following parameter groups were tested: $m = (8,\ 16)$, $efc = (32,\ 64,\ 128)$. The parameters $m = 8$, $efc = 64$ were selected.

The parameters selected for each model were chosen individually by experiment using a specific set of descriptors with a size of 1 million. All defined parameters or combinations of parameters showed the best results. For a different number of descriptors or descriptors from other data sets, these parameters may be different.

Experimental results and discussion

The first experiment compared the use of sequential and parallel WSA algorithms in MDC implementation in RAM.

The search for the top 100 artificially created descriptors with lengths of 32, 128, and 1024 was tested. The number of threads for the parallel implementation of the algorithm was 2, 4, and 8. The number of descriptors in MDC was equal to 1 and 10 million. The MDC parameters were $N = 4$, $k = 10$. This allowed the MDC to be filled evenly and, using 1 search wave, to obtain a specific number of descriptors for comparison with the searched one. For 1 million descriptors, this number is 6,000, and for 10 million, it is 60,000. The speed of comparing candidate descriptors with the searched one is the main operation that is performed either sequentially or in parallel for the found descriptors.

The results of comparing the application of WSA for sequential and parallel variants of the algorithm and 1 million descriptors are shown in Table 1, and for 10 million descriptors in Table 2. The graphical representation of the benefits of using different numbers of threads is shown in Figure 3 for 1 million descriptors and in Figure 4 for 10 million descriptors.

For 10 million descriptors, the length of the 1024 vector was not tested because there were insufficient resources on the computer where the experiment was conducted to store them in memory. The same applies to testing a larger number of low-dimensional descriptors, for example, 100 million descriptors with a dimension of 32.

Table1. *Search time for top 100 results in MDC with 1 million descriptors, seconds*

Count of the threads	Vector length		
	32	128	1024
1	0.00169	0.00238	0.00942
2	0.00168	0.00194	0.00550
4	0.00172	0.00198	0.00368
8	0.00160	0.00175	0.00293

Table 2. *Search time for top 100 results in MDC with 10 million descriptors, seconds*

Count of the threads	Vector length	
	32	128
1	0.02068	0.09020
2	0.01856	0.07378
4	0.00931	0.05805
8	0.00648	0.03163

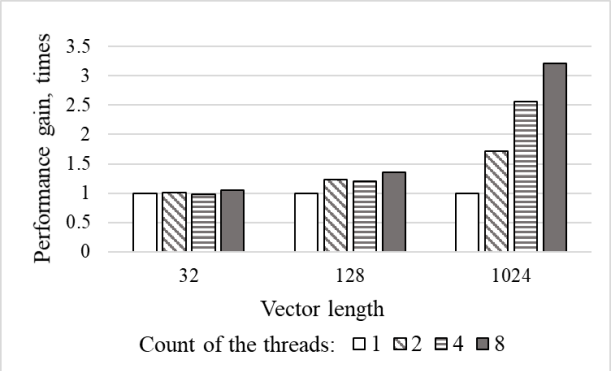


Fig. 3. Time gains when searching among 1 million descriptors using different numbers of threads for descriptors with vectors of different lengths

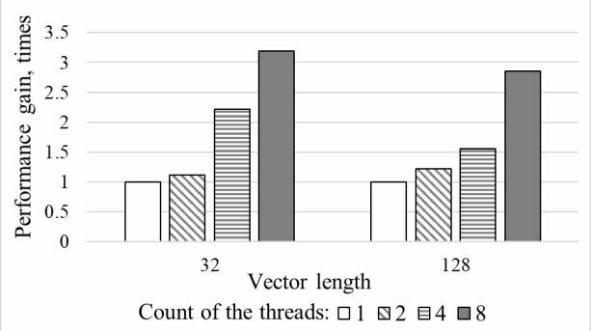


Fig. 4. Time gains in searching among 10 million descriptors when using different numbers of threads for descriptors with vectors of different lengths

The results show that the use of parallel computing when comparing a small number of descriptors (6,000)

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

and vectors of small dimensions (32 or 128) does not make sense, as the time savings are insignificant.

When using higher-dimensional descriptors (1024) or a larger number of descriptors (10 million), it is possible to achieve a gain in using parallel computing up to a 3-fold difference when using 8 threads. This is a significant acceleration, confirming the good adaptability of MDC to the use of parallel computing and the feasibility of its use in real search scenarios.

In the second experiment, MDC in two memory implementations: in a database – MDC (DB) and in RAM – MDC (RAM) with sequential WSA is compared with:

– full search in RAM – BF RAM, which is performed using parallelism at the thread level;

– exhaustive search in a vector database – BF (pgvector);

– a model based on Inverted Multi-Index – IMI;

– a model based on KD-Tree with Manhattan metric – KD-Tree (Mh) and Quadratic Euclidean metric – KD-Tree (Sq);

– a model based on Local-sensitive hashing – LSH;

– Hierarchical Navigable Small World (HNSW) model;

– Reverse index model implemented in Pgvector with Euclidean metric – IVFFlat (pgvector).

– a model based on a reverse index implemented in RAM – IVFFlat (RAM).

For each model, metrics of completeness, labor intensity, search time, and creation time are determined. A search is conducted for the top 10 and top 100 most similar descriptors among 1 million descriptors.

The results of the experiment for searching the top 10 results are shown in Table 3. The results of searching the top 100 results are shown in Table 4. These values are averages for searching 100 selected descriptors after 10 rounds of the experiment.

Table 3. *Average search results for the top 10 results for 100 selected images among 1 million descriptors*

Model	Completeness	Labor intensity	Search time, s	Creation time, s
BF (RAM)	0.97667	1 000 000.00	0.09533	–
BF (pgvector)	0.97667	1 000 000.00	0.06888	–
MDC (RAM)	**0.97000**	**3 881.230**	**0.00129**	**2.623**
MDC (DB)	**0.96000**	**3 505.960**	**0.07668**	**812.562**
IMI	0.87333	4 354.180	0.00268	342.653
KD-Tree (Sq)	0.85600	808.645	0.00043	1.762
KD-Tree (Mh)	0.97667	967 798.782	0.49160	1.827
LSH	0.96567	32 524.953	0.02039	26.085
HNSW	0.97667	–	0.00582	376.410
IVFFlat (RAM)	0.91667	15 289.790	0.00674	3265.941
IVFFlat (pgvector)	0.91600	–	0.00530	2.380

Table 4. *Average search results for the top 100 results for 100 selected images among 1 million descriptors*

Model	Completeness	Labor intensity	Search time, s	Creation time, s
BF (RAM)	0.99333	1 000 000.00	0.09940	–
BF (pgvector)	0.99333	1 000 000.00	0.07259	–
MDC (RAM)	**0.98000**	**3 881.230**	**0.00142**	**2.623**
MDC (DB)	**0.97000**	**3 505.960**	**0.07284**	**812.562**
IMI	0.89333	4 360.230	0.00256	342.653
KD-Tree (Sq)	0.93033	3 091.157	0.00182	1.762
KD-Tree (Mh)	0.99333	989 919.981	0.50460	1.827
LSH	0.97133	40 103.314	0.02567	26.085
HNSW	0.99000	–	0.00661	376.410
IVFFlat (RAM)	0.93667	15 289.790	0.00675	3265.941
IVFFlat (pgvector)	0.92500	–	0.00839	2.380

The results obtained remain consistent when searching for the top 10 and top 100 results, so the results from Tables 3 and 4 are analyzed together. They show that even when using a full search, the completeness metric is not always 100%. This is due to the characteristics of the specific descriptor type selected.

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

The maximum possible completeness result, apart from exhaustive search, was also shown by the KD-Tree model with the Manhattan metric, but the labor intensity of such a search is close to exhaustive search, and the same applies to the search time. This indicates that it is not advisable to use this model with this descriptor similarity metric. The HNSW model also showed the maximum completeness result, demonstrating an average search and model construction time. Next, MDC and LSH showed similar results in terms of completeness. However, MDC (RAM) showed much better performance in terms of search time, model construction time, and labor intensity. The implementation of MDC (DB) showed worse results than BF (pgvector) in terms of both search quality and time, which raises doubts about the advisability of its use.

Other models showed worse search results. IMI showed good results in terms of search time, but poor completeness. KD-Tree with Quadratic Euclidean Metric showed the best results in terms of search time and labor intensity, but the lowest completeness. IVFFlat implementations showed low completeness results because the search was performed in only one cluster, while the search time indicator is high.

In terms of model construction speed, the leaders are MDC (RAM), both KD-Tree and IVFFlat (pgvector) variants. In MDC, this is due to the peculiarities of the structure and the construction process, KD-Tree works fast due to the low dimension of vectors, and IVFFlat (pgvector) has implemented optimizations for clustering, because IVFFlat (RAM) without any optimizations showed the worst result. In MDC (DB), the result is average due to the cost of communication with the database.

After analyzing the results, the conclusion is as follows: exhaustive search demonstrates the best search quality, but requires a lot of time to search; MDC shows some of the best performance in terms of time and quality when placed in RAM; IMI with a small search output is not always effective; KD-Tree is ineffective in terms of either search quality or speed; LSH demonstrates balanced average results, HNSW demonstrates the best results in terms of completeness, but the search and model building time is longer than in MDC, IVFFlat demonstrates average results in terms of completeness and search time, but the structure building time strongly depends on the implementation. Based on the set of indicators, it can be concluded that under these conditions, MDC (RAM) demonstrates the best results among the models considered. However, the use of MDC (DB) is not promising due to results that are worse than those of BF (pgvector).

The load testing results are shown in Table 5. An example of the load testing results for MDC (RAM) is shown in Figure 5. The graph shows that with a constant load of 100 users, the search engine provides stable response times and consistently processes a large number of queries without any noticeable drop in performance.

Table 5. *Load testing results*
(100 users made requests to search for the top 100 results within 1 minute at intervals of 1–2 seconds)

Model	Total number of requests sent	Throughput (requests/second)	Average response time (ms)
BF (RAM)	302	4.45	9 813
BF (pgvector)	2 248	32.81	1 763
MDC (RAM)	**6 377**	**94.13**	**6**
MDC (DB)	**1 495**	**21.92**	**3 018**
IMI	6 403	95.21	4
KD-tree (Sq)	6 418	95.52	6
KD-tree (Mh)	101	1.48	10 758
LSH	5 765	84.76	60
HNSW	6 209	92.04	16
IVFFlat (RAM)	6 386	94.48	15
IVFFlat (pgvector)	6 301	92.19	15

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)
Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

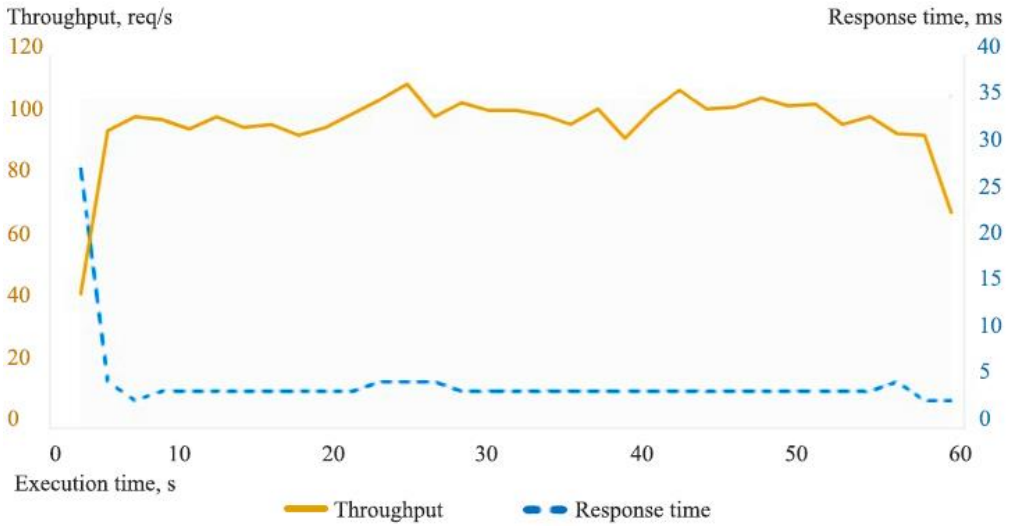


Fig. 5. Load testing results for the MDC (RAM) model

The use of exhaustive search is expected to be inefficient, as is searching in a multidimensional tree with a Manhattan metric, which approximates exhaustive search.

MDC (DB) shows weak results, slightly worse than exhaustive search in a vector database, so the futility of this MDC modification is confirmed by the results of this study.

LSH showed slightly worse results than MDC (RAM), IMI, KD-tree (Sq), HNSW, and IVFFlat in both modifications. These models show similar results in this experiment, which indicates approximately the same throughput of the models. Therefore, the choice between them should be made based on the results in Tables 3 and 4, which include completeness as an indicator of search quality.

Conclusions and prospects for further development

This paper presents a method for searching for similar images based on image descriptor vectors stored in the MDC model, which functions on the basis of the MDC structure and data stored at the descriptor vector processing stage. The algorithm that implements the method is called the Wave-Search Algorithm (WSA). It performs a wave search relative to the cell identified for the desired descriptor. It has adapted versions for placing MDC in RAM and databases.

The algorithm has a parallel version that uses parallelism at the thread level and efficiently utilizes workstation resources. It speeds up the search by up to 3 times when using a large number of descriptors (more than 10 million) and/or descriptors with vectors of considerable size (from 1024).

Compared to other CBIR models, MDC with WSA placed in RAM demonstrates very competitive results. Together with the LSH and HNSW models, it shows a search accuracy of about 97-98% according to the completeness metric. And one of the best search speeds among the tested models, spending an average of 0.0012 seconds when searching for the top 10 and top 100 results among 1 million descriptors. It also has one of the best model construction times, at around 2.5 seconds.

The MDC version stored in a database demonstrates worse results than when using a vector database (pgvector extension for PostgreSQL) in terms of both search time and quality. Therefore, its further development is not promising.

Load testing shows that the model works effectively under load and does not reduce performance, which confirms its suitability for use in real search engines, including in Big Data environments.

The use of MDC with the presented search algorithm is recommended in situations where search quality and speed are important, as well as the time it takes to create or update the content of the repository. In other words, the scope of application is quite broad.

A further direction of research is the creation of versions of the search algorithm that use SIMD operations and a graphics processor to perform parallel vector comparisons. And conducting experiments using a larger number of descriptors and descriptors of higher dimensioni.

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

References

1. American Society for Indexing, "*History of Information Retrieval*", available at: https://asindexing.org/about-indexing/history-of-information-retrieval/ (last accessed 27.04.2025).
2. Maña, N., Babiera, J., Bayloces, K., Palmer, X.-L., Potter, L., Lavilles, R., Velasco, L. (2024), "Information Retrieval Systems: A Methodological Review", *Proceedings of the Future Technologies Conference (FTC) 2024, Volume 3*, Springer Nature Switzerland, Cham, P. 572–591. DOI: 10.1007/978-3-031-73125-9_36
3. Rostami, C., Hosseini, E., Saberi, M. (2021), "Information-seeking behavior in the digital age: use by faculty members of the internet, scientific databases and social networks", *Information Discovery and Delivery*, Vol. 50, No. 1, P. 87–98. DOI: 10.1108/idd-02-2020-0014
4. Jain, R. (2023). "A Comparative Study of Breadth First Search and Depth First Search Algorithms in Solving the Water Jug Problem on Google Colab", *SSRN Electronic Journal*. DOI: 10.2139/ssrn.4402567
5. Li, X., Yang, J., Ma, J. (2021), "Recent developments of content-based image retrieval (CBIR)", *Neurocomputing*, Vol. 452, P. 675–689. DOI: 10.1016/j.neucom.2020.07.139
6. Alsmadi, M. (2020), "Content-Based Image Retrieval Using Color, Shape and Texture Descriptors and Features", *Arabian Journal for Science and Engineering*, Vol. 45, No. 4, P. 3317–3330. DOI: 10.1007/s13369-020-04384-y.
7. Zhang, Q., Canosa, R. (2014), "A comparison of histogram distance metrics for content-based image retrieval", *Imaging and Multimedia Analytics in a Web and Mobile World 2014*, SPIE, Vol. 9027. DOI: 10.1117/12.2042359
8. Li, M., Wang, H., Dai, H., Li, M., Chai, C., Gu, R., Chen, F., Chen, Z., Li, S., Liu, Q., G. Chen. (2024), "A Survey of Multi-Dimensional Indexes: Past and Future Trends", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 36, P. 3635–3655. DOI: 10.1109/tkde.2024.3364183
9. Samoladas, D., Karras, C., Karras, A., Theodorakopoulos, L., Sioutas S. (2022), "Tree Data Structures and Efficient Indexing Techniques for Big Data Management: A Comprehensive Study", *Proceedings of the 26th Pan-Hellenic Conference on Informatics*, ACM, New York, USA, P. 123–132. DOI: 10.1145/3575879.3575977
10. Rakotondrasoa, H.M., Bucher, M., Sinayskiy, I. (2023), "Quantitative Comparison of Nearest Neighbor Search Algorithms", *arXiv*. DOI: 10.48550/arXiv.2307.05235
11. Liu, Q., Li, M., Zeng, Y., Shen, Y., Chen, L. (2025), "How good are multi-dimensional learned indexes? An experimental survey", *The VLDB Journal*, Vol. 34. DOI: 10.1007/s00778-024-00893-6
12. Li, D., Esquivel, J. (2025), "Trust-Aware Hybrid Collaborative Recommendation with Locality-Sensitive Hashing", *Tsinghua Science and Technology*, Vol. 30, No. 4, P. 1421–1434. DOI: 10.26599/tst.2023.9010096
13. Weng, Z., Zhu, Y., Lan, Y., Huang. L.-K. (2019), "A fast online spherical hashing method based on data sampling for large scale image retrieval", *Neurocomputing*, Vol. 364, P. 209–218. DOI: 10.1016/j.neucom.2019.06.053
14. Ryali, C., Hopfield, J., Grinberg, L., Krotov, D. (2020), "Bio-Inspired Hashing for Unsupervised Similarity Search", *arXiv*. DOI: 10.48550/arXiv.2001.04907
15. Jiang, X., Hu, F. (2024), "Multi-scale Adaptive Feature Fusion Hashing for Image Retrieval", *Arabian Journal for Science and Engineering*. DOI: 10.1007/s13369-024-09627-w
16. Liu, R., Zhao, J., Chu, X., Liang, Y., Zhou, W., He, J. (2023), "Can LSH (locality-sensitive hashing) be replaced by neural network?", *Soft Computing*, Vol. 28, P. 1041–1053. DOI: 10.1007/s00500-023-09402-3
17. Malkov, Y., Yashunin D. (2020), "Efficient and Robust Approximate Nearest Neighbor Search Using Hierarchical Navigable Small World Graphs", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 42, No. 4, P. 824–836. DOI: 10.1109/tpami.2018.2889473
18. Dong, W., Moses, C., Li, K. (2011), "Efficient k-nearest neighbor graph construction for generic similarity measures", *Proceedings of the 20th International Conference on World Wide Web*, ACM, New York, USA. DOI: 10.1145/1963405.1963487
19. Weng, S., Fan, Z., Gou, J. (2024), "A fast DBSCAN algorithm using a bi-directional HNSW index structure for big data", *International Journal of Machine Learning and Cybernetics*, Vol. 15, No. 8, P. 3471–3494. DOI: 10.1007/s13042-024-02104-8
20. Zhibing, H. (2024), "Quick and Efficient Large Scale Approximate Nearest Neighbor Search on High-Dimensional Data", *2024 21st International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*, IEEE, P. 1–5. DOI: 10.1109/iccwamtip64812.2024.10873693
21. Zhao, J., Pierre Both, J., Konstantinidis, K. (2024), "Approximate nearest neighbor graph provides fast and efficient embedding with applications for large-scale biological data", *NAR Genomics and Bioinformatics*, Vol. 6, No. 4. DOI: 10.1093/nargab/lqae172.
22. Yousaf, M., Shakoor Khan, M., Ullah, S. (2024), "An Extended-Isomap for high-dimensional data accuracy and efficiency: a comprehensive survey", *Multimedia Tools and Applications*, Vol. 83, No. 38, P. 85523–85574. DOI: 10.1007/s11042-024-19917-y

23. Liu, Y., Pan, Z., Wang, L., Wang Y. (2022), "A new fast inverted file-based algorithm for approximate nearest neighbor search without accuracy reduction", *Information Sciences*, Vol. 608, P. 613–629. DOI: 10.1016/j.ins.2022.06.086
24. Bazdyrev, A. (2023), Semi-supervised inverted file index approach for approximate nearest neighbor search. *System Research and Information Technologies*, No. 4, P. 69–75. DOI: 10.20535/srit.2308-8893.2023.4.05
25. Matsui, Y., Hinami, R., Satoh, S. (2018), "Reconfigurable Inverted Index", *Proceedings of the 26th ACM International Conference on Multimedia*, ACM, New York, USA, P. 1715–1723. DOI: 10.1145/3240508.3240630
26. Jégou, H., Douze, M., Schmid, C. (2011), "Product Quantization for Nearest Neighbor Search", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 33, No. 1, P. 117–128. DOI: 10.1109/tpami.2010.57
27. Babenko, A., Lempitsky, V. (2012), "The inverted multi-index", *2012 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, P. 3069–3076. DOI: 10.1109/cvpr.2012.6248038
28. Qiu, Z., Liu, J., Chen, Y., King, I. (2024), "HiHPQ: Hierarchical Hyperbolic Product Quantization for Unsupervised Image Retrieval", *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 38, No. 5, P. 4614–4622. DOI: 10.1609/aaai.v38i5.28261
29. Gu, L., Liu, J., Liu, X., Wan, W., Sun, J. (2024), "Entropy-Optimized Deep Weighted Product Quantization for Image Retrieval", *IEEE Transactions on Image Processing*, Vol. 33, P. 1162–1174. DOI: 10.1109/tip.2024.3359066
30. Jamalifard, M., Andreu-Perez, J., Hagras, H., López, L. (2024), "Fuzzy Norm-Explicit Product Quantization for Recommender Systems", *IEEE Transactions on Fuzzy Systems*, Vol. 32, No. 5, P. 2987–2998. DOI: 10.1109/tfuzz.2024.3365722
31. pgvector, GitHub. "*pgvector: Open-source vector similarity search for Postgres*", available at: https://github.com/pgvector/pgvector (last accessed 27.04.2025).
32. Danylenko, S., Smelyakov, S. (2025), "Development of a Multidimensional Data Model for Efficient Content-based Image Retrieval in Big Data Storage", *Radioelectronic and Computer Systems*, Vol. 2025, No. 1, P. 137–152. DOI: https://doi.org/10.32620/reks.2025.1.10
33. Sandoz, P. "*JEP 448: Vector API (Sixth Incubator)*", available at: https://openjdk.org/jeps/448 (last accessed 01.05.2025).
34. Deeplearning4j, "*Deeplearning4j Suite Overview*", available at: https://deeplearning4j.konduit.ai/ (last accessed 01.05.2025).
35. jocl.org, "*Java Bindings for OpenCL*", available at: http://www.jocl.org/ (last accessed 01.05.2025).
36. jcuda.org, "*JCuda*", available at: http://www.jcuda.org/jcuda/JCuda.html (last accessed 01.05.2025).
37. Nevliudov, I., Yevsieiev, V., Maksymova, S., Gopejenko, V., Kosenko, V. (2025), "Development of mathematical support for adaptive control for the intelligent gripper of the collaborative robot manipulator", *Advanced Information Systems*, Vol. 9, No. 3, P. 57–65. DOI: https://doi.org/10.20998/2522-9052.2025.3.07
38. COCO, "*Common Objects in Context*", available at: https://cocodataset.org/ (last accessed 02.05.2025).
39. Greg, "*Various Tagged Images*", available at: https://www.kaggle.com/datasets/greg115/various-tagged-images (last accessed 02.05.2025).
40. Hyun, W. "*Amazon Bin Image Dataset (536,434 images, 224×224)*", available at: https://www.kaggle.com/datasets/williamhyun/amazon-bin-image-dataset-536434-images-224x224 (last accessed 02.05.2025).
41. Danylenko, S. "*MDC-2025-WSA*", Available at: https://drive.google.com/drive/folders/1LNgV8MzNkhXC7elWOQemvFFkBddW_wPM?usp=sharing/ (last accessed 02.05.2025).
42. Postman API Platform, "*Postman: The World's Leading API Platform*", available at: https://www.postman.com/ (last accessed 02.05.2025).

Received (Надійшла) 06.06.2025
Accepted for publication (Прийнята до друку) 30.11.2025
Publication date (Дата публікації) 28.12.2025

Відомості про авторів / About the Authors

Danylenko Stanislav – Kharkiv National University of Radio Electronics, PhD Student, Software Engineering Department, Kharkiv, Ukraine; e-mail: stanislav.danylenko@nure.ua, ORCID ID: https://orcid.org/0000-0002-8142-3018; Scopus ID: https://www.scopus.com/authid/detail.uri?authorId=57816229800

Smelyakov Kyrylo – Doctor of Sciences (Engineering), Professor, Kharkiv National University of Radio Electronics, Head of the Software Engineering Department, Kharkiv, Ukraine; e-mail: kyrylo.smelyakov@nure.ua, ORCID ID: https://orcid.org/0000-0001-9938-5489; Scopus ID: https://www.scopus.com/authid/detail.uri?authorId=57203149663

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

Даниленко Станіслав Дмитрович – Харківський національний університет радіоелектроніки, аспірант, кафедра програмної інженерії, Харків, Україна.

Смеляков Кирило Сергійович – доктор технічних наук, професор, Харківський національний університет радіоелектроніки, завідувач кафедри програмної інженерії, Харків, Україна.

МЕТОД ПОШУКУ ЗОБРАЖЕНЬ НА ОСНОВІ ВМІСТУ В БАГАТОВИМІРНІЙ МОДЕЛІ ДАНИХ У МАСШТАБІ ВЕЛИКИХ ДАНИХ

Предметом роботи є метод і алгоритми пошуку зображень за вмістом у моделі багатовимірного куба (MDC). **Мета дослідження** – розробити метод пошуку зображень у моделі MDC на основі векторів дескрипторів зображень та алгоритму, що реалізує цей метод послідовної та паралельної версії. **Завдання статті** передбачають: формування вимог до методу пошуку; аналіз структури MDC та визначення підходу до методу пошуку; розроблення методу й алгоритмів пошуку за умови розміщення моделі в оперативній пам'яті та в реляційній базі даних; упровадження паралельних обчислень в алгоритмі пошуку; аналіз альтернативних моделей, основаних на багатовимірних деревах, графах, гешуванні, зворотному індексі, квантуванні та зворотному мультиіндексі; розроблення метрик і проведення експериментів для порівняння ефективності моделі MDC з альтернативними моделями пошуку. **Методи дослідження**: аналіз можливостей застосування паралельних обчислень на рівні потоків і оптимізацій за допомогою апаратного забезпечення; розроблення методу пошуку, послідовної та паралельної версій алгоритму його реалізації; аналіз сучасних моделей пошуку, зокрема *KD-tree, Locality-sensitive hashing, Hierarchical Navigable Small World, Inverted File with Flat Compression, Inverted Multi-Index*; експерименти з дескрипторами зображень з мережі Інтернет та штучно генерованими для визначення ефективності паралельної версії алгоритму та порівняння результативності пошуку зі згаданими моделями за метриками повноти, часу пошуку й часу створення; проведення експерименту з навантажувального тестування для оцінювання пропускної здатності моделі. **Результати дослідження**. Розроблено новий метод пошуку й алгоритм *Wave-Search*. Його паралельна версія пришвидшує виконання пошуку майже втричі; під час пошуку топ-10 і топ-100 результатів у сховищі з 1 млн дескрипторів MDC демонструє найкращі сумарні результати за метриками серед розглянутих моделей, має гарну продуктивність під навантаженням. **Висновки**: запропоновано метод пошуку й для його реалізації алгоритм *Wave-Search,* що ефективно використовує структуру MDC; модель багатовимірного куба перевищує показники ефективності альтернативних моделей пошуку, демонструє стабільність під навантаженням і має потенціал для подальшого розвитку із застосуванням апаратного прискорення.

Ключові слова: пошук за подібністю; великі дані; алгоритми пошуку; структури даних; пошук зображень на основі вмісту; паралельні обчислення; високопродуктивні обчислення; ефективність пошуку; оптимізація алгоритму.

Бібліографічні описи / Bibliographic descriptions

Даниленко С. Д., Смеляков К. С. Метод пошуку зображень на основі вмісту в багатовимірній моделі даних у масштабі великих даних. *Сучасний стан наукових досліджень та технологій в промисловості*. 2025. № 4 (34). С. 18–31. DOI: https://doi.org/10.30837/2522-9818.2025.4.018

Danylenko, S., Smelyakov, K., (2025), "Content-based image retrieval method in a multidimensional data model at big data scale", *Innovative Technologies and Scientific Solutions for Industries*, No. 4 (34), P. 18–31. DOI: https://doi.org/10.30837/2522-9818.2025.4.018

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

UDC 629.735.3 DOI: https://doi.org/10.30837/2522-9818.2025.4.032

M. YENA, O. POHUDINA

INTEGRATED SIMULATION MODEL OF Swarm CONTROL AND ADAPTIVE ROUTEING OF UAVS IN A CHANGING AIR ENVIRONMENT

Subject matter: the processes of swarm control and adaptive routing of unmanned aerial vehicles (UAVs) in complex and dynamically changing air conditions using adaptive algorithms. **Goal**: to develop an integrated simulation model that combines swarm control methods, adaptive PID control and adaptive routing algorithms to ensure the safety, optimality and efficiency of UAV fleet movement in conditions of a changing air environment. **Tasks**: to analyze existing approaches to swarm control and adaptive routing of UAVs; to develop a mathematical model of an integrated system that takes into account the specifics of interaction between UAVs, collision avoidance and dynamic changes in the air environment; to create a swarm control algorithm based on adaptive PID regulation of UAV movement parameters; to develop and implement an adaptive routing algorithm that responds to changes in traffic, weather conditions and other airspace factors; to implement the integrated model in a simulation environment and test its effectiveness; to conduct a comparative analysis of the efficiency of UAV operation with and without the developed algorithms. **Methods**: use of adaptive PID control methods for dynamic regulation of UAV movement trajectories and ensuring flight accuracy and stability; application of swarm control algorithms (boids-type methods) for synchronization of movement and collision avoidance in UAV groups; nonlinear optimization of routes taking into account dynamically changing conditions, which allows minimizing collision risks, energy consumption and flight time; construction of a graph-theoretic model of airspace for effective route planning and situation forecasting; creation of digital twins of the air environment for conducting simulation experiments. **Results**: an integrated simulation model of swarm control and adaptive routing of UAVs was developed, which takes into account air environment variables; adaptive PID control and swarm control algorithms ensured a reduction in the average positioning error and collision avoidance of UAVs; According to the results of simulation experiments, an increase in the reward of agents by ≈50%, an increase in the successful completion of episodes by ≈50%, and a reduction in agent errors on the way to the goal by ≈10%. **Conclusions**: created integrated model allows for effective management of UAV flotillas in conditions of a changing air environment, significantly increasing the safety and optimality of routes; the use of adaptive algorithms and graph-theoretic models provides high forecasting accuracy and risk minimization; the results of the study confirm the prospects for implementing the developed algorithms for UAV control in urban and regional conditions.

Keywords: unmanned aerial vehicles; swarm control; adaptive routing; PID control; real-time data processing; route optimization.

Introduction

The rapid development of unmanned aerial vehicles (UAVs) and the growth in their use in civil and specialized areas necessitate the creation of effective air traffic management methods. This issue is particularly relevant in urban and complex dynamic environments, where the number of UAVs is rapidly increasing and airspace is characterized by high density, changing weather conditions, and limited resources. Effective management of UAV fleets requires new approaches that take into account group behavior, dynamic response to changes in the environment, and the ability to self-learn [1] (hereinafter, the term "agent" refers to a single unmanned aerial vehicle in the model; these terms are used synonymously in this paper).

There are currently numerous approaches to optimizing UAV routes and coordinating their movement: the use of theoretical graph models, nonlinear optimization methods, evolutionary algorithms, collision avoidance strategies, etc. [2]. However, most of them are either focused on individual trajectory planning or do not take into account the peculiarities of collective fleet dynamics and adaptation to changes in airspace in real time. This significantly limits their application for automated control systems of a large number of UAVs in real operating conditions.

To solve the outlined problems, the paper proposes an integrated simulation model that combines swarm intelligence, adaptive PID control, and dynamic routing algorithms. The developed model makes it possible to simulate the complex collective behavior of a group of UAVs, ensure coordinated avoidance of obstacles and collisions, and dynamically update routes in response to changes in air and weather conditions. The proposed approach is universal for various types of air missions,

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

from monitoring and search and rescue operations to logistics in urban environments.

Research objectives and tasks

The aim of the work is to develop an integrated simulation model of swarm control and adaptive routing for a group of unmanned aerial vehicles, which ensures safe, optimal, and efficient movement of fleets in a changing air environment.

The main criteria for the effectiveness of the model are:

– minimization of the average distance to the target for each agent (UAV);

– avoidance of collisions;

– ability to dynamically adapt to changes in airspace;

– stability and control even in complex and turbulent conditions.

To achieve the goal, the following tasks must be performed:

– analyze modern approaches to swarm control and UAV routing;

– formalize a mathematical model of group movement, taking into account the interaction of agents (UAVs) and dynamic changes in the environment;

– develop and investigate algorithms for collective behavior and adaptive trajectory control;

– implement and test the model in a simulation environment, compare its effectiveness with basic approaches.

Analysis of recent research and publications

Over the past decade, the issue of UAV Traffic Management (UTM) has been attracting increasing attention from the scientific community due to the prospects for the widespread use of unmanned aerial vehicles in urban and regional airspace. Researchers are focusing on developing models for decentralized fleet management [3], route optimization, collision avoidance, and the integration of artificial intelligence into automated systems [4].

Among the most cited works in recent years are those on modeling collective behavior (*swarm intelligence*) for UAVs, which implement the *Boids* approach and its modifications [5, 6]. Such models make it possible to implement coordinated control of the movement of a group of devices, adapt to dynamic obstacles, ensure self-coordination, and avoid collisions.

Special attention is paid to the use of adaptive controllers in UAV trajectory control tasks. Modern research actively implements PID and LQR controllers [7], which are capable of automatically adapting their parameters to changes in the dynamics of the environment [8]. However, most of the work focuses on controlling single UAVs or small groups, while often ignoring the effects of mass interaction and a large number of agents in real airspace.

In the field of dynamic routing, graph models of airspace have become widespread, allowing for effective trajectory planning based on space topology, traffic, and weather conditions [9]. The use of nonlinear optimization methods, evolutionary algorithms (in particular, genetic algorithms, particle swarm algorithms [10], etc.) and hybrid approaches helps to find optimal routes in multifactorial scenarios.

Recently, there have been studies on the use of reinforcement learning (RL) for routing and distributed control of UAVs [11, 12]. RL approaches allow agents to be trained to interact effectively with an unknown environment, but the issues of stability and interpretability of such systems remain open.

A review of the literature shows that, despite significant progress in the development of theory and tools for UAV traffic control, there is still a lack of comprehensive models that combine adaptive routing, swarm control, and the ability to integrate real-time environmental variables. Accordingly, the development and research of integrated simulation models capable of combining these components is a relevant scientific task.

Mathematical model of swarm control and adaptive routing systems

Formalization of the task of controlling groups of UAVs

This section presents a formalization of the problem of collective control of a group of UAVs in complex airspace. The space in which the devices move is modeled as a directed graph $G = (V, E)$, where the set of nodes V corresponds to key geographical or control points (e.g., launch zones, target positions, or obstacle avoidance points), and the edges E define the permissible trajectories of movement between these points [13]. This approach makes it possible to naturally take into account the structure of the environment, the

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

presence of restrictions, and variable movement conditions. Each agent in this model, understood as a separate UAV, is defined by three basic parameters: spatial position, velocity vector, and a set of adaptive PID controller coefficients. Formally, the state of the *i*-th agent at time step *t* is defined as:

– $x_i(t)$ – current position in the corresponding coordinate system;

– $v_i(t)$ – speed of movement;

– $p_i(t) = \left[K_{p,i}, K_{i,i}, K_{d,i}\right]$ – vector of PID parameters, which are adaptively adjusted for each UAV taking into account changes in the environment.

The goal of control is to minimize the average distance to target points for all agents (UAVs) while avoiding collisions and ensuring efficient routing in complex and dynamic conditions. This approach allows not only to coordinate group movement, but also to take into account the individual characteristics of the trajectories and behavior of each device.

The dynamics of agent movement are determined by a system of differential equations that describe changes in position and velocity, taking into account the actions of control signals and random disturbances (e.g., noise or turbulence). This makes it possible to model the behavior of UAVs in realistic conditions:

$$\begin{aligned} x_i(t+1) &= x_i(t) + u_{i,x}(t)dt + \eta_{i,x}(t), \\ y_i(t+1) &= y_i(t) + u_{i,y}(t)dt + \eta_{i,y}(t), \end{aligned} \quad (1)$$

where $u_{i,x}$, $u_{i,y}$ – resulting control actions (determined by the control system); dt – time discretization step; $\eta_{i,*}$ – random component (noise, turbulence).

Swarm Control

Collective behavior is achieved using *swarm* control, which is based on a modified *Boids* algorithm. It takes into account three main components:

– ***cohesion*** – inclination towards the center of mass of neighbors;

– ***separation*** – avoidance of collisions when approaching other agents;

– ***alignment*** – aligning the direction of movement with the nearest neighbors.

The summary control action of the *swarm* component is determined as follows:

$$u_{swarm}(t) = \alpha alignment_i \times \beta cohesion_i \times \gamma separation_i, \quad (2)$$

where α, β, γ – are selected or optimized during simulations, ensuring a balance between group cohesion and individual safety.

Adaptive PID control

To accurately follow the specified trajectories, each UAV is equipped with an adaptive [14] PID controller, which takes into account not only the current error, but also its accumulation and rate of change. The control action of the PID is described by the equation:

$$u(t) = K_p(t)e(t) + K_i(t)\int e(t)dt + K_d(t)\frac{de(t)}{dt}, \quad (3)$$

wher $e(t)$ – the current distance between the agent and the specified target, and the derivative $de(t)/dt$ reproduces the change in this distance over time; $K_p(t)$ – proportional coefficient of the PID controller at time *t*, responsible for responding to the current error; $K_i(t)$ – integral coefficient, compensates for systematic error (accumulated error over time); $K_d(t)$ – differential coefficient, responsible for anticipating and responding to the rate of error change.

These parameters can be adapted, in particular, using *reinforcement learning*, which helps to increase resistance to external influences, turbulence, or changes in the environment.

The proposed formalization makes it possible to model the behavior of a group of UAVs in both calm and changing airspace, ensuring the flexibility, scalability, and adaptability of the control system for solving complex tasks in modern aeronautics.

Routing model and route updates

In a dynamic airspace with variable factors such as traffic, weather conditions, and the emergence of danger zones, it is critical for the system to be able to quickly update UAV routes. To do this, a graph model is used, in which the vertices represent key points in the airspace and the edges represent possible trajectories. At each time step, the optimal route to the target is determined for each agent (UAV) based on current information about the state of the environment. Route optimization consists of finding a path for which the total "cost" of travel is minimal:

$$\pi_i^* \arg\min_{\pi \in P}\left(\sum e \in \pi \omega_e(t)\right), \quad (4)$$

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

where $w_e(t)$ – current "cost" (delay, risk) of the edge e; π – possible route in the set of all permissible routes P.

The "cost" of an edge is determined comprehensively: traffic levels in the area, weather threats, flight restrictions, or *no-fly* zones are all taken into account. With each change in the environment (e.g., the appearance of a new obstacle or a change in weather conditions), the routing model recalculates the optimal path for each agent in real time.

This approach ensures the flexibility of the system, allowing agents to quickly adapt to changing conditions and minimize risks during missions.

Reward function and performance criteria

One of the key elements of the control system is a properly designed reward function that guides agent learning and influences their behavior strategy [15, 16]. The main goal of the reward function is to encourage the agent to reduce the average positional error, avoid collisions [17], effectively achieve goals, and minimize resource expenditure on the route.

Within the current approach, the reward at each step of the simulation is determined by the following formula:

$$r_t = w_\delta(\varepsilon_{t-1} - \varepsilon_t) - w_{step} + r_{success} \times I[\varepsilon_t - \varepsilon_{th}] + r_{collision} I[condition], \quad (5)$$

where ε_t – mean average distance to target (MAE) per step t; w_δ, w_{step} – weight coefficients for progress and penalties per step; $r_{success}$ – additional reward for achieving the goal; $I[condition]$ – indicator equal to 1 if the condition is met, 0 otherwise.

The parameters of the reward function were selected experimentally to achieve a balance between speed and safety. In particular, an increase of w_δ stimulates rapid convergence with the target, but excessively high values can increase the risk of collisions. The penalty w_{step} limits the time to complete the task, while the bonus $r_{success}$ motivates the agent to complete the mission rather than move in endless loops near the target. Typical values of the coefficients used in the experiments: $w_\delta = 400.0$; $w_{step} = 0.2$; $r_{success} = 500.0$.

Analysis of the simulation results showed that this reward function setting not only effectively reduces MAE, but also ensures a high *success rate* (percentage of goals achieved) at an acceptable level of safety.

Next, we will look how the described models and criteria are integrated into a single management system and what the architecture of the simulation model looks like.

Adaptive *Swarm*+PID controller: algorithm

To ensure both coordinated collective behavior and individual adaptation of each UAV in a variable airspace, the proposed model uses a combined controller. It combines the advantages of *swarm* approaches with classic PID control and also includes reinforcement learning mechanisms.

At each step of the simulation, a total control signal is formed for each agent *i*:

$$u_i(t) = u_{swarm,i}(t) + u_{pid,i}(t), \quad (6)$$

where $u_{swarm,i}(t)$ – collective behaviour vector (*alignment, cohesion, separation*);

$u_{pid,i}(t) = K_P e_i(t) + K_i \sum_{r=0}^{t} e_i(r) + K_d \Delta e_i(t)$ – PID-regulation with dynamic adaptation of coefficients.

The controller's adaptability is ensured by dynamic adjustment of PID parameters based on feedback, which allows the system to respond flexibly to changes in turbulence, the appearance of new obstacles, and other external influences.

The logic of the combined controller (algorithm)

The sequence of actions for each simulation step involves several stages.

1. **Obtaining the state.** The RL agent analyzes the current state of the environment (*St*), which contains the positions, speeds, and locations of targets, as well as information about obstacles.

2. **Generation of strategic action.** The RL agent determines the optimal actions for the swarm module, i.e., generates signals for the collective behavior of agents.

3. **Collective coordination.** The swarm module coordinates the interaction between agents, ensures collision avoidance, and maintains group integrity.

4. **Individual adjustment.** For each agent, a PID correction is calculated based on the current position error, after which a total action is formed.

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

5. **Status and reward updates.** The system updates the position of agents in space, calculates the reward function, and proceeds to the next step of the simulation.

This architecture ensures high flexibility and adaptability of the control system, allowing the UAV to operate effectively even in complex and dynamic conditions.

The next important stage of the research was the design and implementation of the simulation model architecture, which combines all the components considered into a single system.

Simulation model architecture

To implement and test the proposed control algorithms, a modular simulation architecture was created that combines all key components of the system into a single integrated environment. This approach provides flexibility in conducting research, facilitates scaling, and allows for in-depth analysis of the interaction of various control elements in a variety of experimental scenarios.

The simulation model contains several specialized modules, each of which plays a specific role in ensuring the effective operation of the entire system. In particular, the **swarm control module** (*Swarm Controller*) is responsible for coordinating the collective behavior of a group of UAVs. Based on a modified *Boids* algorithm [18], it dynamically adjusts the speed and trajectories of agents, promoting coordinated movement and avoiding collisions.

In addition, an adaptive PID controller allows each UAV to individually change control parameters according to the current positioning error, environmental changes, and accumulated experience. This mechanism increases the stability and adaptability of the system to various influences, including turbulence or the appearance of new obstacles.

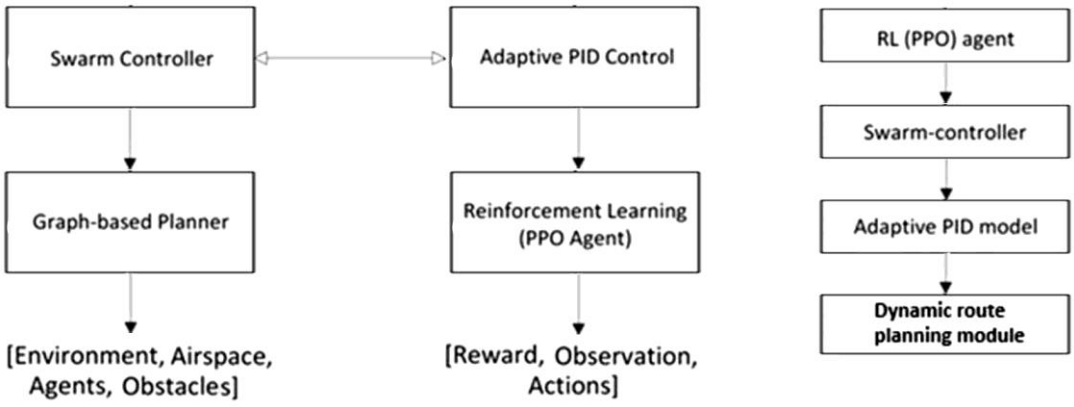
The graph router dynamically updates the optimal routes for each agent, taking into account the current configuration of the airspace, traffic, weather conditions, and possible movement restrictions. This helps the system to respond quickly to changes and optimize trajectories in real time.

An important component of the architecture is the *reinforcement learning agent* (RLA), which acts as the "brain" of the UAV group. This agent analyzes the global state of the system and generates strategic actions for the entire group, constantly improving the policy based on feedback and collective experience.

Interaction between modules is implemented through a main control loop, in which the RL agent generates strategic signals that are sent to the swarm controller. The latter coordinates local actions between agents, after which each UAV individually adjusts its trajectory using a PID controller. At each step, information about the current state of the environment, such as positions, obstacles, turbulence, or other changes in conditions, is sent to all modules. This ensures their synchronization, adaptation, and the integrity of the entire system's functioning.

This modular architecture not only improves the efficiency of group control, but also makes it easy to investigate the impact of individual components or settings on the overall behavior of the fleet in various simulation scenarios.

The model structure is shown in Fig. 1.



A B

Fig. 1. (A) – basic module architecture;
(B) – diagram of interaction between the RL agent, swarm controller, and PID controller in the task of routing a group of UAVs

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

This approach allows for the effective integration of modern artificial intelligence algorithms with classical control methods, which significantly expands the possibilities for modeling and practical application of the system.

The model is configured by selecting hyperparameters for conducting a series of simulation experiments.

Key hyperparameters (PPO, PID, *Swarm*)

When configuring the combined *RL+Swarm+PID* system, special attention was paid to the selection of hyperparameters that determine the quality of training, noise resistance, adaptability, and the effectiveness of UAV group coordination.

In particular, for a reinforcement learning agent using the Proximal Policy Optimization (PPO) algorithm, it was important to select parameters such as **learning_rate, ent_coef, clip_range, net_arch, and total_timesteps**. For example, the selected learning rate value (***learning_rate* = 0.0003**) helped to find the optimal balance between the speed of policy updates and its stability: increasing this parameter led to instability, while decreasing it led to excessively slow convergence.

The entropy coefficient (***ent_coef* = 0.05**) determines the agent's activity level: increasing this parameter above 0.1 made the policy too chaotic, while decreasing it slowed down adaptation to new conditions. The parameter ***clip_range* = 0.25** ensured smooth policy updates, and the selected neural network architecture (***net_arch* = [128, 128]**) made it possible to effectively approximate complex multi-agent strategies.

The PID controller parameters were adjusted according to the need to compromise between quick response to changes and control stability in complex or noisy environments. Additional tests showed that overly aggressive parameters negatively affect the result in turbulent conditions.

The *Swarm* module was modeled using weight coefficients (***alpha, beta, gamma***) that determined the strength of alignment, attraction to the center of the group, and collision avoidance. The introduction of Gaussian noise (***sigma_turbulence* = 0.05**) made it possible to evaluate the robustness of the controller in turbulent conditions.

All these parameters were selected empirically based on numerical experiments with cross-validation. The detailed influence of each parameter on the simulation results is summarized in Table 1, which facilitates a reasonable choice of configuration for different tasks. This allows the system configuration to be selected depending on the tasks and experiment scenarios. It was this combination of hyperparameters that provided the best results in terms of MAE, *success rate*, and convergence speed for different experimental scenarios.

Table 1. *Hyperparameters and their impact*

Parameter	Mean	Impact
learning_rate	0.0003	Speed of agent adaptation
ent_coef	0.05	Level of exploration
clip_range	0.25	Stability of learning
net_arch	[128, 128]	Quality of policy approximation
Kp, Ki, Kd	1.0/0.1/0.01	Stability and speed of PID control
sigma_turb	0.05	Realism of the model (noise testing)
total_timesteps	100_000	Duration of learning for each series of experiments

Training was performed using a train-test scheme with random initial positions to avoid overfitting. Each model was tested on 30 independent episodes. Stability was checked using cross-validation on different seeds.

The choice of parameters was justified by a series of tests that showed that increasing *ent_coef* to 0.1 reduces stability, and smaller networks ([64, 64]) worsen the convergence rate.

The entropy coefficient allows the agent to explore trajectories more actively in complex situations. Increasing the number of neural network layers contributes to better policy approximation in multi-agent cases.

Success criterion: an episode is considered successful if MAE < 0.7 in 40 steps. The MAE threshold value was chosen as a compromise between positioning accuracy and time/resource constraints and was confirmed by experiments.

The following experiments are based on the above-selected system hyperparameters.

Research sequence

The effectiveness of the proposed system was evaluated through a series of numerical experiments implemented in a specialized simulation environment developed on the basis of the ***Gymnasium*** platform. The research was conducted according to a carefully designed scenario that included several interrelated stages.

The first stage involved **preparing the environment**, during which airspace was formed

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

in the form of an oriented graph with specified nodes (start, targets, obstacle zones) and edges denoting possible movement routes. To increase the realism of the simulation, random obstacles, high-risk areas, and so-called no-fly zones were added. This configuration made it possible to test the system's ability to adapt to complex and unpredictable scenarios.

Next, **the agents were initialized**: a group of UAVs was placed in random or predetermined starting positions. Each agent received individual controller parameters, as well as initial speed and orientation values. This ensured diversity of behavior within the group and made it possible to evaluate the model's ability to coordinate flexibly.

The simulation consisted of sequential episodes, in each of which the agents moved according to a combined *RL+Swarm+PID* algorithm. At each step of the trajectory, the trajectories were dynamically updated to account for changes in the environment, interactions between agents, and the presence of turbulence or obstacles. Particular attention was paid to testing the system's behavior under stress conditions, such as the appearance of new obstacles or increased noise.

During the simulation process, **key metrics were collected**: mean absolute error (MAE), success rate, number of collisions, average number of steps to reach the goal, and other performance indicators. Analysis of these indicators provided an objective picture of the system's performance.

To substantiate the advantages of the proposed architecture, the results were **compared with basic approaches**, in particular with classic PID control without swarm control and reinforcement learning elements. This comparative analysis helped to identify specific advantages of the new model in terms of robustness, flexibility, and convergence speed.

A separate stage was **repeating the experiments for different scenarios**: the number of agents (from 1 to 10) was changed, the noise level, the number and topology of obstacles were varied. To avoid overfitting, a *train-test* scheme with random initial conditions and cross-validation was used, which ensured the reliability of the results.

This comprehensive approach made it possible to comprehensively assess the stability and adaptability of the developed control system in various operating conditions, as well as to determine the influence of individual parameters on the overall efficiency of the model's functioning.

At each stage of the experiments, the influence of the key hyperparameters of the RL agent, PID controller, and *swarm* module on the main performance metrics was analyzed separately: average absolute positioning error, frequency of successful episodes, and system robustness to external disturbances. The results of comparing different configurations made it possible to identify the optimal settings for achieving stable model operation in various airspace scenarios.

Experiment results

To evaluate the effectiveness of the proposed approach, numerical simulations were performed in a self-developed environment based on the ***Gymnasium*** platform. The experiment scenarios included both single UAV flights and group interactions of up to 10 agents in environments with varying amounts of obstacles and noise levels. All models were trained for 100,000 steps using pre-selected hyperparameters (see Table 1), which ensured balanced learning dynamics and resistance to noisy conditions.

The criterion for the success of the experiment was to achieve a mean absolute error (MAE) of less than 0.7 over 40 steps. This threshold value was chosen based on previous studies as one that guarantees approach to the target with minimal collision risk and meets realistic requirements for positioning accuracy in multi-component control systems.

Each model underwent a series of 30 independent episodes, which made it possible to evaluate the statistical stability, robustness, and reproducibility of the results. Two approaches were tested to compare their effectiveness: RL+Swarm (a combination of deep reinforcement learning Proximal Policy Optimization with swarm control) and a classic PID controller with fixed coefficients. The evaluation was based on the average MAE, the proportion of successful episodes, the convergence rate, and adaptability to changing environmental conditions.

Key evaluation metrics

A comparison of the total rewards of *RL+Swarm* and *Baseline PID* is shown in Fig. 2.

As can be seen from the graph, *RL+Swarm* demonstrates significantly higher rewards for most episodes.

The MAE values for each episode are shown in Fig. 3.

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

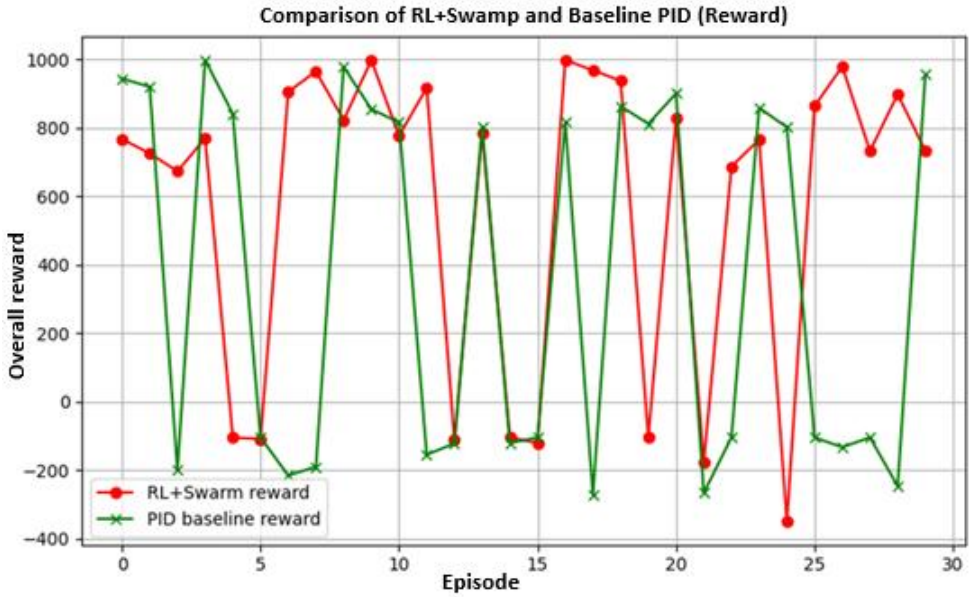


Fig. 2. Reward comparison between algorithms

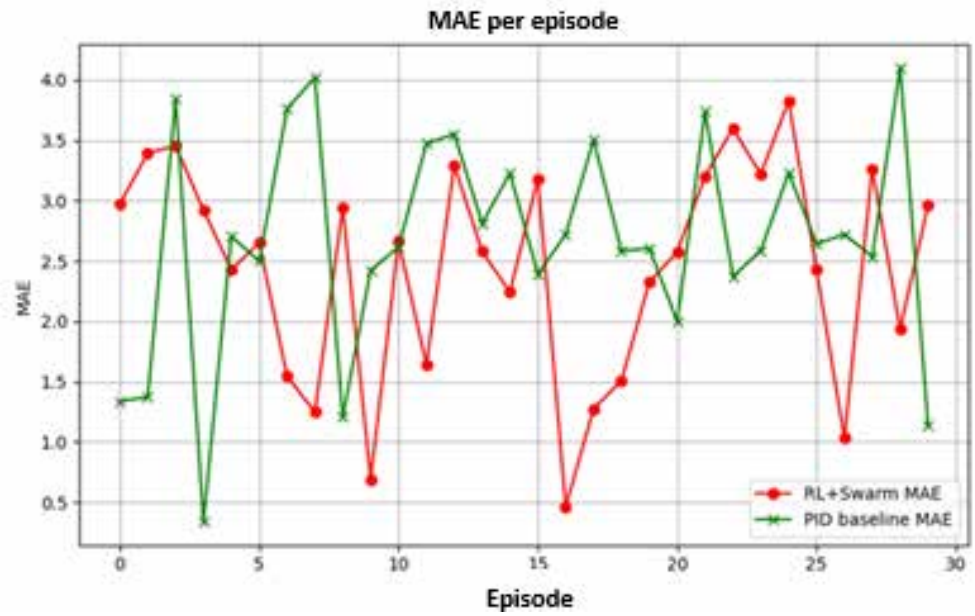


Fig. 3. Comparison of MAE distribution per episode for both algorithms

The value of *Mean RL MAE* is 2.4, and *Mean PID MAE* is 2.67. This means that the number of errors made by agents on the way to the target for the RL algorithm has decreased.

The percentage of successful episodes is shown in Fig. 4.

The goal achievement rate for RL is 73.3%, and for **PID** = 50%. This indicates that the proposed algorithm achieves the goal approximately 50% more often.

The number of steps to the goal is shown in Fig. 5.

The analysis shows that *RL+Swarm* demonstrates better adaptation to dynamic changes and learns more effectively in scenarios with a changing environment. At the same time, *PID baseline* reaches the goal faster in terms of the number of steps, but has a lower success rate.

Fig. 6 shows a comparison of the movement of an agent controlled by *RL+Swarm* (red line, Fig. 6, a) and a classic PID controller (green line, Fig. 6, b) during the first episode of the experiment, where the target is marked as *Goal*.

As can be seen in Fig. 6, *RL+Swarm* demonstrates a more straightforward and efficient path to the target, while the PID controller deviates more often from the shortest trajectory.

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

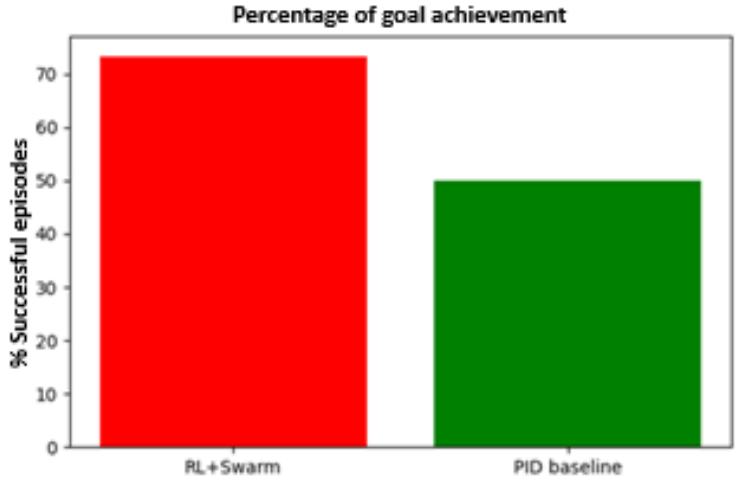


Fig. 4. Percentage of goals achieved

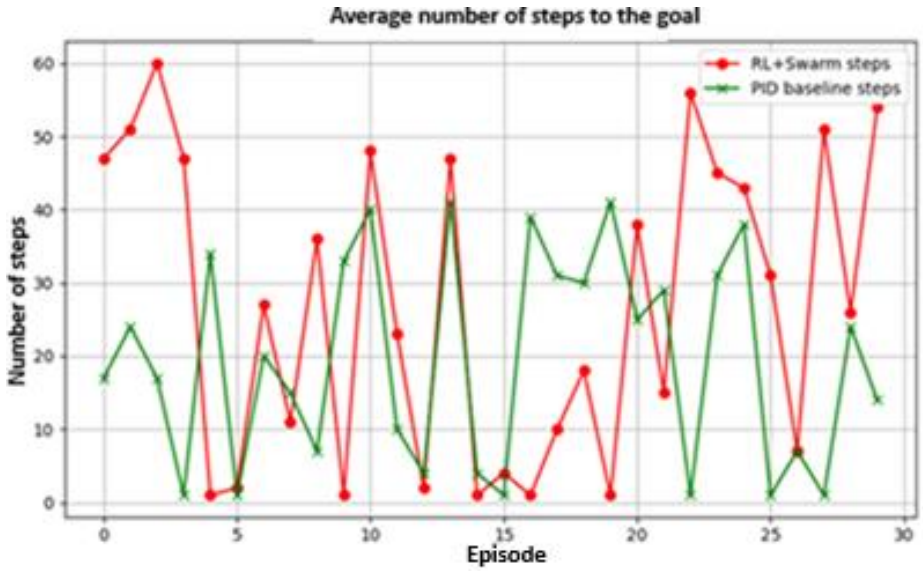


Fig. 5. Average number of steps for each UAV for two algorithms

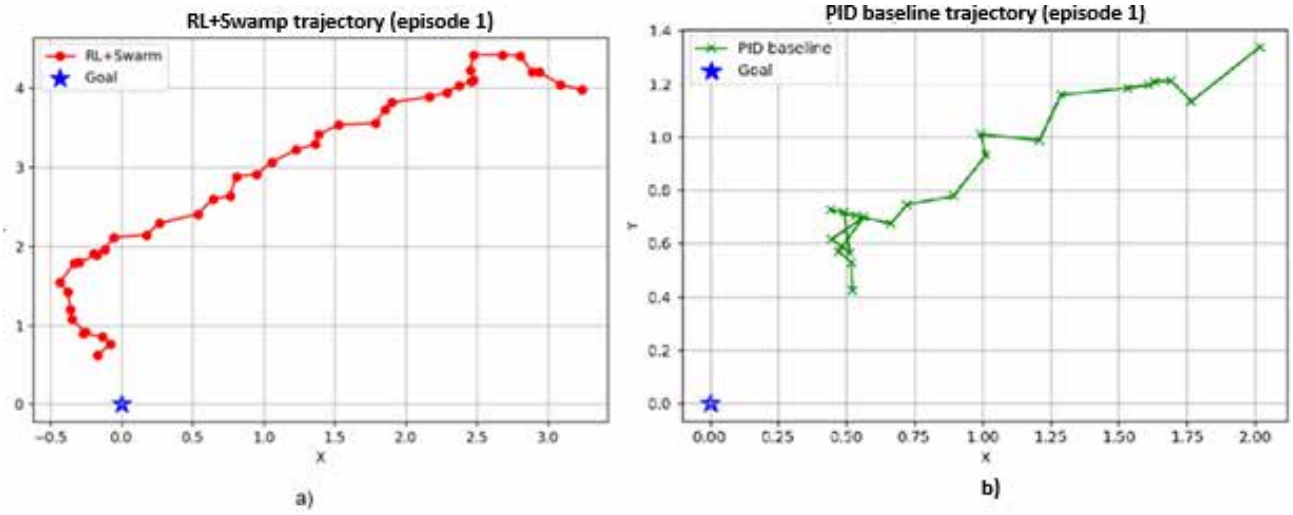


Fig. 6. Typical trajectories of a single agent in an *XY* environment

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

This analysis allows to visually assess the advantage of the trained *RL+Swarm* model over the classical approach, especially in complex or noisy environments.

Research results and their discussion

The results show that *RL+Swarm* demonstrates a significantly higher total reward compared to the PID controller (average 576.6 vs. 357.1). In *MAE* metrics, *RL+Swarm* also outperforms PID: the average MAE value was 2.45 vs. 2.67. The percentage of successful episodes for the RL agent reached 73.3%, while for PID it was only 50%. At the same time, the average number of steps to the target is lower for PID, which is associated with aggressive behavior, but at the same time, there are more unsuccessful episodes (the agent "flies" past the target or hits an obstacle).

The RL agent provides a significantly more stable approach to the target under difficult conditions (obstacles, noise), while the PID controller does not adapt to dynamic changes in the environment. In general, RL+Swarm takes better account of both individual and collective movement characteristics, adapts to dynamic changes and complex topologies, although it sometimes requires more steps to reach the target.

The PID controller is easier to implement but less resistant to obstacles and gets stuck in local minima. Analysis of unsuccessful episodes showed that the shortcomings of *RL+Swarm* are related to strong noise, aggressive starting positions, or insufficient training.

Conclusions

The paper proposes and investigates an integrated simulation model of swarm control and adaptive routing of a group of UAVs in a dynamic air environment. Numerous experiments confirmed the effectiveness of the proposed approach according to a number of metrics: average reward, MAE, and proportion of successful episodes.

Scientific novelty of the work

In this study, for the first time, a hybrid approach is proposed for collective control of a group of UAVs in a dynamic air environment, combining:

– adaptive PID control with parameter optimization through gradient or evolutionary mechanisms;

– *Swarm intelligence* (a collective behavior algorithm modified by Boids) for coordinating movement and avoiding collisions;

– Deep reinforcement learning (PPO) methods used for dynamic real-time adjustment of control strategies.

For the first time:

– RL agent interaction with PID and swarm controller for UAV routing tasks has been implemented;

– A reward function has been introduced that takes into account progress towards the goal, penalties for steps, collisions, and goal achievement, allowing for a balance between speed and safety;

– A comparative analysis with classic PID control was performed, and performance metrics (MAE, reward, % of successful episodes, number of steps to the goal) were presented for representative experimental scenarios.

– The influence of various RL and PID hyperparameters, as well as external factors (noise, obstacles) on the result was investigated.

Comparison with analogues (shown in Table 2) demonstrates that the proposed approach provides:

– an increase in the proportion of successful episodes by 20–35% compared to the baseline PID;

– a reduction in the mean error (MAE) and more stable agent behavior in the event of changes in external conditions.

Table 2. *Comparison of the proposed algorithm with the classic PID*

Metrixes	***RL+Swarm***	***PID Baseline***	Result
Mean reward	576.6	357.1	RL advantage: ≈ +61%
Mean MAE	2.45	2.67	RL advantage: less error on the way to the goal
Success Rate	73.3 %	50.0 %	RL advantage: almost 50% more successful episodes
Steps to goal	26.8	19.4	RL agent takes more steps on average

As can be seen from Table 2, the proposed approach demonstrates a significant increase in performance indicators.

Limitations and prospects

The main limitations of the study are the need for large computing resources for training, as well as the potential complexity of scaling to larger groups of agents in real-world conditions. Further research

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

is planned to focus on optimizing the reward function, hybridization with other planning algorithms, testing on more complex topologies, and conducting field experiments.

Practical significance of the results achieved

Application of the proposed architecture for monitoring systems, logistics, search and rescue operations, as well as in urban mobility scenarios (*Urban Air Mobility*).

References

1. Debnath, D., Vanegas, F., Sandino, J., Hawary, A. F., Gonzalez, F. (2024), "A review of UAV path-planning algorithms and obstacle avoidance methods for remote sensing applications". *Remote Sensing*, № 16 (21), 4019 p. DOI: 10.3390/rs16214019
2. Martins, F. G., Coelho, M. A. N. (2000), "Application of feedforward artificial neural networks to improve process control of PID-based control algorithms". *Computers & Chemical Engineering*, № 24 (2–7). P. 853-858. DOI: https://doi.org/10.1016/S0098-1354(00)00339-2
3. Liu, X., Peng, Z.R., Zhang, L.Y. (2019), "Real-time UAV rerouting for traffic monitoring with decomposition based multi-objective optimization". *Journal of Intelligent & Robotic Systems*, № 94, P. 491–501. DOI: 10.1007/s10846-018-0806-8
4. Almeida, E. N., Campos, R., Ricardo, M. (2022), "Traffic-aware UAV placement using a generalizable deep reinforcement learning methodology". *2022 IEEE Symposium on Computers and Communications (ISCC)*, P. 1–6. DOI: 10.48550/arXiv.2203.08924
5. Madani, A., Engelbrecht, A. Ombuki-Berman, B., (2023), "Cooperative coevolutionary multi-guide particle swarm optimization algorithm for large-scale multi-objective optimization problems". *Swarm and Evolutionary Computation*. № 82. 101262 p. DOI: https://doi.org/10.1016/j.swevo.2023.101262
6. Luo, J., Tian, Y., Wang, Z. (2024), "Research on unmanned aerial vehicle path planning". *Drones*, № 8(2). 51 p. DOI: 10.3390/drones8020051
7. Li, C. Lian, J., (2007), "The Application of Immune Genetic Algorithm in PID Parameter Optimization for Level Control System". *Proceedings of the 2007 IEEE International Conference on Automation and Logistics (ICAL)*, Jinan, China, P. 2670–2674. DOI: https://doi.org/10.1109/ICAL.2007.4338670
8. Yang, F., Lu, Q., Li, R., Xu, Y., Yuan, W., Wu, X. (2023), "Real-time optimal path planning and fast autonomous flight for UAV in unknown environments". *IEEE*. DOI: 10.23919/CCC58697.2023.10240971
9. Li, Q., Li, R., Ji, K., Dai, W. (2015), "Kalman filter and its application". *IEEE*. DOI: 10.1109/ICINIS.2015.35
10. Hooshyar, M., Huang, Y.M. (2023), "Meta-heuristic algorithms in UAV path planning optimization: A systematic review (2018–2022)". *Drones*, № 7(12), 687 p. DOI: 10.3390/drones7120687
11. Li, H., Zhang, Z.-y. (2012), "The application of immune genetic algorithm in main steam temperature of PID control of BP network". *Physics Procedia*, № 25, P. 80-86. DOI: https://doi.org/10.1016/j.phpro.2012.02.013
12. Zhang, M., Liu, Y., Wang, Y., Li, F., Chen, L. (2023), "Real-time path planning algorithms for autonomous UAV". *IEEE*. DOI: 10.1109/CAC57257.2022.10054770
13. Kim, D. H. (2003), "Comparison of PID controller tuning of power plant using immune and genetic algorithms". *The 3rd International Workshop on Scientific Use of Submarine Cables and Related Technologies*, Lugano, Switzerland, P. 358–363. DOI: 10.1109/CIMSA.2003.1227222
14. Zhang, G. (2023), "6G enabled UAV traffic management models using deep learning algorithms". *Wireless Networks*, № 30, P. 6709–6719. DOI: 10.1007/s11276-023-03485-4
15. Bezkorovainyi, V., Binkovska, A., Noskov, V., Gopejenko, V., Kosenko, V. (2025), "Adaptation of logistics network structures in emergency situations", *Advanced Information Systems*, Vol. 9, No. 4, P. 39–50. DOI: https://doi.org/10.20998/2522-9052.2025.4.06
16. Kuchuk, N., Kashkevich, S., Radchenko, V., Andrusenko, Y. Kuchuk, H. (2024), "Applying edge computing in the execution IoT operative transactions", *Advanced Information Systems*. Vol. 8, No. 4, P. 49–59. DOI: https://doi.org/10.20998/2522-9052.2024.4.07
17. Yena, M. (2024), "Optimizing air traffic control: Innovative approaches to collision avoidance in UAV operations". *Integrated Computer Technologies in Mechanical Engineering – 2023 (ICTM 2023)*. P. 543–553. DOI: 10.1007/978-3-031-60549-9_41
18. Saadi, A. A., Soukane, A., Meraihi, Y., Benmessaoud Gabis, A., Mirjalili, S., Ramdane-Cherif, A. (2022), "UAV path planning using optimization approaches: a survey". *Archives of Computational Methods in Engineering*. № 29. P. 4233–4284. DOI: 10.1007/s11831-022-09742-7

Received (Надійшла) 24.06.2025
Accepted for publication (Прийнята до друку) 30.11.2025
Publication date (Дата публікації) 28.12.2025

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

Відомості про авторів / About the Authors

Yena Maksym – National Aerospace University "Kharkiv Aviation Institute", PhD Student, Department of "Information Technology Design", Kharkiv, Ukraine; e-mail: yenamaxim98@gmail.com; ORCID ID: https://orcid.org/0009-0006-0664-3244; Scopus ID: https://www.scopus.com/authid/detail.uri?authorId=59161031800

Pohudina Olha – PhD (Engineering Sciences), Associate Professor, National Aerospace University "Kharkiv Aviation Institute", Associate Professor at the Department of "Information Technology Design", Kharkiv, Ukraine; e-mail: olha.pohudina@poliba.it; ORCID ID: https://orcid.org/0000-0001-5689-2552; Scopus ID: https://www.scopus.com/authid/detail.uri?authorId=57204907264

Єна Максим Вікторович – Національний аерокосмічний університет "Харківський авіаційний інститут", аспірант кафедри інформаційних технологій проєктування, Харків, Україна.

Погудіна Ольга Костянтинівна – кандидат технічних наук, доцент, Національний аерокосмічний університет "Харківський авіаційний інститут", доцент кафедри інформаційних технологій проєктування, Харків, Україна.

ІНТЕГРОВАНА СИМУЛЯЦІЙНА МОДЕЛЬ РОЙОВОГО УПРАВЛІННЯ Й АДАПТИВНОЇ МАРШРУТИЗАЦІЇ БпЛА В УМОВАХ ЗМІННОГО ПОВІТРЯННОГО СЕРЕДОВИЩА

Предмет дослідження – процеси ройового управління й адаптивної маршрутизації безпілотних літальних апаратів (БпЛА) у складних і динамічно змінних повітряних умовах із використанням адаптивних алгоритмів. **Мета** – розробити інтегровану симуляційну модель, що поєднує методи ройового управління, адаптивного PID-контролю та алгоритмів адаптивної маршрутизації для забезпечення безпеки, оптимальності й ефективності руху флотилій БпЛА в умовах мінливого повітряного середовища. **Завдання**: проаналізувати наявні підходи до ройового управління та адаптивної маршрутизації БпЛА; розробити математичну модель інтегрованої системи, яка бере до уваги особливості взаємодії між БпЛА, уникнення зіткнень та динамічні зміни повітряного середовища; створити алгоритм ройового управління, оснований на адаптивному PID-регулюванні параметрів руху БпЛА; розробити та впровадити алгоритм адаптивної маршрутизації, що реагує на зміни трафіку, погодних умов та інших факторів повітряного простору; реалізувати інтегровану модель у симуляційному середовищі та протестувати її ефективність; проаналізувати й порівняти ефективність роботи БпЛА з розробленими алгоритмами та без них. **Методи**: упровадження методів адаптивного PID-контролю для динамічного регулювання траєкторій руху БпЛА та забезпечення точності й стабільності польотів; застосування алгоритмів ройового управління (методи типу *boids*) для синхронізації руху та уникнення зіткнень у групах БпЛА; нелінійна оптимізація маршрутів з огляду на динамічно змінні уми, що дає змогу мінімізувати ризики зіткнень, витрати енергії та час польоту; побудова теоретико-графової моделі повітряного простору для ефективного планування маршрутів і прогнозування ситуацій; створення цифрових двійників повітряного середовища для проведення симуляційних експериментів. **Результати**: розроблено інтегровану симуляційну модель ройового управління та адаптивної маршрутизації БпЛА, яка зважає на зміни повітряного середовища; алгоритми адаптивного PID-контролю та ройового управління забезпечили зменшення середньої похибки позиціювання та уникнення зіткнень БпЛА; за результатами симуляційних експериментів досягнуто збільшення нагороди агентів на ≈50%, збільшення успішного завершення епізодів на ≈50%, а також зменшення помилок агентів на шляху до цілі на ≈10%. **Висновки**: створена інтегрована модель дає змогу ефективно управляти флотиліями БпЛА в умовах мінливого повітряного середовища й водночас значно підвищити безпеку та оптимальність маршрутів; використання адаптивних алгоритмів і теоретико-графових моделей забезпечує високу точність прогнозування та мінімізацію ризиків; результати дослідження підтверджують перспективність упровадження розроблених алгоритмів для управління БпЛА в міських і регіональних умовах.

Ключові слова: безпілотні літальні апарати; ройове управління; адаптивна маршрутизація; PID-регулювання; оброблення інформації в реальному часі; оптимізація маршрутів.

Бібліографічні описи / Bibliographic descriptions

Єна М. В., Погудіна О. К. Інтегрована симуляційна модель ройового управління й адаптивної маршрутизації БпЛА в умовах змінного повітряного середовища. *Сучасний стан наукових досліджень та технологій в промисловості*. 2025. № 4 (34). С. 32–43. DOI: https://doi.org/10.30837/2522-9818.2025.4.032

Yena, M., Pohudina, O. (2025), "Integrated simulation model of swarm control and adaptive routeing of UAVS in a changing air environment", *Innovative Technologies and Scientific Solutions for Industries*, No. 4 (34), P. 32–43. DOI: https://doi.org/10.30837/2522-9818.2025.4.032

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

УДК 004.51 DOI: https://doi.org/10.30837/2522-9818.2025.4.044

Є. Ігнатюк, А. Попов

ЗАСТОСУВАННЯ МЕТОДІВ МАШИННОГО НАВЧАННЯ ДЛЯ АНАЛІЗУ UX/UI-ДАНИХ МАСОВИХ ОПИТУВАНЬ КОРИСТУВАЧІВ

Предметом дослідження є впровадження методів машинного навчання для інтерпретації UX/UI-даних, зібраних способом масових опитувань користувачів освітніх онлайн-платформ. У роботі розглянуто гіпотезу про те, що узгоджене застосування різних моделей класифікації дає змогу виокремити поведінкові патерни, які мають прогностичне значення для оцінки використання окремих функцій продукту. **Метою** є порівняльний аналіз точності таких моделей на прикладі реального UX/UI-опитування. **Методика дослідження** передбачає етапи підготовки даних, кодування ознак, нормалізації, кластеризації та побудови моделей шести типів: дерева рішень, випадкового лісу, градієнтного бустингу, багатошарової нейромережі, логістичної регресії та методу *k*-найближчих сусідів. Основну увагу зосереджено на тому, як кожна з моделей поводиться в умовах UX/UI-даних, що мають обмежений обсяг, складну структуру й множину змішаних ознак. **Результати** моделювання демонструють, що навіть у межах невеликих вибірок моделі можуть виявляти значущі залежності між соціально-демографічними факторами, типами користувачів і функціональною активністю. **Висновки.** Застосування машинного навчання до UX/UI-даних є перспективним підходом до аналізу поведінки користувачів з можливістю подальшої інтеграції таких підходів у системи підтримки рішень. Запропонований підхід дає змогу виявляти стійкі поведінкові патерни користувачів на основі структурованих UX/UI-даних і має потенціал для подальшого розвитку методів прогнозування функціональної активності. Це відкриває перспективи інтеграції таких рішень у системи підтримки рішень в онлайн-освіті, охороні здоров'я, державному управлінні та інших сферах, де аналіз користувацьких даних відіграє критичну роль.

Ключові слова: UX/UI-аналітика; анкетні дані; машинне навчання; поведінкові сценарії; нейронні мережі; ансамблеві моделі; цифрові сервіси; досвід взаємодії (UX).

Вступ і постановка проблеми

У сучасному цифровому середовищі дизайн інтерфейсів відіграє ключову роль не лише в естетичному сприйнятті, а й у зручності використання цифрових сервісів [1] – від освітніх платформ і банківських застосунків до систем електронного врядування. Залежно від того, наскільки зручною, зрозумілою та послідовною є взаємодія користувача з інтерфейсом, безпосередньо залежить ефективність його застосування, рівень залученості аудиторії та показники утримання. Усе більше компаній – від стартапів до великих платформ – інтегрують UX/UI як окрему стратегічну функцію в процес розроблення [2].

Чимало дослідників і практиків вивчають оптимізацію та систематизацію покращення інтерфейсів та їх окремих аспектів, зокрема візуальних або структурних. Так, наприклад, у роботі *Heil* та *Gaedke* [3] запропоновано систему автоматичного аналізу візуальної складності вебінтерфейсів, що поєднує методи комп'ютерного зору, розпізнавання тексту й просторового структурування для обчислення метрик складності, які потім корелюються з оцінками користувачів. Автори дослідження *Computational Layout Perception using Gestalt Laws* [4] алгоритмічно змоделювали сприйняття інтерфейсу користувача (UI), упроваджуючи гештальт-принципи як евристику для автоматизованого групування елементів. Це дало змогу сформувати ієрархічну структуру, яка в 90% випадків узгоджувалась з очікуваною візуальною організацією інтерфейсу, що відкриває потенціал для застосування в завданнях адаптивного дизайну або автоматизованого UX-аудиту.

Використовуючи згорткові нейронні мережі, *Wang, Zhang, Du* та ін. у дослідженні [5] запропонували модель кількісного оцінювання якості кольорового оформлення інтерфейсів. Модель витягує та аналізує багатовимірні ознаки зображень інтерфейсів (насиченість, яскравість, відтінок) і зіставляє їх з оцінками користувачів, щоб уникнути суб'єктивності й забезпечити стабільну, формалізовану оцінку візуальної якості інтерфейсів.

Проте, щоб створювати не лише красиві, а й ефективні інтерфейси, дизайнери мають спиратися на дослідження користувацького досвіду, додатково беручи до уваги UI-складник. UX/UI-дослідження дають змогу не лише з'ясувати проблеми поточної взаємодії, а й виявити потреби, мотивації та бар'єри користувачів. У цьому сенсі вони стають аналітичною опорою для дизайнерських і продуктових рішень [6]:

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

допомагають сформулювати гіпотези, перевірити сценарії використання, виявити закономірності та обґрунтувати зміни. Дослідження бувають різними за форматом і обсягом [7] – від глибинних інтерв'ю та масових опитувань до п'ятисекундних тестів та A/B-тестування. І кожен із цих методів дає унікальний зріз реального користувацького досвіду. Однак розмаїття форматів і збільшення обсягів дослідницької інформації ускладнює її оброблення: результати можуть бути структурованими або ні, містити відкриті текстові відповіді, числові шкали, категоріальні змінні, поведінкові залежності, бути у форматі неструктурованого діалогу, містити відеозаписи користування інтерфейсом тощо. Показовим прикладом є актуальне дослідження, де *Zhang* та команда використали тематичне моделювання для автоматизованого аналізу тисяч текстових UX-відгуків на маркетплейсі, виявивши ключові сценарії поведінки користувачів і тренди взаємодії [8]. Ця та подібні роботи демонструють, що аналіз UX-даних активно розвивається, і дослідження в цій сфері з'являються протягом останніх 2–3 років. Водночас більшість з них зосереджені на неструктурованих або слабкоформалізованих джерелах, текстах, відгуках, відео, де застосування NLP або *vision*-підходів є найбільш очевидним. Формалізовані ж UX/UI-дані, зокрема результати опитувань, що мають структуру шкал, категорій і числових змінних, залишаються менш дослідженим полем. Утім, саме вони так само мають значний потенціал для порівнюваного, стандартизованого аналізу, що особливо цінно в умовах масштабування досліджень.

Особливо гостро це питання постає під час спроби працювати з повторюваними дослідженнями або вводити UX-аналітику як постійний процес, а не разову активність. Тоді на перший план виходять методи структурування, інтерпретування та порівняння інформації, незалежно від її масштабів. У реальних умовах усе це часто лягає на плечі дизайнера, який одночасно має будувати вигляд інтерфейсу та проєктувати досвід взаємодії [9], але не завжди володіє достатніми ресурсами для складного оброблення даних. Це створює запит на побудову систем, здатних частково або повністю брати на себе ці функції, знижуючи навантаження на команду та забезпечуючи стабільність аналізу з часом.

У науковій літературі, яку ширше розглянуто в наступному розділі, відзначено висхідну потребу в нових підходах до аналізу UX/UI-даних, зокрема в умовах повторюваних вимірювань, збільшенні проєктів і роботи з різними форматами даних (текстовими, візуальними чи поведінковими). Попри зростання кількості досліджень, що інтегрують UX-підхід з автоматизованим аналізом, більшість з них зосереджуються на нестандартизованих джерелах інформації. Натомість формалізовані UX/UI-дані, які дають змогу системно збирати стандартизовану інформацію про досвід користувачів, досі залишаються малодослідженим полем. Саме ця лакуна відкриває можливість для системного вивчення потенціалу формалізованих джерел інформації, наприклад UX/UI-опитування, для обчислювального аналізу, що може закласти підґрунтя для програмних стабільних підходів до UX-аналітики в цифрових продуктах. А загальна зацікавленість дизайн-спільноти у використанні нейромереж [10] підсилює ці дослідження, що формує засновок їх застосування для аналізу UX-даних зокрема.

У попередній науковій публікації [11] запропоновано концептуальну архітектуру комплексного рішення для оброблення результатів UX/UI-досліджень. У ній окреслено модулі роботи з різними форматами інформації (структурованими, візуальними й текстовими) й описано методичний підхід до поєднання аналітики з дизайнерським процесом. Зокрема порушено питання про необхідність моделей, здатних працювати на повторюваних наборах даних різного походження, що збираються протягом життєвого циклу сервісу.

Однак, попри наявні теоретичні основи досліджень, що демонструють роботу таких моделей на реальних UX/UI-даних, наприклад, з опитувальників, майже немає. Саме цей розрив між концептуальною можливістю та практичною реалізацією і визначає наукову новизну цієї роботи. Відповідно, аналітика анкетних UX/UI-даних із застосуванням моделей машинного навчання – потенційна можливість стандартизації UX-аналітики як повторюваного процесу.

Ця стаття присвячена одному з напрямів – формалізованій інформації з масових опитувань. В її основі лежить спроба опрацювати реальні анкетні UX/UI-дані користувачів цифрової освітньої платформи, трансформувати їх у формат, придатний для обчислювального аналізу, й оцінити ефективність різних моделей машинного навчання в завданні

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

прогнозування поведінкових сценаріїв. Отже, ця робота не лише розглядає конкретний приклад використання анкетної інформації, але й продовжує реалізацію загального дослідницького підходу до роботи з UX/UI-даними в межах запропонованої структури.

Аналіз останніх досліджень і публікацій

Зі зростанням складності цифрових інтерфейсів і варіативності користувацьких сценаріїв UX/UI-дослідження набули статусу системної практики, необхідної для створення інтерфейсів, що дійсно відповідають потребам користувачів. Сучасні дослідження у сфері UX-фахової аналітики охоплюють широкий спектр підходів – від методів збору інформації до їх інтерпретації та подальшого застосування в проєктних рішеннях. У межах цієї роботи увагу привертають наукові спроби поєднати результати UX-досліджень з алгоритмами машинного навчання, зокрема для побудови адаптивних або персоналізованих систем взаємодії.

У згаданому раніше дослідженні *Zhang, Chen* і *Huang* демонструють використання моделей для передбачення потреб нових користувачів у дизайнерських середовищах. Автори доводять, що навіть обмежені за обсягом UX-дані можуть бути корисними для побудови практичних предикативних сценаріїв. Це свідчить про потенціал простих опитувальних даних у завданнях індивідуалізації досвіду. *Zender, Humm* і *Holzheuser* [12] вивчали інтеграцію UX-принципів у дизайн взаємодії з AI-системами. Результати дослідження вказують на те, що адаптивні інтерфейси, побудовані на основі користувацьких шаблонів поведінки, підвищують довіру та задоволеність від системи. Зібрана інформація є не лише аналітичним ресурсом, а й основою для проєктування сценаріїв взаємодії. У роботі аналізується застосування пояснюваного штучного інтелекту (XAI) для UX-оцінювання. Запропонований підхід не лише класифікує проблеми користувацького досвіду, а й пояснює ці рішення, що сприяє зростанню довіри з боку кінцевих користувачів.

Duan, Chen та *Li* [13] запропонували застосовувати методи комп'ютерного зору для автоматичного виявлення елементів інтерфейсу у відеозаписах. Це стандартизує UX-аналіз на ранніх етапах без потреби в повній залученості людини як експерта. Показовим є приклад *Čejka, Smutný, Vondrák* та колег [14], які використовували *YOLOv7* для побудови динамічних теплових карт взаємодії. Модель відстежує події в реальному часі, аналізуючи зміну елементів інтерфейсу на відео. Ці підходи демонструють прогрес у напрямі мультимодального UX-аналізу.

Водночас, як свідчить публікація *Batch, Ji, Fan*, дослідники визнають важливість переходу від суто відео- або аудіоаналізу до комбінованих методів, здатних працювати з простішими типами даних. У роботі [15] запропоновано систему *uxSense*, що поєднує технології комп'ютерного зору, оброблення природної мови та базове класифікування на основі UX-метрик, хоча застосування нейронних мереж виявилось обмеженим. А в огляді [16] у журналі *Human-Centric Intelligent Systems* узагальнено сучасні напрями застосування нейромереж для поліпшення користувацького досвіду. Поряд із мультимодальними підходами, у роботі наголошено на недостатності досліджень, спрямованих на структури на кшталт анкет чи табличних даних, які залишаються найбільш доступними в умовах обмежених ресурсів.

Dou, Zheng і *Sun* у праці [17] запропонували модель *Webthetics* – глибоку нейронну мережу, розроблену для обчислювального оцінювання естетичності вебсторінок. Ця система навчається на реальних оцінках користувачів і застосовує трансферне навчання зі стилістичної класифікації зображень для підвищення точності. Замість ручних ознак, таких як колір чи симетрія, система самостійно виокремлює багатовимірні репрезентації візуального оформлення, що відповідають оцінкам користувачів. Це уможливлює об'єктивне порівняння варіантів UI-дизайну на основі глибоко інтегрованих параметрів. Хоча модель працює з візуальною інформацією, ключовим є те, що саме формалізовані оцінки користувачів були застосовані як навчальні дані.

Використанню простих, але формально структурованих UX-даних у поєднанні з машинним навчанням присвячено дослідження *Zhang et al*. [8], у якому запропоновано інструмент для автоматизованої класифікації користувацьких потреб на основі моделі *Kano*. Автори об'єднують анкетні UX-дані, глибинні інтерв'ю та зібрані онлайн-відгуки, щоб сформувати повноцінний навчальний набір для моделей глибокого навчання. Найуспішнішою виявилась модель RCNN (*Recurrent Convolutional Neural*

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

Network), яка дала змогу класифікувати текстові фрагменти за типами потреб. Важливо, що результатом стала не лише класифікація, а й інтерактивний інструмент з графічним інтерфейсом, орієнтований на початківців у UX. Тобто аналітика складних моделей стала доступною навіть у навчальному контексті.

Проте питання автоматизації UX-дизайну не зводиться лише до передбачення реакцій користувача. У праці [18] порушено глибшу проблему: UX-дизайн є процесом із непередбачуваним результатом, який містить не лише логіку, а й інтуїцію, досвід, гнучке реагування на контекст. Автор пропонує розглядати оптимізацію як підтримку на етапах зворотного зв'язку та генерації варіантів для створення нових рішень або аналізу даних. Однак ці моделі мають уміти адаптуватись до неоднозначних сценаріїв і коригуватися людиною в разі нових або нестандартних запитів. Навіть за умови успішного дизайну, на думку автора, користувацький досвід може розвиватися несподівано.

Власна попередня публікація присвячена архітектурі модульної системи для підтримки продуктового та дизайнерського прийняття рішень на основі UX/UI-досліджень. Запропонована структура передбачає інтеграцію моделей, що працюють з різними форматами інформації: текстовий (наприклад, відкриті відповіді), числовий (наприклад, анкети), візуальний (відеозаписи інтерв'ю чи з екранів) та поведінковий (переходи, теплові карти, перегляди). Для кожного з цих типів передбачається використання спеціалізованих моделей машинного навчання, які можуть бути інтегровані в єдину систему через центральний шар логіки. У роботі наголошено на важливості формування бази знань, яка не лише проводить аналітику системи, а й формулює рекомендації щодо покращення інтерфейсних рішень.

Попри висхідну кількість досліджень на перетині UX-аналітики та машинного навчання, переважає увага до аналізу мультимедійних або складноструктурованих даних [19]. Натомість можливості нейромережевих моделей щодо оброблення табличних результатів масових UX/UI-опитувань залишаються недостатньо вивченими. У цій роботі висувається гіпотеза про те, що навіть в умовах обмежених, але формалізованих UX-даних можливо досягти ефективної класифікації користувачів із застосуванням сучасних моделей машинного навчання. Наукова новизна полягає в апробації цього підходу на реальних UX/UI-даних опитування, що відтворюють взаємозв'язок між користувачами, поведінковими шаблонами й важливими для продукту метриками. Такий підхід дає змогу розширити інструментарій UX-досліджень у ресурсно обмежених проєктах, де недоступні складні або коштовні методи збирання та оброблення інформації. Він також може бути екстрапольований на інший освітній або сервісний продукт суміжних галузей і застосовуватись за схожим принципом, який описано в наступних розділах.

Мета й завдання дослідження

У межах загальної концепції побудови інструментарію для аналізу UX/UI-досліджень у цій статті необхідно вивчити можливості використання моделей машинного навчання для роботи з формалізованими UX/UI-даними, отриманими внаслідок масових опитувань користувачів. Особливий інтерес становить дослідження потенціалу цих моделей у контексті обмежених вибірок, властивих для цифрових сервісів із вузькопрофільною аудиторією, а також для стартапів і сервісів з короткостроковими функціоналами, де час і обсяг інформації обмежені. На відміну від мультимодальних чи поведінкових джерел, анкети становлять один з найбільш доступних форматів UX-досліджень, а тому потребують адаптації сучасних інструментів аналізу. Ключовою метою роботи є оцінювання ефективності застосування моделей на основі дерев рішень і перцептронів до структурованих UX/UI-даних масових опитувань для їх класифікації та визначення придатності для долучення до модульної системи UX-аналітики.

Для досягнення поставленої мети необхідно розв'язати такі завдання:

1) підготувати UX/UI-дані для подальшого аналізу, зокрема з перейменуванням полів, кодуванням ознак, заповненням пропущених значень, бінаризацією відповідей і нормалізацією числових шкал;

2) здійснити попередній дослідницький аналіз даних, зокрема визначити основні взаємозв'язки за допомогою кореляційного аналізу, оцінити залежності між змінними, а також виконати кластеризацію за поведінковими характеристиками;

3) визначити ключову цільову змінну, що може бути використана як прогнозний індикатор у моделі, та обґрунтувати її вибір;

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

4) побудувати моделі машинного навчання, як базові (*Decision Tree*), так і ансамблеві (*Random Forest,* XGBoost), а також модель на основі багатошарового перцептрона (MLP) для подальшого порівняння їх результативності;

5) оцінити ефективність цих моделей з використанням метрик точності, а також зробити порівняльний аналіз продуктивності на вибірці даних.

Поставлені завдання визначають логіку наступного розділу, в якому поетапно подано реалізацію кожного з етапів – від формування вибірки до побудови й тестування моделей.

Методика проведення UX/UI-опитування

У сфері дослідження користувацького досвіду застосовується широкий спектр підходів – від якісних глибинних інтерв'ю до поведінкового відстежування та аналізу метрик взаємодії. Ці дослідження розрізняють [20, 21] за різними критеріями: за метою (оцінювання ставлення, поведінки чи очікувань), типом даних (якісні, кількісні), джерелом (пряме опитування, непряме спостереження, автоматизований запис взаємодії), методом збору тощо. У цьому дослідженні обрано кількісний підхід, а саме масове анкетування, спрямоване на вивчення ставлення користувачів до обраної платформи та їхнього досвіду взаємодії. Попри потенційні обмеження в глибині відповідей, цей підхід є цінним інструментом для пошуку загальних закономірностей у поведінці користувачів і прийнятті рішень.

Опитування було реалізовано в межах цифрової платформи CASES [22] – екосистеми для навчання, кар'єрного розвитку та пошуку роботи у сфері креативних індустрій. CASES має власну навчальну платформу, живі та записані курси, додаткові навчальні формати, базу студентських профілів, можливості спілкування та пошуку роботи. Платформа орієнтована на українську аудиторію, зокрема молодих фахівців і фахівців середнього рівня, які навчаються або підвищують кваліфікацію у сфері дизайну, медіа, маркетингу та суміжних напрямів. UX/UI-дослідження в такому контексті є важливим інструментом підтримки гіпотез щодо поведінки, зацікавлень і мотивацій користувачів, зокрема для оптимізації підписок, структури курсів, інтерфейсних рішень і контентного подання матеріалу.

Анкета масового опитування була сформована через сервіс форм *Google* та поширена за допомогою соціальних мереж платформи у *Facebook, Telegram, Instagram* та електронної пошти. В опитуванні взяли участь понад 170 респондентів. Було запропоновано відповісти на серію запитань щодо їхнього стилю навчання, частоти використання платформи, професійного контексту, рівня доступу до техніки, самооцінки досвіду та інтересу до платних форматів. Відповіді були попередньо підготовлені, після цього експортовані у форматі CSV і завантажені в середовище для оброблення, що підтримує мову програмування *Python*, і саме з цього етапу було розпочато аналіз, оброблення та побудову моделей.

Попереднє оброблення інформації

Отримані результати були трансформовані в структуру табличного типу для подальшого опрацювання. У процесі попереднього оброблення неопрацьований набір даних поступово був трансформований у формат, придатний для машинного аналізу. У вихідній таблиці було 22 ознаки. Серед них можна було виокремити кілька ключових груп, перелік яких поданий у табл. 1. Інформація містила окремі пропущені значення (до 8–10% у певних стовпцях), а також комбіновані формати відповідей, що вимагали структурного поділу.

Усі етапи оброблення було реалізовано на *Python*. Розглянемо їх.

1. Розбиття складних змінних на окремі логічні ознаки. Наприклад, ознаку *active-subscription*, що відповідала одразу на два запитання (Чи була колись у користувача підписка? Чи є вона наразі?), було трансформовано у дві. Це ознаки *ever_had_subscription* та *currently_has_subscription*, які відтворюють відповіді на ці запитання окремо. Аналогічно ознаку *auditor-mode* було поділено на *knows_auditor_mode* та *used_auditor_mode*, а *employment-status* – на *is_employed* і *works_in_creative*.

2. Кодування впорядкованих категорій (*Ordinal Encoding*). Для ознак на кшталт *income-range, age, usage-time, skill-level* було зафіксовано логічну послідовність значень і закодовано їх як числові ранги. Наприклад, для ознаки *income-range* – послідовність від "не маю доходу" до "понад 40 тис. грн".

3. Кодування невпорядковних категорій за підходом *One-hot Encoding*. Ознаки *preferred-learning, preferred-online-courses, marital-status* тощо були

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

перетворені на множину бінарних змінних. Це усуває ризик того, що модель помилково сприйматиме номінальні значення як кількісні.

Наприклад, ознака *preferred-learning*, що містила значення на кшталт "самостійне навчання за власним графіком" або "навчання в групі з дедлайнами під наглядом куратора", була перетворена на кілька окремих бінарних стовпців, кожен з яких позначає наявність певного варіанта у відповіді.

Таблиця 1. *Групи ознак в табличній структурі даних*

№	Групи	Назва	Опис
1	Кількісні	*platform-usability*	Оцінка зручності користування
		search-difficulty	Оцінка складності пошуку
		feedback-quality	Оцінка якості зворотного зв'язку
2	Категоріальні впорядковані	*skill-level*	Рівень навичок у відповідній галузі
		age	Вік користувача
		income-range	Оцінка доходу
		usage-time	Час проведення на платформі
		earning-frequency	Як часто навчається на платформі?
3	Категоріальні невпорядковані	*preferred-online-courses*	Типи вподобаних онлайн-курсів
		archetype-purpose	Мотивація і мета користування платформою (кар'єра, хобі, перекваліфікація)
		gender	Стать користувача
		marital-status	Сімейний стан
		preferred-learning	Вподобаний формат навчання
		employment-status	Статус зайнятості
4	Складні комбіновані	*active-subscription*	Чи була колись активною підписка на освітній контент і чи активна підписка наразі?
		auditor-mode	Чи знає користувач про функцію "вільний слухач" та чи послуговується нею?
5	Мультивибіркові поля	*used-features*	Функції платформи, які активно використовуються
		three-main-functions	Основні функції платформи, які застосовує користувач
6	Бінарні	*platform-issues*	Наявність проблем користування
		children	Наявність дітей
		job-search	Чи перебуває в пошуку роботи?
		freelance-projects	Чи має фриланс-проєкти?

4. Кодування мультивибіркових полів (*MultiLabel Binarization*). Для полів *used-features* і *three-main-functions* використано *MultiLabelBinarizer*.

5. Перетворення бінарних змінних. Ознаки на зразок *platform-issues, job-search, freelance-projects* були уніфіковані до числового формату 0 чи 1.

6. Заповнення пропущених значень. Для логічних змінних, де *NaN* означало "ні / не знаю / не користуюсь", значення було замінено на 0. Кількісні змінні заповнювалися середнім по стовпцю. Це виконано для збереження всіх спостережень без вилучення рядків.

Після всіх трансформацій таблиця розширилась з 22 до 47 стовпців, що свідчить про значне зростання інформаційної щільності. Усі змінні набули цифрової форми – цілочисельної (*int*8, *int*64), логічної (*bool*) або з рухомою комою (*float*64) – і були повністю готові до моделювання.

Дослідницький аналіз (EDA)

На етапі дослідницького аналізу (*exploratory data analysis*, EDA) ставилась мета – попередньо виявити залежності між ознаками, оцінити якість і структуру зібраних UX/UI-даних, а також перевірити припущення щодо обраної цільової змінної. Саме тут аналітик може знайти неочевидні зв'язки між поведінковими характеристиками

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)
Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

користувачів і результатами їхньої взаємодії з платформою, а також сформулювати попередні гіпотези для подальшого моделювання. На цьому етапі виконано роботу зі складними структурами респондентських відповідей, які важко оцінити вручну. Деякі зв'язки є слабкими або прихованими, інші залежать від багатьох змінних одночасно.

Перший крок – аналіз зв'язків між числовими змінними. Було перевірено, чи наявні залежності між такими параметрами, як, наприклад, суб'єктивна оцінка платформи, тривалість використання та кількість функцій, з якими користувач мав досвід. У процесі кореляційного аналізу виявлено лише окремі помірні залежності: оцінка якості платформи була слабко пов'язаною з частотою навчання, а тривалість використання – із впевненістю у власному рівні. Загальний рівень кореляцій між числовими змінними був низьким, що є типовим для UX-даних, де прямі лінійні зв'язки, як правило, не виражені. З огляду на обмежену інформативність такого аналізу подальший етап було зосереджено на категоріальних змінних. Для цього застосовано коефіцієнт Крамера (*Cramér's V*), який дає змогу оцінити силу асоціації між парами номінативних змінних. Сила зв'язку подана числами за шкалою від 0 (відсутній зв'язок) до 1 (сильний зв'язок). Цей підхід вважається доцільним у разі, коли аналізу підлягають поведінкові установки, мотивації або переваги, властиві для UX-опитувань. Значущі зв'язки між категоріальними ознаками було візуалізовано за допомогою матриці значень *Cramér's V*, що зображена на рис. 1. Цей і наступні візуальні матеріали статті подані у вигляді знімків з екрана безпосередньо з робочого середовища дослідження, що дає змогу зберегти точність відображення елементів та уникнути спотворень.

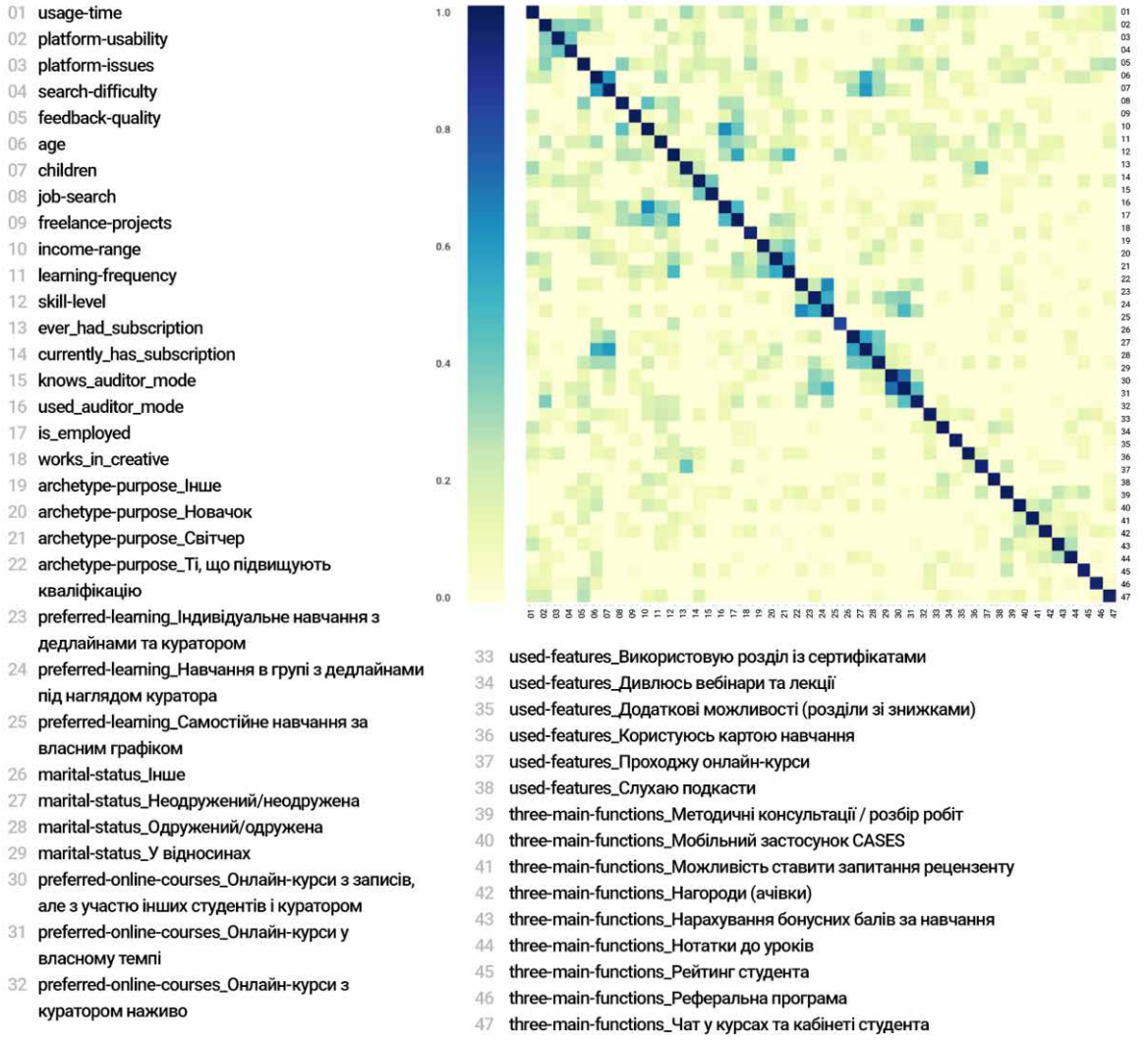


Рис. 1. Матриця Крамера для категоріальних ознак

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

На основі матриці виявлено кілька суттєвих залежностей, зокрема між частотою навчання й типом обраних курсів, архетипом і роллю користувача на платформі, а також між форматом зайнятості та участю у фриланс-проєктах. Отже, навіть в умовах обмеженої вибірки вдалося виявити значущі поведінкові шаблони, які не визначалися в межах числового аналізу й потребували категоріального підходу.

У фокусі аналізу була цільова змінна *currently_has_subscription*, яка відтворює, чи має користувач активну передплату на платформі. Це бінарна ознака, яка розглянута як така, що позначає загальну залученість користувача. Змістовно вона добре репрезентує основну мету – зрозуміти, що відрізняє платних користувачів від безплатних. На рис. 2 зображено частотний розподіл цієї змінної. Очікувано, що більшість користувачів не мають активної підписки. Такий дисбаланс вплине на метрики моделювання, тому його враховано надалі у виборі стратегії тренування.

Щоб зрозуміти потенційні сегменти користувачів без попередньої прив'язки до цільової змінної, здійснено кластеризацію методом *KMeans*. Завдання кластеризації – виділити архетипи на основі реальної поведінки й контекстів використання, не нав'язуючи готових категорій. Щоб визначити оптимальну кількість кластерів, застосовано два підходи: *Elbow*-метод та оцінку якості за *Silhouette Score*. Результати обох визначень подано на рис. 3.

Обидва графіки демонструють, що оптимальна кількість кластерів – два. Це означає, що в популяції користувачів справді є чіткий бінарний розподіл: умовно активні / пасивні або залучені / ознайомчі. Після побудови кластеризації та зниження розмірності (через PCA) отримано візуалізацію на рис. 4, де видно окремі хмари точок.

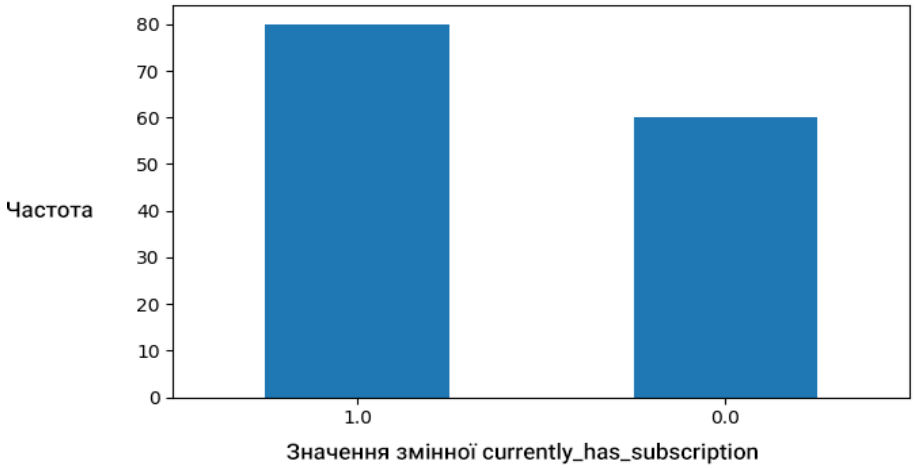


Рис. 2. Частотний розподіл активної підписки

Унаслідок кластеризації користувачів на основі числових ознак виявлено ключові розбіжності між двома групами. Найсильніше на формування кластерів вплинули змінні, пов'язані з користувацькою активністю, типом взаємодії з платформою та залученістю. Зокрема одна з найважливіших ознак досвіду наявності передплати: у першому кластері значно більше користувачів, які вже мали або мають активну підписку. Це вказує на більшу готовність платити за освітній сервіс. Подібним чином позначилась і змінна *currently_has_subscription*, яка ще більше закріплює кластер "залучених" користувачів. Також сильний вплив мало те, чи працює респондент у креативній індустрії. Ця змінна часто корелює з вищим інтересом до додаткових функцій платформи та самоосвіти. Водночас показовими стали ознаки використання конкретних функцій платформи, таких як режим вільного слухача *used_auditor_mode* та застосування сертифікатів, через що можна припустити, що активні користувачі глибше досліджують можливості та прагнуть підтвердження свого навчання. Висока частка користувачів із досвідом фриланс-проєктів також була притаманна кластеру з більшою залученістю. Нарешті, навіть суб'єктивна оцінка зручності платформи виявилася вищою в тій групі, де загальна активність і оплата були більшими, а отже, помітний причинно-наслідковий зв'язок між зручністю та конверсією.

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

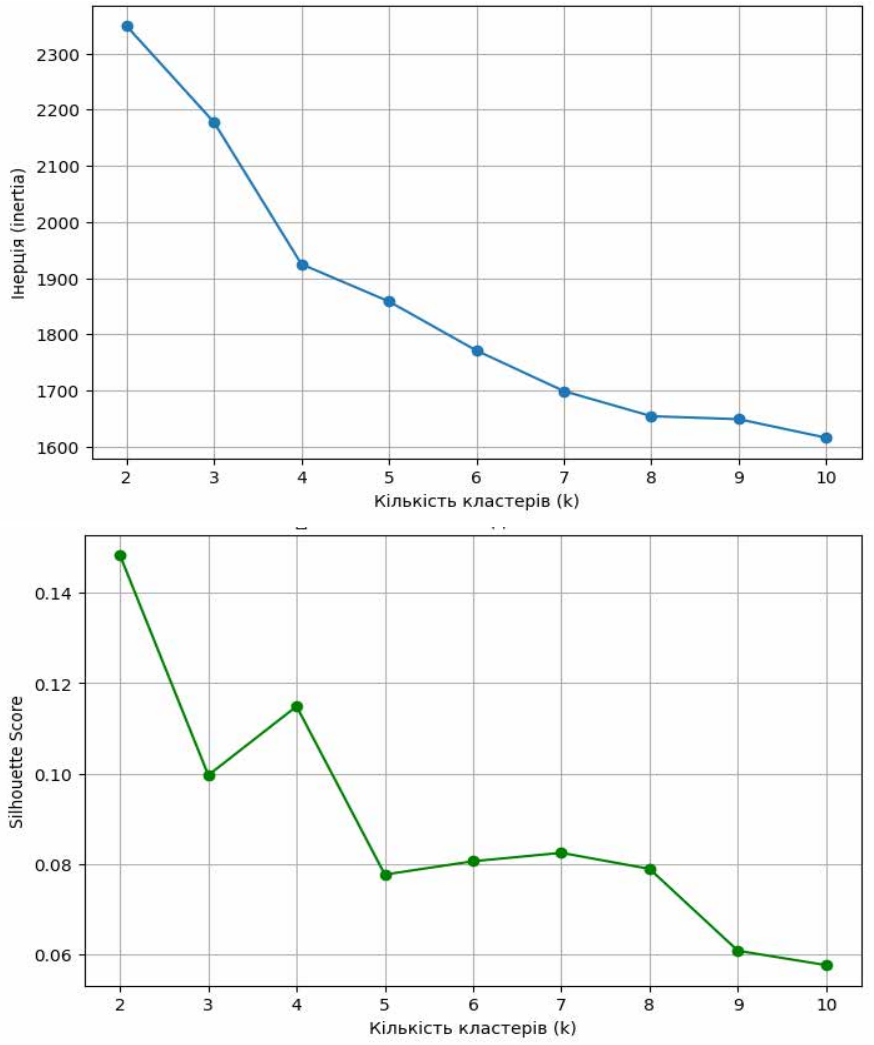


Рис. 3. *Elbow*-метод і *Silhouette Score* для вибору кількості кластерів

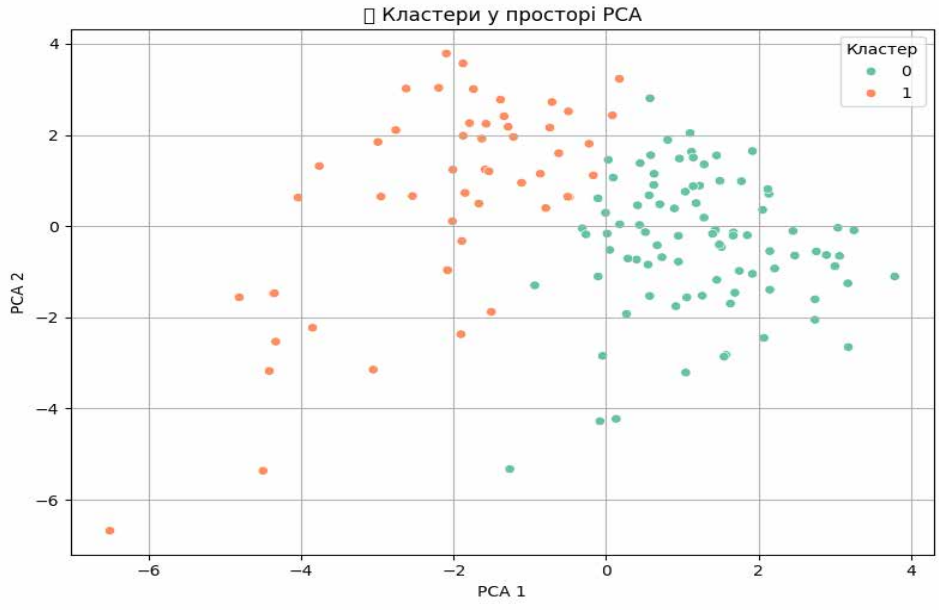


Рис. 4. Кластери користувачів у просторі PCA

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

Така сегментація не лише виявляє латентні шаблони поведінки, а й задає напрями для подальшої персоналізації комунікацій і ціннісних пропозицій. Дослідницький аналіз показав, що навіть на невеликій вибірці можна виявити сильні залежності та розбіжності між користувачами. Комбінація кореляційного аналізу, методу Крамера та кластеризації потрібна для виявлення залежності між ознаками, які не завжди очевидні, та підтвердження релевантності вибору цільової змінної. Після цього стала можливою сегментація для персоналізованої аналітики.

Цей підхід може бути масштабований на будь-який інший освітній або сервісний продукт, який збирає змішані UX-дані – опитування, поведінкові записи, підписки. У наступному розділі розглянуто реалізацію моделей для автоматичного прогнозування залученості користувача.

Обґрунтування цільової змінної

У подальшому аналізі як цільову ознаку використано змінну *currently_has_subscription*, що позначає, чи має користувач активну підписку. Ця ознака є бінарною, однозначно інтерпретованою та має практичне значення: саме вона безпосередньо репрезентує один із ключових поведінкових результатів взаємодії. З погляду прикладного застосування, це критично важливий індикатор для прогнозування платоспроможності, готовності до купівлі або продовження передплати. Крім того, *currently_has_subscription* має збалансовану вибірку: у наборі достатньо прикладів як позитивного, так і негативного класу, через що можна уникнути типових проблем у навчанні моделей. Під час EDA також було підтверджено, що ця змінна пов’язана з низкою інших характеристик (зокрема оцінкою зручності, частотою навчання та використанням функцій), що додатково доводить її аналітичну цінність як результативного показника в нашому моделюванні.

Побудова моделей

У межах цього етапу дослідження побудовано й оцінено кілька моделей, мета яких – передбачення наявності активної підписки користувача (*currently_has_subscription*) на основі структурованих UX/UI-даних. Моделювання дає змогу виявити чинники, що впливають на рішення користувача оформити або продовжити передплату, а також описати потенційні способи їх використання на практиці, наприклад для побудови механізмів персоналізованих рекомендацій або підтримки в інтерфейсі. Моделі добирались з огляду на їх здатність працювати з табличними, гетерогенними UX/UI-даними. Для навчання й тестування дані були розподілені в пропорції 80/20. Точність моделей оцінювалась за чотирма метриками: *accuracy, precision, recall* і *F1-score*, – що комплексно аналізує поведінку класифікаторів в умовах помірного дисбалансу класів. Для їх порівняння сформовано консолідовану таблицю результатів (табл. 2), що містить основні показники алгоритмів.

Перша з моделей є дерева рішень. У цьому дослідженні використано дерево завглибшки до 6 з балансуванням класів для врахування нерівномірного розподілу ознаки-підписки. За результатами оцінювання дерево досягло точності 0.61, демонструючи помірну збалансованість між *recall* і *precision*, але слабшу загальну точність, якщо порівнювати з ансамблевими методами. Зі структури дерева було помітно, що найважливішими є критерії, які розподіляють сталі ознаки на кшталт *used-features*__Користуюсь картою навчання, *age, feedback-quality* тощо. На основі цього можна не лише прогнозувати, а й інтерпретувати логіку рішення для кожного випадку.

Інша ансамблева модель – "випадковий ліс" (*Random Forest*) – поєднує значну кількість дерев рішень і зменшує ризик перенавчання, притаманний окремим деревам. Завдяки використанню випадкового підбору ознак та агрегації результатів ця модель забезпечує високу стійкість до шуму й варіативності даних. У поточному дослідженні застосовано 100 дерев завглибшки до 8. За результатами тестування модель (за шкалою від 0 до 1) продемонструвала точність 0.88 та високі значення *F1-score* для обох класів. Вона також аналізує важливість ознак (рис. 5). Найбільш значущими стали *ever_had_subscription, used-features*__Користуюсь картою навчання, *income-range, usage-time, learning-frequency* – основні поведінкові та соціальні предиктори продовження підписки. Зокрема підписку частіше продовжують ті, хто вже мав її раніше, активно користується інструментами навчання (особливо мапами навчання), має вищий рівень доходу, знайомий більше з платформою та навчається регулярно.

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)
Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

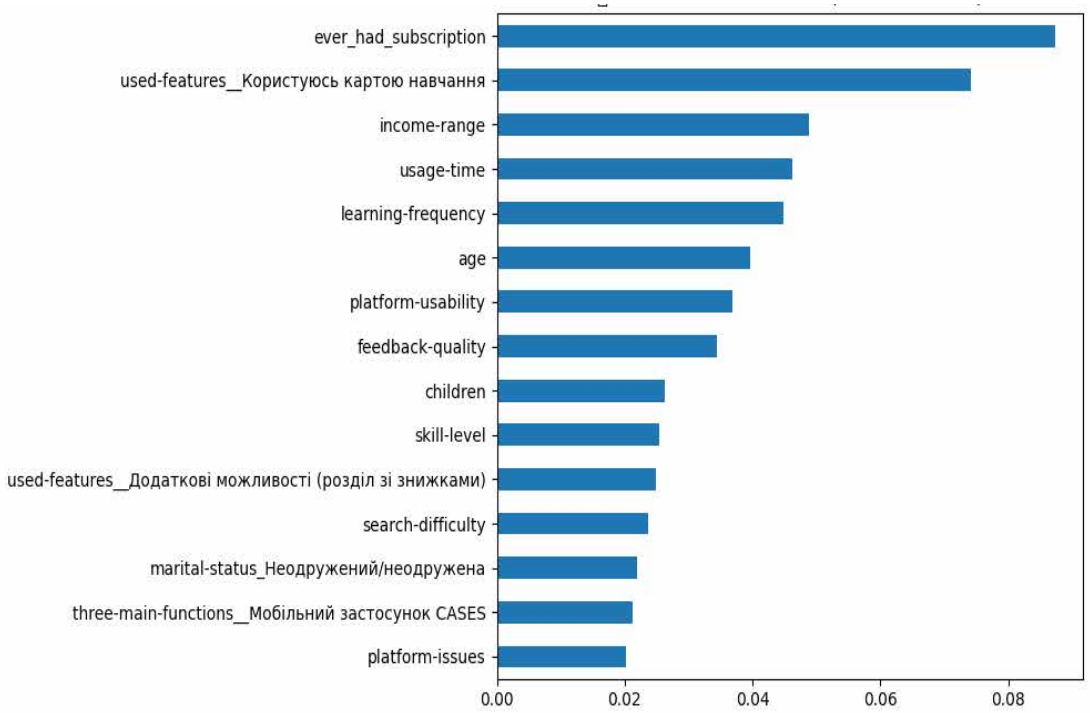


Рис. 5. Важливість ознак для *Random Forest*

Таблиця 2. *Порівняльна таблиця метрик моделей*

Модель	Метрики			
	Accuracy	**Precision (avg)**	**Recall (avg)**	**F1-score (avg)**
Decision Tree	0.61	0.61	0.61	0.6
Random Forest	0.88	0.86	0.85	0.86
XGBoost	0.87	0.85	0.3	0.84
MLP	0.89	0.89	0.88	0.87
kNN	0.61	0.62	0.6	0.59

На основі цього профілю можна сформувати цільові сегменти для маркетингової чи продуктової стратегії. Також ці ознаки допомагають ідентифікувати групи ризику, а саме користувачів, які з меншою ймовірністю продовжать підписку. Для них доцільно впроваджувати персоналізовані механіки утримання: тригери, нагадування, освітні рекомендації чи бонуси.

Застосовано ще один ансамблевий підхід XGBoost (*Extreme Gradient Boosting*). Через послідовність навчання дерев це підвищує точність у задачах з високою складністю структури ознак. У моделі використано 100 дерев завглибшки до 5 та *learning rate* 0.1. За результатами класифікації точність моделі становила 0.82, зі значеннями *F1-score* до 0.87, особливо добре працюючи з позитивним класом (наявність підписки). Також було сформовано графік важливості ознак, де лідирують *age, children, is_employed, learning-frequency*, що підтверджує перехресну значущість з *Random Forest*.

Крім того, створено базовий тип штучної нейронної мережі MLP (*Multilayer Perceptron*). Ця модель була обрана, бо здатна виявляти складні нелінійні залежності, що часто не доступні для простіших алгоритмів, зокрема дерев рішень. У нашому випадку модель містить два приховані шари по 128 і 64 нейрони відповідно, використовує функцію активації *ReLU* і регуляризацію через параметр *alpha* зі значенням 0.0005 для запобігання перенавчанню. Модель навчалась на тих самих розбитих даних, що й інші моделі. Результати показали найвищу точність серед усіх протестованих підходів 0.89 з досить симетричними значеннями *precision* та *recall*. Цей результат підтверджує ефективність MLP у задачах, де є багатофакторні зв'язки та помірна кількість ознак. Високі показники точності моделі свідчать про її здатність брати до уваги складні зв'язки, які не розпізнаються лінійними або ієрархічними моделями.

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

Ще один алгоритм – пам'яткова модель *k*-ближніх сусідів (kNN). Модель не вимагає етапу навчання, але сильно залежить від масштабу ознак. Саме тому перед класифікацією застосовано *StandardScaler*. Використано п'ять сусідів, але результати були найнижчими серед усіх моделей: точність 0.61, що вказує на слабшу здатність kNN до генералізації в умовах високої кількості ознак і невеликого обсягу вибірки. Така модель може бути корисною в інших контекстах, але в нашій задачі вона поступилася іншим підходам.

Для додаткової перевірки якості протестовано кілька моделей на випадковому прикладі з тестової вибірки. Моделі MLP, *Random Forest* та XGBoost класифікували один і той самий випадок. У цьому прикладі лише модель XGBoost давала неправильні відповіді, тоді як решта мали позитивний результат. Варто зазначити, що такий підхід не є заміною загальній оцінці, але перевіряє, наскільки узгоджені або суперечливі висновки моделей на індивідуальних рівнях.

Отримані результати підтверджують науковий засновок про те, що навіть прості, формалізовані UX/UI-дані з масових опитувань можуть слугувати надійним джерелом для побудови точних і стабільних моделей класифікації користувацької поведінки. У контексті сучасних досліджень, що здебільшого зосереджені на мультимедійних або складних структурованих даних, це становить помітну наукову новизну. Зокрема на тлі зростання зацікавленості в застосуванні нейромереж до відео, текстів і мультимодальних сигналів тема автоматизованого аналізу опитувальних таблиць залишається малоопрацьованою. Отже, це дослідження пропонує вагоме доповнення до наявної парадигми, доводячи, що високоякісні моделі можуть бути побудовані й на базі простих, дешево отримуваних UX-даних без необхідності в складних сенсорних системах або великих обсягах мультимедійної інформації. Це відкриває шлях до створення адаптивних, масштабованих і пояснюваних систем підтримки дизайну, здатних працювати в реальних прикладних умовах.

Висновки

Проведене дослідження підтверджує гіпотезу про те, що навіть у межах невеликої, частково заповненої, але структурованої UX/UI-вибірки можна досягти достатньої точності в задачах передбачення поведінкових сценаріїв користувачів. Аналіз на основі ознаки *currently_has_subscription* дав змогу виявити змінні, що корелюють із користувацькою залученістю та готовністю сплачувати підписку за контент. Цей підхід продемонстрував свою валідність і в технічному, і в практичному вимірах.

Моделі на основі ансамблевих алгоритмів (*Random Forest,* XGBoost) забезпечили найвищі метрики точності та збалансованості, водночас даючи змогу інтерпретувати важливість окремих ознак. До основних предикторів у цьому наборі даних належать: попередній досвід підписки, активне застосування навчальних функцій, рівень доходу, тривалість користування та частота навчання. Ці ознаки можуть лягти в основу сценаріїв для сегментації аудиторії та побудови систем утримання. Зокрема виявлені шаблони допомагають ідентифікувати як активних користувачів, так і групи ризику, до яких можуть бути застосовані персоналізовані стратегії втримання. Нейронна мережа (MLP) також показала високі результати: попри вищу чутливість до оброблення даних, вона підтвердила свою придатність до задач із комплексними ознаками. Водночас моделі на кшталт дерева рішень або kNN можуть застосовуватись у разі, коли пріоритетом є простота або інтерпретованість.

Наукова новизна полягає в демонстрації застосовності глибинних і ансамблевих моделей до табличних UX/UI-даних, отриманих із масових опитувань, а також у створенні відтворюваного підходу до роботи з UX/UI-опитуваннями, який можна масштабувати на інші домени. Зокрема методика може бути використана в освітніх платформах для виявлення ознак відтоку студентів, прогнозування інтересу до нових курсів або автоматичного формування персоналізованих рекомендацій на основі змін у поведінці. У сфері онлайн-медіа цей підхід дає змогу аналізувати й передбачати падіння залученості читачів до певних рубрик або форматів, а також сегментувати аудиторію за ймовірністю підписки. У сервісах підписного типу – допомагає вчасно ідентифікувати користувачів із ризиком відмови та запустити адаптивні механіки утримання. Отже, модель дає змогу не лише аналізувати наявні шаблони, але й проактивно впливати на розвиток користувацьких сценаріїв і поведінкову динаміку в автоматизованому режимі.

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

Описаний у статті підхід є лише частиною більш широкої дослідницької роботи, спрямованої на розроблення системного підходу до оброблення, аналізу та інтерпретації UX/UI-даних. На наступних етапах заплановано розширити класифікаційний підхід на мультимодальні джерела інформації (відкриті відповіді, неструктуровані природні дані, поведінкові патерни, RNN), побудувати рекомендаційні механізми та формалізувати систему як окремий модуль для UX-аналітики. Цей напрям буде розгорнуто в наступних статтях й дослідницьких напрацювань серії.

References

1. Farooqui T., Rana T., Jafari F. Impact of Human-Centered Design Process (HCDP) on Software Development Process. *Proceedings of the 2019 2nd International Conference on Communication, Computing and Digital Systems*. 2019. DOI: https://doi.org/10.1109/C-CODE.2019.8680978
2. Berni A., Borgianni Y., Basso D., Carbon C.-C. Fundamentals and issues of user experience in the process of designing consumer products. *Design Science*. 2023. DOI: https://doi.org/10.1017/dsj.2023.8.
3. Heil S., Gaedke M. Auto-Extraction and Integration of Metrics for Web User Interfaces. *Journal of Web Engineering*. 2018. DOI: https://doi.org/10.13052/jwe1540-9589.17676
4. Koch J., Oulasvirta A. Computational Layout Perception using Gestalt Laws. *CHI Extended Abstracts*. 2016. DOI: https://doi.org/10.1145/2851581.2892537
5. Wang S., Zhang R., Du J., Hao R., Hu J. A Deep Learning Approach to Interface Color Quality Assessment in HCI. 2025. DOI: https://doi.org/10.48550/arXiv.2502.09914
6. Stige Å., Zamani E.D., Mikalef P., Zhu Y. Artificial intelligence (AI) for user experience (UX) design: a systematic literature review and future research agenda. *Information Technology & People*. 2024; 37(6):2324–2352. DOI: https://doi.org/10.1108/ITP-07-2022-0519
7. Rohrer C. When to Use Which User-Experience Research Methods. Nielsen Norman Group. URL: https://www.nngroup.com/articles/which-ux-research-methods (дата звернення: 27.04.2025).
8. Zhang Z., Chen H., Huang R., Zhu L., Ma S., Leifer L., Liu W. Automated Classification of User Needs for Beginner User Experience Designers: A Kano Model and Text Analysis Approach Using Deep Learning. *Information*. 2024. DOI: https://doi.org/10.3390/ai5010018
9. Perrig S.A.C., Aeschbach L.F., Scharowski N., von Felten N., Opwis K., Brühlmann F. Measurement practices in user experience (UX) research: a systematic quantitative literature review. *Frontiers in Computer Science*. 2024. DOI: https://doi.org/10.3389/fcomp.2024.1368860
10. Ігнатюк Є. О., Попов А. В. Методико-інструментальні засоби автоматизованого аналізу даних результатів UX-досліджень з використанням систем штучного інтелекту. *Системи управління, навігації та зв'язку*. 2024. DOI: https://doi.org/10.26906/SUNZ.2025.1.96-106
11. Kuchuk, N., Kashkevich, S., Radchenko, V., Andrusenko, Y., Kuchuk, H. (2024), "Applying edge computing in the execution IoT operative transactions", *Advanced Information Systems*. Vol. 8, No. 4. P. 49–59. DOI: https://doi.org/10.20998/2522-9052.2024.4.07
12. Zender A., Humm B.G., Holzheuser A. Successfully Improving the User Experience of an Artificial Intelligence System. *Proceedings of the 19th Conference on Computer Science and Intelligence Systems (FedCSIS)*. 2024. DOI: https://doi.org/10.15439/2024F2707
13. Duan P., Chen C.-Y., Li G., Hartmann B., Li Y. UICrit: Enhancing Automated Design Evaluation with a UI Critique Dataset. *Proceedings of the 2024 CHI Conference*. ACM. 2024. DOI: https://doi.org/10.1145/3654777.3676381
14. Čejka M., Masner J., Jarolímek J., Šimek P. UX and Machine Learning – Preprocessing of Audiovisual Data Using Computer Vision to Recognize UI Elements. *Agris on-line Papers in Economics and Informatics*. 2023; 15(3):35–44. DOI: https://doi.org/10.7160/aol.2023.150304
15. Batch A., Ji Y., Fan M., Zhao J., Elmqvist N. uxSense: Supporting User Experience Analysis with Visualization and Computer Vision. *IEEE*. 2023. DOI: https://doi.org/10.1109/TVCG.2023.3241581
16. Fan M., Yang X., Yu T.T., Liao Q.V., Zhao J. Human-AI Collaboration for UX Evaluation: Effects of Explanations and Synchronization. 2021. DOI: https://doi.org/10.48550/arXiv.2112.12387
17. Dou Q., Zheng X.S., Sun T., Heng P.-A. Webthetics: Quantifying webpage aesthetics with deep learning. *International Journal of Human-Computer Studies*. 2018. DOI: https://doi.org/10.1016/j.ijhcs.2018.11.006
18. Choudhury N. Can Artificial Intelligence be Used to Improve Productivity by Automating Elements of the User Experience Design Processes? 2023. DOI: https://doi.org/10.20944/preprints202202.0057.v2
19. Bezkorovainyi, V., Kolesnyk, L., Gopejenko, V., Kosenko, V. (2024), "The method of ranking effective project solutions in conditions of incomplete certainty", *Advanced Information Systems*, Vol. 8, No. 2, P. 27–38. DOI: https://doi.org/10.20998/2522-9052.2024.2.04

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

20. Nielsen J. Usability 101: Introduction to Usability. Nielsen Norman Group. 2012. URL: https://www.nngroup.com/articles/usability-101-introduction-to-usability (дата звернення: 27.04.2025).
21. Malyeyeva, O., Lytvynenko, D., Kosenko, V., Artiukh, R. Models of harmonization of interests and conflict resolution of project stakeholders. *Ceur Workshop Proceedings*, 2020. 2565. P. 24–35. URL: https://www.scopus.com/pages/publications/85082134474?origin=resultslist (дата звернення: 27.04.2025)
22. CASES. Соцмережа та EdTech для креативних індустрій. URL: https://cases.media (дата звернення: 11.05.2024).

Received (Надійшла) 01.02.2025
Accepted for publication (Прийнята до друку) 30.11.2025
Publication date (Дата публікації) 28.12.2025

Відомості про авторів / About the Authors

Ігнатюк Євгеній Олексійович – Національний аерокосмічний університет "Харківський авіаційний інститут", аспірант кафедри комп'ютерних наук та інформаційних технологій, Харків, Україна; e-mail: y.o.ihnatiuk@khai.edu; ORCID ID: https://orcid.org/0009-0002-0568-829X

Попов Андрій В'ячеславович – кандидат технічних наук, доцент кафедри комп'ютерних наук та інформаційних технологій, Національний аерокосмічний університет "Харківський авіаційний інститут", Харків, Україна; e-mail: a.popov@khai.edu; ORCID ID: https://orcid.org/0000-0001-8984-731X; Scopus ID: https://www.scopus.com/authid/detail.uri?authorId=58558782100

Ihnatiuk Yevhenii – National Aerospace University "Kharkiv Aviation Institute", PhD Student, Department of Computer Science and Information Technologies, Kharkiv, Ukraine.

Popov Andrii – PhD (Engineering Sciences), Associate Professor, Department of Computer Science and Information Technologies, National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine.

APPLICATION OF MACHINE LEARNING METHODS FOR ANALYSIS OF UX/UI DATA FROM MASS USER SURVEYS

The subject of this article is the application of machine learning methods to the interpretation of UX/UI data collected through mass user surveys on digital platforms. The paper explores the hypothesis that coordinated use of various c lassification models allows for the identification of behavioral patterns that hold predictive value for assessing users' interactions with product features. **The goal** is to perform a comparative analysis of classification accuracy using real-world UX/UI survey data. **The methodology** includes data preprocessing, feature encoding, normalization, clustering, and training of six model types: decision trees, random forest, gradient boosting, multilayer perceptron (MLP), logistic regression, and k-nearest neighbors. Particular attention is paid to how these models perform on small-scale, mixed-type UX/UI datasets. The modeling results demonstrate that even with limited data, it is possible to uncover significant relationships between socio-demographic variables, user types, and feature usage. **These findings suggest** that machine learning can be a promising approach for analyzing user behavior, with the potential for further integration into decision support systems. This approach can be adapted to various domains where structured user data is available, including online education, healthcare, public administration, urban services, and internal organizational platforms.

Keywords: UX/UI analytics; survey data; machine learning; behavioral scenarios; neural networks; ensemble models; digital services; user experience (UX).

Бібліографічні описи / Bibliographic descriptions

Ігнатюк Є. О., Попов А. В. Застосування методів машинного навчання для аналізу UX/UI-даних масових опитувань користувачів. *Сучасний стан наукових досліджень та технологій в промисловості.* 2025. № 4 (34). С. 44–57. DOI: https://doi.org/10.30837/2522-9818.2025.4.044

Ihnatiuk, Y., Popov, A. (2025), "Application of machine learning methods for analysis of UX/UI data from mass user surveys", *Innovative Technologies and Scientific Solutions for Industries*, No. 4 (34), P. 44–57. DOI: https://doi.org/10.30837/2522-9818.2025.4.044

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

UDC 004.8:519.876.5:004.6 DOI: https://doi.org/10.30837/2522-9818.2025.4.058

L. MELNIKOVA, O. LINNYK, S. SHTANGEI, A. MARCHUK

OPTIMIZATION OF MOBILE FLOW ROUTING IN A WIRELESS SENSOR NETWORK USING HEURISTIC ALGORITHMS

The subject of the study is a wireless sensor network (WSN) with a mobile sink. **The purpose of the work** is to improve the performance of the WSN, increase its lifetime and functionality by reducing the data transmission delay time in the process of polling routers by optimizing the mobile sink route using the most efficient algorithm. To achieve this goal, the following **tasks** must be performed: optimize the route of the WSN mobile stock by solving the traveling salesman problem using the branch and bound method and comparing the conditional average route length of a set of solutions without optimization and with optimization using the Robbins–Monroe procedure; conduct a comparative analysis of the exact solution of the traveling salesman problem obtained by the branch and bound method and the approximate solution obtained by heuristic methods; formulate practical recommendations for the selection of algorithms for optimizing the mobile flow route depending on the size of the sensor network. The following **methods** were used: simulation modeling, optimization methods, mathematical data processing. **Results achieved.** The solution of the mobile flow route optimization problem in BSM using heuristic algorithms was investigated in order to formulate practical recommendations for selecting mobile flow route optimization algorithms depending on the size of the sensor network. A comparative analysis was performed of the exact solution of the traveling salesman problem, performed using the branch and bound method, and the approximate solution, performed using heuristic methods. To obtain an approximate solution, two heuristic algorithms were implemented: the ant colony optimization (ACO) algorithm and the simulated annealing (SA) algorithm. These algorithms were implemented for the traveling salesman problem with specific coordinates for each problem. The effectiveness of the algorithms is evaluated on networks of various sizes, from 10 to 500 nodes. The simulation results show that ACO is highly effective on small and medium-sized networks (up to 50 nodes), providing shorter routes and faster computation times. SA is determined to be the best scalable on large networks (100 nodes and more), offering stable performance under high computational load. **Conclusions.** It has been demonstrated that introducing optimization in the selection of the mobile flow route in BSM leads to a reduction in the length of the mobile flow bypass contour in the range of 30–40% depending on the network size and the distances between routers. Reducing the polling time of routers in a sensor network leads to an increase in the residual power of power supplies, and thus extends the life of the network. It has been proven that the use of heuristic algorithms is only appropriate when a high speed of calculating a new mobile flow route is required. If the speed of calculating a new route is not critical, then it is better to use accurate calculation algorithms. For each algorithm, parameters must be selected depending on the task at hand, since these parameters affect the speed of the algorithm and can reduce the range of possible routes that can be obtained during calculations. The study proves the importance of individual parameter tuning of algorithms to improve the accuracy and adaptability of solutions in mobile flow routing tasks.

Keywords: heuristic algorithms; mobile flow; optimization; traveling salesman problem; wireless sensor network.

Introduction. Analysis of recent publications

The key indicator of wireless sensor networks (WSN) that determines their practical application is their lifetime, so increasing it remains a pressing issue [1, 2]. The simplest methods for increasing the lifetime of WSNs are to improve the hardware characteristics of devices: reducing the power consumption of individual components, optimizing their placement on a chip or printed circuit board, or increasing battery capacity. Recent research in the field of miniature alternative energy converters (MEH, Micro-Energy Harvesters) has opened up a number of opportunities for creating fully autonomous sensor network nodes while maintaining their small size. There are known ready-made solutions for connecting sensor nodes to miniature solar cells, vibration energy converters, and thermogenerators based on the Peltier element [3]. However, none of the solutions for collecting and converting alternative energy has yet been widely used in real data collection networks containing hundreds of nodes. This can be explained primarily by high costs and significant expenses for regular maintenance [4].

Sensor networks are primarily designed for data collection. This means that there are one or more dedicated nodes to which information from the entire network flows. Such nodes (sinks) usually have a constant power source, interfaces for connecting to local and global networks, or more powerful computing devices. Therefore, there is a predominant direction of useful traffic in a sensor network. This results in a significantly higher volume of traffic passing through the routing nodes located near the sink. The more data

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

passes through a wireless network node, the higher its power consumption. It is known that in event-driven sensor networks operating according to the asynchronous access to the transmission medium algorithm, routers are an obstacle in the life of the network. This is because in order to deliver event information in a timely manner, the router must remain in receiver mode at all times [5]. As a result, the network experiences a problem of energy consumption imbalance [6]. This leads to autonomous elements located near the central data collection node (nodes) failing earlier than others due to the discharge of their own batteries, and, as a result, the autonomous operation time of the sensor network is reduced. Various energy balancing methods are used to equalize the power consumption of all network nodes.

A promising balancing method is the use of the mobility of individual network components. Some studies [6–8] have shown that the mobility of flow can potentially provide the greatest advantage in terms of increasing the autonomous operation time of the network. In a stationary flow, it is obvious that in some areas the nodes consume almost no energy and, in the event of failure, have more than 90% of their initial energy. Even random flow movement significantly improves the distribution of residual energy. Compared to a stationary data collection node, there is a more uniform energy consumption. The paper [9] provides the following ratios: the average power of the router in transmission mode is 42 mW, the average power of the router in reception mode is 52 mW, the average power of the router in processing mode is 20 mW, and the power of the router in standby mode is 0.03 mW. The addition of a mobile node significantly reduces the power consumption of the router, as it reduces the time spent on transit transmissions. However, changing the network configuration during the information collection process leads to an increase in network latency. If the set of tasks performed by the node is not critical, there may be a decline in the quality of network service.

One way to solve this problem is to add a mobile sink, which can be an autonomous robotic system [10]. The addition of a mobile sink significantly reduces the router's energy consumption by reducing the time spent on transit transmissions. The route optimization problem can be modeled as a variant of the classic traveling salesman problem (TSP), a widely recognized combinatorial optimization problem [5]. Although exact solutions for TSP, such as those obtained using the branch and bound method, guarantee optimality, they are computationally complex for large networks, as the complexity increases factorially with the number of nodes. In scenarios where immediate and efficient solutions are required, exact algorithms may be impractical due to the high computational load. As a result, heuristic and metaheuristic approaches have gained popularity as practical alternatives, offering near-optimal solutions with significantly less computation time [11]. The article discusses the application of two heuristic algorithms – the ant colony optimization (ACO) algorithm and the simulated annealing (SA) algorithm – to solve the traveling salesman problem in the context of mobile stock route optimization. These algorithms were chosen because of their well-documented ability to perform complex optimization tasks under various constraints. The ant colony algorithm, inspired by the behavior of ants when searching for food, is known for its effectiveness in finding shorter paths in smaller networks. Annealing simulation, based on the physical process of metal annealing, offers a robust search mechanism that is suitable for complex solution spaces but typically requires more computation time than ACO.

Goals and objectives

The addition of mobile traffic can change the network configuration and, consequently, routing, leading to increased network latency. Therefore, within the scope of the study, it is necessary to select algorithms for optimizing mobile traffic routing depending on the size of the network and computational requirements. For each algorithm, parameters must be selected according to the task at hand, as these parameters affect the speed of the algorithm and can reduce the range of possible routes that can be obtained during calculations. It is important to evaluate the effectiveness of each algorithm in networks of different sizes and in conditions typical for the deployment of autonomous robotic systems. By analyzing key performance indicators such as route length, computation time, and scalability, it is necessary to determine the conditions under which each algorithm provides the best performance. The resulting metrics are intended to assist in selecting appropriate route optimization methods for autonomous robotic systems operating in real-world conditions where efficiency and adaptability are of paramount importance.

The objective of this work is to study the solution of the mobile stream route optimization problem in

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

a wireless sensor network using heuristic algorithms to formulate practical recommendations for selecting mobile stream route optimization algorithms depending on the size of the sensor network.

Materials and methods

The mobile stick must survey all routers in the network only once and return to the starting point. Its route should minimize the total cost of the distance traveled. This task can be formulated as a traveling salesman problem [1].

If there are no specific requirements for the network topology, then the set of routers $N, |N| = n$, where n is the number of routers, can be considered as a set of vertices of a fully connected undirected graph of dimension n. The link of the mobile stick route between the i-th and j-th routers $(i, j) \in E$, $|E| = n^2$ will be an edge of the graph $G(N, E, C)$. The weight matrix $C = \|c_{ij}\|$ of the arcs $(i, j), (i = \overline{1,n}, j = \overline{1,n}; i \neq j)$ is known and will be called the route link cost matrix. Depending on the functionality of the network and the requirements for quality of service, the cost of a route link may vary. The article solves the problem of minimizing the length of the path traveled by the mobile flow, and the distance between routers is used as the cost of the route link. The variable of the traveling salesman problem is a Boolean variable x_{ij}, which takes the value 1 if the arc (i, j) is part of the traveling salesman's closed route, and takes the value 0 if the arc (i, j) is not part of the traveling salesman's closed route.

Then the model of the traveling salesman problem will look like this

$$Z = \min \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} x_{ij} \quad (1)$$

subject to such restrictions:

$$\sum_{i=1}^{n} x_{ij} = 1, \quad j = \overline{1,n}, \quad (2)$$

$$\sum_{j=1}^{n} x_{ij} = 1, \quad i = \overline{1,n}, \quad (3)$$

$$u_i - u_j + n x_{ij} \leq n - 1, \quad i, j = 2, \ldots, n, \quad i \neq j, \quad (4)$$

$$u_i \geq 0, \quad i \in N \quad (5)$$

$$x_{ij} \in \{0,1\}, \; i = \overline{1,n}, \; j = \overline{1,n}. \quad (6)$$

Constraints (2)–(5) create a so-called Hamiltonian cycle. Constraints (2) and (3) reflect the fact that the salesperson must visit each point only once. Since the graph is undirected, two constraints are necessary. Constraint (2) states that each vertex has only one incoming arc. Constraint (3) reflects the fact that each vertex has only one outgoing arc. It is necessary that the salesman's route has one cycle. This requirement is ensured by conditions (4) and (5), where u_i is the number of the step at which the i-th point is visited.

Network modeling design and configuration

Models with 10 to 500 nodes were created to simulate the conditions in which autonomous robotic systems might operate. Each node is a point of interest or a task location that the mobile stack must visit, and the distance between nodes is calculated using Euclidean metrics. Two approaches were used to distribute distances: uniform distribution (1–100 m) and normal distribution with a mathematical expectation of 5 m and a variance of 1 m^2, which allows for the simulation of different spatial configurations.

All simulations were performed in MATLAB, and each experiment was repeated 500 times to achieve statistically reliable results [12].

Implementation of algorithms

Ant-colony algorithm. ACO was implemented with parameters selected based on previous studies and adjusted for each network size:

- number of iterations: 55 for networks with 20 nodes; 3900 for networks with 50 nodes;
- ant population: 10 ants per generation;
- pheromone influence (α): 1, indicating moderate sensitivity to pheromone trails;
- heuristic influence (β): 2, emphasizing the preference for the shortest route;
- local pheromone decay (p): 0.1, decreasing the intensity of pheromones with each pass;
- global pheromone decay (e): 0.1, applied after each cycle is completed;
- selection probability (q): 0.9, which shifts the ants' selection towards optimal paths.

At each iteration, the ants independently explore routes, placing pheromones on shorter paths. Routes with

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

higher pheromone concentrations become increasingly likely to be selected in subsequent iterations. The algorithm performs iterations until convergence or the maximum number of iterations is reached, after which the shortest path found is recorded.

Simulated annealing algorithm. SA was implemented using parameters related to temperature control:

- initial temperature: a value of 1000 for smaller networks and 100,000 for larger networks;
- final temperature: 0.1, which ensures gradual cooling;

cooling schedule: the temperature was reduced using an exponential decay function:

$$T_{k+1} = T_k \alpha, \quad (7)$$

where $\alpha = 0.99$ – cooling coefficient;

- iterations: 250,000 for both network sizes to ensure the research result;
- decision-making probability function:

$$P(\Delta E) = e^{-\Delta E/T}, \quad (8)$$

where ΔE – change in route length.

The algorithm begins by selecting a random route and sequentially explores new routes by rearranging nodes. Route changes that improve the overall length are accepted immediately, while less optimal solutions are accepted with a probability that the route length decreases in proportion to the decrease in temperature. This allows the algorithm to avoid local minima [13].

Performance indicators and their analysis

The main indicators for evaluating the performance of algorithms were:

- route length – the total distance traveled by the mobile stack to complete the route;
- computation time – the period required for the algorithm to converge to a solution;
- error rate – the percentage deviation of each heuristic solution from the known optimal route, obtained using the exact branch and bound method for smaller networks (10 and 30 nodes).

Performance indicators were analyzed for each network size, and the Robbins–Monroe procedure was used to verify the average route length values. Based on recursive estimates, ACO and SA were compared to determine the most effective algorithm depending on the network size and the required accuracy [14, 17].

Research results and their discussion

The results of this work provide insight into the comparative performance of ant colony optimization (ACO) and simulated annealing (SA) algorithms for solving the traveling salesman problem (TSP) with the aim of optimizing the route of a mobile stream. Key performance indicators, including route length, computation time, and error rate, are evaluated for networks of different sizes (10, 20, 50, 100, and 500 nodes).

Performance in small-scale networks (10 and 20 nodes)

For small networks, both ACO and SA achieved near-optimal route lengths with minimal error rates.

- Route length: ACO achieved an average route length of 357.45 units for a network with 20 nodes, outperforming SA's 367.57 units.
- Computation time: ACO demonstrated faster computation time, averaging 0.3112 s per simulation compared to 0.5230 s for SA.
- Error rate: Both algorithms had an error rate of less than 3%, with ACO recording an average error of 0.0319% and SA recording an average error of 2.8659%.

The simulation results are presented in Table 1. The minimum paths for 10 and 20 points are shown in Figures 1 and 2.

Table 1. *Modeling results for 10, 20, and 50 nodes*

№	Ant colony algorithm			Annealing simulation algorithm		
	average path length	average time	mean error	average path length	average time	mean error
	units	seconds	%	units	seconds	%
			10 points			
1	356.900	0.5595	0	355.244	0.5595	0
			20 points			
2	356.9009	0.2102	0.4662	355.2446	0.5595	0
			50 points			
3	571.8715	27.1727	0.5264	572.6818	42.6522	0.6689

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)
Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

These results demonstrate that ACO is more efficient for small networks, offering shorter routes and faster computation times than SA. The low error rate confirms the suitability of both algorithms for scenarios with a low number of nodes, with ACO having a slight advantage.

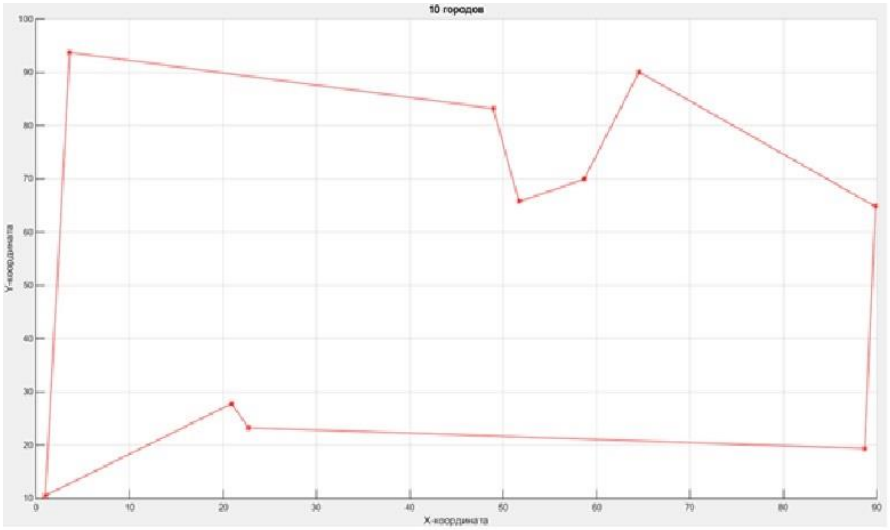


Fig. 1. Minimum route for 10 points

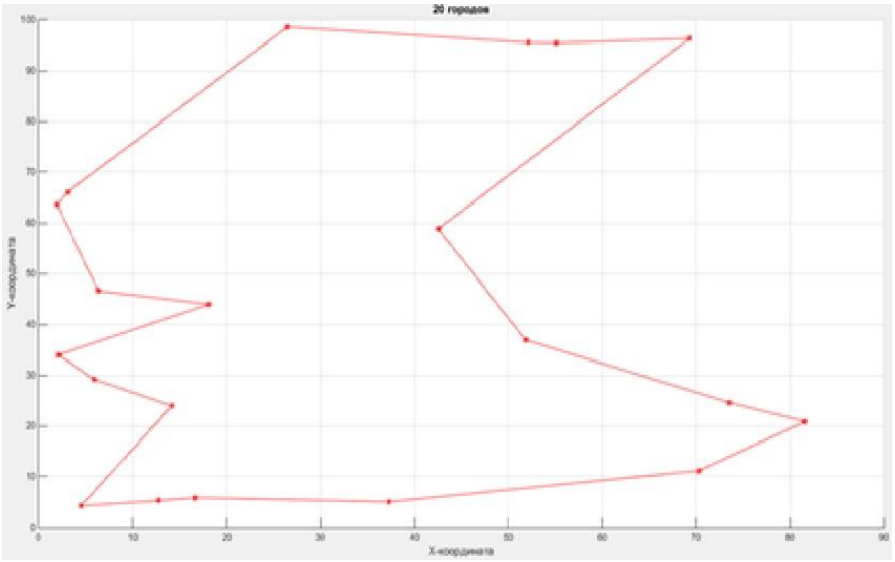


Fig. 2. Minimum route for 20 points

Productivity in medium-sized networks (50 nodes)

The simulation results are presented in Table 1. The minimum path for 50 points is shown in Figure 3.

As the network size increased to 50 nodes, ACO continued to outperform SA, although both algorithms experienced a slight increase in route length and computation time.

- Route length: ACO recorded an average length of 571.87 units, while SA created a similar but slightly longer route length of 572.68 units.
- Calculation time: ACO took 27.17 seconds, which is approximately 36% faster than the 42.65 seconds required by SA.
- Error rate: ACO maintained a low error rate of 0.5264%, while SA's error rate was 0.6689%, which is within the acceptable range for medium-sized networks.

The results show that ACO continues to be more efficient in terms of selection time for route optimization in medium-sized networks, while SA offers comparable accuracy but with higher computational costs.

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

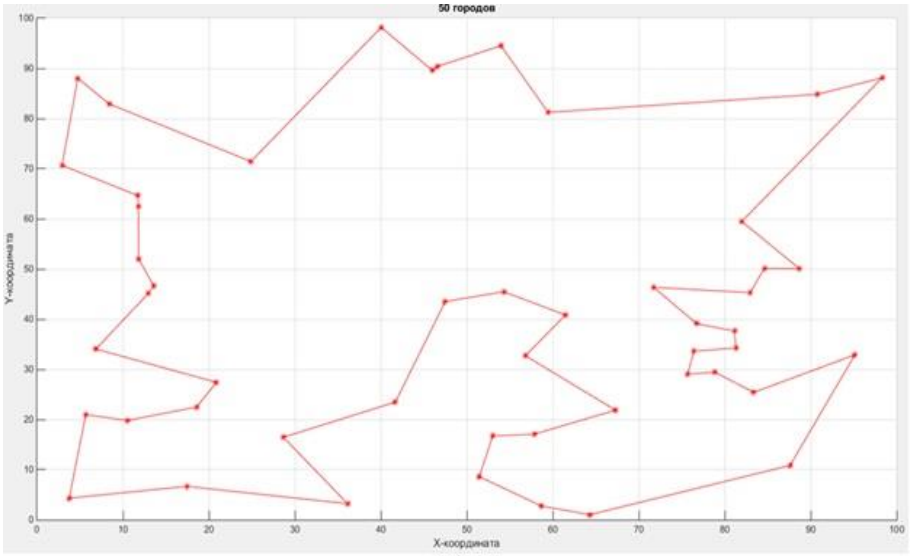


Fig. 3. Minimum route for 50 points

Performance in large-scale networks (100 and 500 nodes)

The simulation results are presented in Table 2. The minimum paths for 100 and 500 points are shown in Figures 4 and 5.

For larger networks, both algorithms encountered scalability issues, although the efficiency of ACO decreased significantly with an increase in the number of nodes [2, 14, 15].

Table 2. *Simulation results for 100 and 500 nodes*

Sample №	Ant colony algorithm		Annealing simulation algorithm	
	route length	**time**	**route length**	**time**
	units	seconds	units	seconds
100 nodes				
1	821.4988	189.513869	817.5347	42.239714
2	816.6737	186.088802	801.4941	43.638596
3	799.4715	140.772230	800.1126	48.471429
500 nodes				
1	∞	∞	1836.4	79.681766
2	∞	∞	1843.1	83.506554

1. *Network with 100 nodes*

- Route length: ACO created an average route length of 821.5 units, while SA's average route length was 817.5 units.
- Computation time: ACO dramatically increased this metric to 189.5 seconds, while SA worked faster at 42.2 seconds.
- Error rate: ACO recorded an error rate of approximately 2.6% compared to 2.3% for SA.

2. *Network with 500 nodes*

- Route length and time: Due to high computational loads, ACO's performance lagged significantly, and execution time increased non-linearly. SA, although slower than in smaller networks, was able to generate feasible solutions in a more practical time frame.

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)
Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

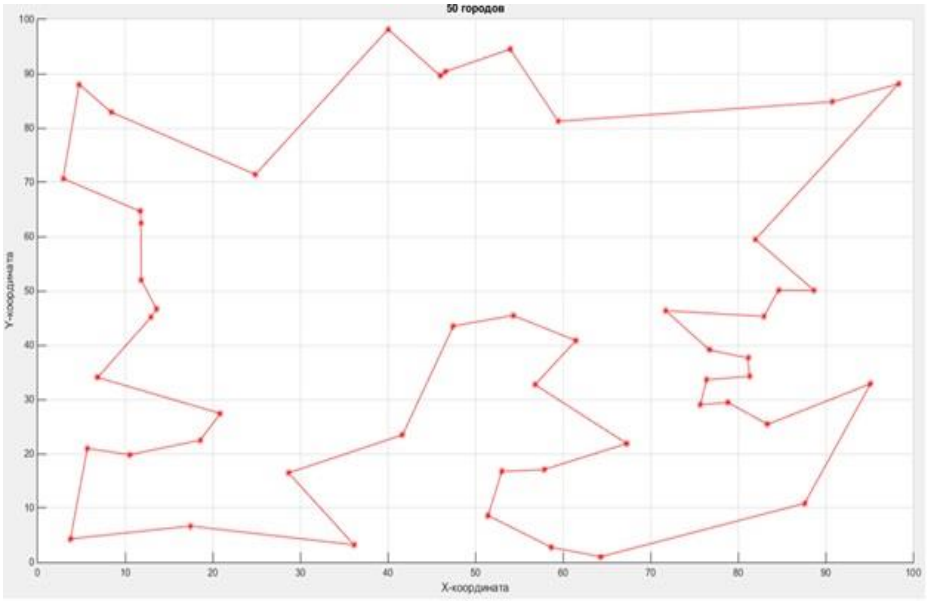


Fig. 4. Minimum route for 100 points

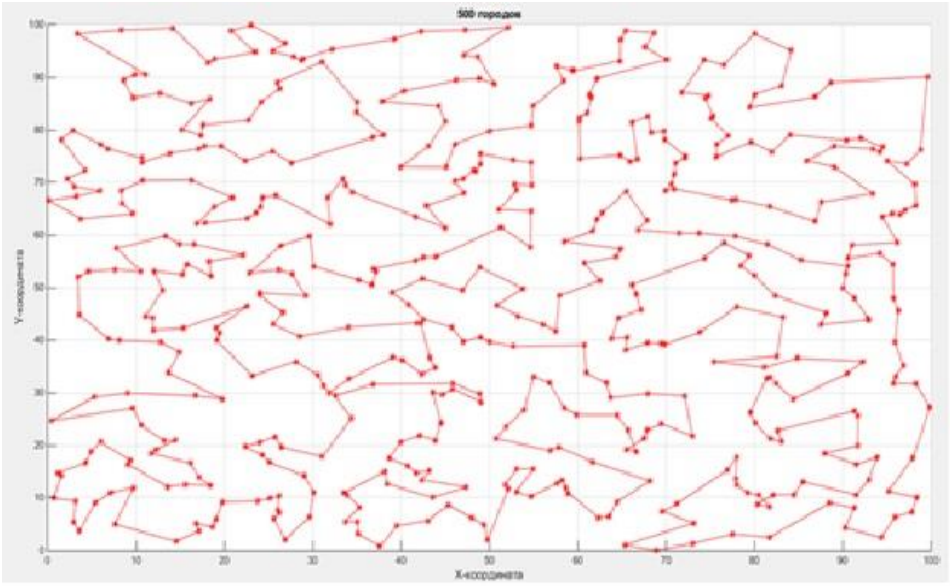


Fig. 5. Minimum route for 500 points

Brief description of algorithm productivity

The overall performance of ACO and SA for all tested network sizes indicates that ACO is better for small and medium-sized networks (up to 50 nodes) because it provides shorter routes and faster processing times. For networks with 100 nodes or more, SA becomes a significantly more scalable solution, offering more stable performance without an excessive increase in computation time.

Table 1 shows the average route lengths, computation times, and error rates for each network size and algorithm, providing a clear comparison of their suitability for different network scales.

This study contains a comparative analysis of the ant colony optimization (ACO) algorithm and the simulated annealing (SA) algorithm for solving the traveling salesman problem in the context of optimizing mobile flow routes in a wireless sensor network. The results demonstrate the advantages and disadvantages of each algorithm on networks of different sizes, providing valuable information for application in sensor networks.

The results show that ACO is very effective for small and medium-sized networks (up to 50 nodes),

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

providing high route efficiency and faster convergence time compared to SA. The ability of ACO to find shorter paths with fewer iterations makes it suitable for time-sensitive applications where mobile stacks need to find optimal routes with limited computational resources.

The advantage of ACO is due to its iterative search mechanism using pheromones, which allows it to find short routes and make optimal decisions quickly. This makes it more effective for networks with a small number of nodes, where the decision space can be explored without significant computational resources.

In contrast, the simulated annealing (SA) algorithm has shown better scalability on large networks (100 nodes and more), where ACO faces a sharp increase in computation time and errors. SA uses a probabilistic decision-making mechanism that allows it to explore the solution space more broadly, avoiding premature convergence, making it suitable for complex and large networks. Although SA requires more iterations to converge on smaller networks, its performance remains stable as the network size increases, confirming the reliability of this algorithm in large environments typical of complex sensor networks.

The results obtained are consistent with previous studies, which also demonstrate the effectiveness of ACO for small, very dense networks and the performance of SA in working with larger, complex networks. Scalability confirms that the computation time for ACO increases exponentially with increasing network size, which has also been observed in similar applications.

Based on these results, practical recommendations can be formulated for selecting mobile sink route optimization algorithms depending on the size of the sensor network. In cases where the mobile stream polls a limited number of sensors and conditions require frequent route recalculations (e.g., rapid response sensor networks for emergency services, fire departments, or medical services), then the faster convergence of ACO is a significant advantage [13, 15]. For navigation tasks in large-scale environments, such as surveillance systems, the scalability of SA ensures reliable performance, allowing routes to be calculated in a reasonable time even with a high level of network complexity.

The main limitation of this study is the use of modeling that does not take into account dynamic environmental factors that can affect mobile stream navigation in real-world conditions. Future work plans to add adaptive real-time algorithms to account for changing obstacles or operational constraints. In addition, hybrid approaches can be developed that combine the fast convergence of ACO with the broader search capabilities of SA to optimize both computational efficiency and reliability of solutions in networks of various scales.

Conclusions and prospects for further research

An evaluation and comparative analysis of the performance of the ant colony optimization (ACO) algorithm and the simulated annealing (SA) algorithm for solving the traveling salesman problem in mobile flow route optimization tasks in wireless sensor networks has been performed. The results demonstrate that ACO is effective for small and medium-sized networks, where it provides shorter routes and faster computation times compared to SA.

For large networks (100 nodes and more), SA has demonstrated the ability to provide acceptable solutions within practical time constraints. This makes it suitable for complex environments where autonomous robotic systems need to navigate large spatial structures or dynamically respond to changing route requirements.

The results of the study offer guidelines for selecting mobile stream route optimization algorithms depending on network size and computational requirements.

Promising areas of research may include the development of hybrid approaches that combine the advantages of ACO and SA to improve both efficiency and scalability [16]. In addition, the implementation of these algorithms in real sensor networks could provide practical information about their effectiveness in dynamic conditions, which will help to further refine their use in robotic navigation and route planning tasks.

References

1. Hannan, M.A., Hossain Lipu, M.S., Mahmuda, Akhtar, Begum, R.A., Md Abdullah Al Mamun, Hussain, Aini, Mia, M.S., Basri, Hassan (2020), "Solid waste collection optimization objectives, constraints, modeling approaches, and their challenges toward achieving sustainable development goals". *Journal of Cleaner Production*, Vol 277, 2020, 123557. ISSN 0959-6526. DOI: 10.1016/j.jclepro.2020.123557
2. Bondarenko, O., Ageyev, D., Mohammed, O. (2019), "Optimization Model for 5G Network Planning", *2019 IEEE 15th International Conference on the xperience of Designing and Application of CAD Systems (CADSM)*, Polyana, Ukraine, 2019, P. 1–4, DOI: 10.1109/CADSM.2019.8779298

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

3. Löppenberg, M., Yuwono, S., Mochammad Rizky Diprasetya, Schwung, A. (2024), Dynamic robot routing optimization: State-space decomposition for operations research-informed reinforcement learning", *Robotics and Computer-Integrated Manufacturing*", Vol.90, 102812. DOI:10.1016/j.rcim.2024.102812
4. Chekubasheva, V., Glukhov, O., Kravchuk, O., Levchenko, Y., Linnyk, E., Rohovets, V. (2022), "Possibility of Creating a Low-Cost Robot Assistant for Use in General Medical Institutions During the COVID-19 Pandemic". *Optics and Its Applications. Springer Proceedings in Physics*, Vol 281. Springer, Cham., P. 203–213. DOI:10.1007/978-3-031-11287-4_16
5. Melnikova, L., Linnyk, E., Ageyev, D., Melnikova, O., Kryvoshapka, N., Barsouk, V. (2019), "Minimizing the Route of Sink Node in Wireless Sensor Network," *2019 IEEE International Scientific Practical Conference Problems of Infocommunications, Science and Technology, PIC S&T*, Kyiv, Ukraine, 2019, P. 861–864, DOI: 10.1109/PICST47496.2019.9061563
6. Anufriiev, V., Kravchuk, O., Levchenko, Y., Hlukhova, H., Linnyk, E., Glukhov, O. (2024), "Perspective of Creating a Low-Cost Medical Assistant Robot Based on the Waffle PI4 Platform with a Palm Vein Pattern Scanner". *Preprints*, 2024031696. DOI:10.20535/RADAP.2024.98.46-54
7. Sasan, Z., Shokrnezhad, M., Khorsandi, S., Taleb, T. (2024), "Joint Network Slicing, Routing, and In-Network Computing for Energy-Efficient 6G", *IEEE Wireless Communications and Networking Conference (WCNC),* Dubai, United Arab Emirates, P. 1–6, DOI: 10.1109/WCNC57260.2024.10571186
8. Mezni, H., Yahyaoui, H., Elmannai, H. et al. (2025), "Personalized service recommendation in smart mobility networks". *Cluster Comput*, Vol. 28, № 3. DOI: https://doi.org/10.1007/s10586-024-04694-y
9. Rema, C.; Costa, P.; Silva, M.; Pires, E.J.S. (2025), "Task Scheduling with Mobile Robots a Systematic Literature Review". *Robotics*, Vol. 14, DOI: https://doi.org/10.3390/robotics14060075
10. Teja GK, Mohanty PK, Das S. (2025), "Review on path planning methods for mobile robot". Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science. 2025;239(14). P. 5547–5580. DOI: 10.1177/09544062251330083
11. Christensen, J., Bastien, C. (2016), "Heuristic and Meta-Heuristic Optimization Algorithms", *Nonlinear Optimization of Vehicle Safety Structures*, P. 277–314, 2016, DOI: 10.1016/b978-0-12-417297-5.00007-9
12. Melnikova, L., Shtangei, S., Linnyk, O. (2025), "Heuristic algorithms for optimization of mobile flow route using structured databases". *Zenodo*. DOI: https://doi.org/10.5281/zenodo.16903056
13. Obeidat, A., Shalabi, M. (2022), "An Efficient Approach towards Network Routing using Genetic Algorithm". *International Journal of Computers, Communications & Control (IJCCC).* 17(5). DOI: 10.15837/ijccc.2022.5.4815
14. Malavalli, A., Sama, K., Chhabra, J., Bassin, P., Srinivasa, S. (2025), "Engineering Resilience: An Energy-Based Approach to Sustainable Behavioural Interventions". *Multiagent Systems*. DOI: 10.48550/arXiv.2506.16836
15. Christopher Bayliss, Djamila Ouelhadj (2024), "Mobility as a Service: An exploration of exact and heuristic algorithms for a new multi-modal multi-objective journey planning problem", *Applied Soft Computing,* Vol. 162, 111871, DOI: https://doi.org/10.1016/j.asoc.2024.111871
16. Frías, N., Franklin J., Carlos V. (2023), "Hybrid Algorithms for energy minimizing vehicle routing problem: integrating clusterization and ant colony optimization." *IEEE Access*. P.: 125800-125821. DOI:10.1109/ACCESS.2023.3325787

Received (Надійшла) 13.09.2024
Accepted for publication (Прийнята до друку) 30.11.2025
Publication date (Дата публікації) 28.12.2025

Відомості про авторів / About the Authors

Melnikova Liubov – PhD (Engineering Sciences), Associate Professor, Kharkiv National University of Radio Electronics, Associate Professor at the Department of Infocommunication Engineering V. V. Popovsky, Kharkiv, Ukraine; e-mail: liubov.melnikova@nure.ua; ORCID ID: https://orcid.org/0000-0003-0439-7108; Scopus ID: https://www.scopus.com/authid/detail.uri?authorId=57216486212

Linnyk Olena – PhD (Engineering Sciences), Associate Professor, Kharkiv National University of Radio Electronics, Associate Professor at the Department of Physical Foundations of Electronic Engineering, Kharkiv, Ukraine; e-mail:elena.linnyk@nure.ua; ORCID ID: https://orcid.org/0000-0002-4906-3796; Scopus ID: https://www.scopus.com/authid/detail.uri?authorId=57190438040

Shtangei Svitlana – PhD (Engineering Sciences), Associate Professor, Kharkiv National University of Radio Electronics, Associate Professor at the Department of Infocommunication Engineering V. V. Popovsky, Kharkiv, Ukraine; e-mail: Svitlana.shtanhei@nure.ua; ORCID ID: https://orcid.org/0000-0002-9200-3959; Scopus ID: https://www.scopus.com/authid/detail.uri?authorId=56485954400

Marchuk Artem – PhD (Engineering Sciences), Associate Professor, Kharkiv National University of Radio Electronics, Associate Professor at the Department of Infocommunication Engineering V. V. Popovsky, Kharkiv, Ukraine; e-mail: artem.marchuk@nure.ua, ORCID ID: https://orcid.org/0000-0002-2720-3954; Scopus ID: https://www.scopus.com/authid/detail.uri?authorId=56485457100

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

Мельнікова Любов Іванівна – кандидат технічних наук, доцент, Харківський національний університет радіоелектроніки, доцент кафедри інфокомунікаційної інженерії ім. В. В. Поповського, Харків, Україна.

Лінник Олена Вячеславівна – кандидат технічних наук, доцент, Харківський національний університет радіоелектроніки, доцент кафедри фізичних основ електронної техніки, Харків, Україна.

Штангей Світлана Вікторівна – кандидат технічних наук, доцент, Харківський національний університет радіоелектроніки, доцент кафедри інфокомунікаційної інженерії ім. В. В. Поповського, Харків, Україна.

Марчук Артем Володимирович – кандидат технічних наук, доцент, Харківський національний університет радіоелектроніки, доцент кафедри інфокомунікаційної інженерії ім. В. В. Поповського, Харків, Україна.

ОПТИМІЗАЦІЯ МАРШРУТУ МОБІЛЬНОГО СТОКУ В БЕЗДРОТОВІЙ СЕНСОРНІЙ МЕРЕЖІ З ВИКОРИСТАННЯМ ЕВРИСТИЧНИХ АЛГОРИТМІВ

Предметом дослідження є бездротова сенсорна мережа (БСМ) з мобільним стоком. **Мета роботи** – підвищити працездатність БСМ, збільшити тривалість її життя та функціювання завдяки зменшенню часу затримки передачі даних у процесі опитування маршрутизаторів унаслідок оптимізації маршруту мобільного стоку способом обрання найбільш ефективного алгоритму. Для досягнення окресленої мети необхідно виконати такі **завдання**: оптимізувати маршрут мобільного стоку БСМ за допомогою розв'язання задачі комівояжера методом гілок і меж та порівняння умовної середньої довжини маршруту множини рішень без оптимізації та з оптимізацією за допомогою процедури Роббінса – Монро; провести порівняльний аналіз точного розв'язку задачі комівояжера, отриманого методом гілок і меж, і наближеного рішення, отриманого евристичними методами; сформулювати практичні рекомендації щодо вибору алгоритмів оптимізації маршруту мобільного стоку залежно від розміру сенсорної мережі. Застосовано такі **методи:** імітаційне моделювання, методи оптимізації, математичне оброблення даних. **Досягнуті результати.** Досліджено розв'язання задачі оптимізації маршруту мобільного стоку в БСМ з використанням евристичних алгоритмів з метою формулювання практичних рекомендації щодо вибору алгоритмів оптимізації маршруту мобільного стоку залежно від розміру сенсорної мережі. Проведено порівняльний аналіз точного розв'язку задачі комівояжера, виконаного методом гілок і меж, і наближеного рішення, виконаного евристичними методами. Для отримання наближеного рішення реалізовано два евристичні алгоритми: мурашиний алгоритм (ACO) і алгоритм імітації відпалу (SA). Зазначені алгоритми було реалізовано для задачі комівояжера з конкретними координатами для кожної задачі. Ефективність алгоритмів оцінюється на мережах різного масштабу – від 10 до 500 вузлів. Результати моделювання демонструють, що ACO має високу ефективність на малих і середніх мережах (до 50 вузлів), забезпечуючи більш короткі маршрути та швидкий час обчислень. SA визначається кращою масштабованістю на великих мережах (100 вузлів і більше), пропонуючи стабільну продуктивність за високого обчислювального навантаження. **Висновки.** Продемонстровано, що введення оптимізації у виборі маршруту мобільного стоку в БСМ приводить до зменшення довжини контуру обходу мобільного стоку в діапазоні 30–40% залежно від розмірності мережі та відстаней між маршрутизаторами. Скорочення часу опитування маршрутизаторів у сенсорній мережі сприяє збільшенню залишкової потужності блоків живлення, а отже, продовжує тривалість життя мережі. Доведено, що використання евристичних алгоритмів доцільне тільки тоді, коли необхідна висока швидкість розрахунку нового маршруту мобільного стоку. Якщо швидкість розрахунку нового маршруту не є критичною, тоді краще використовувати точні алгоритми обчислення. Для кожного алгоритму потрібно підбирати параметри залежно від поставленого завдання, оскільки ці параметри впливають на швидкість роботи алгоритму й здатні зменшити діапазон можливих маршрутів, які можна отримати внаслідок розрахунків. Дослідження доводить значущість індивідуального налаштування параметрів алгоритмів для підвищення точності й адаптивності рішень у задачах маршрутизації мобільного стоку.

Ключові слова: евристичні алгоритми; мобільний стік; оптимізація; задача комівояжера; бездротова сенсорна мережа.

Бібліографічні описи / Bibliographic descriptions

Мельнікова Л. І., Лінник О. В., Штангей С. В., Марчук А. В. Оптимізація маршруту мобільного стоку в бездротовій сенсорній мережі з використанням евристичних алгоритмів. *Сучасний стан наукових досліджень та технологій в промисловості*. 2025. № 4 (34). С. 58–67. DOI: https://doi.org/10.30837/2522-9818.2025.4.058

Melnikova, L., Linnyk, O., Shtangei, S., Marchuk, A. (2025), "Optimization of mobile flow routing in a wireless sensor network using heuristic algorithms", *Innovative Technologies and Scientific Solutions for Industries*, No. 4 (34), P. 58–67. DOI: https://doi.org/10.30837/2522-9818.2025.4.058

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

УДК 005.93:005.71-021.131 DOI: https://doi.org/10.30837/2522-9818.2025.4.068

Ю. Полупан, О. Малєєва

МОДЕЛІ ЛОГІСТИКИ ЗБУТУ ТЕХНІЧНОЇ ПРОДУКЦІЇ ОБОРОННОГО ПРИЗНАЧЕННЯ В УМОВАХ РИЗИКІВ ВОЄННОГО ЧАСУ

Предметом дослідження статті є процеси організації та оптимізації логістики збуту технічної продукції оборонного призначення в умовах воєнного часу, зокрема побудова мережевої структури логістики з огляду на реверсивні потоки, ризики й обмежені ресурси. **Мета роботи** – підвищення ефективності логістики збуту технічної продукції оборонного призначення за умов воєнного часу способом побудови адаптивної структури логістичних потоків і розроблення математичної моделі оптимізації відповідно до ризиків і ресурсних обмежень. У статті необхідно розв'язати такі **завдання**: аналіз особливостей логістики збуту технічної оборонної продукції в умовах бойових дій; розроблення структурної моделі логістичних потоків з огляду на прямі й зворотні зв'язки та повторне використання компонентів; побудова математичної моделі оптимізації логістичних процесів, зважаючи на часові витрати, ризики затримок та ресурсні обмеження; обґрунтування застосування сценарного аналізу та адаптивного планування для підвищення стійкості та безперервності постачання в умовах невизначеності. Застосовано такі **методи**: системний підхід, структурна декомпозиція логістичних процесів, ризик-орієнтований підхід, математичне моделювання з обмеженнями, сценарний аналіз і адаптивне планування. **Досягнуті результати**. Побудовано структурну модель логістичних потоків у забезпеченні ЗСУ, яка бере до уваги прямі й зворотні потоки та повторне використання компонентів; розроблено математичну модель оптимізації логістики збуту з цільовою функцією мінімізації часу та обмеженнями на обсяг постачання, сукупний ризик і кількість маршрутів. Показано, що взяття до уваги ризиків значно впливає на вибір оптимальних маршрутів і стратегій. Обґрунтовано доцільність застосування сценарного аналізу й адаптивного планування для підтримки безперервності постачання в умовах бойових дій. **Висновки.** Запропоновані структурна та математична моделі дають змогу зменшити час логістичного циклу, підвищити стійкість логістичної системи до ризиків і забезпечити оперативне реагування на зміни обстановки. Упровадження реверсивної логістики та повторного використання компонентів сприяє зниженню витрат і дефіциту ресурсів. Сценарний аналіз і адаптивне планування є ефективними інструментами управління логістикою збуту технічної продукції в умовах невизначеності та воєнних ризиків.

Ключові слова: логістика збуту; технічна продукція; забезпечення ЗСУ; структурна модель; математична модель; оптимізація; управління ризиками.

Вступ

У контексті забезпечення ефективного функціювання підприємства, що спеціалізується на виробництві технічної продукції оборонного призначення, важливе місце посідає логістика збуту. Збутова логістика охоплює всі процеси, пов'язані з переміщенням готової продукції – від місця виробництва до кінцевого споживача (наприклад, військових частин). Її основною метою є забезпечення вчасної, надійної та безпечного доправлення продукції, зокрема модернізованих зразків технічного обладнання, таких як безпілотні авіаційні комплекси [1].

В умовах бойових дій і високої невизначеності особливо актуальним є формування адаптивних моделей логістики збуту, що беруть до уваги ризики руйнування інфраструктури, блокування маршрутів, обстріли й кібератаки. Особливість технічної оборонної продукції зумовлює необхідність дотримання спеціальних умов зберігання, транспортування та обліку, що додатково ускладнює логістичні процеси.

Для розв'язання окреслених питань необхідно розробити математичні моделі та цифрові рішення, що зважають як на прямі, так і зворотні потоки, зокрема ремонт, утилізацію та повторне використання компонентів.

Тому актуальним є завдання створення комплексної математичної моделі логістики збуту технічної продукції оборонного призначення, яка дасть змогу мінімізувати загальний час виконання логістичного циклу, знизити ризики затримок і водночас зберегти належний рівень безпеки й функціональності постачання.

Аналіз останніх досліджень і публікацій

Аналіз публікацій з окресленої теми свідчить про значне зацікавлення проблематикою реверсивної

ISSN 2522-9818 (print)
Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34) ISSN 2524-2296 (online)

логістики та управління ризиками як у науковому, так і в професійному середовищі.

Сучасні дослідження активно вивчають реверсивну логістику та її роль у сталому розвитку. У роботі [2] систематизовано види зворотних потоків, фактори, бар'єри та витрати, що впливають на управлінські рішення в цій сфері. У статті [3] проаналізовано співпрацю цивільного й військового секторів у реверсивній логістиці для підтримки циркулярної економіки ЄС, запропоновано модель інтеграції сталого ланцюга постачання. У праці [4] подано інформаційну систему для управління військовим ланцюгом постачання на основі аналізу інцидентів, що підвищує стійкість операцій.

Автор статті [5] досліджує застосування військових логістичних дронів для швидкого доправлення ресурсів у складних умовах, наголошує на їх автономності та гнучкості. У роботі [6] проаналізовано методи оптимізації військової логістики в Україні, зокрема цифрові технології та безпілотники. У дослідженні [7] запропоновано математичні моделі та геоінформаційні системи для оптимізації логістичних маршрутів харчових підприємств під час воєнного стану.

Статтю [8] присвячено застосуванню алгоритму A* для оптимізації військових маршрутів із використанням штучного інтелекту. У роботі [9] проаналізовано стратегії управління логістичними ланцюгами в умовах конфлікту для підвищення стійкості бізнесу. Автори статті [10] визначили економічні пріоритети воєнного часу, зокрема йдеться про розвиток високоточної зброї та співпрацю з Польщею.

У студіях [11, 12] запропоновано алгоритми й системи для підвищення ефективності матеріально-технічного забезпечення військ, автоматизації управління потоками та прогнозування потреб. У статті [13] подано агентну модель логістики постачання озброєння з огляду на ризики, що сприяє військовому паритету.

Дослідження [14] присвячено реверсивній логістиці перероблення електронних відходів у воєнних умовах для зниження витрат і дефіциту сировини. Автори роботи [15] запропонували методи прогнозування й оптимізації реверсивної логістики із застосуванням логістичної регресії та дерев рішень для підтримки циркулярної економіки.

У статті [16] проаналізовано загрози реверсивної логістики під час перероблення відпрацьованих автомобілів, а також упроваджено нечіткий аналіз для їх мінімізації. У праці [17] доведено, що сучасна логістика орієнтована на сервіс, стійкі відношення та створення цінності, а не на простий обмін товарами. Роботу [18] присвячено ключовим вузлам мереж зворотної логістики з реманufактурингом, наголошено на недостатності досліджень для медичних і меблевих товарів, а також на ролі нових технологій.

Автори статті [19] запропонували концепцію інтеграції прямої та зворотної логістики в будівництві для циркулярної економіки, зосереджуючи увагу на подоланні бар'єрів у замкненні матеріальних циклів.

З огляду сучасних студій можна зробити висновок про актуальність дослідження інтеграції реверсивної логістики та управління ризиками в умовах воєнного часу, зокрема для виробництва й збуту технічної оборонної продукції. Недоліками проаналізованих підходів є вузька спеціалізація моделей (переважно на окремих етапах або видах продукції), відсутність комплексної уваги до зворотних потоків, ризиків і дефіциту ресурсів в єдиній системі, а також недостатня орієнтація на практичні інструменти сценарного планування й адаптації логістичних рішень у реальному часі. Водночас наявні роботи створюють основу для подальших досліджень, оскільки демонструють значення цифрових технологій, штучного інтелекту й агентного моделювання для підвищення стійкості логістичних систем, особливо в умовах конфлікту.

Мета й завдання роботи

Метою дослідження є підвищення ефективності логістики збуту технічної продукції оборонного призначення за умов воєнного часу способом побудови адаптивної структури логістичних потоків; розроблення математичної моделі оптимізації з огляду на ризики й ресурсні обмеження.

У статті передбачено виконання таких завдань:

1) аналіз особливостей логістики збуту технічної оборонної продукції в умовах бойових дій;

2) розроблення структурної моделі логістичних потоків, зважаючи на прямі та зворотні зв'язки й повторне використання компонентів;

3) побудова математичної моделі оптимізації логістичних процесів відповідно до часових витрат, ризиків затримок і ресурсних обмежень;

4) обґрунтування впровадження сценарного аналізу й адаптивного планування для підвищення стійкості та безперервності постачання в умовах невизначеності.

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

Матеріали й методи

У дослідженні впроваджено комплексний підхід, що поєднує аналіз теоретичних основ логістики збуту технічної оборонної продукції в умовах бойових дій, а також відомості про функціювання відповідних логістичних систем. Основними методами стали системний аналіз логістичних потоків [20], розроблення структурної моделі взаємодії виробничих підприємств, складів, ремонтних і утилізаційних пунктів з огляду на прямі та зворотні зв'язки [21].

Для оптимізації логістичних процесів застосовано математичне моделювання, зважаючи на часові витрати, ризики затримок і ресурсні обмеження. Також використано сценарний аналіз і адаптивне планування [22], що дають змогу забезпечити безперервність і стійкість постачання в умовах високої невизначеності та загроз бойових дій.

Розглянемо особливості побудови й функціювання системи логістики збуту технічної продукції оборонного призначення в сучасних умовах. Збутова логістика такої продукції розрізняється складною структурою, оскільки поєднує прямі й зворотні потоки, потребує високого рівня цифровізації та гнучкості для швидкого реагування на зміну обстановки. Вона має брати до уваги високі вимоги до зберігання, транспортування та обліку, що випливають з особливостей самої продукції: складності конструкції, чутливості до зовнішніх впливів, потреби в гарантійному обслуговуванні та високої вартості компонентів.

Крім того, система має бути стійкою до ризиків, зокрема руйнування транспортної інфраструктури, загроз обстрілів або кібератак, нестачі ресурсів і зміни безпекової ситуації [23]. Для цього передбачаються децентралізовані логістичні пункти, альтернативні маршрути доправлення та гнучке управління запасами, а також упровадження сучасних інформаційних систем, що допомагають контролювати рух продукції в реальному часі.

Ще однією ключовою особливістю є наявність розвиненої реверсивної логістики, що забезпечує повернення техніки з військових частин для ремонту чи утилізації та повторне використання придатних компонентів. Це дає змогу зменшити загальний обсяг витрат і прискорити логістичний цикл, що є критично важливим в умовах інтенсивної експлуатації техніки. Розглянемо детальніше основні етапи функціювання такої логістичної системи.

Першим етапом процесу є організація складування, яке може бути як централізованим на основному складі, так і децентралізованим – у логістичних пунктах ближче до споживача. Це підвищує оперативність доставки, скорочує час реагування на потреби замовника й знижує ризики транспортування в зоні бойових дій. Для ефективного зберігання необхідно дотримуватися відповідних умов (температура, волога, пил), забезпечити фізичну та інформаційну безпеку, охоронну інфраструктуру та під'єднання до логістичних IT-систем. Продукція має бути правильно маркована, упакована та збережена відповідно до технічних регламентів. Упровадження WMS-систем дає змогу вести точний облік, контролювати залишки та оптимізувати маршрути відвантаження.

Наступний етап – оброблення замовлень: отримання заявок, перевірка наявності продукції, формування документів, перевірка комплектності та технічного стану. Цифрове оброблення охоплює інтеграцію із системами електронного документообігу та логістичного контролю. Після цього виконується пакування з огляду на характеристики техніки та підготовка до транспортування [24].

Ключовим етапом є транспортування відповідно до зони доставки, інфраструктури та безпекової ситуації. Найчастіше використовується автомобільний транспорт, але для об'ємних або віддалених вантажів – залізничний чи авіаційний. Необхідно передбачати альтернативні маршрути. Доправлення здійснюється безпосередньо до військових частин або через логістичні хаби. За потреби залучають сторонніх логістичних операторів за умови належного контролю [25].

Після прибуття організовується приймання продукції: перевірка цілісності, відповідності технічним вимогам і підписання акта. За потреби відбувається тестування на місці. Завершальний етап – зворотний зв'язок: збирання відгуків, виявлення недоліків, оновлення інформації в IT-системах для оптимізації наступних циклів.

Реверсивна логістика охоплює повернення техніки: БпЛА – на ремонт, пошкоджені апарати – на утилізацію, придатні компоненти – на повторне використання. В умовах війни цей процес набуває критичного значення через високу частоту зносу техніки. Повернення передбачає реєстрацію, транспортування, діагностику, ремонт або заміну компонентів і повторне впровадження в експлуатацію.

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

Швидкість і ефективність ремонту напряму впливають на бойову спроможність.

Утилізація здійснюється щодо техніки, яка не підлягає ремонту. Вона передбачає транспортування до спеціалізованих пунктів, демонтаж і знищення згідно з екологічними стандартами. Компоненти, що підлягають повторному використанню, обліковують та зберігають.

Зворотна логістика у воєнний час стикається з труднощами: небезпека маршрутів, руйнування інфраструктури, нестача ресурсів і недостатня координація між учасниками. Це призводить до затримок та ускладнень ремонту й утилізації техніки.

Результати дослідження

1. Структурна модель логістичних потоків

Загальна схема логістики збуту продукції оборонного призначення містить такі типи об’єктів (пунктів): виробничі підприємства, склади (підприємства, оптові, волонтерські), пункти ремонту (підприємства, волонтерські, 3D-друку), пункти утилізації. Між вказаними об’єктами мають реалізовуватися як прямі зв’язки, так і зворотні (до пунктів ремонту). Структурна модель логістичних потоків у забезпеченні ЗСУ зображена на рис. 1.

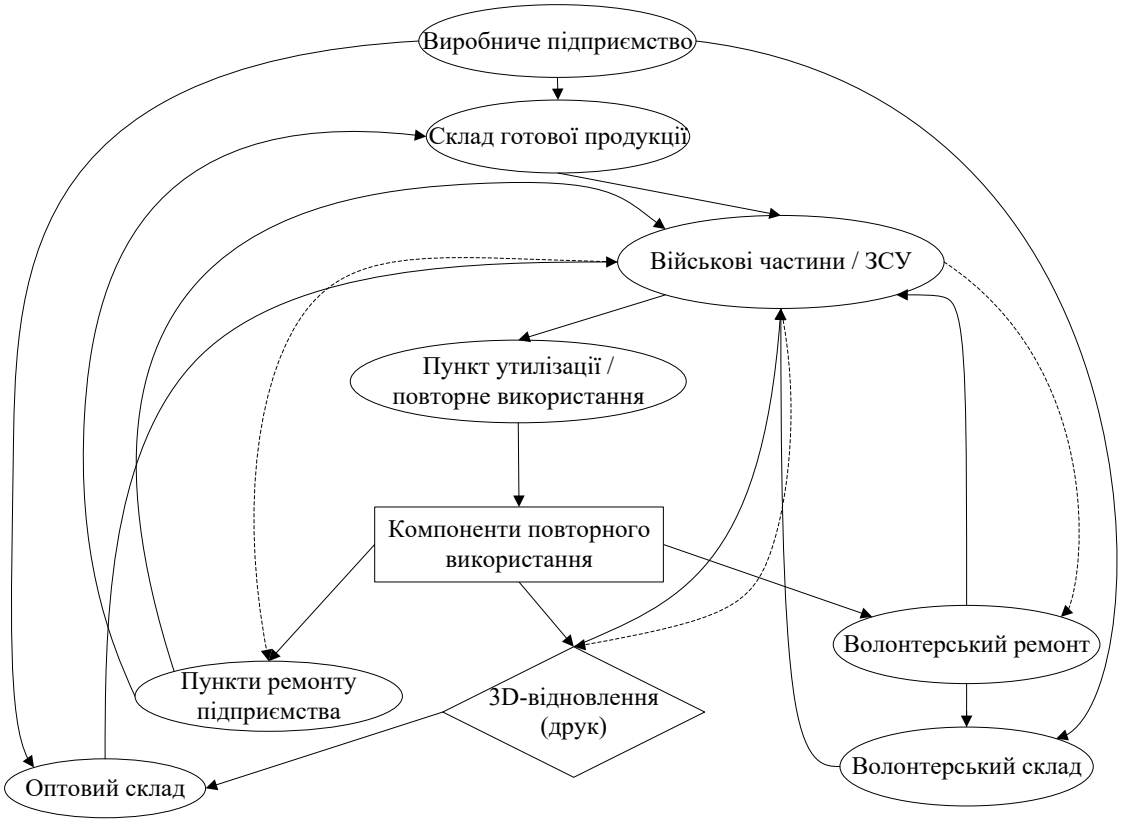


Рис. 1. Структурна модель логістичних потоків у забезпеченні ЗСУ технічною продукцією

Центральне місце в ланцюгу належить виробничому підприємству, що постачає готову продукцію до складів різного типу: складу готової продукції, оптового й волонтерського складів. Зі складів продукція надходить до військових частин, де використовується за призначенням. У процесі експлуатації техніка може частково втратити функціональність або повністю вийти з ладу. У такому разі вона повертається до ремонтних пунктів: виробничих, волонтерських або 3D-відновлення. Після відновлення техніка може бути відправлена як безпосередньо до військових частин, так і на склади для подальшого розподілу.

У разі неможливості відновлення техніка направляється до пунктів утилізації, де її розбирають. Частина непридатних компонентів утилізується, а інша частина, що зберегла функціональність, може бути використана повторно. Водночас система передбачає застосування компонентів, які були вироблені раніше, не потребують нового проєктування

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

та можуть бути застосовані в поточному виробництві або ремонті – так звані компоненти повторного використання. Вони надходять до ремонтних пунктів або до 3D-виробництва, а також можуть інтегруватися у виробничий процес, що забезпечує економію ресурсів і прискорення логістичних операцій.

Особливу роль у схемі відіграють зворотні логістичні зв'язки, позначені пунктирними лініями. Вони відтворюють переміщення техніки з військових частин до ремонтних пунктів або пунктів утилізації, а також циркуляцію компонентів повторного використання між утилізацією, ремонтними точками та виробництвом. Така гнучка структура дає змогу мінімізувати втрати, скоротити час обігу техніки й оперативно реагувати на потреби війська.

В умовах бойових дій ефективність логістичних процесів визначається їх адаптивністю та цифровізацією: відстеженням вантажів у реальному часі, автоматизованим управлінням, обміном інформацією та прогнозуванням потреб. Використання дронів для доправлення техніки до військових частин знижує ризики для персоналу й підвищує мобільність логістичних операцій.

У цьому контексті все більшої актуальності набуває питання скорочення часу логістичного обігу. У сучасних умовах ведення бойових дій завдання оптимізації логістичних процесів у часі стає критично важливою. Основною метою є мінімізація загального часу виконання логістичного циклу – від складування до приймання та реверсивної логістики – за умови забезпечення надійності, безпеки й відповідності технічним і організаційним вимогам.

Особливу увагу в цьому разі необхідно приділити етапу транспортування, адже саме він є найбільш вразливим до впливу зовнішніх ризиків, таких як затримки на маршрутах через руйнування інфраструктури, блокування, зміну безпекової ситуації, а також втрат часу, пов'язаних із переміщенням небезпечними територіями. Так, наприклад, перевезення техніки прифронтовими районами може потребувати попереднього погодження маршрутів, супроводу або доправлення малими партіями, що суттєво впливає на загальний час.

Додатковим чинником, що ускладнює виконання логістичних завдань, є інформаційний складник: затримки в її обробленні, збої в IT-системах, відсутність вчасного оновлення відомостей про запаси або замовлення, а також слабка координація між логістичними ланками можуть призводити до суттєвого збільшення загального часу виконання операцій. У цьому контексті важливою є цифровізація логістичних процесів: застосування WMS-систем, інтеграція з електронним документообігом, автоматизоване відстеження вантажів у реальному часі, що дає змогу зменшити ризики затримок і підвищити адаптивність системи. Наприклад, упровадження RFID-міток для ідентифікації техніки на кожному етапі дає змогу зменшити кількість помилок і покращити прозорість постачання.

2. Математична модель оптимізації часу доправлення в логістичному ланцюзі збуту в умовах ризиків

Оптимізаційна задача передбачає розгляд усіх етапів логістичного ланцюга як послідовних компонентів із відповідними часовими витратами та ризиками затримки. Для математичного моделювання запропоновано використовувати таку формулу обчислення загального часу логістичного процесу з огляду на загрози:

$$T_{total} = \sum_{i=1}^{n}\sum_{j=1}^{n} t_{ij}\left(1+r_{ij}\right)x_{ij},$$

де T_{total} – загальний час логістичного процесу;

t_{ij} – номінальний час транспортування або оброблення вантажу між вузлами i та j;

r_{ij} – коефіцієнт ризику затримки на ділянці $i-j$;

n – загальна кількість логістичних вузлів;

$x_{ij} \in \{0,\ 1\}$ – змінна вибору маршруту.

Ця функція відтворює загальний час проходження логістичного ланцюга, зважаючи на ризики, і є цільовою функцією оптимізаційної задачі, яка мінімізується. Проте за відсутності обмежень задача має тривіальний розв'язок, що не відтворює реальні умови збуту, оскільки формально її мінімум може бути досягнутий за нульового значення змінних x_{ij}, що не відповідає реальним умовам логістики. Тому математична модель доповнюється такими обмеженнями.

1. Обмеження на мінімальний обсяг постачання. Це гарантує, що система логістичних маршрутів задовольнить щонайменше мінімально необхідну потребу в доправленні продукції оборонного призначення. У контексті забезпечення військових частин це може бути визначене значенням $\left(Q_{\min}\right)$,

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

що встановлюється залежно від потреб технічного або оборонного характеру:

$$\sum_{i=1}^{n}\sum_{j=1}^{n} q_{ij} x_{ij} \geq Q_{min},$$

де q_{ij} – імовірний обсяг вантажу, який може бути перевезено за маршрутом $i-j$;

$x_{ij} \in \{0,\ 1\}$ – змінна вибору маршруту;

$Q_{\min}$ – мінімальний необхідний обсяг доставки.

Це обмеження виконує роль гарантії функціональної спроможності логістичної мережі. Воно запобігає вибору лише "найбільш швидких" або "найменш ризикованих" маршрутів, які, однак, не забезпечують необхідного обсягу постачання. Отже, модель змушена долучити маршрути, що забезпечують потрібну пропускну здатність.

2. Обмеження на сукупний ризик. Оскільки в процесі доправлення продукції оборонного призначення до військових частин маршрути можуть суттєво різнитися за рівнем безпеки, важливо обмежити допустимий рівень ризику для всієї логістичної системи. Надмірне накопичення маршрутів із високими ризиками може призвести до зриву постачання або втрати вантажів. Для обмеження загроз обираються маршрути, що задовольняють умову:

$$\max_{i}(1-\prod_{j=1}^{n} r_{ij} x_{ij}) \leq R_{max},$$

де r_{ij} – імовірність ризикових подій на маршруті $i-j$ (наприклад, можливого обстрілу, блокування, пошкодження);

$x_{ij} \in \{0,\ 1\}$ – змінна вибору маршруту;

$R_{\max}$ – гранично допустиме значення сукупного ризику.

Це обмеження діє як фільтр, що не дає змогу долучити в логістичну систему надмірно небезпечні маршрути навіть у разі їх високої швидкості або вантажомісткості. Отже, воно зберігає баланс між ефективністю та безпекою.

3. Обмеження на кількість задіяних маршрутів. Це додаткове обмеження є корисним у разі, коли існують обмеження на кількість одночасно активних маршрутів через нестачу ресурсів: транспорту, персоналу, пального або організаційної спроможності.

$$\sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij} \leq M,$$

де $x_{ij} \in \{0,\ 1\}$ – змінна вибору маршруту;

M – максимальна кількість маршрутів, що може бути активована одночасно.

Це обмеження забезпечує реалістичність моделі у разі, коли ресурси на транспортні операції обмежені, наприклад, через кількість доступних вантажівок, залізничних платформ чи через потребу уникати надмірного розпорошення зусиль.

З практичного погляду воно може бути особливо корисним у застосуванні адаптивного планування, коли доводиться розглядати кілька сценаріїв із різними значеннями M залежно від обстановки на момент прийняття рішень.

3. Застосування сценарного аналізу й адаптивного планування в умовах невизначеності

Запропонована математична модель дає змогу формалізувати задачу вибору логістичних маршрутів з огляду на ключові обмеження на обсяг постачання, рівень ризику та доступні ресурси. Однак у реальних умовах бойових дій ефективність логістичних рішень залежить не лише від початкової оптимізації, а й від здатності системи адаптуватися до змін середовища в режимі реального часу. Постійні загрози руйнування маршрутів, блокування або зміни безпекової ситуації вимагають динамічного підходу до управління логістикою, що передбачає сценарний аналіз і адаптивне планування.

Перевагою запропонованої моделі є можливість оцінювання впливу загроз на загальний час реалізації логістичного процесу. Якщо деякі етапи мають підвищені ризики (наприклад, транспортування через зону бойових дій або оброблення замовлень без цифрової автоматизації), їх внесок у загальний час суттєво зростає. У такому разі доцільно аналізувати альтернативні маршрути або впроваджувати додаткові ІТ-рішення, щоб зменшити коефіцієнти r_i на критичних ділянках.

Оптимізаційна задача зводиться до мінімізації T_{total} способом вибору таких маршрутів, технологій оброблення, способів транспортування та стратегій організації зберігання, що знижують як базовий час t_{ij}, так і ризики затримок r_{ij}.

Отже, оптимізація логістичних процесів у часі в умовах бойових дій має комплексний характер і потребує уваги як до матеріально-технічних, так і інформаційних параметрів. Вона має розв’язуватися

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

в межах адаптивної, динамічної моделі управління, здатної реагувати на зміни обстановки за допомогою цифрових інструментів, прогнозування та мережевої координації між усіма учасниками логістичного ланцюга.

Зважаючи на високу динамічність бойових дій та численні ризики, просте планування логістичних процесів заздалегідь часто є недостатнім. Для ефективного управління часом логістичного циклу необхідно зважати на можливі зміни обстановки та невизначеності, що можуть призвести до раптових затримок або руйнувань критичних ланок. Саме тому важливо застосовувати сценарний аналіз із розробленням альтернативних планів дій, що дає змогу адаптувати логістичну модель до різних рівнів загроз і надзвичайних ситуацій.

Такий підхід реалізується завдяки створенню моделей, які в реальному часі оновлюються з огляду на нові показники, що надходять від інформаційних систем, моніторингу й розвідки. Це допомагає не лише прогнозувати можливі затримки, а й оперативно коригувати маршрути й організацію постачання, забезпечуючи безперервність логістичних процесів навіть у складних умовах бойових дій.

Сценарний аналіз і адаптивне планування є ключовими елементами управління логістичними процесами в умовах військових дій, коли ситуація швидко змінюється, а ризики надзвичайно високі. Для забезпечення безперервності збуту продукції та мінімізації втрат часу розробляються кілька альтернативних планів дій, що відповідають різним рівням загроз. Зазвичай це декілька сценаріїв: план *A* – оптимальний за сприятливих умов, план *B* – із додатковими заходами безпеки, план С – резервний, що передбачає обхідні маршрути в разі руйнування або блокування основних логістичних ланок. Такі плани формуються на основі графових моделей логістичної мережі, де вершинами є склади та інші логістичні вузли, а ребрами – маршрути транспортування. Для кожного сценарію визначається відповідна підмножина маршрутів з огляду на часові витрати й коефіцієнти ризику.

Одним із найважливіших завдань є моделювання випадків раптового руйнування логістичної ланки, наприклад виходу з ладу складу, знищення транспортної інфраструктури або втрати транспортних засобів. У такій ситуації логістична модель оновлюється внаслідок вилучення пошкоджених вузлів або шляхів з графа, після чого перераховується оптимальний маршрут, зважаючи на новий стан системи. Це дає змогу оперативно адаптуватися до змін і продовжувати постачання навіть за значних ускладнень.

Для зручності прийняття рішень результати сценарного аналізу можуть візуалізуватися на географічних картах із виділенням доступних і недоступних маршрутів, оцінкою часу доправлення та рівня ризику на кожному етапі. Візуалізація допомагає операторам швидко орієнтуватися в поточній ситуації та обирати найбільш доцільні варіанти дій.

Отже, сценарний аналіз і адаптивне планування забезпечують комплексний підхід до управління логістикою в складних умовах бойових дій. Вони дають змогу не лише готуватися до можливих негативних сценаріїв, але й швидко реагувати на зміни, підтримуючи стійкість і ефективність логістичних процесів. Упровадження таких підходів у поєднанні із сучасними цифровими технологіями перетворює логістичну систему на живу, самовідновлювану структуру, здатну забезпечувати безперервність постачання навіть у найскладніших умовах.

Для реалізації сценарного підходу розроблено застосунок, що дає змогу будувати маршрути доправлення з огляду на небезпечні ділянки. У побудові маршруту забезпечується мінімізація часу, але основним обмеженням є сукупний ризик (надійність доправлення).

На рис. 2 подано результат побудови маршруту між вказаними пунктами. Серед вхідних показників для моделювання вказується динамічна інформація – рівень небезпеки в кожній з областей України (на рисунку позначено відповідним кольором). Він обчислюється на підставі відомостей про кількість небезпечних подій (обстрілів, влучань ракет, КАБів, дронів тощо) за певний період часу.

Було визначено найкоротший маршрут доправлення продукції оборонного призначення з Ужгорода до Херсона: Ужгород → Львів → Тернопіль → Хмельницький → Вінниця → Кропивницький → Миколаїв → Херсон, довжина якого дорівнює 943 км, приблизний час маршруту становить 13 год 30 хв. Необхідно було брати до уваги проміжні склади у Львові та Кропивницькому. У другому сценарії зважали на ризики маршруту з підвищеним рівнем небезпеки (індекс небезпечності – 0.56) і побудовано більш безпечний маршрут: Ужгород → Львів → Луцьк → Рівне → Житомир → Вінниця → Кропивницький → Миколаїв → Херсон, довжина якого дорівнює 1 114 км, приблизний час доправлення 15 год 50 хв.

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

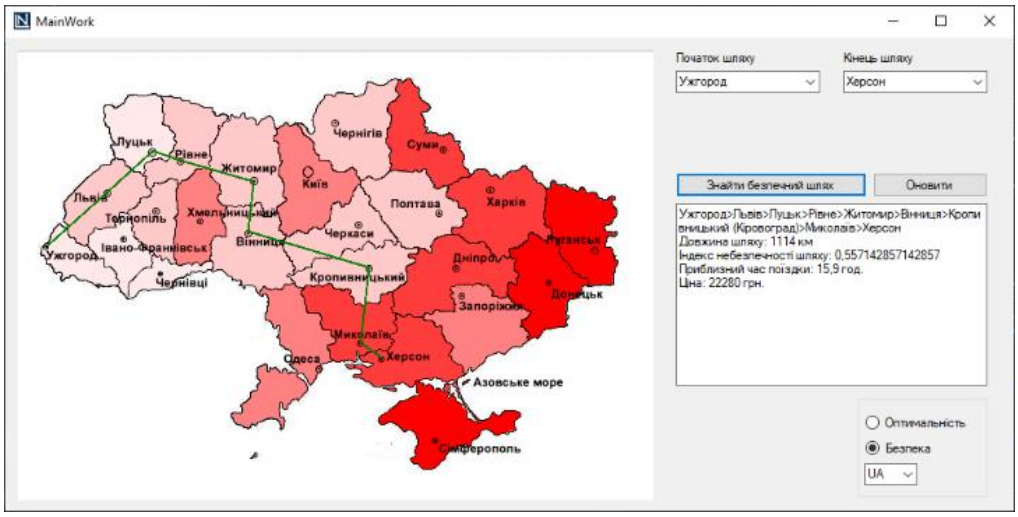


Рис. 2. Приклад побудови маршруту збуту відповідно до рівнів небезпечності ділянок

Отже, обмеження за рівнем ризику зумовило вибір субоптимального рішення щодо часу доправлення.

Висновки

У статті досліджено особливості побудови логістики збуту технічної продукції оборонного призначення в умовах воєнного часу. Показано, що для забезпечення ефективності та стійкості логістичної системи потрібна комплексна мережева структура, яка об'єднує прямі та зворотні потоки, а також високий рівень цифровізації для гнучкого реагування на ризики обстрілів і руйнування інфраструктури.

Розроблено структурну модель логістичних потоків, що містить виробничі підприємства, склади різного типу, ремонтні пункти, пункти утилізації та військові частини. Особливу увагу приділено повторному використанню компонентів, що сприяє зменшенню витрат і підвищенню швидкості обігу техніки.

Для оптимізації логістичних процесів запропоновано математичну модель, яка бере до уваги часові витрати, ризики затримок на різних етапах ланцюга постачання, а також обмеження на мінімальний обсяг постачання, допустимий рівень загрози й доступність ресурсів. Це дає змогу формувати оптимальний набір маршрутів і технологій оброблення з огляду на реальну ситуацію бойових дій.

Наукова новизна дослідження полягає у створенні комплексної адаптивної моделі управління логістикою збуту, яка поєднує оптимізацію часу, управління ризиками та сценарний аналіз з використанням цифрових технологій для моніторингу й оперативного коригування планів у режимі реального часу.

Практичне впровадження розробленої моделі дає змогу підвищити стійкість логістичної системи, скоротити час логістичного циклу та забезпечити безперервність постачання навіть в умовах високої невизначеності й динамічної безпекової ситуації.

Напрямами подальших досліджень є деталізація механізмів взаємодії між учасниками логістичного ланцюга й розроблення цифрових рішень для автоматизації сценарного аналізу з огляду на зворотні цикли.

References

1. De Koster, R., Le-Duc, T., Roodbergen, K. J. (2007), "Design and control of warehouse order picking: A literature review", *European Journal of Operational Research,* Vol. 182, No. 2, P. 481–501. DOI: 10.1016/j.ejor.2006.07.009
2. Savchenko, L., Hryhorak, M. (2018), "Conceptual foundations for the development of reverse logistics in the circular economy", Pryazovskyi Economic Bulletin, P. 14–32. URL: http://www.pev.kpu.zp.ua/journals/2018/5_10_uk/5_10_2018.pdf
3. Vasiliauskas, A. V., Ivanauskas, S., Čižiūnienė, K. (2024), "Challenges of ensuring reverse logistics in a military organization using outsourced services", *Sustainability,* Vol. 16, No. 11, Article 4569. DOI: https://doi.org/10.3390/su16114569
4. Cantelmi, R., Di Gravio, G., Patriarca, R. (2020), "Learning from incidents: A supply chain management perspective in military environments", *Sustainability*, Vol. 12, No. 14, Article 5750. DOI: https://doi.org/10.3390/su12145750
5. Minculete, G. (2025), "Military logistics drones: The innovative solution for transportation challenges on the battlefield", *Review of the Air Force Academy*, P. 342–354. DOI: https://doi.org/10.2478/raft-2025-0033

6. Hres, O. M. (2024), "Optimization of logistic processes for supplying military units during active combat: the Ukrainian experience", *Scientific Bulletin of the International Humanitarian University. Series: Jurisprudence*, No. 71. P. 4–7. DOI: https://doi.org/10.32782/2307-1745.2024.71.1
7. Zahorianska, O., Ziabrev, V., Kyshchyk, D. (2025), "Logistical support of enterprise operations under martial law: problems and ways of optimization", Herald of Khmelnytskyi National University. Economic sciences, Vol. 342, No. 3(2), P. 43–48. DOI: https://doi.org/10.31891/2307-5740-2025-342-3(2)
8. Burkivskyi, O. S. (2025), "Optimization of military logistics under combat conditions using the A* algorithm", *Natural and Technical Sciences*, Vol. 1, No. 17, P. 168–172. DOI: https://jvestnik-sss.donnu.edu.ua/article/view/17323
9. Tereshchenko, S. I., Yevtushenko, A. M. (2023), "Supply chain: management and optimization", *Journal of Strategic Economic Studies*, No. 6(17), P. 207–214. DOI: https://doi.org/10.30857/2786-5398.2023.6.21
10. Havrys, P., Nesterenko, R., Havrys, M. (2022), "Prospects and logistics support for the creation of high-precision weapons in a wartime economy", *The collection of scientific works of the Military Institute of Internal Troops of the Ministry of Internal Affairs of Ukraine*. Vol. 1, No. 39. P. 89–97. DOI: https://doi.org/10.33405/2409-7470/2022/1/39/263373
11. Dachkovskyi, V. O., Sampir, O. M. (2019), "Algorithm of the functioning of the logistics support system", *Modern Information Technologies in the Sphere of Security and Defence*, No. 2(35), P. 87–92. DOI: https://doi.org/10.33099/2311-7249/2019-35-2-87-92
12. Havryliuk, I. Yu., Matsko, O. Y., Dachkovskyi, V. O. (2019), "Conceptual foundations of flow management in the logistics support system of the Armed Forces of Ukraine", *Modern Information Technologies in the Sphere of Security and Defence*, Vol. 34, No. 1, P. 37–44. DOI: 10.33099/2311-7249/2019-34-1-37-44
13. Fedorovych, O. Ye., Uruskyi, O. S., Pronchakov, Yu. L., Rybka, A. V., Leshchenko, Yu. O. (2022), "Modeling logistics for achieving military parity of forces in combat zones", *Aerospace technic and technology*, No. 4. DOI: https://doi.org/10.32620/aktt.2022.4.07
14. Pereira, N., Antunes, J., Barreto, L. (2023), "Impact of management and reverse logistics on recycling in a war scenario", *Sustainability*, Vol. 15, No. 4, Article 3835. DOI: 10.3390/su15043835
15. Abba Dabo, A.-A., Hosseinian-Far, A. (2023), "An integrated methodology for enhancing reverse logistics flows and networks in Industry 5.0", *Logistics*, Vol. 7, No. 4, Article 97. DOI: 10.3390/logistics7040097
16. Omosa, G. B., Numfor, S. A., Kosacka-Olejnik, M. (2023), "Modeling a reverse logistics supply chain for end-of-life vehicle recycling risk management: a fuzzy risk analysis approach", *Sustainability*, Vol. 15, No. 3, Article 2142. DOI: 10.3390/su15032142
17. Carvalho, J. C. de, Vilas-Boas, J., O'Neill, H. (2014), "Logistics and supply chain management: an area with a strategic service perspective", *American Journal of Industrial and Business Management*, Vol. 4, No. 1. DOI: 10.4236/ajibm.2014.41005
18. Zhang, X., Zou, B., Feng, Z., Wang, Y., Yan, W. (2022), "A review on remanufacturing reverse logistics network design and model optimization", *Processes,* Vol. 10, No. 1, Article 84. DOI: 10.3390/pr10010084
19. Ding, L., Wang, T., Chan, P. (2023), "Forward and reverse logistics for circular economy in construction: a systematic literature review", *Journal of Cleaner Production*, Vol. 388, Article 135981. DOI: 10.1016/j.jclepro.2023.135981
20. Bezkorovainyi, V., Binkovska, A., Noskov, V., Gopejenko, V., Kosenko, V. (2025), "Adaptation of logistics network structures in emergency situations", *Advanced Information Systems*. Vol. 9, No. 4, P. 39–50. DOI: https://doi.org/10.20998/2522-9052.2025.4.06
21. Malyeyeva, O., Malieieva, Yu., Kosenko, V., Artiukh, R. (2020), "Formalized Models of Processes and Optimization of Indicators for Complex Equipment Recycling Project", *2020 IEEE International Conference on Problems of Infocommunications. Science and Technology (PIC S&T),* Kharkiv, Ukraine, P. 583–586. DOI: 10.1109/PICST51311.2020.9467933
22. Petrovska, I., Kuchuk, H. (2023), "Adaptive resource allocation method for data processing and security in cloud environment", *Advanced Information Systems*. Vol. *7*, No. 3. P. 67–73, DOI: https://doi.org/10.20998/2522-9052.2023.3.10
23. Bezkorovainyi, V., Kolesnyk, L., Gopejenko, V., Kosenko, V. (2024), "The method of ranking effective project solutions in conditions of incomplete certainty", *Advanced Information Systems*, Vol. 8, No. 2, P. 27–38. DOI: https://doi.org/10.20998/2522-9052.2024.2.04
24. Díaz-Madroñero, M., Peidro, D., Mula, J. (2015), "A review of tactical optimization models for integrated production and transport routing planning decisions", *Computers & Industrial Engineering,* Vol. 88, P. 518–535. DOI: 10.1016/j.cie.2015.06.010
25. Litvinenko, D., Malyeyeva, O. (2020), "Models of stakeholders management at the stages of the life cycle of projects of transport systems' development", *Radioelectronic and Computer Systems*, No. 3, P. 97–107. DOI: 10.32620/reks.2020.3.10

Надійшла (Received) 25.05.2024
Accepted for publication (Прийнята до друку) 30.11.2025
Publication date (Дата публікації) 28.12.2025

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

Відомості про авторів / About the Authors

Полупан Юрій Володимирович – Національний аерокосмічний університет "Харківський авіаційний інститут", аспірант кафедри комп'ютерних наук та інформаційних технологій, Харків, Україна; e-mail: yuriypolupan6@gmail.com; ORCID: http://orcid.org/0009-0009-3030-8448; Scopus ID: https://www.scopus.com/authid/detail.uri?authorId=60146024600

Малєєва Ольга Володимирівна – доктор технічних наук, професор, Національний аерокосмічний університет "Харківський авіаційний інститут", професор кафедри комп'ютерних наук та інформаційних технологій, Харків, Україна; e-mail: o.maleyeva@khai.edu; ORCID: http://orcid.org/0000-0002-9336-4182; Scopus ID: https://www.scopus.com/authid/detail.uri?authorId=60146226200

Polupan Yuriy – PhD Student at the Department of Computer Science and Information Technologies, National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine.

Malyeyeva Olga – Doctor of Sciences (Engineering), Professor, National Aerospace University "Kharkiv Aviation Institute", Professor at the Department of Computer Science and Information Technologies, Kharkiv, Ukraine.

MODELS OF DISTRIBUTION LOGISTICS OF TECHNICAL DEFENSE PRODUCTS UNDER WARTIME RISK CONDITIONS

The subject of this article is the processes of organization and optimization of the distribution logistics of technical defense products under wartime conditions, in particular, the development of a network logistics structure considering reverse flows, risks, and limited resources. **The purpose** of the study is to improve the efficiency of distribution logistics of technical defense products in wartime by constructing an adaptive structure of logistics flows and developing a mathematical optimization model that accounts for risks and resource constraints. The article addresses the following **tasks**: analysis of the specific features of distribution logistics for technical defense products in combat conditions; development of a structural model of logistics flows that considers direct and reverse links as well as the reuse of components; construction of a mathematical model for optimizing logistics processes with consideration of time costs, delay risks, and resource constraints; justification of the application of scenario analysis and adaptive planning to enhance the resilience and continuity of supply under uncertainty. The following **methods** are applied: a systems approach, structural decomposition of logistics processes, risk-based approach, constrained mathematical modeling, scenario analysis, and adaptive planning. **The results** obtained include the construction of a structural model of logistics flows for supplying the Armed Forces of Ukraine, which takes into account direct and reverse flows as well as component reuse; the development of a mathematical model for optimization of distribution logistics with an objective function of minimizing time and constraints on supply volume, total risk, and the number of routes. It is demonstrated that risk consideration significantly affects the choice of optimal routes and strategies. The feasibility of using scenario analysis and adaptive planning to ensure continuity of supply in combat conditions is substantiated. **Conclusions.** The proposed structural and mathematical models make it possible to reduce the logistics cycle time, increase the resilience of the logistics system to risks, and ensure rapid response to changing circumstances. The use of reverse logistics and component reuse helps reduce costs and mitigate resource shortages. Scenario analysis and adaptive planning are effective tools for managing the distribution logistics of technical products under uncertainty and wartime risks.

Keywords: distribution logistics; technical products; supply of the Armed Forces of Ukraine; structural model; mathematical model; optimization; risk management.

Бібліографічні описи / Bibliographic descriptions

Полупан Ю. В., Малєєва О. В. Моделі логістики збуту технічної продукції оборонного призначення в умовах ризиків воєнного часу. *Сучасний стан наукових досліджень та технологій в промисловості.* 2025. № 4 (34). С. 68–77. DOI: https://doi.org/10.30837/2522-9818.2025.4.068

Polupan, Y., Malyeyeva, O. (2025), "Models of distribution logistics of technical defense products under wartime risk conditions", *Innovative Technologies and Scientific Solutions for Industries*, No. 4 (34), P. 68–77. DOI: https://doi.org/10.30837/2522-9818.2025.4.068

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

UDC 004.412.032:004.056:004.77 DOI: https://doi.org/10.30837/2522-9818.2025.4.078

V. TKACHOV, I. RUBAN

INTEGRAL SURVIVABILITY METRIC OF AN INFORMATION SYSTEM ON A MOBILE PLATFORM UNDER FUNCTIONAL CASCADING AND SECONDARY FAILURES

The subject of the study is an integral metric of the survivability of an information system on a mobile platform under intermittent connectivity, partial observability, and cascading and secondary failures. The system is presented as a multilevel "data–processes–resources" structure. **The goal of the work** is to develop an integrated survivability metric that takes into account time deviations from requirements, their propagation through a dependency graph, and hidden violations; to propose a single-pass algorithm and prove its properties on scenarios. The article solves the following **tasks**: to formalize service requirements and structure projection; to build a metric with risk-oriented aggregation of deficits, cascading correction, and systematic consideration of secondary failures; to develop single-pass calculations in "availability windows"; to prove monotonicity, scale invariance, and resistance to omissions; to define rules for parameter tuning and configuration comparison; to perform experimental verification and comparison with baseline indicators. The following **methods** are used: projection of services at the level of data, processes, and resources; use of conditional average excess as risk-oriented aggregation; cascading correction by depth and width; organization of secondary failures and desynchronization fixation; normalization in "availability windows"; single-pass updates close to linear complexity. The following **results** were obtained: an integrated metric of survivability was proposed and formally defined; its monotonicity in terms of parameters, boundedness, invariance to scaling, and resistance to omissions were proven; the difference from the average deficit was shown – the proposed metric amplifies the contribution of rare deep failures and responds to cascading, while the average values are almost constant; in the absence of service deficits, a positive level is maintained due to the detection of latent secondary failures; scenarios yield consistent families of curves and a three-dimensional surface that demonstrate controllable sensitivity tuning and stable ranking of configurations for industrial mobile platform operating conditions. **Conclusions:** The proposed metric provides a service-consistent assessment of the state of the information system, while taking into account time deficits, cascading propagation, and secondary failures. It is suitable for sequential computing in resource-constrained environments, enhances early risk detection, and supports monitoring, localization, and survivability.

Keywords: integrated metric; survivability; information system; cascading fault; risk-oriented monitoring.

Introduction

Over the past years, information systems on mobile platforms have increasingly found functional applications in industry – autonomous mobile robots in manufacturing, inspection drones for oil, gas, and energy infrastructure, and field service teams with onboard computing – operate in a changing environment with intermittent connectivity, partial observability, and severe constraints on computing and energy resources [1–3]. Data exchange takes place within the windows of availability of information system elements, processing queues and routing are reorganized in dynamic conditions of a rapidly changing external environment, and data and process synchronization is complicated by the absence of stable data transmission channels. Under such conditions, the failure of a single element of the information system rarely remains local: the disruption spreads through chains of dependencies "data–processes–resources", changing the quality of services at adjacent nodes and levels [4]. The complex interaction of individual elements of an information system or entire subsystems amplifies cascading effects and generates new requirements for risk management. Therefore, there is a need to develop a service-semantic survivability metric that aggregates time deficits from the service level objective (SLO) while taking into account rare deviations and inter-level dependencies in the information system architecture.

In addition to primary failures (hardware or channel), cascade effects often occur when the unavailability or quality degradation of one element of the information system induces degradation in other elements due to structural dependencies [5]. Additionally, secondary functional failures are observed – situations when the service deteriorates without physical failure of the mobile platform components, for example, due to violations of the cause-time consistency of data, accumulation of outdated information, shifts in priorities, or breaks in service chains. The cumulative effect of such phenomena manifests itself at the level of service

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

indicators and directly affects the system's ability to perform its target functions within the allotted time limits.

Well-known indicators of reliability [6], availability [7], average downtime of an information system [8], and average quality deficits [9] are classic approaches to forming the concept of information system survivability, but they do not reflect the integral picture in the presence of cascading and secondary effects. Thus, average values "smooth out" short but critical failures, and binary assessments of "function performed/not performed" do not take into account the depth and duration of failures. Graph models of fault propagation are usually not combined with service level requirements due to the excessive complexity of working with them in the process of solving more complex failure prediction problems. In addition, given the mobility of the platform, these limitations are exacerbated by interruptions in internal system connections in horizontal or vertical planes (fragmentary observations, confirmation delays) and system resource constraints.

In this regard, there is a scientific challenge to develop and study an integrated metric of information system survivability on a mobile platform that is applicable at the service level and takes into account the temporal structure of deviations from specified requirements, quantitatively accounts for the spread of violations along the dependency graph and induced (secondary) degradations, is capable of identifying rare but significantly harmful deviations, and is suitable for calculation on a stream of events in a mode of limited resources and unstable communication between internal elements.

In general terms, the task can be interpreted as follows. There is a set of services with specified quality requirements. For each service, time-based quality values are observed and events are recorded that reflect the structure of dependencies between elements of the information system. It is necessary to determine a coordinated integral assessment that aggregates temporally weighted deficits relative to requirements, adjusts them to account for the depth and breadth of the spread of violations, and amplifies the contribution of peak failures. The resulting assessment should be used to compare configurations, select localization/recovery policies, and monitor the status of information system services in dynamic conditions.

Statement of the problem

An information system on a mobile platform with a multi-level structure of "data–processes–resources" is considered. In this representation, services are an observable cross-section of this structure. SLOs are set for services, and disruptions propagate through chains of dependencies within the "data–processes–resources" structure. An information system on a mobile platform is characterized by unstable connections between elements, partial observability, and tight resource budgets for processing, storage, and data exchange channels. Local disruptions cause cascading effects (spreading through the dependency structure) and secondary functional failures (induced degradation, such as data desynchronization, priority shifts, and service chain breaks). In addition, information system indicators may malfunction when typical averaged or binary indicators do not reflect the integral impact of the situations described above.

Therefore, an integrated service-level survivability metric is needed that aggregates temporally weighted deviations from SLOs, corrects them for the spread of failures in the "data–processes–resources" model and the contribution of secondary failures, amplifies the impact of rare but critical episodes through conditional value at risk (CVaR), and is calculated sequentially in real time with incomplete observability.

The task can be solved by decomposing it into subtasks and solving each of them step by step:

– formalization of service indicators, SLO, time weights, and service projection at the data, process, and resource levels;

– description of dependencies in the "data–processes–resources" model and event classes (cascades, secondary failures) with signs of their detection;

– design of risk-sensitive aggregation of deviations (conditional value at risk) with normalization and prioritization of services;

– development of a tool for correcting the spread of violations, reflecting the depth/width of the cascade and marking secondary effects;

– development of a single-pass algorithm for calculating metrics on the event stream, resistant to data omissions and/ing the peculiarities of availability windows.

– research of metric properties and its differences from baseline indicators in representative scenarios.

Analysis of recent studies and publications

Recent years have seen the emergence of new approaches to solving this problem in four areas: developing metrics and approaches to the survivability of information systems in cyberspace, modeling cascading

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

failures in dependency networks, ensuring SLO-level requirements in cloud, microservice, and fog computing environments, and ensuring the survivability of information systems with unstable connections between elements.

Thus, in [10], a fault tolerance metric for UAV swarms is proposed, based on the relationship between the speed of reaching a new coordinated state and the spectral characteristics of the graph connectivity (algebraic connectivity) that describes the information system. The approach focuses primarily on the control dynamics and topology of the UAV swarm, rather than on system service deficits, and does not model secondary failures at the data or process level. Thus, the work [10] lacks integration with SLO requirements and risk-sensitive aggregation of deviations.

Studies devoted to the analysis of cascading failures in interdependent information systems [11], using the example of data exchange networks, describe the mechanisms of failure propagation in directed and undirected interdependent graphs and systematically classify system vulnerabilities at different levels. In study [11], the model correctly describes the structural propagation of failures. At the same time, the connection with service semantics and a unified numerical metric of service status, which simultaneously takes into account the depth, duration, and cascading amplification of deviations from SLO, is not considered in study [11].

The approach to building the architecture of distributed systems that are tolerant to delays and frequent communication interruptions in environments with unstable data transmission channels (relevant to mobile platforms) is analyzed [12]. The advantages and limitations of such architectures for processing and transmitting data with high delays and significant losses are shown. At the same time, the indicators used remain transport-network-based, and the single service metric of survivability, which integrates cascade/secondary effects and the risk of "heavy tails", is considered superficially in [12] and does not provide a complete picture of its applicability.

An overview of the principles of cyber resilience of information systems systematizes organizational and technical strategies, tools, and approaches to building and evaluating indicators [13]. The need for operationally viable quantitative characteristics that link service level with system behavior during failures and attacks is emphasized. At the same time, the results are generalized and difficult to apply in applied systems.

An analysis of studies related to the use of SLA/SLO indicators for cloud environments, in particular in [14], shows the generality of approaches to formalizing the semantics of service requirements. The main focus is on centralized cloud scenarios and binary availability/unavailability indicators. Risk-sensitive tail allocation (in particular, through conditional value at risk, CVaR) and the connection with cascade effects in complex dependencies are considered superficially.

An integrated quantitative framework for assessing the stability of multi-level cyber-physical systems is presented in [15]: the cyber level is described by probabilistic anomalies (Markov chains), and the physical level is described by models of degradation and recovery of performance. The approach [15] includes parameter normalization and coordination of cyber-physical subsystems with the construction of stability curves (time profiles of performance changes under the influence of disturbances) for different scenarios. The framework demonstrates a consistent quantitative assessment procedure with clear steps and output in the form of an interpreted curve/index; However, direct modeling of SLO deficits, communication interruptions, and secondary failures in service-oriented information systems in [15] is not the main goal, although the results are interpreted from the perspective of service semantics.

The analysis [10–15] made it possible to identify systemic gaps – the lack of a single service-oriented integrated metric that takes into account temporally weighted deviations from SLO, cascading propagation, and secondary failures under conditions of intermittent communication; to clarify the requirements for the desired assessment (normality, sensitivity to deviation "tails", resistance to incomplete observations); make a formal definition of an integrated survivability metric, introduce the concept of a corrective rule for the propagation of failures, a procedure for evaluating the flow of events, and a set of representative scenarios for comparison with baseline indicators. Taken together, this justifies the feasibility of building an integral metric of information system survivability on a mobile platform, taking into account cascading and secondary functional failures.

Terminological basis and formalization of the system of level quality indicators and interlevel invariants

The study will use the presentation of an information system on a mobile platform as a multi-level structure of "data–processes–resources" [16]. To construct an integrated survivability metric, we propose to establish the following set of tools: an oriented

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

graph of logical dependencies between objects of three levels, which determines the paths of propagation of violations and the contexts of secondary failures; event logs of failures and recoveries and time series of quality indicators aggregated in availability windows; a service slice as a reflection of each service on a subset of data, processes, and resources with SLO requirements assigned to services; inter-level invariants of cause-time consistency, the violation of which is marked as a secondary failure; weighting coefficients for the importance of services and levels, and normalization rules that enable configuration comparisons. These tools are used as input ports for metrics, i.e., level indicators form the basic contributions, graphs and event data form the corrective component, and the service view forms the interpretation of the result in relation to the requirements imposed on the information system or its individual elements. Subsequently, these tools simultaneously serve as inputs for an integrated assessment – level measurements and invariant violations form its basic and corrective parts.

Next, it is necessary to determine the observed indicators at the "data–processes–resources" levels and inter-level invariants that mark secondary failures and set the context for cascading propagation. It is assumed that the triad "data–processes–resources" records observable temporal indicators and conditional event logs. Then the service slice is a reflection of each service on subsets of data, processes, and resources, and SLOs are used as reference boundaries when causal dependencies and violations occur at all levels of the information system. The assessment is carried out in availability windows (intervals where data is guaranteed to be consistent), between which observation gaps are allowed.

Data level indicators include: timeliness (the time the data exists within the window), causal-temporal consistency (no "jumping" of events and conflicting timestamps), completeness (the proportion of useful data), and version stability (no collisions in the presence of duplicates). In real-world conditions, acceptability thresholds are also set, which, for example, determine the maximum allowable lifetime of data for a single version and tolerances for time shifts between related events.

Process level indicators include: timeliness of execution (proportion of operations completed within the acceptable time), delay (quantiles of execution time), goal achievability (proportion of successfully completed transactions), queue status (average length and spikes), routing stability (frequency of process chain reorganization between information system elements).

Resource level indicators include: quota utilization (proportion of allocated budgets), overload level (budget overruns within the window), availability fluctuations (short dips or anomalies).

Using groups of triad indicators, it is possible to form conditions of correctness (invariants), the violation of which we interpret as secondary functional failures in the information system:

Inv_T – time consistency invariant (Inv_T). The point is that if a process uses data with a timestamp t_d, the completion of the associated operation with the timestamp t_s must satisfy the logic t_s and t_d, and $t_s - t_d \le \Delta_{\max}$ in critical intervals. Otherwise, we have a synchronization violation;

– resource load preservation invariant ( ), based on the logic that in each time window, the volume of incoming requests is equal to the sum of processed, irretrievably lost, and in critical intervals. Otherwise, we have a synchronization violation;

– resource load preservation invariant (Inv_R), based on the logic that in each time window, the volume of incoming requests is equal to the sum of processed, irretrievably lost, and queue growth requests. A systematic deviation from this balance indicates either telemetry errors or hidden failures that are not recorded;

– resource budget invariant (Inv_B), which uses logic whereby all process requests within the window do not exceed the available resource (taking into account reserves), and their regular exceeding is an indicator of induced degradation;

– invariance of dependency continuity (Inv_I), when deviations in the "parent" element or level of the information system must be reflected in the dependent element (level) with no more than a specified delay, therefore the absence of reflection indicates a break in the process chain;

– recovery invariant Inv_{rec}, when, after a recovery event, the level indicators return to the acceptable range no later than a predetermined time. Exceeding this time is an indicator of a hidden secondary problem in the information system.

Violation of any of the invariants $(\mathrm{Inv}_T$, Inv_R, Inv_B, Inv_I, $\mathrm{Inv}_{rec})$ is recorded as a secondary failure event with reference to the level and dependencies within which it was detected. The set of level indicators and invariants forms the observable basis for the further

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)
Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

introduction of an integral metric of system survivability: level indicators determine basic deviations, invariants form signs of induced degradation, and the dependency diagram provides the context for taking into account cascading propagation.

Development of an integrated survivability metric

Let us assume that a multi-level information system operates on the observation horizon $[t_0, t_1]$. Let us introduce the notation of a set of services $S' = \{S_i\}_{i=1}^{m}$, time windows of availability $\{W_r'\}_{r=1}^{R}$ (intervals within which observations are consistent), and an oriented graph of dependencies $G = (V, E)$ between levels of the information system (nodes V – abstractions of "data/processes/resources", edges E – causal/technological connections). For each service S_i, a service level requirement (SLO) is specified – a threshold function $\theta_i(t)$ and a quality indicator $q_i(t)$ (both values are normalized in the interval $[0,1]$).

Let's formalize the basic deviations in the process of performing the main function by the information system. The service deficit can be presented as:

$$d_i(t) := [\theta_i(t) - q_i(t)]_+ = \max\{0, \theta_i(t) - q_i(t)\} \in [0,1]. \quad (1)$$

At (1) $d_i(t) = 0$ – if SLO are satisfied, and $d_i(t) > 0$ – if there are deviations.

Next, we formalize the representation of temporal weighting. Let $w(t) \geq 0$ be a time priority profile (a penalty multiplier for a deficit at time t). It will be greater in critical intervals $w(t)$ and smaller outside them. To suppress short-term fluctuations, we will apply a linear smoothing filter W with a non-negative and normalized kernel $k(\tau)$:

$$(W * x)(t) = \int_0^\infty k(\tau) x(t-\tau) d\tau, \quad k(\tau) \geq 0, \quad \int_0^\infty k(\tau) d\tau = 1. \quad (2)$$

Based on (2), a temporally weighted deficit is formed:

$$x_i(t) := w(t)(W * d_i)(t). \quad (3)$$

Next, the concepts of a cascade corrector and a secondary failure indicator are introduced. Since the previously specified oriented dependency graph G defines the paths of propagation of failures between levels of the information system, it is advisable to formalize the adjacency matrix, which describes the non-negative weights on the arcs, in the form $A \in \mathbb{R}_+^{|V| \times |V|}$, and the service projection in the form of a binary vector $\Pi_i \in [0,1]^{|V|}$, which marks the elements V involved in the service S_i.

To formalize the cascade corrector, it is first necessary to formulate the logic of cascade amplification. Let $\lambda \in (0,1]$ be the propagation depth parameter, then we define the cascade operator as:

$$\Phi_\lambda := \sum_{\ell=0}^{\infty} \lambda^\ell A^\ell = (I - \lambda A)^{-1}, \quad \rho(\lambda A) < 1, \quad (4)$$

where $\rho(\bullet)$ is the spectral radius; A is the adjacency matrix of the dependency graph.

In (4), the operator Φ_λ accumulates the contributions of all oriented paths in G with attenuation along the length (depth of the cascade).

Next, we form a cascade-corrected deviation vector (corrector). Let $z(t) \in \mathbb{R}_+^{|V|}$ be the level-localized deficits (matched to $[0,1]$) on the elements V originating from service deviations $d_i(t)$ through a known matching mechanism (for example, distribution of service deficit across the involved nodes [17]). Then the corrector is set as:

$$\tilde{z}(t) := \min\{\Phi_\lambda z(t), \mathbf{1}\}. \quad (5)$$

After correction via Φ_λ, some components may exceed $\mathbf{1}$. Therefore, in (5) we apply a component-wise upper bound of one (projection onto the hypercube $[0,1]^{|V|}$) using the unit vector $\mathbf{1}$ to ensure that all components belong to the interval $[0,1]$.

We form the secondary failure indicator as follows. Let $I' = \{I_u\}_{u=1}^{U}$ be a set of inter-level invariants Inv_T, Inv_R, Inv_B, Inv_I, Inv_{rec}. Then the degree of violation of the invariant I_u is denoted as:

$$\sigma_u(t) \in [0,1] \quad (6)$$

If $\sigma_u(t) = 0$, then no violation is detected. Therefore, taking into account the interval (6), the secondary failure indicator is as follows:

$$\sigma(t) = (\sigma_u(t))_{u=1}^{U}. \quad (7)$$

Next, we move from level-localized deficits to risk-oriented systemic aggregation with time weighting and CVaR. For the service S_i, we will highlight the

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

part of the cascading corrected deviations that falls on the nodes of its projection:

$$\tilde{d}_i(t) := \frac{\langle \Pi_i, \tilde{z}(t) \rangle}{\max\{1, \langle \Pi_i, \mathbf{1} \rangle\}} \in [0,1], \qquad (8)$$

where $\langle \bullet, \bullet \rangle$ is a scalar product; Π_i is the projection of the service onto the elements V.

To highlight rare cases, such as "heavy tails," it is necessary to consider discretization by availability windows $\{W'_r\}_{r=1}^{R}$. Let:

$$Y_{i,r} := \int_{W'_r} w(t)\left(W * \tilde{d}_i\right)(t)\,dt \in \left[0, |W'_r|\right], \qquad (9)$$

and the distribution on $\{Y_{i,r}\}$ will be determined by specific weights $\omega_r := |W'_r| \Big/ \sum_{q=1}^{R} |W'_q|$ or another agreed rule (for example, proportionally to the number of events). Then $\{Y_{i,r}\}$, taking into account (9), form a sample for risk-oriented aggregation.

For risk-oriented allocation of "tails" in the sample $\{Y_{i,r}\}$, we further introduce the metric of conditional average excess (CVaR) at the confidence level $\alpha \in (0,1)$, which estimates the average value from the largest $(1-\alpha)\%$ observations:

$$\mathrm{CVaR}_\alpha(Y_i) = \min_{u \in \mathbb{R}} \left\{ u + \frac{1}{1-\alpha} \sum_{r=1}^{R} \omega_r \left[Y_{i,r} - u \right]_+ \right\}. \qquad (10)$$

In the context of system integration, the vectors $\sigma(t)$ indicate secondary failures, and their integral effect in the window W' can be estimated by:

$$\Sigma_r = \int_{W'_r} \langle \eta, \sigma(t) \rangle\, dt, \quad \eta \in \mathbb{R}_+^U, \qquad (11)$$

where η are the weights of the significance of the invariants Inv_T, Inv_R, Inv_B, Inv_I, Inv_{rec}.

Therefore, taking into account (5, 8, 9, 11), the following formalization of the integral metric of survivability ISM_α is proposed:

$$\mathrm{ISM}_\alpha = \sum_{i=1}^{m} \beta_i\, \mathrm{CVaR}_\alpha\left(\{Y_{i,r}\}_{r=1}^{R}\right) + \underbrace{\gamma \sum_{r=1}^{R} \omega_r \Sigma_r}_{\text{secondary crops}}, \qquad (12)$$

where $\beta_i \geq 0$ is the weight of service importance; $\lambda \in [0,1)$ is the cascade depth parameter in the graph G; γ is the weight of secondary failures (contribution of invariant violations).

In (12) $\sum_{i=1}^{m}(\bullet)$ – risk-oriented service component (through CVaR from temporally weighted deficits (3), already corrected by a cascade operator). $\sum_{r=1}^{R}(\bullet)$ – system penalty for secondary failures, independent of a specific service (because secondary nature is an inter-level phenomenon). Accordingly, we assume that the lower the value of ISM_α, the higher the survivability of the information system.

Next, the study considers the formalization of the main properties of ISM_α (12) with a demonstration of their applicability. As a basis for constructing proofs of properties, it is advisable to first introduce our own statements, which will be used as an evidence base.

Lemma 1 on the monotonicity of non-negative operators. If $x(t) \leq y(t)$ for all t, then:

– $w(t)x(t) \leq w(t)y(t)$;

– $(W * x)(t) \leq (W * y)(t)$;

– for $u, v \geq 0$ and the matrix $M \geq 0$, we have $Mu \leq Mv$;

– for $\Phi_\lambda = \sum_{\ell \geq 0} \lambda^\ell A^\ell$, $A \geq 0$: $\Phi_\lambda u \leq \Phi_\lambda v$.

Each point of Lemma 1 follows from the fact that all operations with non-negative coefficients preserve order: multiplication by a non-negative function does not reduce the smaller part relative to the larger one; smoothing with a non-negative kernel is averaging with non-negative weights, which also preserves order; multiplication by a matrix with non-negative elements is a non-negative linear combination of components and therefore does not reduce values; and the sum of such monotonic mappings with non-negative coefficients remains monotonic.

Lemma 2 formalizes the basic properties of CVaR for a finite weighted sample:

– monotonicity: if $y_r \leq \hat{y}_r$ for all r, then $\mathrm{CVaR}_\alpha(\{y_r\}) \leq \mathrm{CVaR}_\alpha(\{\hat{y}_r\})$;

– shift: $\mathrm{CVaR}_\alpha(\{y_r + c\}) = \mathrm{CVaR}_\alpha(\{y_r\}) + c$ for any stable c;

– homogeneity: for

$$c \geq 0,\ \mathrm{CVaR}_\alpha(\{c\, y_r\}) = c\, \mathrm{CVaR}_\alpha(\{y_r\});$$

– resistance to small changes: if $|y_r - \hat{y}_r| \leq \varepsilon$ for all r, then $\left|\mathrm{CVaR}_\alpha(\{y_r\}) - \mathrm{CVaR}_\alpha(\{\hat{y}_r\})\right| \leq \varepsilon$.

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

The known properties follow from the equivalent representation (10) and the monotonicity/convexity of the mapping $y \mapsto [y-u]_+$.

Taking into account the introduced lemmas, the following properties of ISM_α are formulated:

1. Zero normalization property. If $d_i(t) \equiv 0$ for all i, t and $\sigma_i(t) \equiv 0$, then $\mathrm{ISM}_\alpha = 0$. The base level without deviations gives a zero metric.

Proof: For $d_i(t) \equiv 0$, for all i, each element of the construction is zero $z(t) \equiv 0 \Rightarrow \Phi_\lambda z(t) \equiv 0 \Rightarrow \tilde{z}(t) \equiv 0 \Rightarrow \tilde{d}_i(t) \equiv 0 \Rightarrow Y_{i,r} \equiv 0 \Rightarrow \mathrm{CVaR}_\alpha = 0$. The second component is also zero $\sigma(t) \equiv 0 \Rightarrow \Sigma_r = 0$. Therefore, $\mathrm{ISM}_\alpha = 0$.

2. Monotonicity property for temporal deviations. If, over a certain period of time, the deviations $d_i(t)$ increase by at least $\varepsilon > 0$ (for fixed values λ, β_i, w, W), then ISM_α does not decrease $\left(\mathrm{ISM}_\alpha \geq \mathrm{ISM}_\alpha\right)$. Larger deviations at any interval do not decrease the metric.

Proof: First, Lemma 1 gives the monotonicity of all links: $d_i(t)\uparrow \Rightarrow z\uparrow \Rightarrow \Phi_\lambda z\uparrow \Rightarrow \tilde{z}\uparrow \Rightarrow \tilde{d}_i\uparrow \Rightarrow Y_{i,r}\uparrow$; second, Lemma 2 gives $\mathrm{CVaR}_\alpha\left(\{Y_{i,r}\}\right)\uparrow$ for each i; third, adding non-negative weights β_i and a fixed penalty for secondary failures (the same failure), we obtain $\mathrm{ISM}_\alpha \geq \mathrm{ISM}_\alpha$.

3. Cascade sensitivity property. For $\lambda \in [0,1]$ and non-negative A and $z(t)$, we have $\partial \mathrm{ISM}_\alpha / \partial \lambda \geq 0$. Amplifying cascade propagation does not reduce the risk estimate.

Proof: First, (4) has a derivative:

$$\frac{d\Phi_\lambda}{d\lambda} = (I-\lambda A)^{-1} A (I-\lambda A)^{-1} = \Phi_\lambda A \Phi_\lambda \geq 0, \qquad (13)$$

where all elements are non-negative, because $A \geq 0$; secondly, since $\Phi_\lambda z(t)$ increases component-wise according to λ, then after applying the component-wise upper bound by one (operations $v \mapsto \min\{v,1\}$), the obtained values also do not decrease with the increase of λ; and thirdly, the monotonicity chain, as in the second property, gives an increase in each $\{Y_{i,r}\}$ and service component; the penalty for secondary failures does not depend on λ. Therefore, ISM_α does not decrease with λ.

4. Sensitivity to the appearance of "tails." For the same "area" of deviations, the metric is higher when they are more concentrated in large peaks (CVaR's value is higher). The metric emphasizes rare but large violations.

Proof: since CVaR_α is the average value in the upper part of the sample (by weight $1-\alpha$), the redistribution of the population from the average values to the largest ones increases the upper $(1-\alpha)$ quantile and the average value in the "tail." Formally, this is a well-known fact of "dominance by increasing convex functionals" (Hardy–Littlewood–Polya logic [18–19]): if the sample $\{Y_{i,r}\}$ has no less partial sums after sorting in descending order than $\{Y_{i,r}\}$ (majorization), then for all convex and increasing φ (in particular, $\varphi(y) = [y-u]_+$ in the representation CVaR we have $\sum \omega_r \varphi(Y_{i,r}) \geq \sum \omega_r \varphi(Y_{i,r})$. By minimizing u the inequality will be achieved for any $\alpha \in (0,1)$.

5. Property of large-scale consistency of measurement units. Linear change of units $q_i(t) \mapsto a q_i(t) + b$, $\theta_i(t) \mapsto a\theta_i(t) + b$ at $a > 0$ does not change the ranking of configurations if the normalization $d_i(t)$ and $z(t)$ are consistent. The transition between measurement units does not affect the comparison.

Proof: with a linear change in units, the difference between the plan and actual values scales identically – the new difference is equal to a (previous difference). Therefore, the "deficit to normalization" at each moment is simply multiplied by the same positive coefficient. Accordingly, the entire subsequent path (smoothing over time, weighting factors, projection onto service, cascading propagation onto the graph, etc.) is linear with non-negative coefficients, so it also simply multiplies the time series of values by the same coefficient a up to the normalization stage.

According to the assumption of "coordinated normalization" $d_i(t)$ and $z(t)$ are strictly increasing transformations that remove the global scale (for example, by dividing by the corresponding normalization factor and truncating to $[0,1]$). Therefore, applying this same normalization to two time series of

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

values that differ only by a common positive factor preserves the component order and returns the same (or equivalent for comparison) normalized values. Thus, all subsequent time integrations/aggregations (including CVaR) receive the same normalized inputs, and the final value for each configuration remains unchanged. Accordingly, the ranking of configurations remains the same.

6. Property of stability to small perturbations. For any $\varepsilon > 0$ there exists $\delta > 0$ such that if $\sup_t \|q_i(t) - \hat{q}_i(t)\| \le \delta$ and $\sup_t \|\theta_i(t) - \theta_i(t)\| \le \delta$ for all i, then $\left|\mathrm{ISM}_\alpha(q,\theta) - \mathrm{ISM}_\alpha(\hat{q},\hat{\theta})\right| < \varepsilon$.

Proof:

– changing (q_i, θ_i) within the limits of δ changes $d_i(t) = [\theta_i - q_i]_+$ by no more than δ;

– Lemma 1 and the boundedness of kernels/weights give $\left|(W * d_i)(t) - (W * \hat{d}_i)(t)\right| \le \delta$, $|w(t)(\bullet) - w(t)(\bullet)| \le \|w\|_\infty \delta$;

– since $\Phi_\lambda = \sum_{\ell \ge 0} \lambda^\ell A^\ell$ has a finite norm over $\rho(\lambda A) < 1$, the difference $\left(\tilde{z} - \hat{\tilde{z}}\right)$ is bounded by a constant multiple of δ;

– Integration in windows W_r' scales the error by no more than their length; as a result, there exists $C > 0$ (a constant independent of the data, which estimates the increase in integrals in windows), such that $\left|Y_{i,r} - \hat{Y}_{i,r}\right| \le C\delta$ for all i and r;

– applying Lemma 2 to each service and linearity by weight β_i, we have:

$$\left|\mathrm{ISM}_\alpha - \mathrm{ISM}_\alpha\right| \le \left(\sum_i \beta_i\right) C\delta + \left|\sum_i \omega_r \left(\Sigma_r - \hat{\Sigma}_r\right)\right|. \quad (14)$$

Since $\sigma_u(t)$ are determined from invariants as functions of the same observations, they change by no more than $C'\delta$, and therefore the penalty is no more than $C''\delta$. By choosing $\delta = \dfrac{\varepsilon}{\left(C'' + \sum_i \beta_i C\right)}$, this property is proven.

7. Additivity property by services. ISM_α is the weighted sum of service components and penalties for secondary failures, which allows changing β_i according to the weight of services without losing comparability. The result is formed as a weighted aggregation of services.

Proof: At the system aggregation step for each service S_i, its own risk component R_i is first calculated (by constructing $\{Y_{i,r}\}$ and CVaR on this sample). The final metric is defined as a linear combination of these components with non-negative weights β_i (usually normalized to $\sum_i \beta_i = 1$). Therefore, ISM_α is the weighted sum of service components. Changing β_i only reorients the priorities between services; if we fix the same vector β for all compared configurations, we obtain a well-defined scalar value, so comparability is preserved (the scale is unified by normalizing the weights).

If, in some scenarios, certain components of $\Phi_\lambda z(t)$ reach 1 and are "truncated," all proofs are preserved in a weak form: monotonicity and sensitivity to λ become non-strict (13) (the metric does not decrease and may not increase with small changes (14)), which is consistent with the logic: after reaching the "maximum" local degradation, further strengthening does not affect the evaluation beyond the established limit.

If necessary, a structural penalty for an overly connected graph (for overly connected configurations) can be added using the regularizer $\rho\left((I - \lambda A)^{-1}\right)$ or $-\log\det(I - \lambda A)$ with a coefficient, but the basic definition is already sufficiently sensitive due to the cascade operator (4).

Procedure for calculating integral metrics in real time

The organization of real-time integral metric calculations consists of dividing the calculation process into four modules that operate in an event stream and are synchronized outside the availability windows:

– module A (localization of deficits at graph nodes): $d_i(t) \to z(t)$;

– module B (cascade correction (graph broadcast)): $d_i(t) \to \tilde{z}(t)$ via local transmissions along graph edges with decay λ;

– module C (temporally weighted aggregation at the service level): $\tilde{d}_i(t) \to \tilde{z}(t) \to Y_{i,r}$ (exponential convolution followed by integration in a window);

– module D (risk-oriented aggregation and secondary failures): online evaluation CVaR_α in sequence $\{Y_{i,r}\}$; and accumulation Σ_r.

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)
Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

Let us first consider module A. The main goal of the module is to convert service deficits $d_i(t)$ into a node vector $z(t)$.

The distribution rule is formed as follows. For node $v \in V$, we determine its localized deficit:

$$z_v(t) = \min\left\{1, \sum_{i=1}^{m} \underbrace{\frac{\Pi_i(v)\beta_i}{\sum_{u \in V} \Pi_i(u)}}_{\text{service share } S_i \text{ at the node } v} \cdot d_i(t)\right\}. \quad (15)$$

In (15): $\Pi_i(v) \in \{0,1\}$ – a sign of belonging to the service node S_i; normalization divides the service deficit between the involved nodes; β_i allows you to strengthen important services. Other schemes agreed upon with the subject area (for example, taking into account the local weights of the nodes) are also acceptable – the algorithm below does not change.

When the event "d_i updated at t" occurs, we enumerate only those components z_v where $\Pi_i(v) = 1$.

Next, let us consider module B. Since direct calculation of $\tilde{z}(t) = (I - \lambda A)^{-1} z$ in an information system is excessively resource-intensive, it is proposed to use local message propagation along the arcs of the graph instead (analogous to the Neumann series with truncation of small terms [20]).

Formalization of the propagation rule (B1). Let us assume that at time t, an increment Δz_v is recorded at node v of the graph. We create a message queue Q with initialization $Q \leftarrow \{(v, \Delta z_v, 1)\}$, where the third argument is the path coefficient c (initially 1). Next, iteratively:

– extract $(u, \Delta, c) \in Q$. Add to the accumulated cascade deviation of the node u according to the scheme $\tilde{z}_u(t) \leftarrow \{1, \tilde{z}_u + \Delta\}$;

– for the outgoing edge $(u \to w) \in E$ with weight A_{uw}, we form a new message $(w, \lambda c A_{uw}\Delta, c' = \lambda c A_{uw})$ and add it to Q only if $c' \geq \varepsilon$ and $\hat{z}_w < 1$.

Heuristic explanation B1: the contribution along paths of length ℓ has weight $\lambda^{\ell} \prod A$; the threshold ε cuts off paths in the graph that are too long or too weak.

Formalization of the computational complexity limits of module B when processing a single event (B2). Let $\Delta_{\max}$ be the maximum change in component z per event. The number of messages generated per event is geometrically limited (geometric series logic):

$$N_{msg} \lesssim \sum_{\ell \geq 0} \left(\bar{d}\lambda\bar{A}\right)^{\ell} = \frac{1}{1 - \bar{d}\lambda\bar{A}}, \quad (16)$$

where $\bar{d}$ is the average outgoing degree of a node; $\bar{A}$ is the average weight of an arc (after scaling) under the condition $\bar{d}\lambda\bar{A} < 1$.

In (16), geometricity is understood as the construction of a geometric series from the contributions of all edge lengths.

In practice, ε further restricts ℓ, ensuring quasi-linear cost with respect to the number of tangent edges.

Next, we consider the C module, whose purpose is to convert the current normalized deficits $\tilde{d}_i(e)$ into representative time statistics $Y_{i,r}$ in the availability windows W_r for further risk-oriented aggregation.

Next, in submodule C1, we consider recursive smoothing with an exponential kernel. We use a one-sided exponential convolution with a kernel $k(\tau) = \rho e^{-\rho\tau} (\rho > 0)$, which gives greater weight to previous values and ensures causality and normalization $\left(\int k = 1\right)$. For each service S_i, we maintain a state $s_i(t) \approx \left(W * \tilde{d}_i\right)(t)$, which evolves according to

$$\dot{s}_i(t) = \rho\left(\tilde{d}_i(t) - s_i(t)\right), \quad (17)$$

and is updated recursively in discrete time

$$s_i \leftarrow e^{-\rho\Delta t} s_i + \left(1 - e^{-\rho\Delta t}\right)\tilde{d}_i. \quad (18)$$

In (17–18), the parameter ρ determines the "memory" of the filter $1/\rho$. The update is stable, non-negative, and does not change the scale of the time series of values, because the kernel is normalized.

Next, in submodule C2, we determine the integration in the availability window. At the same time, we sum up the time-averaged values, multiplying each by the importance coefficient for the corresponding moment in time:

$$S_i \leftarrow S_i + w(t) s_i(t) \Delta t. \quad (19)$$

When the window W_r closes, we record $Y_{i,r} := S_i$, after which $S_i \leftarrow 0$ (if necessary, normalized to the window duration. The rule (taking into account (19)) is the same for all services). Thus, $Y_{i,r}$ forms a sample of deficit episodes taking into account previous values

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

(via s_i) and calendar priority (via w), which is further used in CVaR aggregation.

Next, let's look at module D. Submodule D1 provides two evaluation options: CVaR_α.

$$u_i \leftarrow u_i - \eta_r\left(1-\frac{1}{1-\alpha}\mathbf{1}\{Y_{i,r}\ge u_i\}\right),\quad c_i \leftarrow (1-\zeta_r)c_i + \zeta_r\left(u_i+\frac{1}{1-\alpha}\left[Y_{i,r}-u_i\right]_+\right), \tag{20}$$

where $\eta_r, \zeta_r > 0$ and ζ_r are small, decreasing (decaying) steps; for example, $\eta_r = \eta_0/(1+r)^\kappa$.

Current estimate $\mathrm{CVaR}_\alpha \approx c_i$. Advantage: no buffer required (zero memory). Disadvantage: careful selection of the step schedule is required.

In option D1.2, we use a quantile buffer (fixed memory M) – we maintain a buffer of the latest values $B_i := \{Y_{i,r}\}$ of size M (discarding the oldest ones) and keep an approximate α quantile $\hat{q}_\alpha$. Then:

$$\mathrm{CVaR}_\alpha = \frac{1}{(1-\alpha)}\cdot\frac{1}{|T_i'|}\sum_{y\in T_i'} y,\qquad T_i' = \{y\in B_i' : y\ge \hat{q}_\alpha\}. \tag{21}$$

The advantage of (21) over (20) is stability and controlled memory, while the disadvantage is a slight delay due to the buffer. Both modes are consistent with the definition CVaR (10) and preserve single-passability.

In submodule D2, we aggregate secondary failures within the window W_r. First, we integrate

$$\Sigma_r \leftarrow \Sigma_r + \langle \eta, \sigma(t)\rangle \Delta t, \tag{22}$$

where $\sigma(t)$ is the degree of violation of invariants (7) with weights $\eta \ge 0$.

After closing the window W_τ, add the contribution $\omega_r\Sigma_r$ to ISM_α and reset $\Sigma_r \leftarrow 0$ (22).

Before moving on to evaluating the cost and settings, let's record how the structure behaves when observations are missing and when working with availability windows:

– between windows, we do not update the value $s[i]$ (or we update only the decrease $e^{-\rho\Delta t}$ without new observations); we do not increase the integral $S[i]$ – the absence of data is not considered zero quality;

– when opening a new window, we initiate the time weight $\omega(t)$ according to the mission schedule, and it is advisable to reduce the parameter s η_r and ζ_r for CVaR for small samples (short windows).

In option D1.1, we use a stochastic gradient (without a buffer). To do this, we enter a threshold estimate u_i (analogous to VaR_α) and an estimate CVaR c_i for the service S_i. Then, for each new $Y_{i,r}$, we perform

Since the practical applicability of the method is determined not only by accuracy but also by resource cost, we formulate the asymptotic complexity per event/step and the corresponding memory costs for each module (Table 1).

Table 1 uses the following symbols: m – number of services; $|V|$ – number of nodes; $|E|$ – number of edges; $|\mathrm{supp}\,\Pi_i|$ – number of service nodes; N_{msg} – number of messages in a cascade per event; M – buffer capacity.

Table 1. *Asymptotic complexity and memory consumption for each module*

Module	**Complexity**	**Memory**
A	$O\left(\sum_i \lvert\mathrm{supp}\,\Pi_i\rvert\right)$ per deficit change event	$O(1)$ – additional; $O(\lvert V\rvert)$ – shared
B	$O(N_{msg})$ on the event	$O(\lvert V\rvert+\lvert E\rvert)$
C	$O(m)$ per event	$O(m)$
D (gradient mode)	$O(m)$	$O(m)$
D (buffer mode)	$O(m\log M)$	$O(m\cdot M)$

The complexity limits obtained (Table 1) define the working ranges of parameters. Below are practical guidelines for selecting them in real conditions:

– tail sensitivity (α) determines what proportion of the largest observations to take into account CVaR. For severe peak risks, it is advisable to use $\alpha = 0.9...0.99$; for more balanced scenarios, use $\alpha = 0.7...0.9$. Higher values shift the metric toward rare large episodes.

– λ (cascade depth) – selected so that the stability condition is met $\bar{d}\lambda\bar{A} < 1$ (in practice, it is necessary to calibrate based on historical cascade episodes and keep a reserve to the limit so as not to distort long graph paths);

– ε (cutoff threshold) – a regulator of the compromise between accuracy and speed in the propagation module; accordingly, we cut off very weak/long contributions that have almost no effect on the result. Typical values are $10^{-3}\ldots10^{-2}$;

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)
Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

– ρ (smoothing rate) – corresponds to the filter's "memory time" $\tau \approx 1/\rho$, respectively: the higher the ρ, the shorter the memory and sensitivity to recent changes; the lower the ρ, the stronger the averaging. In real conditions, it is necessary to match τ with the characteristic duration of episodes in the given data;

– β, η, γ are weight parameters extracted from the priorities of the information system's objective function: β_i is the importance of services in the final aggregation; η is the weight of secondary failures; γ is the structural penalty coefficient (if necessary). Changing these weights does not destroy the properties of metric (12), but orders the configurations according to the selected priorities.

CVaR mode is selected as follows: under severe memory constraints, we use the gradient option (D1.1) – single-pass, without a buffer. When additional stability and reproducibility of reports are required, we use the buffer option (D1.2) with a fixed capacity M.

Configuring parameters and practical modes of use of the integrated survivability metric

The following protocol for applying the metric is proposed (12).

Step 1. Ensure that "local" deviations $(z(t)$ components) lie within $[0,1]$. This fixes the scale and makes comparisons between configurations correct.

Step 2. We divide the event stream into intervals where data consistency is guaranteed. If the target function of the information system has critical phases, we set a higher $w(t)$ during these periods (for example, stepwise or bell-shaped [21]).

Step 3. Determine the "response horizon" for subsidence. For short peaks, take a smaller $\tau_{1/2}$ (i.e., a larger ρ) so as not to stretch the peaks. For slow drifts, take a larger $\tau_{1/2}$. Typical recommendations:

– reactive mode $\tau_{1/2} = 10...30$ s;

– balanced monitoring $\tau_{1/2} = 60...300$ s;

– background monitoring $\tau_{1/2} = 600...1800$ s.

Step 4. We reproduce the empirical reachability of the cascade. From event logs, we estimate the average "length" of the effect by calibrating λ so that the simulated propagation matches this length (with a margin to the stability limit $\rho(\lambda A) \le 1$), i.e., we essentially work with $\lambda \in (0, 0.9]$. We select the threshold ε according to the calculation budget using the cutoff scheme for contributions that change $\tilde{z}$ by less than $10^{-3}...10^{-2}$. And we check the impact on (12) – the difference should be within the selected reporting accuracy.

Step 5. Based on practical guidelines for selecting parameter values α, we link α to the frequency of acceptable failures. If significant deviations are allowed in no more than 5% of windows, we take $\alpha = 0.95$. For supercritical services, we take $\alpha = 0.99$; for ordinary services, we take $\alpha = 0.8...0.9$.

Step 6. We build a "priority budget" $\sum_i \beta_i = 1$. By β_i, we mean the proportional values of the service in the target function of the information system or the expected losses from its degradation. A simple way is to standardize ξ points p_i according to the selected scheme.

Step 7. First, we normalize η so that $\sum_u \eta_u = 1$. Next, we select a reference scenario: "one complete violation of the invariant I_u within T_{ref} seconds in 1 window". We select γ so that the contribution $\gamma\omega_r\Sigma_r$ in this scenario is equal to, say, the median CVaR_α for the base service. This level curve makes the scales comparable.

Step 8. Perform sensitivity analysis. Change each parameter to ±10% and check whether the configuration ranking remains the same. If the ranking "jumps", reduce α (reduce peak dominance) or λ (reduce cascade amplification), or increase $\tau_{1/2}$ (for greater stability).

Step 9. Perform a backtest on previously recorded incidents. We expect (12) to rise to the peak of the incident, remain there during the incident, and decrease after recovery to the baseline level. This is the criterion for the correctness of the settings.

In this protocol, ξ points are understood as preliminary numerical assessments of service priority, which are then converted into weights β by simple normalization:

$$\beta_i = \xi_i \Big/ \sum_j \xi_j . \qquad (23)$$

Such ξ_i can be set in several complementary ways:

1. Uniform a priori method:

$$\xi_i \equiv 1 \quad \Rightarrow \quad \beta_i = 1/m, \qquad (24)$$

where m is the number of services. This sets the initial priority.

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

2. Method based on service load level:

$$\xi_i = \overline{\text{req}}_i \cdot v_i, \quad \xi_i = \tilde{\xi}_i \Big/ \sum\nolimits_j \tilde{\xi}_j . \qquad (25)$$

where $\overline{\text{req}}_i$ is the average intensity of requests to S_i; v_i is the "usefulness" coefficient (if absent, then $v_i = 1$).

3. Risk-oriented method based on observation history:

$$\tilde{\xi}_i = \wp_i \cdot \text{CVaR}_\alpha\left(\{Y_{i,r}\}_r\right), \quad \xi_i = \tilde{\xi}_i \Big/ \sum\nolimits_j \tilde{\xi}_j , \qquad (26)$$

where $\wp_i$ – frequency of "problematic" windows for S_i; $\text{CVaR}_\alpha\left(\{Y_{i,r}\}_r\right)$ – conditional average of the largest $(1-\alpha)\%$ values $Y_{i,r}$.

In (26), frequent and "heavy" availability windows increase W_r.

4. Method based on structural influence in the graph component:

$$s_i = \frac{\langle \Pi_i, \Phi_\lambda \mathbf{1} \rangle}{\langle \Pi_i, \mathbf{1} \rangle}, \quad \tilde{\xi}_i = s_i, \quad \xi_i = \frac{\tilde{\xi}_i}{\sum_j \tilde{\xi}_j}, \qquad (27)$$

where $\Pi_i \in \{0,1\}^{|V|}$ – indicator vector of service nodes S_i.

The value of s_i (27) is the average "reachable influence" of service nodes and reflects its potential cascading.

5. SLO method:

$$\tilde{\xi}_i = 1/E_i, \quad \xi_i = \tilde{\xi}_i \Big/ \sum\nolimits_j \tilde{\xi}_j , \qquad (28)$$

where E_i – acceptable budget for failures/deficits for S_i (a smaller budget gives higher priority).

If necessary, several methods (24–28) can be combined for a single assessment. An example of such a combination:

$$\tilde{\xi}_i = a \cdot \text{req}_i + b \cdot \text{CVaR}_\alpha\left(Y_i\right) + c \cdot \hat{s}_i, \quad a + b + c = 1,$$

where the circumflex sign means the preliminary scaling of each indicator to $[0,1]$. After that, the usual normalization of $\xi_i = \tilde{\xi}_i \Big/ \sum\nolimits_j \tilde{\xi}_j$ is performed and the weights β_i (23) are calculated.

To facilitate initial setup, we provide practical profiles – ready-made sets of coordinated parameters for typical operating modes. They do not replace calibration, but they provide the correct "entry point" and explain why these particular values work.

1. "Fast-moving critical target functions" profile. In this profile, it is important to catch rare peak deviations in time. Therefore, we take a high "tail" $\alpha = 0.95$ (the metric focuses on the upper tail), short memory $\tau_{1/2} = 10...30$ c (large ρ; peaks are not "blurred"), moderate cascade depth $\lambda = 0.3...0.6$ (the cascade is taken into account, but does not "overheat"), low cutoff threshold $\varepsilon \approx 10^{-3}$ (more precisely, if the resource allows). We make the time weight $w(t)$ pulsed during critical phases, and increase the weight of secondary failures γ – during maneuvers, secondary failures often become decisive.

"Overheating" of the cascade refers to a situation where propagation parameters (large λ and/or dense graph A) overly amplify cascade paths, causing minor local deviations to spread throughout the information system and artificially causing an unjustified increase in metrics (12).

2. Profile "Continuous service flows." The balance between peaks and averages is more important than "chasing tails." Therefore, $\alpha = 0.8...0.9$ (compromise), memory is longer $\tau_{1/2} = 60...300$ c (less sensitivity to small fluctuations), the cascade is deeper $\lambda = 0.4...0.7$, and the cutoff can be made coarser $\varepsilon \approx 10^{-2}$ to save resources. $w(t)$ is almost uniform with slight amplifications according to the schedule; γ is average, i.e., secondary failures have an impact but do not dominate.

3. Profile "Environments with dense dependencies (dense graphs)". The main risk is cascade amplification. This occurs when the network is so dense or the λ is so large that even a small local deviation quickly amplifies along many paths of the graph. Therefore, we first scale A so that $\rho(A) \le 1$, and keep λ restrained $(0.2...0.5)$ to avoid overcounting long paths. We make the threshold ε slightly higher, controlling the error ISM (12); we set $w(t)$ according to the parameters of the target function of the information system. If necessary, you can increase the structural penalty γ to add a penalty for excessive graph connectivity to the metric.

4. Profile "Intermittent connection and long 'dead zones'". When telemetry comes in fragments, the first thing to do is to extend or adapt W_r (so as not to merge short bursts into noise (short, irregular metric deviations caused not by service degradation, but by side effects of measurement and information system functioning)) and increase $\tau_{1/2}$ – this reduces random jumps between windows. In $w(t)$, we underestimate the weight of periods without confirmed data. We leave other parameters (λ, ε, α) from the base profile and adjust them after a quick check of the existing historical data.

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

After selecting the start profile, it is recommended to perform a short calibration on previous metric runs – specify (λ, ρ, ε, α) under the graph provided by the information system, data rhythm (telemetry every second, logs every minute, spikes under load, downtime, time peaks) and the computing budget, and then check the stability of the ranking and the limits of the system resources.

After fixing the parameters and profiles, we move on to the presentation of the methodological principles of interpreting the integrated survivability metric, which regulate the rules for reading the metric in a production environment and the procedures for making decisions based on it.

Thus, when determining the scales and thresholds of the integral survivability metric, it is advisable to set two levels – warning and critical – with numerical values established either by empirical percentiles (e.g., 80th and 95th) or by target risk budgets (probability of exceeding the limit). When ensuring the stability of the metric indicator, it is recommended to adjust only the filtration parameters: if the metric is "noisy," it is necessary to slightly increase $\tau_{1/2}$ or decrease α; if the metric is inertial, it is necessary to decrease $\tau_{1/2}$ or strengthen $w(t)$ in critical intervals (if necessary, slightly adjust ε), while preserving the semantics of the metric.

Below is a summary of typical errors in configuring the integrated survivability metric, with an analysis of the causes and practical steps to resolve them.

1. Error: too high λ in dense graphs (cascade "overheating" effect).

Problem analysis: when $\rho(\lambda A)$ or $\bar{d}\lambda\bar{A}$ approaches 1, the contribution of long paths increases sharply: the number of messages increases exponentially, most $\hat{z}_v$ are cut to 1, and the metric is artificially inflated.

Solution: scale A so that $\rho(A) \le 1$; reduce λ; increase the cutoff threshold ε if necessary, limit the path length ℓ or add a structural penalty to γ.

2. Error: too large α with a poor sample (few windows).

Problem analysis: the tail estimate becomes unstable: CVaR_α is very volatile because only a few observations remain in the upper $(1-\alpha)\%$.

Solution: reduce α (more tail averaging) or switch to buffer mode D1.1 with a larger M (stabilizes the tail).

3. Error: abnormally small or abnormally large ρ (smoothing rate).

Problem analysis: too small ρ approximates peaks – the system lags behind in performing the function; too large ρ makes the metric abnormally active – the reaction occurs to noise fluctuations.

Solution: focus on the required response time of the target function of the information system: select $\tau_{1/2} \sim 1/\rho$ so that you can see relevant peaks, but not noise; for "noisy" metrics, slightly increase $\tau_{1/2}$, and for "inertial" metrics, decrease it.

4. Error: β_i and γ are not comparable (dominance of one indicator).

Problem analysis: if the weights of services β_i or the weight of secondary failures γ are set without coordination, one component (service or secondary) begins to dominate regardless of the data.

Solution: apply a link to the level curve (sixth step of the metric application protocol), in particular, select γ so that in the reference scenario, the contribution of $\gamma\omega_r\Sigma_r$ is equal to, for example, the median CVaR_α of the base service; normalize β_i $\left(\sum_i \beta_i = 1\right)$ and check the sensitivity (±10%) – the ranking should be preserved.

The relationship between the integral metric of survivability and the operational area and the rules for comparing configurations

Next, it is necessary to formalize when the integral survivability metric certifies that the information system is in the operational zone and how to correctly compare architectural configurations and recovery policies based on it.

Taking the above into account, let us summarize some concepts:

– $\mathcal{I}_i := \mathrm{CVaR}_\alpha\left(\{Y_{i,r}\}\right)$ – service "tail" assessment for the service S_i;

– $\mathcal{J} := \sum_{r=1}^{R} \omega_r \Sigma_r$ – the integral contribution of secondary failures.

Next, we set the reference limits:

– $\kappa_i > 0$ – acceptable tail degradation of the S_i service (at the level of CVaR_α);

– $\tau > 0$ – acceptable integral level of secondary failures;

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

– $\Gamma > 0$ – global limit of the integral survivability metric.

Therefore, the trajectory of the information system lies within the operational range if three budget conditions are met simultaneously:

$$\mathcal{I}_i \le \kappa_i \ \forall i, \quad \mathcal{J} \le \tau, \quad \mathrm{ISM}_\alpha \le \Gamma. \tag{29}$$

Condition (29) is explained by the fact that local service "tails" and the aggregate contribution of secondary failures fit within their budgets, and their weighted sum does not exceed the global risk budget.

A simple sufficient condition for operability follows from the linear structure of ISM :

$$\sum_{i-1}^{m} \beta_i \kappa_i + \gamma \cdot \tau \le \Gamma \Rightarrow \left(\mathcal{I}_i \le \kappa_i \ \forall i \ \wedge \ \mathcal{J} \le \tau\right) \Rightarrow \mathrm{ISM}_\alpha \le \Gamma. \tag{30}$$

Condition (30) provides a mechanism for reconciling local limits (κ_i, τ) with global limits Γ. Accordingly, if local budgets are chosen so that their weighted sum does not exceed the global limit, then the fulfillment of local constraints guarantees the fulfillment of the global constraint.

The algebraic difference between the limit and the current value of the metric (margin to limit) allows you to conveniently track the margin to threshold:

$$\Delta := \Gamma - \mathrm{ISM}_\alpha. \tag{31}$$

Then, at $\Delta > 0$ (31) shows that the system is within the operating range (there is a reserve); $\Delta = 0$ – at the limit; $\Delta < 0$ – outside the range (exceeding $|\Delta|$). For operational control, it is also convenient to use a normalized reserve:

$$\zeta := 1 - \frac{\mathrm{ISM}_\alpha}{\Gamma} \in (-\infty, 1], \quad (31) \textit{ shows that the system is within the operating range } (\textit{there is a reserve}); \tag{32}$$

which is a dimensionless indicator: $\zeta = 1$ means zero risk, $\zeta = 0$ – threshold point, $\zeta < 0$ – shows the percentage of "exceeding" the global limit.

To calibrate the limits (κ_i, τ, Γ), we first calculate empirical values $\mathcal{I}_i$ for each service and $\mathcal{J}$ for secondary failures at stable intervals; then we set local thresholds as high percentiles: $\kappa_i := q_p(\mathcal{I}_i)$, $\tau_i := q_p(\mathcal{J}_i)$ from $p \in [0.90, 0.95]$. After that, we reconcile them with weights and set a global limit so that with a given margin $\Delta_{\min} > 0$ the sufficient condition is satisfied:

$$\sum_{i=1}^{m} \beta_i \kappa_i + \gamma \cdot \tau \le \Gamma - \Delta_{\min} \tag{33}$$

Any Γ from the interval $\left[\sum_{i=1}^{m} \beta_i \kappa_i + \gamma \cdot \tau + \Delta_{\min}, \infty\right)$ satisfies condition (33). To fix a specific value, it is convenient to take the minimum allowable:

$$\Gamma := \sum_{i=1}^{m} \beta_i \kappa_i + \gamma \cdot \tau + \Delta_{\min}. \tag{34}$$

If necessary, the reserve can be specified in relative form $\zeta_{\min} > 0$, then from (34) we have:

$$\Gamma = \frac{\sum_i \beta_i \kappa_i + \gamma \cdot \tau}{1 - \zeta_{\min}}. \tag{35}$$

When the mission priorities change (weights β_i; if necessary, κ_i), we check the higher condition with the new values and rebind Γ (35) or $\Delta_{\min}$ to maintain the target reserve.

Next, we can formalize the rules for comparing configurations. We assume that the configurations/policies Ψ_1 and Ψ_2 are evaluated using the same weights (β_i, γ), limits (κ_i, τ, Γ) and windows W_r. Then:

Rule R1. A configuration is acceptable if and only if conditions (29) are satisfied. Configurations that do not satisfy these conditions are discarded without further ranking.

Rule R2. Among acceptable configurations, preference is given to the one with less ISM_α. For practical stability of comparison, the dominance margin $\delta > 0$ is used: if $\mathrm{ISM}_\alpha(\Psi_1) \le \mathrm{ISM}_\alpha(\Psi_2) - \delta$, then Ψ_1 dominates Ψ_2.

Rule R3. If $\left|\mathrm{ISM}_\alpha(\Psi_1) - \mathrm{ISM}_\alpha(\Psi_2)\right| < \delta$ (tie according to R2), preference is given to the configuration with a smaller $\mathcal{J}$ (fewer invariant violations), because secondary failures increase the structural risk of cascades.

Rule R4. If the previous criteria are equal, compare the vector $\mathcal{I}_i$ in descending order β_i. The configuration with lower $\mathcal{I}_i$ for the highest priority services wins.

Rule R5. For strategic planning (before selecting β_i, γ), it is useful to construct a Pareto front [22] along two coordinates: $\left(\sum_i \beta_i \mathcal{I}_i, \mathcal{J}\right)$. Alternatives that lie

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

above/to the right of the front (worse on both criteria) are rejected as dominated (alternatives for which there is another alternative that is not worse on all criteria and is better on at least one).

To make the comparison of configurations resistant to estimation errors (since $\mathcal{I}_i$ and $\mathcal{J}$ are calculated based on a finite number of windows W_r), we take a confidence interval for each value: $\mathcal{I}_i \in \left[\underline{\mathcal{I}}_i, \overline{\mathcal{I}}_i\right]$, $\mathcal{J} \in \left[\underline{\mathcal{J}}, \overline{\mathcal{J}}\right]$.

Then we formulate the rules:

Rule 1. The configuration is considered acceptable if, even under the worst-case scenario, the budget system is fulfilled (robust acceptance):

$$\overline{\mathcal{I}}_i \le \kappa_i \ \forall i, \quad \overline{\mathcal{J}} \le \tau, \quad \sum_i \beta_i \overline{\mathcal{I}}_i + \gamma \overline{\mathcal{J}} \le \Gamma. \qquad (36)$$

Those who do not satisfy condition (36) are cut off without further ranking.

Rule 2. Let us compare two admissible configurations Ψ_1 and Ψ_2. Suppose that Ψ_1 dominates Ψ_2 if its worst generalized indicator is less than the best indicator of Ψ_2:

$$\sum_i \beta_i \overline{\mathcal{I}}_i^{\Psi_1} + \gamma \overline{\mathcal{J}}^{\Psi_1} < \sum_i \beta_i \underline{\mathcal{I}}_i^{\Psi_2} + \gamma \underline{\mathcal{J}}^{\Psi_2}. \qquad (37)$$

This guarantees an advantage even in the "worst case for Ψ_1" versus the "best case for Ψ_2". If necessary, we add a small margin of $\delta > 0$ to the right side of (37) as a tolerance for rounding.

In practice, if the intervals overlap and the inequality is not satisfied with a margin, and the conclusion is "undetermined," then we either increase the number of windows (more data), or use the buffer mode CVaR with a larger M, or transfer the decision to the profile level (rules R3–R5).

To transform the integrated survivability metric from a simple diagnostic figure into an operational tool, two aspects need to be fixed: how to act for different values of the indicator, and how to make correct comparisons between different target functions of the information system.

To this end, an algorithm of actions for the operational zone is proposed. For this purpose, a standardized indicator (32) will be used as a single "traffic light":

– emerald zone: $\zeta \ge 0.2$ – normal operation of the information system;

– yellow zone: $0 < \zeta < 0.2$ – warning mode: limit non-critical loads, increase the weight of $w(t)$ in critical intervals, activate preventive localization policies;

– red zone: $\zeta \le 0$ – information system outside the operational zone: cascade cut-off policies must be applied immediately – reduce λ and/or increase ε; forcibly align invariants, increase resource reserves.

At the same time, we monitor the contribution structure. If the share of secondary failures in ISM_α exceeds ~ 40–50%, this is an indicator of the secondary nature of the degradation. In this case, the priority should be to restore invariants $\left(\mathrm{Inv}_T,\ \mathrm{Inv}_R,\ \mathrm{Inv}_B\right)$.

To correctly compare configurations for different target functions of the same information system (different critical interval schedules, different W_r'), it is necessary to ensure the invariance of settings: maintain the same weights (β_i, γ) and the same level of α; normalize $w(t)$ so that $\int_{W_r} w(t)dt = 1$ in each target function of the information system (then $Y_{i,r}$ – comparable in scale); ensure the same $\tau_{1/2}$.

Under these conditions, ISM_α becomes comparable in terms of the results of performing different target functions by the same information system.

Analytical comparison of the integral metric of survivability with basic indicators

Next, we will demonstrate through analytical examples (scenarios) how ISM_α provides a more informative and stable assessment than common basic indicators such as the proportion of time out of demand (binary indicator), average deficit, information system availability, average downtime, etc.

First, let's look at the basic indicators that will be compared with ISM_α. So, let's say that for the service S_i, the quality indicator $q_i(t) \in [0,1]$ and the requirement $\theta_i(t) \in [0,1]$ are observed on the horizon $[t_0, t_1]$. The deficit is determined by formula (1). Then:

1. Binary indicator of violations (the proportion of time when $d_i(t) > 0$):

$$\Upsilon_i := \frac{1}{T}\mu\left(\left\{t \in [t_0, t_1] : d_i(t) > 0\right\}\right), \quad T = t_1 - t_0, \qquad (38)$$

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

where $\mu(\bullet)$ is the Lebesgue metric (the proportion of time during which the deficit $d_i(t)$ was positive) [23].

2. Average deficit indicator:

$$\vartheta_i := \frac{1}{T}\int_{t_0}^{t_1} d_i(t)dt. \quad (39)$$

3. Availability indicator:

$$o_i := 1 - \Upsilon_i. \quad (40)$$

These indicators (38–40) do not see either the topology of dependencies or the tail structure of the distribution of deviations, and therefore may give the same values for significantly different risks.

Next, let's move on to the scenarios themselves.

Scenario S1 ("heavy tails"). Let's construct two deficit trajectories for a single service at $[0,1]$:

– trajectory $\mho_1$ with rare but significant declines in service quality: $d_i(t) = 1$ in the range of ε, and in the rest of the interval $d_i(t) = 0$;

– trajectory $\mho_2$ with frequent but minor dips in service quality: $d_i(t) = \varepsilon$ throughout the entire interval.

Both trajectories have the same average deficit:

$$\vartheta_i^{(\mho_1)} = \varepsilon\cdot 1 + (1-\varepsilon)\cdot 0 = \varepsilon, \quad \vartheta_i^{(\mho_2)} = (1)\cdot\varepsilon, \quad \vartheta_i^{(\mho_1)} = \vartheta_i^{(\mho_2)}.$$

Similarly, $\Upsilon_i^{(\mho_1)} = \varepsilon$, $\Upsilon_i^{(\mho_2)} = 1$ (for any $\varepsilon > 0$), i.e., the binary indicator even ranks them in reverse: in $\mho_2$, the indicator records constant violations, while in $\mho_1$, they are rare. However, in terms of peak risk (operational loss, emergency shutdowns), the $\mho_1$ scenario is worse.

Let $w(t) \equiv 1$ and $\lambda = \gamma = 0$ (i.e., ignore cascading and secondary). Then $Y_{i,r} = \int_0^1 (W * d_i)(t)\,dt$ equals (by unit convolution) ϑ_i. But CVaR_α acts on a sample of windows, and for $\alpha > (1-\varepsilon)$ we will have:

$$\mathrm{CVaR}_\alpha^{(\mho_1)} \approx 1, \quad \mathrm{CVaR}_\alpha^{(\mho_2)} \approx \varepsilon.$$

Therefore, $\mathrm{ISM}_\alpha^{(\mho_1)} > \mathrm{ISM}_\alpha^{(\mho_2)}$ for any α in the tail $(\alpha > (1-\varepsilon))$. Thus, the average indicators do not distinguish $\mho_1$ from $\mho_2$, and the binary ones are misleading; ISM_α correctly reflects the danger of short critical peaks.

Scenario S2 (cascading). Consider a graph with two nodes $V = \{1,2\}$ and an arc $1 \to 2$, $A_{12} = 1$, the rest of the values are 0. The service S is projected onto both nodes: $\Pi = e_1 + e_2$.

Let's compare two patterns of localized deficits $z(t)$ (5) of equal amplitude $\delta > 0$ (at equal time intervals):

– pattern ι_1 ("deficit at the top" (the primary node pulls the dependent node)): $z = [\delta, 0]^\top$;

– pattern ι_2 ("deficit down" (final node)): $z = [0, \delta]^\top$.

Cascading correction gives:

$$\tilde{z}^{\iota_1} = (I - \lambda A)^{-1} z = \begin{bmatrix} 1 & \lambda \\ 0 & 1 \end{bmatrix}\begin{bmatrix} \delta \\ 0 \end{bmatrix} = \begin{bmatrix} \delta \\ \lambda\delta \end{bmatrix},$$

$$\tilde{z}^{\iota_2} = \begin{bmatrix} 1 & \lambda \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ \delta \end{bmatrix} = \begin{bmatrix} \lambda\delta \\ \delta \end{bmatrix}.$$

Service deficit in each pattern:

$$\tilde{z}^{\iota_1} = \frac{\delta + \lambda\delta}{2}, \quad \tilde{z}^{\iota_2} = \frac{\lambda\delta + \delta}{2}.$$

That is, at the level of instantaneous values, they are the same. However, the reason is different: in ι_1, the violation "pulls" the dependent node (cascade), and in ι_2, the primary node is "at the end." When secondary violations (violations of interlevel invariants) are marked in event logs for ι_1, then $\sigma(t)$ takes on a positive value, while for ι_2, it does not. Thus, with the same $\tilde{d}\,\mathrm{ISM}_\alpha$, ι_1 is greater due to the addition of $\gamma\sum_r \omega_r \Sigma_r$. The basic indicators ϑ, Υ, o do not distinguish between these structurally different cases, while ISM_α separates configurations through the component of secondary failures.

Scenario S3 (hidden desynchronization). Let $q_i(t)$ perform the requirement $q_i(t) \ge \theta_i(t) \Rightarrow d_i(t) \equiv 0$ over the entire interval. At the same time, the invariant Inv_T is violated on a large scale, so $\sigma_{u^*}(t) \approx 1$ on a significant portion of the windows. Then:

$$\vartheta_i = 0,\ \Upsilon_i = 0,\ o_i = 1,$$

but

$$\mathrm{ISM}_\alpha = \underset{\text{service part}}{0} + \gamma\sum_r \omega_r \Sigma_r > 0.$$

That is, ISM_α warns of a hidden systemic risk (high probability of secondary failures and cascades when dependent processes are activated), which basic indicators do not notice.

Scenario S4 (observation gaps and availability windows). With intermittent communication, part of the telemetry is missing from the information system. Basic

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)
Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

indicators either exclude these intervals (overestimating "availability") or consider the deficit to be zero (underestimating the problem). In the approach proposed in this work, aggregation is performed only within the agreed windows $\{W_r'\}_{r=1}^{R}$ and integrals are normalized by definition – the metric becomes resistant to gaps and comparable between target functions of information systems with different observation "densities."

Thus, the scenarios presented show that ISM_α is the basis for achieving increased survivability of an information system on a mobile platform with multi-level dependencies, while basic indicators remain useful only as simple diagnostic overviews without a guarantee of correct configuration ranking.

Experimental verification of the integral survivability metric

As part of this study, a reproducible Python testbed [24] was developed that implements ISM_α and demonstrates its behavior on synthetic but representative scenarios. The experiment was conducted in a software environment using Python 3.11.8, NumPy 1.24.0, Pandas 1.5.3, and Matplotlib 3.6.3.

Initial data:

– dependency graph $A \in \mathbb{R}_+^{|V|\times|V|}$: random sparse directed graph $(|V| = 30)$ without loops, edge weights in $[0.2; 0.8]$;

– services $m = 3$ with projections onto nodes via matrices B (distribution of deficit onto nodes) and C (reverse service cut). Each service covers a subset of nodes, normalization guarantees $[0; 1]$;

– time: $T = 300$ steps (discretization in 1 s). Availability windows: mostly 30 s each; in scenario S4, every third window is missing (interrupted connection);

– parameter values: $\alpha \in \{0.8, 0.9, 0.95, 0.99\}$, $\lambda \in \{0, 0.2, 0.5, 0.8\}$, exponential smoothing with half-life $\tau_{1/2} \approx \ln 2/\rho$, service weights $\beta \in (0.5, 0.3, 0.2)$. The penalty for secondary failures is implemented as $\sum_r \omega_r \Sigma_r$ with a weight vector of invariants $\eta = (0.05, 0.02, 0.01)$, i.e. γ absorbed in η for ease of comparison);

– cascade correction: approximation of the Neumann series $\tilde{z} = \min\left\{\sum_{\ell=0}^{L} (\lambda A)^\ell z, \mathbf{1}\right\}$ with $L = 6$ (fast convergence for selected λ;

– CVaR : empirical weighted calculation over sets $Y_{i,r}$, where weights are proportional to window duration.

Scenarios:

– S1: S_0 service – rare but deep failures; S_1 – constant minor failures; S_2 – nominal;

– S2: two spikes (the first one is at the top of the chain with a secondary marking (invariance violation), the second one is more local);

– S3: $q_i(t) \geq \theta_i(t)$ always, but three intervals violate the invariants $(\mathrm{Inv}_T, \mathrm{Inv}_R)$, so $\Sigma_r > 0$;

– S4: similar to S1, but every third window is unavailable; time weights are amplified in available windows.

For analytical validation, along with ISM_α, basic indicators were collected on the services: average deficit ϑ and proportion of time outside demand Υ.

The dataset of the experiment results is available in the Zenodo repository [25]. Below are graphs showing the results obtained by (Figs. 1–5).

Fig. 1 shows the family of dependencies of the integral survivability metric ISM_α on the tail sensitivity parameter α for a fixed cascading depth $\lambda = 0.5$ for four scenarios: S1 – rare deep failures versus frequent minor failures; S2 – cascading failures with marked secondary nature; S3 – only secondary failures without service deficits; S4 – operation with intermittent connection, taking into account availability windows. The curves demonstrate three typical modes: monotonous growth followed by saturation at large α when CVaR focuses on the worst windows; increased metrics in the presence of cascades and secondary failures; practically α – invariant behavior in the absence of a service component ($\mathcal{J}$ dominates).

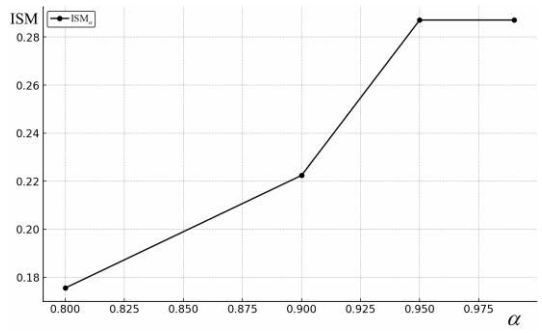
Fig. 1, *a*) shows that as the parameter α increases from 0.8 to ≈ 0.95, the metric value increases significantly and then levels off; this is due to the fact that at higher α, CVaR gives greater weight to rare but deep service failures – a manifestation of "tail" sensitivity.

In Fig. 1, *b*), the steepest growth is to $\alpha \approx 0.95$, followed by a plateau; tail sensitivity amplifies the contribution of windows with cascading/secondary disturbances.

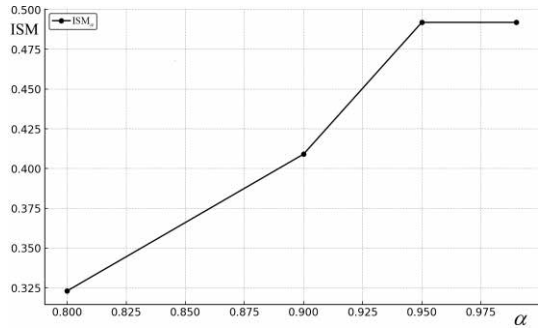
In Fig. 1, *c*) growth to $\alpha \approx 0.9$ and subsequent saturation; missing windows do not distort the estimate, since aggregation is performed only on available intervals.

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

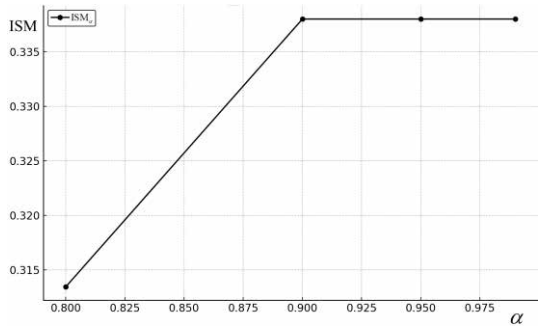
It should be noted that when α increases above the threshold that "covers" the worst windows, CVaR_α averages the same tail of the distribution, so ISM_α stops growing. At $\vartheta \approx 0$, the curve is practically horizontal – the service part has no effect, and the system penalty $\mathcal{J}$ does not depend on α; the metric remains positive, fixing the latent risk.



a)



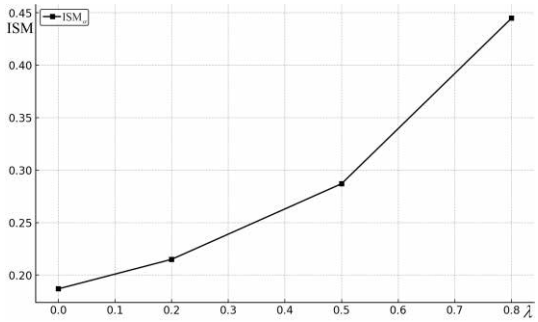
b)



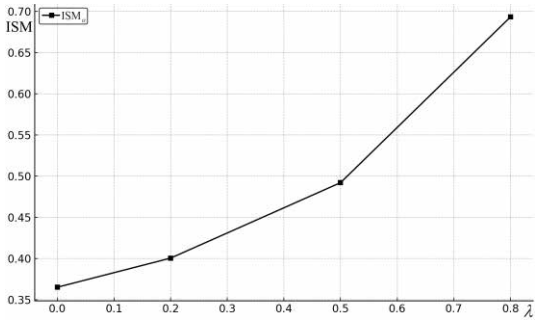
c)

Fig. 1. Dependence of the integral metric on the level of "tail" sensitivity (ISM_α) from α for a fixed $\lambda = 0.5$: a) scenario S1; b) scenario S2; c) scenario S4

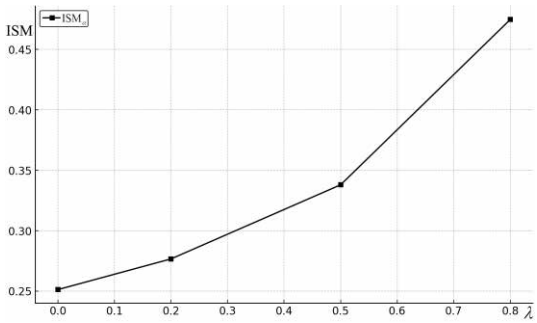
Fig. 2 shows the dependence of the integral metric on the depth of cascade correction (ISM_α) from λ for a fixed $\lambda = 0.95$ for four scenarios. The parameter $\lambda = [0, 0.8]$ scales the impact of propagation of failures along the dependency graph $(I - \lambda A)^{-1}$; thus, an increase in λ amplifies the contribution of cascades, while secondary failures are accounted for through $\mathcal{J}$. The curves show the structural sensitivity of the metric: in the presence of cascades/secondary failures, the slope increases, and in the absence of service deficits, the curve remains almost unchanged.



a)



b)



c)

Fig. 2. Dependence of the integral metric on the depth of cascade correction (ISM_α on λ) for a fixed $\lambda = 0.95$: a) scenario S1; b) scenario S2; c) scenario S4

Fig. 2, *a*) shows that the curve increases monotonically and is convex at the top; the largest increase is observed in the interval $\lambda = [0.5, 0.8]$, where cascade correction "pulls up" the influence of rare deep drops in dependent nodes.

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)
Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

Fig. 2, *b*) shows that the steepest gradient among all scenarios: an increase in λ sharply increases ISM_α due to the combined action of cascade amplification and $\mathcal{J}$ from violated invariants.

Fig. 2, *c*) shows a steady, smooth curve growth without jumps; the largest increase is at high λ. Working exclusively in accessible windows makes the assessment robust to telemetry omissions, so cascading manifests itself without artificial artifacts.

If the service component is zero and $\mathcal{J}$ does not depend on λ, then ISM_α remains almost constant.

The example of scenario S1 (rare deep troughs S0 versus frequent small troughs S1) (Fig. 3) shows the discrepancy between the integral metric and the average deficit. With a fixed $\alpha = 0.95$ and an increase in the cascade parameter λ, the value of ISM_α monotonically increases (from approximately 0.19 to 0.45), while ϑ for services remains almost unchanged $(S_0 \approx 0.012,$ $S_1 \approx 0.005)$. This illustrates that ϑ does not respond to either the depth of the cascade or the "tails" of the distribution, and therefore may misrank situations. Instead, ISM_α correctly increases the risk assessment with the growth of λ, accumulating the cascade amplification of rare but critical episodes.

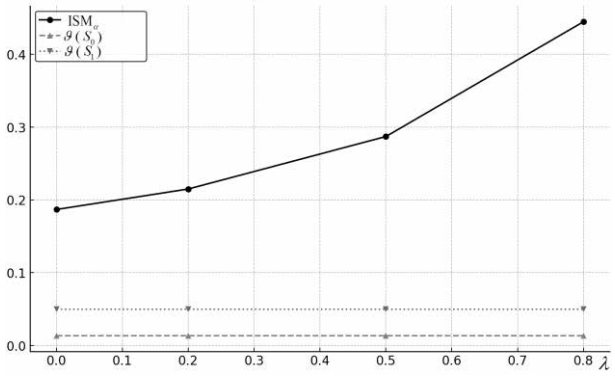


Fig. 3. Discrepancy between ISM_α and ϑ in scenario S1 – dependence on the cascade parameter λ $(\alpha = 0.95)$

Fig. 4 shows that in scenario S3, the average deficit ϑ is practically zero for all α, since $q_i(t) \geq \theta_i(t)$ (there are no service deviations). At the same time, ISM_α maintains a stable level (~0.166) and is almost independent of α, because the value of the metric is determined by $\mathcal{J}$, which accumulates invariant violations (secondary failures) and is not included in CVaR – aggregation of service deficits. Thus, the metric signals a latent risk when the SLO is in the emerald zone, which remains invisible to basic indicators such as ϑ.

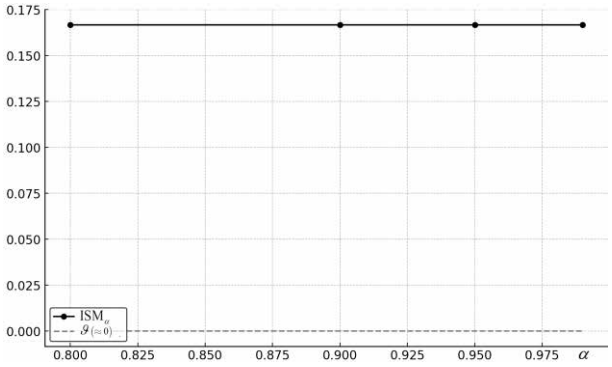


Fig. 4. Detection of secondary failures in the absence of service deficits $(\lambda = 0.5)$

Fig. 5 demonstrates the coordinated growth of ISM_α with both an increase in "tail" sensitivity α and a deepening of the cascade λ. The largest increase occurs in the direction of λ (the cascade component "pulls up" the deficits from the upper levels), while the increase in α amplifies the contribution of the worst windows; in the $\alpha \gtrsim 0.95$ zone, a plateau is observed, since CVaR averages the same tail of the distribution. The surface ridge in the area of large α and λ reflects the combined effect of cascading and secondary failures and sets the working parameter range for the target functions of an information system on a mobile platform with increased risk and survivability requirements.

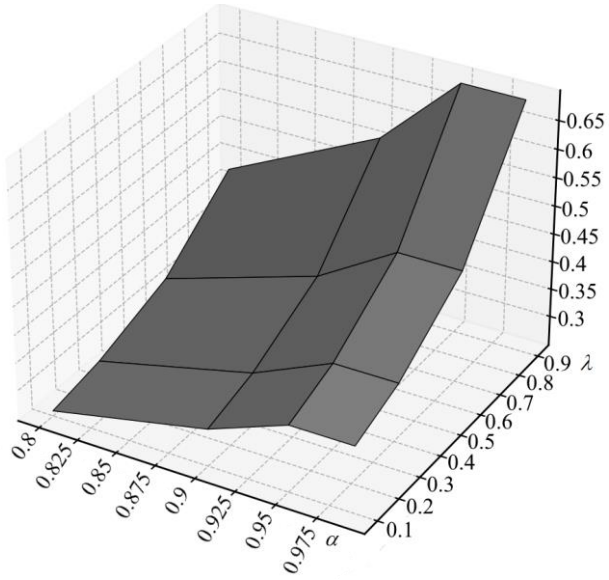


Fig. 5. Three-dimensional surface – $\mathrm{ISM}_\alpha(\alpha, \lambda)$ s for cascade disturbances with secondary markers

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

The stand confirmed the key properties of the metric: monotonicity in cascade depth, tail sensitivity to rare critical episodes, structural informativeness (distinction between primary and secondary degradations), and robustness to observation gaps due to availability windows. Thanks to the developed procedure for calculating the integral metric in real time, the calculations are lightweight and suitable for on-board applications when the information system is deployed on a mobile platform, and the parameters considered earlier in the study are interpreted and controlled under the target function of the information system.

Scope of application of the integrated survivability metric

The proposed integrated survivability metric is designed for information systems on a mobile platform and reflects their specific conditions (availability windows, partial observability, cascades of "data–processes–resources"). Outside this class of information systems, the metric can be applied by reconfiguring SLO thresholds, dependency graph topology, and aggregation parameters, but empirical validation of such transfers is not the subject of this study.

Conclusions

The paper develops an integrated metric of information system survivability on a mobile platform (ISM_α) as a single scalar functional operating across the entire multilevel structure of "data–processes–resources" and reconciles temporally weighted deficits in relation to service requirements with cascading propagation of failures and the contribution of secondary functional failures. A single-pass real-time calculation procedure is proposed based on the localization of deficits at nodes, cascade correction with messages according to the dependency graph, exponential time convolution in availability windows, and risk-sensitive aggregation through conditional average excess (CVaR), which makes the metric suitable for use with intermittent connectivity and limited resources. The limitations, monotonicity, and resistance – ISM_α s to small parameter perturbations and incomplete observations are analytically justified; the concept of an operability zone with agreed thresholds and configuration comparison rules is introduced, which ensures controllability in decision-making. Experimental scenarios confirmed the correct ranking of configurations in cases where basic indicators (proportion of time outside demand, average deficit) are uninformative, and also demonstrated tail sensitivity to rare critical episodes and the ability to distinguish primary disturbances from induced secondary ones. A method for adjusting parameters is proposed separately, and the reproducibility of the results is demonstrated (all related FAIR/O data are added).

The scientific novelty lies in the fact that the paper proposes for the first time an integral metric of the survivability of an information system on a mobile platform, which simultaneously aggregates temporally weighted service deficits, corrects them taking into account their cascading propagation in the "data–processes–resources" graph, introduces a system penalty for secondary failures due to consistency invariants, and amplifies the contribution of rare but critical episodes using CVaR, with the entire procedure implemented in a single pass with an operation on availability windows. For the first time in such a coordinated format, service semantics, structural cascading, and secondary nature are combined in a single functionality suitable for comparing configurations and managing a viable area, as well as for working with intermittent communication; Additionally, robust comparison rules are proposed, taking into account the uncertainty of evaluation, which ensures the stability of conclusions with a finite sample of windows.

The proposed metric provides researchers with a single indicator of the state of the information system with a decomposition of the contribution of services and secondary failures, serves as an early warning tool and allows prioritizing recovery actions – reducing cascading through redistribution of dependencies and loads or eliminating secondary failures through restoration of time, integrity, and queue invariants. ISM_α It supports the selection and validation of configurations according to agreed survivability thresholds, ensures comparability between the target functions of the information system through weight normalization, and can be used in a mode of minimal use of additional information system resources without heavy matrix operations, which facilitates integration into industrial and embedded information systems. Working with availability windows makes the assessment resistant to intermittent communication and telemetry gaps, ensuring support for the mobility conditions of the platform on which the information system can be deployed.

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

Further work involves identifying the structure of dependencies and weight parameters of telemetry metrics, developing adaptive schemes for selecting levels α and time weights $w(t)$ in real time, integrating ISM_α with predictive and control strategies for selecting localization and recovery policies, deepening the dictionary of invariants for more accurate detection of secondary failures and reduction of false positives, studying robustness to shifts in environment distributions (in highly mobile platforms), scaling the approach to groups of information systems (megasystems) with shared risk budgeting [26–27], as well as verification on real event logs of operation with calibration of survivability thresholds and assessment of the effect of changes in architectural and operational decisions in load balancing tasks of multiprocessor computer systems [28] and virtual heterogeneous data centers [29].

References

1. Dodonov, A., Lande, D. (2021), "Modeling the survivability of network structures", *CEUR Workshop Proceedings*, Vol. 2859, P. 1–10. Available at: https://ceur-ws.org/Vol-2859/paper1.pdf (accessed 01 October 2025).
2. Andriulo, F. C. et al. (2024), "Edge computing and cloud computing for Internet of Things: a review", *Informatics*, Vol. 11, No. 4, 71. DOI: https://doi.org/10.3390/informatics11040071.
3. Coletta, A. et al. (2024), "A 2-UAV: application-aware resilient edge-assisted UAV networks", *Computer Networks*, Vol. 256, 110887. DOI: https://doi.org/10.1016/j.comnet.2024.110887
4. Tkachov, V. et al. (2021), "Method to determine fault-tolerant performance probability of high-survivability computer network based on mobile platform", *2021 IEEE 8th International Conference on Problems of Infocommunications, Science and Technology (PIC S&T)*, Kharkiv, Ukraine, 5–7 October 2021, P. 1–5. DOI: https://doi.org/10.1109/picst54195.2021.9772202
5. Fesenko, H. et al. (2024), "Methods and software tools for reliable operation of flying LiFi networks in destruction conditions", *Sensors*, Vol. 24, No. 17, 5707. DOI: https://doi.org/10.3390/s24175707
6. Buckley, I. A., Fernandez, E. B. (2023), "Dependability patterns: a survey", *Computers*, Vol. 12, No. 10, 214. DOI: https://doi.org/10.3390/computers12100214
7. Mehdi, I., Boudi, E. M., Mehdi, M. A. (2024), "Reliability, availability, and maintainability assessment of a mechatronic system based on timed colored Petri nets", *Applied Sciences*, Vol. 14, No. 11, 4852. DOI: https://doi.org/10.3390/app14114852
8. Su, H., Li, Y., Cao, Q. (2024), "A stochastic model of preventive maintenance strategies for wind turbine gearboxes considering the incomplete maintenance", *Scientific Reports*, Vol. 14, 5700. DOI: https://doi.org/10.1038/s41598-024-56436-0
9. Zhao, Y. et al. (2024), "Failure dependence and cascading failures: a literature review and investigation on research opportunities", *Reliability Engineering & System Safety*, Vol. 256, 110766. DOI: https://doi.org/10.1016/j.ress.2024.110766
10. Hu, T. et al. (2025), "Dynamic recovery and a resilience metric for UAV swarms under attack", *Drones*, Vol. 9, No. 8, 589. DOI: https://doi.org/10.3390/drones9080589
11. Xu, X., Fu, X. (2023), "Analysis on cascading failures of directed – undirected interdependent networks with different coupling patterns", *Entropy*, Vol. 25, No. 3, 471. DOI: https://doi.org/10.3390/e25030471
12. Elewaily, D. I. et al. (2023), "Delay/disruption-tolerant networking-based the integrated deep-space relay network: state-of-the-art", *Ad Hoc Networks*, Vol. 152, 103307. DOI: https://doi.org/10.1016/j.adhoc.2023.103307
13. AlHidaifi, S. M., Asghar, M. R., Ansari, I. S. (2024), "A survey on cyber resilience: key strategies, research challenges, and future directions", *ACM Computing Surveys*, Vol. 56, No. 8, P. 1–48. DOI: https://doi.org/10.1145/3649218
14. Qazi, F. et al. (2024), "Service level agreement in cloud computing: taxonomy, prospects, and challenges", *Internet of Things*, Vol. 25, 101126. DOI: https://doi.org/10.1016/j.iot.2024.101126
15. Cao, Z. et al. (2025), "A resilience quantitative assessment framework for cyber-physical systems: mathematical modeling and simulation", *Applied Sciences*, Vol. 15, No. 15, 8285. DOI: https://doi.org/10.3390/app15158285
16. Ruban, I., Tkachov, V. (2025), "Formalizing the survivability target function of information system on mobile", *Eighth International Scientific and Technical Conference on Computer and Information Systems and Technologies*, Kharkiv, Ukraine, 9–10 October 2025, P. 32–34. Available at: http://csitic.com/images/data/CSITIC.2025.pdf

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

17. Samet, D., Tauman, Y., Zang, I. (1984), "An application of the Aumann–Shapley prices for cost allocation in transportation problems", *Mathematics of Operations Research*, Vol. 9, No. 1, P. 25–42. DOI: https://doi.org/10.1287/moor.9.1.25
18. Babenko, V. F., Parfinovych, N. V., Skorokhodov, D. S. (2022), "The best approximation of closed operators by bounded operators in Hilbert spaces", *Carpathian Mathematical Publications*, Vol. 14, No. 2, P. 453–463. DOI: https://doi.org/10.15330/cmp.14.2.453-463
19. Marshall, A. W., Olkin, I., Arnold, B. C. (2011), *Inequalities: Theory of Majorization and Its Applications*, Springer, New York, 909 p. DOI: https://doi.org/10.1007/978-0-387-68276-1
20. Hasler, D., Koberstein, J. (2024), "On the expansion of resolvents and the integrated density of states for Poisson distributed Schrödinger operators", *Complex Analysis and Operator Theory*, Vol. 18, No. 5, 128. DOI: https://doi.org/10.1007/s11785-024-01546-w
21. Baricz, Á., Pogány, T. K. (2023), "Probabilistic and analytical aspects of the symmetric and generalized Kaiser–Bessel window function", *Constructive Approximation*, Vol. 58, P. 713–783. DOI: https://doi.org/10.1007/s00365-023-09627-3
22. Audet, C. et al. (2020), "Performance indicators in multiobjective optimization", *European Journal of Operational Research*, Vol. 292, No. 2, P. 397–422. DOI: https://doi.org/10.1016/j.ejor.2020.11.016
23. Weir, A. J. (1973), *Lebesgue Integration and Measure*, Cambridge University Press, Cambridge, 281 p.
24. Tkachov, V. (2025). ISM (integrated survivability metric) – experimental package. GitHub. Available at: https://github.com/tikey/ism-integrated-survivability-metric (accessed: 16 October 2025).
25. Tkachov, V. and Ruban, I. (2025). ISM experimental results (S1–S4), v1.0. Zenodo. DOI: https://doi.org/10.5281/zenodo.17390902
26. Lemeshko, O., Papan, J., Yeremenko, O., Yevdokymenko, M., Segec, P., (2021). "Research and Development of Delay-Sensitive Routing Tensor Model in IoT Core Networks", *Sensors*, Vol. 21, No. 11, 3934. DOI: https://doi.org/10.3390/s21113934
27. Papan, J., Segec, P., Yeremenko, O., Bridova, I., Hodon, M., (2020). "Enhanced Multicast Repair Fast Reroute Mechanism for Smart Sensors IoT and Network Infrastructure", Sensors, Vol. 20, No. 12, 3428. DOI: https://doi.org/10.3390/s20123428
28. Kuchuk, N., Zakovorotnyi, O., Pyrozhenko, S., Radchenko, V., Kashkevich, S., (2025). "A Method for Redistributing Virtual Machines of Heterogeneous Data Centres". *Advanced Information Systems*, Vol. 9, no. 1, P. 80–85. DOI: 10.20998/2522-9052.2025.1.09
29. Kuchuk, N., Zakovorotnyi, O., Radchenko, V., Andrusenko, Y., Lysytsia, D., (2025). "Load Balancing of a Multiprocessor Computer System Using the Method Particle Swarm Optimization", *Advanced Information Systems*. Vol. 9, no. 4, P. 82–88. DOI: 10.20998/2522-9052.2025.4.11

Received (Надійшла) 01.10.2025
Accepted for publication (Прийнята до друку) 30.11.2025
Publication date (Дата публікації) 28.12.2025

Відомості про авторів / About the Authors

Tkachov Vitalii – PhD (Engineering Sciences), Associate Professor, Kharkiv National University of Radio Electronics, Doctoral Candidate at the Department of Electronic Computers, Kharkiv, Ukraine; e-mail: vitalii.tkachov@nure.ua; ORCID ID: https://orcid.org/0000-0002-6524-9937; SCOPUS ID: https://www.scopus.com/authid/detail.uri?authorId=56485859400

Ruban Ihor – Doctor of Sciences (Engineering), Professor, Kharkiv National University of Radio Electronics, Professor at the Department of Electronic Computers, Kharkiv, Ukraine; e-mail: ihor.ruban@nure.ua; ORCID ID: https://orcid.org/0000-0002-4738-3286; SCOPUS ID: https://www.scopus.com/authid/detail.uri?authorId=7004018101

Ткачов Віталій Миколайович – кандидат технічних наук, доцент, Харківський національний університет радіоелектроніки, докторант кафедри електронних обчислювальних машин, Харків, Україна.

Рубан Ігор Вікторович – доктор технічних наук, професор, Харківський національний університет радіоелектроніки, професор кафедри електронних обчислювальних машин, Харків, Україна.

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

ІНТЕГРАЛЬНА МЕТРИКА ЖИВУЧОСТІ ІНФОРМАЦІЙНОЇ СИСТЕМИ НА МОБІЛЬНІЙ ПЛАТФОРМІ В УМОВАХ КАСКАДНИХ І ВТОРИННИХ ФУНКЦІОНАЛЬНИХ ЗБОЇВ

Предметом дослідження є інтегральна метрика живучості інформаційної системи на мобільній платформі за переривного зв'язку, часткової спостережуваності та каскадних і вторинних збоїв. Систему подано як багаторівневу структуру "дані – процеси – ресурси". **Мета роботи** – розробити інтегральну метрику живучості, яка бере до уваги часові відхилення від вимог, їх поширення графом залежностей і приховані порушення; запропонувати однопрохідний алгоритм і довести його властивості на сценаріях. Для досягнення окресленої мети необхідно виконати так **завдання**: формалізувати сервісні вимоги й проєкцію структури; побудувати метрику з ризик-орієнтованим агрегуванням дефіцитів, каскадною корекцією та системним урахуванням вторинних збоїв; розробити однопрохідні обчислення у "вікнах доступності"; довести монотонність, інваріантність масштабу й стійкість до пропусків; визначити правила налаштування параметрів і порівняння конфігурацій; експериментально перевірити й зіставити з базовими індикаторами. Упроваджено такі **методи:** проєкція сервісів на рівні даних, процесів і ресурсів; використання умовного середнього перевищення як ризик-орієнтованої агрегації; каскадна корекція за глибиною та шириною; організація фіксації вторинних збоїв і розсинхронізації; нормування у "вікнах доступності"; однопрохідні оновлення близької до лінійної складності. **Досягнуті результати:** запропоновано й формально визначено інтегральну метрику живучості; доведено її монотонність за параметрами, обмеженість, інваріантність до масштабування й стійкість до пропусків; продемонстровано відмінність від середнього дефіциту – розроблена метрика підсилює внесок рідкісних глибоких провалів і реагує на каскадність, тоді як середні значення майже сталі; за відсутності сервісних дефіцитів зберігається додатний рівень завдяки виявленню латентних вторинних збоїв; на сценаріях отримано узгоджені сімейства кривих і тривимірну поверхню, які демонструють кероване налаштування чутливості та стійке ранжування конфігурацій для промислових умов експлуатації мобільних платформ. **Висновки:** запропонована метрика забезпечує сервісно узгоджену оцінку стану інформаційної системи, одночасно зважаючи на часові дефіцити, каскадне поширення та вторинні збої; інтегральна метрика придатна для послідовних обчислень у ресурсно обмежених умовах, підсилює раннє виявлення ризиків і підтримує моніторинг, локалізацію, забезпечує живучість.

Ключові слова: інтегральна метрика; живучість; інформаційна система; каскадний збій; ризик-орієнтований моніторинг.

Бібліографічні описи / Bibliographic descriptions

Ткачов В. М., Рубан І. В. Інтегральна метрика живучості інформаційної системи на мобільній платформі в умовах каскадних і вторинних функціональних збоїв. *Сучасний стан наукових досліджень та технологій в промисловості*. 2025. № 4 (34). С. 78–100. DOI: https://doi.org/10.30837/2522-9818.2025.4.078

Tkachov, V., Ruban, I. (2025), "Integral survivability metric of an information system on a mobile platform under functional cascading and secondary failures", *Innovative Technologies and Scientific Solutions for Industries*, No. 4 (34), P. 78–100. DOI: https://doi.org/10.30837/2522-9818.2025.4.078

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

UDC 004.358:681.518
DOI: https://doi.org/10.30837/2522-9818.2025.4.101

V. FILATOV, O. ZOLOTUKHIN, M. KUDRYAVTSEVA

INTELLECTUAL DATA ANALYSIS IN RELATIONAL INFORMATION AND ANALYTICAL SYSTEMS

The subject of the study is the methods of intellectual analysis, namely the construction of a decision tree, associative analysis, the identification of patterns between related events based on data presented by a relational model. **The purpose of the study** is to analyze the features of information units and data structures, using the example of relational systems that affect the technology of knowledge extraction. **Tasks:** the article solves the following tasks: to consider the relational data model as the most popular and effective data structure used in intelligent information systems for data processing and storage; to analyze the operations of relational algebra, the operational component of the relational data model regarding the application of aggregate functions; to develop a general formal statement of the problem of knowledge extraction from a relational database; to consider the concept of functional associative rules; the ID3 decision tree generation algorithm focused on data processing in relational systems is analyzed. **The following methods are implemented**: modern view and trends in the field of data mining; features of building information systems based on relational databases, relational algebra, theory of normalization of relations; analysis of literature on the topic of research; comparative analysis. **Results achieved:** the relational data model is considered as the most effective data structure used in intelligent information systems for data processing and storage. A group of aggregate functions of relational databases is identified and analyzed with respect to key attributes of the relation, which makes it possible to build logical dependencies between information units of the subject area being analyzed. The task of extracting knowledge from the database is formally formulated. The concept of functional associative rules is introduced. The ID3 decision tree generation algorithm focused on data processing in relational systems is carefully analyzed. The semantic network (SN), built on the basis of the proposed approach, allows to increase the efficiency of decision support systems. **Conclusions:** the universal approach proposed in the article to build a relational data model of an information system for searching for associative patterns in data allows to solve a whole class of typical tasks in which objects are connected by a "many-to-many" relationship or $M \rightarrow N$. The relational database model is proposed as a universal information structure for solving associative analysis tasks and presenting knowledge in the form of a semantic network. The examples given in the article confirm the effectiveness of the developed and considered approaches to solving the problem of data mining in the environment of relational systems. Solving the problem of identifying knowledge in data will allow to improve the quality of management decisions made.

Keywords: relational model, data mining, associative dependencies, database, decision tree, semantic network, information system.

Introduction

The main trends in the development of information systems and databases (DB) are, on the one hand, decentralization and distribution of information management resources, and on the other hand, heterogeneity of data storage structures of information systems. Thus, distributed automated systems become the basis of the information structure for data storage, processing, and management and ensure the informatization of society – the gradual creation of a unified information space. Given the expansion of regional, national, and international integration processes in the economy and business, the demands on the performance of distributed intelligent automated systems are growing [1].

The development of computing technology and the increase in the volume of stored information have led to the need to separate database technology into a separate field of intensive scientific research. Over more than 40 years of history, databases have evolved from automated workstations, so-called ARMs, file and client-server technologies to information spaces. Information management is undergoing changes in the following aspects: storage and access models, the scale of projected systems – from databases to intelligent decision support systems. New areas of research are rapidly developing in this field: integrated information systems and technologies, information spaces and communities, systems based on computational intelligence; issues related to information analysis remain relevant. The development of effective methods for analyzing large volumes of primary, or so-called "raw", data will enable the acquisition of new knowledge and the making of informed decisions.

There is currently a rapid growth in the number of software products that use new technologies, as well as in the types of tasks where their application provides

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)
Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

significant economic benefits. Elements of automatic data processing and analysis, known as *Data Mining* (knowledge extraction), are becoming an integral part of the concept of databases, electronic data warehouses, and the organization of intelligent computing [2, 3].

However, traditional mathematical statistics, which has long claimed to be the main tool for analyzing information, is effective for solving real-life complex problems. It operates with averaged sample characteristics, which are often fictitious values (such as the average temperature of patients in a hospital). Therefore, mathematical statistics methods are mostly useful for testing pre-formulated hypotheses. *Data Mining* is increasingly becoming a multidisciplinary field that has emerged on the basis of achievements in various sciences [4–6]. Hence, there are a significant number of methods and algorithms implemented in various existing *Data Mining* systems. Many of them integrate several approaches. However, each system usually has a key element on which the main focus is concentrated.

In the development of this direction, the use of such mathematical apparatus in database technology as aggregate functions is of interest. Aggregate functions are functions that determine the number of records in a table, count the number of values in a column or find the minimum and maximum values for it, and also sum up the data. Aggregate functions include *COUNT, SUM, MAX, MIN, AVG*, and probably others that may be offered by the developer of a particular system. To apply aggregate functions in calculations of a group of identical values, the *GROUP BY* parameter is used. It "compresses" identical values of a given attribute into a single row of summary results. For example, to find the average price of a part, you can formulate a query in *SQL* [7, 8].

This article is devoted to the study of information systems that combine relational databases as a source of primary information and means of intelligent analysis.

The purpose of the study

The purpose of the study is to analyze the relational data model, namely the impact of the normalization level of the relational database schema on knowledge extraction technology using the example of developing a semantic model of knowledge representation and associative data analysis. The results achieved will make it possible to formulate conditions for effective collaboration with analytical processing systems based on *Data Mining* methods.

Relational database as an information storage environment in intelligent analytical systems

A relational DB is based on the relational model. This model is defined by a simple data structure, a user-friendly tabular representation, and the ability to use the formal apparatus of relational algebra and relational calculus for information processing [9–12].

Traditional means of specifying the relational data model are used to construct the structural diagram of databases.

The basic structural unit of data in the relational model is an n-ary relation, which is a finite subset of the Cartesian product of domains, i.e., sets of atomic values of data elements – attributes of the relation.

Let R be a finite subset of the names of relations in the database;

D – $\{D_1,...,D_i\}$ – a set of domains where each domain is a named set of atomic values of data elements;

A – final set of attribute names of a relation;

dom – reproduction from A to D, which determines from which domain the attribute values are selected.

A pair $\langle A_i, domA_i \rangle$, where $A_i \in A$, is called an attribute.

The structural diagram S_i of the relation $R_i (R_i \in R)$ can be presented in the form $R_i (A_1,..., A_n)$, in which all A_i are different. The relation r_i can be defined as an extension of the diagram S_i: $r_i \subseteq domA_1 \times ... \times domA_n$.

Reordering attributes in the schema does not generate a new extension, and the set $\{A_1,...A_n\}$ of relation attributes R_i defines the relation type. The expression $R_i = A_1...A_n$ is used to specify the composition of the medium. The structural schema U of a relational database is a specification of the form $(R_1,..., R_p)$, where $R_i \in R$ and all R_i are different [13].

The task of identifying knowledge in relational databases is considered, the effective solution of which will improve the quality of management decisions.

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

Building a decision tree based on data presented by a relational model

Formally, the task of extracting knowledge from a DB can be demonstrated as follows. The subject area is reproduced in the form of a relational model, presented as a universal relation R, as a subset of tuples of the Cartesian product

$$R = \left(DX_1, DX_2, ..., DX_n, DY_1, ..., DY_m\right) = \\ = \left\{< x_1, ..., x_n, y_1, ..., y_m >\right\},$$

where x_i – values of input attributes X_i from the domain DX_i; y_i – values of output attributes Y_j from the domain DY_j; $P\left(x_1, ..., x_n, y_1, ..., y_m\right)$ – predicate as a condition for reproducing a specific subject area in a tuple of attribute values $< x_1, ..., x_n, y_1, ..., y_m >$.

The objective of the study is to develop a set of rules:

$$\left\{X_1, .X_2, ..., X_n\right\} \to \left\{Y_1, Y_2, ..., Y_m\right\},$$

which assign a set of target values $\left\{y_j = DY_j, j = \overline{1,m}\right\}$ to each incoming set of values

$$\left\{x_i = DX_i, i = \overline{1,n}\right\}.$$

The obtained functional dependencies

$$Y_j = F_j\left(X_1, X_2, ..., X_n\right), j = \overline{1,m}$$

must be correct for relation tuples and can be used in the process of finding output attributes Y_j for new input attribute values $X_i, i = \overline{1,n}$.

If the database is presented in the form of a strongly typed relation, for example, in third normal form, then for subsequent calculations it is necessary to switch to a universal relation. To transform the model, you can use the join operation contained in the basic operations of relational algebra [14].

The join is equivalent to the following sequence of relational operations:

– renaming of identical attributes in relations A and B;

– Cartesian product of relations A and B;

– selection by corresponding attribute values that had the same names in relations A and B;

– projection with the removal of duplicate attributes from relations A and B;

– renaming attributes of relations A and B, returning them to their original names.

The universal relation obtained by such a transformation is a set of facts, each of which is described by a finite set of discrete attributes [15].

For further data analysis, one of the *Data Mining* methods can be applied – building *Decision Trees*. Decision Trees are currently the most common approach to identifying and reproducing logical patterns in data. Among them are:

- ID3 (*Interactive Dichotomizer*);
- CART (*classification and regression trees*);
- CHAID (*chi square automatic interaction detection*).

Let's take a closer look at the process of building decision trees using the ID3 system as an example. The data table fully meets the requirements for the structure of the relational model discussed above, namely: the key attribute ***Day*** is selected, the tuples are unique, and the attributes of the relation are not repeated.

The ID3 decision tree generation algorithm is an algorithm that builds a tree from the root to the leaves, selecting the attribute that best classifies the data at each node. The input parameters of the algorithm are: Examples – the current training example – the target attribute, Attributes – a set of candidate attributes.

Let's consider the general scheme of the algorithm's operation using the example of a universal relation presented in Table 1. We select the target attribute – this is *Play Tennis*.

At this stage, the root node of the tree is formed. The *InformationGain* is calculated for each candidate attribute, and the one with the *InformationGain* that best matches expressions (1) and (2) is selected.

$$Gain\left(S, A\right) = Entropy\left(S\right) - \sum_{v \in Values(A)} \frac{\left|S_v\right|}{\left|S\right|} Entropy\left(S_v\right) \quad (1)$$

$$Entropy(S) \equiv -p_+ \log_2 p_+ - p_- \log_2 p_-, \quad (2)$$

where p_+ – number of positive examples; p_- – number of negative examples; S – a set of attributes, $S_v = \left\{s \in S \middle| A(S) = v\right\}$.

Table 1 presents a sample containing 14 tuples. To determine the root node, it is necessary to calculate the entropy of all the parameters under study – the attributes of the relation. Among all the attributes, the information threshold is highest in *Outlook*, which is defined on the set of values *Sunny, Overcast*, and *Rain*. According to the values of parameters (1) and (2), a tree letter is formed (Fig. 1).

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

Table 1. *Weather conditions*

Day	Outlook	Temperature	Humidity	Wind	Play Tennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

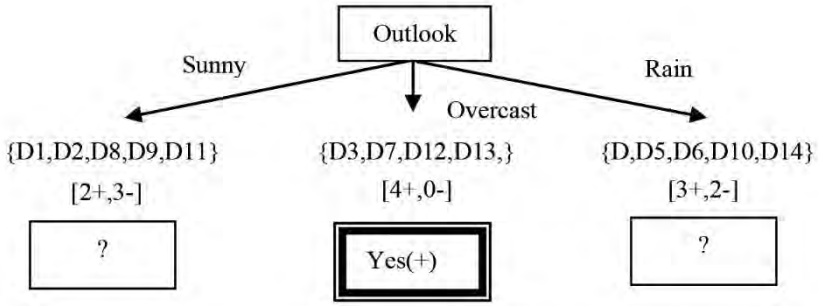


Fig. 1. Example of constructing a decision tree

We perform similar calculations until all the leaves of the tree are formed.

As a result, we obtain the desired decision tree, shown in Fig. 2.

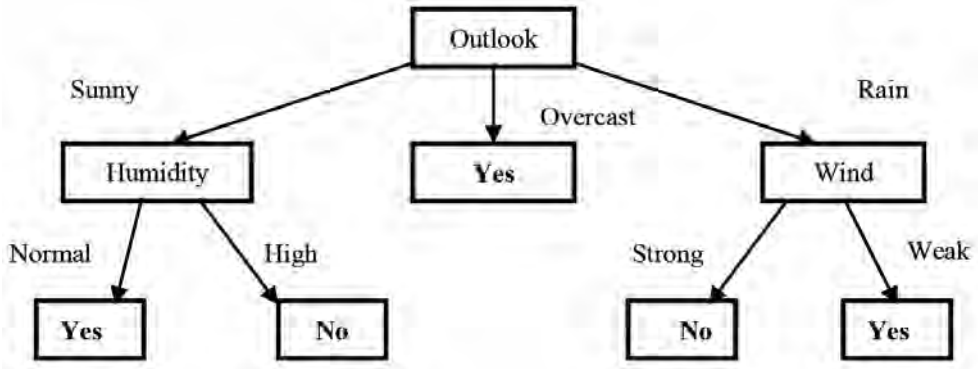


Fig. 2. Decision tree

The above algorithm for extracting and representing knowledge has a number of advantages, among which the following can be highlighted:

– high speed of the knowledge discovery process;

– generation of rules for subject areas in which it is not possible to obtain a formal model of knowledge representation by other methods;

– intuitively understandable classification model of the subject area.

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

The results obtained can be interpreted in the form of one of the classical models of knowledge representation – a semantic network. A semantic network is a directed graph whose vertices are concepts and whose edges are the relationships between them. Concepts are usually abstract or concrete objects, and relations are connections such as "genus-species", "part-whole", "class-subclass", etc. A distinctive feature of semantic networks is the mandatory presence of three types of relationships: "class – class element", "property – value", and "example – class element".

A semantic network provides the following basic functions:

– storage of information about objects and the relationships between them;

– search for objects by various properties;

– replenishment and correction of the system's knowledge during its training;

– implementation of various procedures for generalizing and specifying knowledge.

In general, a semantic network is understood as an expression in the form of $S = \langle O, R \rangle$, where $R = \{R_j, j = \overline{1,k}\}$, $O = \{O_i, i = \overline{1,n}\}$, $O_i, i = \overline{1,n}$ a set of objects in a specific subject area; $R_j, j = \overline{1,k}$ – a set of relations between objects; j – type of relationship.

Let's imagine a decision tree obtained by applying the ID3 algorithm in the form of a semantic network. The root of the tree, *Outlook*, is linked to the child nodes by the relationships *Sunny, Overcast*, and *Rain*. By analyzing all the components of the decision tree – links and relationships – we can represent a set of objects in the subject domain as follows:

Outlook (Sunny, Overcast, Rain)
Humidity (Normal, High)
Wind (Strong, Weak)

Thus, for example, the problem situation "how weather affects tennis" can be described as a semantic network, the structure of which is shown in Fig. 3.

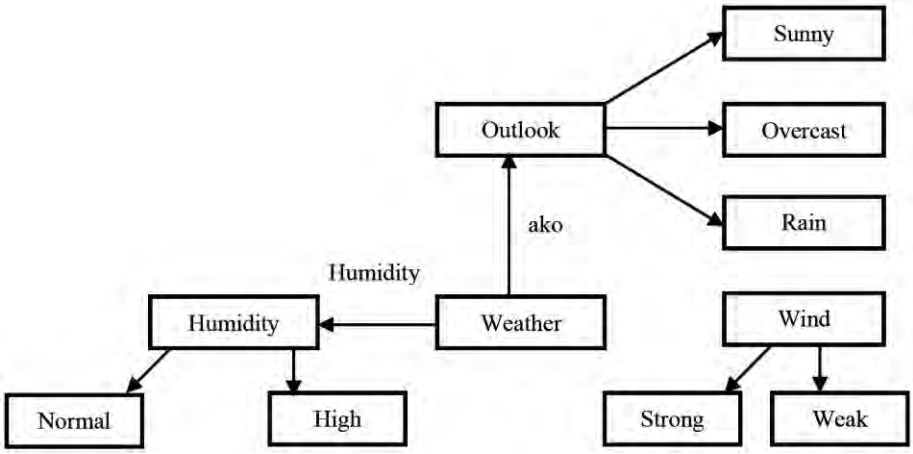


Fig. 3. Semantic network

With the help of a semantic network, it is possible to provide a formal decision-making procedure and pre-record network elements in the form of predicates. From the analysis of the decision tree obtained using the ID3 algorithm, we can draw the disappointing conclusion that three weather parameters follow from this decision: *Outlook, Humidity, Wind*. The predicate will look like this: let's go play (yes/no) (*Outlook, Humidity, Wind*).

So,

If we go play

Yes (*Outlook = Sunny, Humidity = Normal*)

Yes (*Outlook = Rain, Wind = Weak*)

Yes (*Outlook = Overcast*)

If we are not going to play

No (*Outlook = Sunny, Humidity = High*)

No (*Outlook = Rain, Wind = Strong*)

Based on the predicates formulated above, a semantic network can be developed, the appearance of which is shown in Fig. 4.

It should be noted that the approach to intelligent data analysis and knowledge representation in the form of a semantic network proposed in this article can be applied to large-scale databases, the schema of which may contain more than 1,000 attributes.

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)
Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

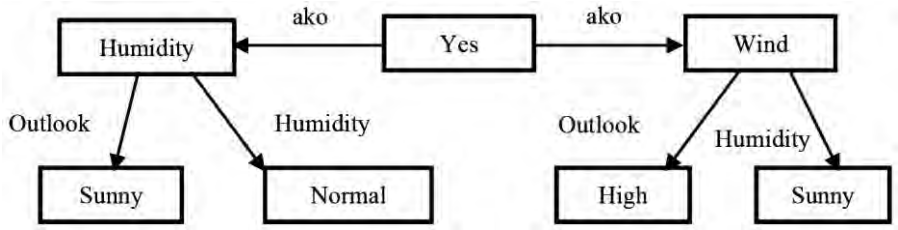


Fig. 4. Semantic network construction

Searching for associative rules in data represented by a relational model

Most often, the method of evaluating the associative properties of data is used in the following areas:

– retail (determining which products should be promoted together; choosing the location of products in the store; analyzing the consumer basket; forecasting demand);

– marketing (searching for market segments, trends in consumer behavior);

– customer segmentation (identifying common characteristics of the company's customers, identifying buyer groups);

– catalog design, sales campaign analysis, determining customer purchase sequences (which purchase will follow the purchase of product A);

– analysis of web logs.

One of the most frequently cited examples of associative rule mining is the Market-Basket Problem. The essence of this problem is to identify products that people buy together. This is necessary so that marketing specialists can place these products in the store in an appropriate manner to increase sales, as well as make other effective decisions. It is the ability to discover hidden rules that makes associative rule mining valuable and contributes to knowledge discovery.

Let us consider this task in a generalized form (3)–(11). Let us denote the objects that make up the sets under study by the set

$$I = \{i_1, i_2, ...i_r, ...i_m\}, \quad (3)$$

where i – objects contained in the sets being analyzed; n – total number of objects.

Transactions are called sets of objects located in the database that will be analyzed. Let's describe a transaction as a subset I:

$$T = \{i_j \mid i_j \in I\}. \quad (4)$$

The set of transactions for which information is available for analysis will be presented as a set D

$$D = \{T_1, T_2, ..., T_r, ..., T_m\}, \quad (5)$$

where m – number of transactions available for analysis.

Set of transactions to which i_j-object belongs,

$$D = \{T_r \mid i_j \subseteq T_r; j = 1...n; r = 1...m\} \subseteq D. \quad (6)$$

We will define the general set of objects as follows:

$$F = \{i_j \mid i_j \in I; j = 1...n\}. \quad (7)$$

The set of transactions to which the set F belongs will be represented as

$$D_F = \{T_r \mid F \subseteq T_r; r = 1...m\} \subseteq D. \quad (8)$$

The ratio of the number of transactions to which the set F belongs to the total number of transactions is defined as the *support* of the set and denoted by $Supp(F)$:

$$Supp(F) = |D_F| / |D|. \quad (9)$$

During the search process, the analyst can specify the minimum support value for the sets that interest him $Supp_{\min}$. A set is called large (*large itemset*) if its support value exceeds the minimum support value specified by the user:

$$Supp(F) > Supp_{\min}. \quad (10)$$

Therefore, when searching for associative rules, it is necessary to find a set in the sets:

$$L = \{F \mid Supp(F) > Supp_{\min}\}. \quad (11)$$

Development of a database model for solving associative analysis problems

Let us consider the options for relational database (RDB) structures that meet the requirements of tasks (3)–(11) for searching for associative rules [16]. Expression (3) describes the objects of intellectual analysis. According to the rules of RDB design, the following model can be considered for representing objects from different subject areas: the "Objects"

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

relation has a key attribute "**Object ID**" and a list of attributes $A_1,...,A_n$. An example of the relationship between "Objects" in general terms and the practical "SUPPLY" is given.

Fig. 5 shows a diagram of the "Objects" relationship. As an example, a fragment of a database for displaying a list of goods is given, which in theoretical-set form looks like this: $SUPPLY = \{ID, NAME, PRICE\}$. In the methodology discussed above (3)–(11), a transaction is understood to be an object or set of second-level objects that are determined by the integration properties of previously located objects. Therefore, the second-level object relationship diagram will be identical to the diagram shown in Fig. 6.

"Objects"

Object ID
Attribute 1
Attribute 2
Attribute 3
Attribute n

"SUPPLY"

ID	Name	Price
101101	LG HD Phone	1452
101102	PCHP321	5687
101103	PRINTER	540

Fig. 5. Example of the "Objects" relations

Tax invoices

ID_N	Delivery	Supplier
101	10-11-2009	MKS
102	11-11-2009	FOX
103	12-11-2009	EPICENTER

Fig. 6. Relation "Tax invoices"

Let us determine the type of relationships between first- and second-level objects. Note that a single transaction may contain several first-level objects and, accordingly, there cannot be duplicate objects in a single transaction.

The type of relationship considered corresponds to the many-to-many multiple type, or $M \to N$. In a relational data model, this type of relationship is implemented through an additional relation [17]. An example of a transaction is a real second-level object with integration properties – a delivery note containing objects (goods). The RDB schema "Waybills_Goods" is shown in Fig. 7.

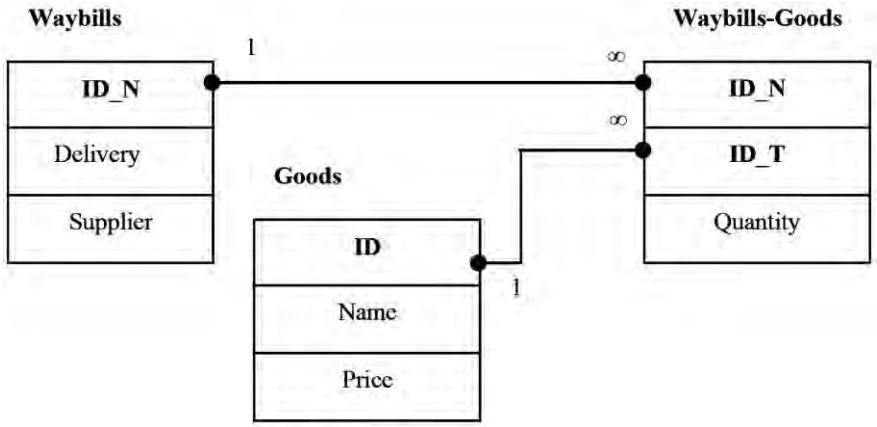


Fig. 7. Database "Waybills_Goods" schema

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

The "Waybills_Goods" database schema contains three relations: "Waybills", "Goods", and "Waybills Goods". Key attributes are marked in bold. A double composite key is used for "Waybills_Goods". This solution allows us to maintain the integrity constraint formulated at the stage of setting the task of synthesizing the database schema: one item can be added twice to one waybill. An example of the "Waybills_Goods" relation is shown in Fig. 8.

Waybills_Goods

ID_N	Supplier	**Goods ID**	Name	Quantity
101	MKS	**101102**	PC HP321	1
101	MKS	**101103**	PRINTER	1
102	FOX	**101103**	PRINTER	2
103	EPICENTER	**101101**	LG HD Phone	1
103	EPICENTER	**101102**	PC HP321	1
103	EPICENTER	**101103**	PRINTER	1

Fig. 8. Database "Waybills_Goods" schema

The above relationship "Invoices_Goods" meets the normalization conditions: goods PRINTER and PC HP321 are added to invoice No. 101, invoice No. 102 contains one item PRINTER, and invoice No. 103 contains three goods: PRINTER, PC HP321, and LG HD phone.

An arbitrary set of objects added to the analyzed ones forms a set for associative analysis. This parameter is variable and is a subset of the set:

$$I = \{LG\ HD\ phone, PC\ HP321, PRINTER\}.$$

Let $F = \{LG\ HD\ phone\}$, then the content of the associative analysis task (9)–(11) will be as follows: determine how many times an object $LG\ HD\ phone$ occurs in the total number of goods received according to all invoices $Supp(LG\ HD\ phone) = 1/6$.

The universal approach to building a relational data model for an information system that searches for associative patterns in data, as proposed in this article, makes it possible to solve a whole class of typical problems in which objects are related by a "many-to-many" relationship, or $M \rightarrow N$ [18]. Examples of such tasks include: a polyclinic (this task involves two entities in a $M \rightarrow N$ relationship: "Patients" and "Doctors"); library ("Readers" and "Literature"); audio and video media rental ("Customers" and "Discs"), etc.

As an example, let's take the task of searching for associative patterns in the subject area "POSTGRADUATE STUDIES" of a higher education institution (HEI) in Ukraine. Let's consider one of the most important tasks performed by this department. PhD and doctoral dissertations are defended by applicants in specialized councils, which include scientists who are leading experts in scientific fields in their specialties. A specialized council may consist of 5 (PhD defense) to 25 specialists, if it is a specialized doctoral council. In addition, according to the requirements, one scientist-specialist in a specific field of research cannot participate in the work of more than two specialized councils. The universal relationship of the relational database is shown in Fig. 9.

Specialist_Council_Specialty

id	Last Name	HEI	Tel	**Council**	Address	Specialty
1	Ivchenko	NURE	33-22	Д.01.001	KhPI	05.13.06
1	Ivchenko	NURE	33-22	Д.02.011	KhAI	05.13.06
2	Petrenko	NURE	33-22	Д.01.001	KhPI	05.13.23
3	Skidan	KhPI	11-44	Д.01.001	KhPI	05.13.06
3	Skidan	KhPI	11-44	Д.02.011	KhAI	05.13.23

Fig. 9. Relationship "Specialist_Council_Specialty"

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

As a result of normalizing the universal relation shown in Fig. 9, the relational DB schema can be transformed and take the form shown in Fig. 10.

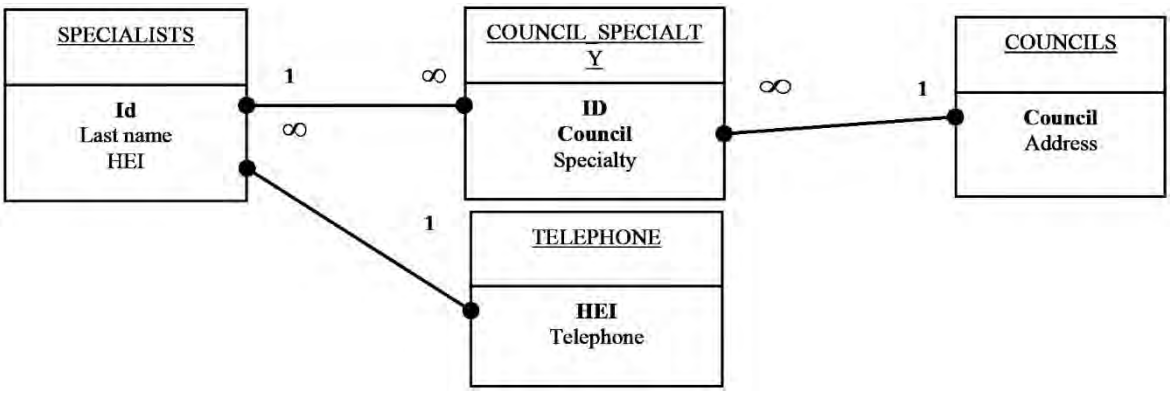


Fig. 10. Schema of the relational DB "Specialist_Council_Specialty"

The COUNCIL-SPECIALITY relation fully satisfies the requirements for a data structure oriented towards the application of associative pattern search technology (3)–(11).

The example considered confirms that normalized relations of the relational data model, reduced to 3 normal form, containing a double composite key, allow for the most effective application of associative data analysis methods [19].

Conclusions

The scientific novelty of the research lies in a comprehensive analysis of the relational data model, namely, identifying the impact of the normalization level of the RDB schema on knowledge extraction technology using the example of developing a semantic model of knowledge representation and associative data analysis. The results obtained make it possible to formulate conditions for effective collaboration with analytical processing systems based on *Data Mining* methods.

The article examines the features of relational algebra operations regarding the application of aggregate functions and considers the concept of functional associative rules, which contributes to the implementation of the classic ID3 decision tree generation algorithm focused on information processing in relational systems. The universal approach to building a relational data model of an information system for searching for associative patterns in data, proposed in the article, makes it possible to solve a whole class of typical tasks in which objects are related by a "many-to-many" relationship, or $M \to N$. The semantic network, built on the basis of the proposed approach, helps to improve the efficiency of decision support systems.

The relational DB model, proposed as a universal information structure for solving associative analysis tasks, will improve the efficiency of the analyzed approach through its implementation in modern information systems.

The examples given in the article confirm the effectiveness of the developed and considered approaches to performing the task of intelligent data analysis in the environment of relational systems. Solving the task of identifying knowledge in data will contribute to improving the quality of management decisions.

Theoretical and practical issues of automatic or automated construction (generation) of the structure and model of an expert system knowledge base, based, for example, on a production model and, therefore, the integration of information system data designed on a relational data model and an expert system, can be considered as a prospect for further development of this area of research.

References

1. Xuanhe, Z., Chengliang, C., Guoliang, L., Ji, S. (2022), "Database Meets Artificial Intelligence: A Survey". *IEEE Transactions on Knowledge and Data Engineering,* Vol. 34, No. 3, P. 1096–1116. DOI: 10.1109/TKDE.2020.2994641
2. Chen, J., Sun, J., Wang, G. (2022), "From Unmanned Systems to Autonomous Intelligent Systems". *Engineering,* 12(5), P. 16–19. DOI: 10.1016/j.eng.2021.10.007

3. Zhang, F., Yuan, N. J., Lian, D., Xie, X., Ma, W.-Y. (2016), "Collaborative knowledge base embedding for recommender systems". *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, P. 353–362. DOI: 10.1145/2939672.2939673
4. Avrunin, O., Vlasov, O., Filatov, V. (2020), "Model of semantic integration of information systems properties in relay database reengineering problems". *Innovative Technologies and Scientific Solutions for Industries*, 4 (14), P. 5–12. DOI: 10.30837/itssi.2020.14.005
5. Cappuzzo, R., Papotti, P., Thirumuruganathan, S. (2020), "Creating embeddings of heterogeneous relational datasets for data integration tasks". *In Proceedings of the 2020 ACM SIGMOD international conference on management of data*. P. 1335 – 1349. DOI: 10.1145/3318464.3389742
6. Shi, C., Li, Y., Zhang, J., Sun, Y., Yu, P. S. (2017), "A Survey of Heterogeneous Information Network Analysis". *IEEE Transactions on Knowledge and Data Engineering*, 29(1), P. 17–37. DOI: 10.1109/TKDE.2016.2598561
7. Parciak, M., Weytjens, S., Hens, N., Neven, F., Peeters, L. M., Vansummeren, S. (2025), "Measuring approximate functional dependencies: a comparative study". *The VLDB Journal*, 34(4). DOI: 10.1007/s00778-025-00931-x
8. Filatov, V., Doskalenko, S. (2018), "On the Approach to Searching for Functional Dependences of Data in Relational Systems". *Innovative Technologies and Scientific Solutions for Industries*, 1 (3), P. 54–58. DOI: 10.30837/2522-9818.2018.3.054
9. Filatov, V., Semenets, V., Zolotukhin, O. (2019), "Synthesis of Semantic Model of Subject Area at Integration of Relational Databases". *2019 IEEE 8th International Conference on Advanced Optoelectronics and Lasers (CAOL).* DOI: 10.1109/caol46282.2019.9019532
10. Filatov V., Kovalenko A. (2019), "Fuzzy Systems in Data Mining Tasks". *Advances in Spatio-Temporal Segmentation of Visual Data. Studies in Computational Intelligence*, Vol 876. Springer, Cham P. 243–274. DOI: 10.1007/978-3-030-35480-0_6
11. Glava, M., Malakhov, V. (2018), "Information Systems Reengineering Approach Based on the Model of Information Systems Domains". *International Journal of Software Engineering and Computer Systems (IJSECS).* Vol. 4, P. 95–105. DOI: 10.15282/ijsecs.4.1.2018.8.0041
12. Codd, E. F. (1983), "A relational model of data for large shared data banks". *Communications of the ACM*. Vol. 26, No. 1. P. 64–69. DOI: 10.1145/357980.358007
13. Maier, D. (1983), *The theory of relational databases*. London: Pitman, 637 p.
14. Wan, X., Han, X., Wang, J., Li, J. (2024), "Efficient Discovery of Functional Dependencies on Massive Data". *IEEE Transactions on Knowledge and Data Engineering*, 36(1), P. 107–121. DOI: 10.1109/tkde.2023.3288209
15. Wang, Y. (2025), "Design and Implementation of a General Data Collection System Architecture Based on Relational Database Technology. In: Xu, Z., Alrabaee, S., Loyola-González, O., Ab Rahman, N.H. (eds) *Cyber Security Intelligence and Analytics. CSIA 2024. Lecture Notes in Networks and Systems,* Vol 1351. Springer, Cham. DOI: 10.1007/978-3-031-88287-6_53
16. Date, C. J. (2003), *Introduction to database systems.* Pearson Education, Limited. 1024 p.
17. Hector Garcia-Molina, Jennifer Widom, Jeffrey D. Ullman (2013), *Database systems the complete book*. Pearson India Education, 1139 p.
18. Sliusarenko, T., Filatov, V. (2023), "Relational vs non-relational databases". *Grail of science.* No. 23. P. 269–271. DOI: 10.36074/grail-of-science.23.12.2022.41
19. Filatov, V. O., Yerokhin, A. L., Zolotukhin, O. V., Kudryavtseva, M. S. (2019), "Information space model in tasks of distributed mobile objects managing". *Information Extraction and Processing*, 2019(47), P. 80–86. DOI: 10.15407/vidbir2019.47.080

Received (Надійшла) 06.06.2025
Accepted for publication (Прийнята до друку) 30.11.2025
Publication date (Дата публікації) 28.12.2025

Відомості про авторів / About the Authors

Filatov Valentin – Doctor of Sciences (Engineering), Professor, Kharkiv National University of Radio Electronics, Professor of Artificial Intelligence Department, Kharkiv, Ukraine; e-mail: valentin.filatov@nure.ua; ORCID ID: https://orcid.org/0000-0002-3718-2077; Scopus ID: https://www.scopus.com/authid/detail.uri?authorId=56911938100

Zolotukhin Oleh – PhD (Engineering Sciences), Associate Professor, Kharkiv National University of Radio Electronics, Dean of Computer Science Faculty, Kharkiv, Ukraine; e-mail: oleg.zolotukhin@nure.ua; ORCID ID: https://orcid.org/0000-0002-0152-7600; Scopus ID: https://www.scopus.com/authid/detail.uri?origin=resultslist&authorId=57207774022

Kudryavtseva Maryna – PhD (Engineering Sciences), Associate Professor, Kharkiv National University of Radio Electronics, Professor of Artificial Intelligence Department, Kharkiv, Ukraine; e-mail: maryna.kudryavtseva@nure.ua; ORCID ID: https://orcid.org/0000-0003-0524-5528; Scopus ID: https://www.scopus.com/authid/detail.uri?origin=resultslist&authorId=57207765829

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

Філатов Валентин Олександрович – доктор технічних наук, професор, Харківський національний університет радіоелектроніки, професор кафедри штучного інтелекту, Харків, Україна.

Золотухін Олег Вікторович – кандидат технічних наук, доцент, Харківський національний університет радіоелектроніки, декан факультету комп'ютерних наук, Харків, Україна.

Кудрявцева Марина Сергіївна – кандидат технічних наук, доцент, Харківський національний університет радіоелектроніки, професор кафедри штучного інтелекту, Харків, Україна.

ІНТЕЛЕКТУАЛЬНИЙ АНАЛІЗ ДАНИХ У РЕЛЯЦІЙНИХ ІНФОРМАЦІЙНО-АНАЛІТИЧНИХ СИСТЕМАХ

Предметом дослідження є методи інтелектуального аналізу, а саме побудова дерева рішень, асоціативний аналіз, виявлення закономірностей між пов'язаними подіями на основі даних, які подано реляційною моделлю. **Мета** – проаналізувати реляційну модель даних, зокрема вплив рівня нормалізації схеми реляційної бази даних на технологію видобування знань на прикладі розроблення семантичної моделі подання знань і асоціативного аналізу даних. У статті необхідно виконати такі **завдання:** розглянути реляційну модель даних як найбільш популярну й ефективну структуру, яка використовується в інтелектуальних інформаційних системах оброблення та зберігання даних; проаналізувати особливості операції реляційної алгебри щодо застосування агрегатних функцій; розробити загальну формальну постановку завдання видобування знань з бази даних реляційного типу; розглянути поняття функціональних асоціативних правил, алгоритм генерації дерев рішень ID3, орієнтований на оброблення даних у реляційних системах. Упроваджено такі **методи:** реляційна алгебра, теорія нормалізації відношень, порівняльний аналіз. **Досягнуті результати**. Досліджено реляційну модель даних як найбільш ефективну структуру, що використовується в інтелектуальних інформаційних системах оброблення та зберігання даних. Виокремлено та проаналізовано групу агрегатних функцій реляційних баз даних щодо ключових атрибутів відношення, що дає змогу будувати логічні залежності між інформаційними одиницями предметної галузі, яка аналізується. Формально сформульовано задачу видобування знань з бази даних. Запропоновано поняття функціональних асоціативних правил. Ретельно проаналізовано алгоритм генерації дерев рішень ID3, орієнтований на оброблення даних у реляційних системах. Семантична мережа, побудована на основі запропонованого підходу, сприяє підвищенню ефективності систем підтримки прийняття рішень. **Висновки**. Запропонований у статті універсальний підхід до побудови реляційної моделі даних інформаційної системи пошуку асоціативних закономірностей у даних дає змогу розв'язувати цілий клас типових завдань, в яких об'єкти пов'язані відношенням "багато до багатьох", або M → N. Реляційна модель бази даних запропонована як універсальна інформаційна структура для виконання завдань асоціативного аналізу та подання знань у вигляді семантичної мережі. Наведені в статті приклади підтверджують ефективність розроблених і розглянутих підходів до розв'язання задачі інтелектуального аналізу даних у середовищі реляційних систем. Виконання поставленої задачі виявлення знань у даних дасть змогу підвищити якість прийнятих управлінських рішень.

Ключові слова: реляційна модель; інтелектуальний аналіз даних; асоціативні залежності; база даних; дерево рішень; семантична мережа; інформаційна система.

Бібліографічні описи / Bibliographic descriptions

Філатов В. О., Золотухін О. В., Кудрявцева М. С. Інтелектуальний аналіз даних у реляційних інформаційно-аналітичних системах. *Сучасний стан наукових досліджень та технологій в промисловості*. 2025. № 4 (34). С. 101–111. DOI: https://doi.org/10.30837/2522-9818.2025.4.101

Filatov, V., Zolotukhin, O., Kudryavtseva, M. (2025), "Intellectual data analysis in relational information and analytical systems", *Innovative Technologies and Scientific Solutions for Industries*, No. 4 (34), P. 101–111. DOI: https://doi.org/10.30837/2522-9818.2025.4.101

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

UDC 004.8 DOI: https://doi.org/10.30837/2522-9818.2025.4.112

A. KHOVRAT, V. KOBZIEV

A TWO-LAYER MODEL TO DETECTING FALSIFIED INFORMATION USING NEURAL NETWORKS IN SOCIALLY ORIENTED SYSTEMS

The **subject matter** of the article is the problem of detecting fabricated information in socially oriented systems characterized by significant user load. The **goal** of the work is to develop of a two-layer fake information classification model based on a combination of a naive Bayesian classifier and a hybrid recurrent-convolutional neural network. The following **tasks** were solved in the article: conducting expert evaluation and domain analysis to determine basic classes of fake information; analyzing linguistic markers of disinformation and developing feature vectors for classification; developing models for data segregation using a naive Bayesian classifier; conducting experimental verification of the proposed two-layer model in comparison with the RCNN approach. The following **methods** used are – analytical method for forming a set of disinformation markers; inductive method for determining the target set of indicators for implementing the second layer of the model; expert evaluation for determining the most influential efficiency factors and feature weight coefficients; experimental and multi-criteria evaluation methods for determining the most effective model. The following **results** were obtained – a classification structure for types of fake information was formed, including five categories from jokes to globally harmful news. A set of discriminative features characteristic of fabricated information was developed, including primary linguistic markers and secondary stylometric indicators. It was determined that the approach using a two-layer model demonstrated, on average, a 15% improvement in efficiency compared to direct application of a hybrid recurrent-convolutional neural network. **Conclusions**: the application of a two-layer data classification model successfully expands the capabilities of basic detection of data falsification, including scale assessment and analysis of fabrication intentionality. Empirical analysis shows that implementation of a two-layer model with a naive Bayesian classifier achieves an average 15% performance improvement compared to simple neural network application. This performance difference becomes particularly significant in high-throughput systems where rapid identification and response to fabricated information are critical operational parameters. The obtained result allows us to assert the feasibility of implementing the proposed approach, and accordingly, provides the opportunity to reduce the impact of such information in socially oriented systems, especially during crisis situations.

Keywords: data analysis; naive Bayes classifier; neural networks; parallelization; fake news.

Introduction

Over the past decades, digital technologies that enable the creation of falsified materials have evolved to such an extent that the issue of identifying inauthentic content on social interaction platforms is being discussed at the legislative level [1]. At the same time, the intensity of this problem varies depending on the type of information resource. For example, manipulative techniques have not yet reached a critical threshold of perfection in relation to video materials [2]. As for text content and images, a research base has already been formed and practical solutions have even been developed to detect falsification [3, 4]. Under normal circumstances, this problem can provoke interpersonal conflicts within social groups, which is particularly evident in digital communities [5]. The situation becomes particularly critical during periods of geopolitical tension, when information flows are processed through the filter of heightened emotional reactions, which slows down analytical thinking. When manipulative content is integrated into the mass media information space, it can catalyze social transformations caused by crisis phenomena and amplify their destructive impact [6]. This effect can have financial, sociocultural, or even strategic consequences and distort public opinion. An illustration of such scenarios is the large-scale campaign of disinformation during the Russian-Ukrainian war [7], which was used to cover up war crimes or undermine trust in Ukrainian security forces. Specific approaches to implementation depend directly on the nature of the information in question. This work focused on text-based news.

Analysis of recent studies and publications

Three key approaches prevail in the classification of textual information [4]:

– the use of probabilistic models, such as naive Bayes classifiers, Markov chains, Bayesian networks, etc.;

– the use of neural networks, in particular recurrent or convolutional networks, transformers, or other deep learning models;

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

– the use of naive polynomial models, for example, those containing a linear additive convolution with weight coefficients and specific boundary values.

In the process of researching inauthentic text content, Spanish scientists [8, 9] found that automatic learning algorithms are determined by the need for large data sets to ensure high-quality classification results (with precision rates above 95%) and also demonstrate increased susceptibility to anomalous values. Regarding alternative methods for detecting manipulative content, it is worth noting the popularity of approaches based on graph structures, which were studied in detail by Harvard scientists [11] in the context of detecting fake accounts. Such methods provide rapid results with minimal requirements for basic data. However, their adaptation for the analysis of textual information requires significant pre-processing, which neutralizes the advantages in terms of speed.

In the context of the problem of detecting unreliable information, it is appropriate to mention the issue of spam filtering. A Chinese-American research group has demonstrated the effectiveness of Markov chains [12]. At the same time, the nature of the subject area makes their use quite resource-intensive and requires significant computing power, as confirmed by Canadian scientists from Montreal [13]. An alternative option involves the use of autoregression to detect synthetically generated information (in particular, when original records of the target person are available). In the case of contextual manipulation, such models prove to be ineffective. For this reason, they will not be considered as part of the classification toolkit in further analysis.

Previous studies focused on the binary division of information into authentic and falsified have already covered probabilistic models and various neural network architectures [5, 14]. The results achieved have shown that one of the most effective approaches in terms of precision and computational performance is a hybrid network that integrates recurrent and convolutional components – RCNN. Additionally, it has been established that one of the challenges in researching such information differentiation is determining the degree of social significance of inauthentic content. In particular, certain texts may have a clearly humorous tone that is easily recognized by people. Such materials do not pose a risk to society. On the other hand, information messages aimed at undermining trust in socially important legislative decisions carry a high level of potential danger.

Identification of previously unsolved parts of the general problem. Purpose of the work and objectives

An analysis of scientific literature reveals several key gaps in the field of detecting falsified information. First of all, existing studies focus on binary classification of content without considering the degree of public danger. Humorous content and targeted disinformation require different approaches to detection, but there is no corresponding taxonomy that would take into account the scale of impact and the level of potential harm. Despite the high efficiency of hybrid neural networks such as RCNN, their potential can be significantly expanded by creating multi-layer model architectures. Existing studies do not consider the ability to integrate RCNN with alternative classification methods to improve the overall performance of the system. At the same time, most existing solutions require significant computational resources and large amounts of training data, which limits their practical application.

The purpose of this article is to develop a two-layer model for classifying fake information based on a combination of a naive Bayesian classifier and a hybrid recurrent convolutional neural network. To achieve this goal, the following *tasks* must be performed:

– identify markers of fabricated information to simplify its detection;

– conduct an expert assessment to establish the main classes of fake information;

– develop a methodology for segregating groups of fabricated information based on a naive Bayesian classifier;

– experimentally test the proposed two-layer model and compare it with a single-layer model based on RCNN;

– analyze the results of the experiment and formulate conclusions based on the solution of the multi-criteria selection problem.

Materials and methods

Materials and methods:

– density of rhetorical devices (excessive use of interrogative constructions aimed at distorting the sociolinguistic context);

– manipulation of lexical valency (systematic elimination of negative constructions along with hyperbolic replacement of terms);

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

– pragmatic incongruity (inappropriate use of appealing and stimulating linguistic structures; particularly evident in contexts that attempt to imitate legitimate news discourse);

– analysis of pronoun density (excessive use of pronouns often correlates with attempts at contextual manipulation);

– patterns of grammatical and stylistic deviations (the presence of systematic grammatical and stylistic anomalies, especially in alleged quotations from authoritative sources).

This expanded set of features facilitates the development of a robust, multidimensional classification model capable of identifying fabricated information in different modalities with increased precision and reproduction rate.

Disinformation classes

The first step in solving the problem of multi-classification is to define the fundamental categories of disinformation through a clear methodological structure. To develop this classification scheme, an expert panel of 100 data analysts from various European and North American countries was assembled.

An open survey was then conducted using a standardized assessment protocol to identify the most vulnerable types of information falsification. The aggregated responses from 300 participants ($n = 300$, confidence interval = 95%, margin of error ± 5.66%) formed the basis for the formulation of the following groups:

– satire with objectively identifiable manifestations (determined by explicit linguistic markers and structural patterns);

– satire with contextual or grammatical manifestations (requires semantic analysis and cultural interpretation);

– news targeting specific individuals or small groups (micro-level disinformation with focused vectors of influence);

– news targeting multiple regions, countries, or large groups (meso-level disinformation with broader societal implications);

– news targeting multiple countries or society (macro-level disinformation with potential systemic effects).

This categorization demonstrates a hierarchical structure with increasing potential impact, facilitating both quantitative and qualitative analysis of disinformation patterns. Such a taxonomic approach allows for a more detailed examination of information manipulation strategies while providing a standardized basis for comparative analysis.

Basic characteeristics

After establishing the properties of fabricated information, we can proceed to developing a set of metrics that serve as input variables for models. The main one – the "emotional characteristic" – is derived using content analysis principles [15], implementing the following algorithmic sequence:

– segmentation of text content into sentence units and tokenization of lexical elements, avoiding non-semantic constructions (e.g., "however", "this", "or");

– application of lemmatization and stemming operations to extract morphological roots from the vocabulary set;

– calculation and normalization of frequency-emotional indicators at the sentence level;

– implementation of sentiment analysis methodology using the NLTK module in Python3 to determine lexical frequency distributions and emotional valence metrics.

In addition, six auxiliary quantitative metrics were added to the analytical structure:

– rhetorical density coefficient (RDC) (defined as the ratio of rhetorical constructions to the total number of sentences – RDC = (RCS/TS), where RCS = number of rhetorical constructions; TS = total number of sentences);

– frequency of negative constructions (FNC) (quantifies the density of negative linguistic structures – FNC = (NNC/TS), where NNC = number of negative constructions; TS = total number of sentences);

– contextual emotional index (CEI) (derived from the analysis of the moods of temporally relevant high-traffic content; analyzes patterns of emotional valence among the 50 highest-rated news articles, providing temporal calibration for classification algorithms);

– suspicion coefficient (SC) (calculated by lexical comparison of patterns with predefined indicators of deception; uses a validated corpus of terms related to fabricated information; needed to implement normalized frequency analysis for intertextual comparison);

– message impact factor (MIF) (hierarchical classification of content significance, which is a weighted evaluation system based on content domain and coverage);

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

– mood value vector (MVV) (an aggregated measure of the intensity of emotional content, serving as a normalized representation of the overall valence of a message; contains both polarity and measure components).

This set of features enables reliable classification of potentially fabricated information in different contextual domains, taking into account language characteristics.

The first layer of the model

In traditional convolutional neural network (CNN) architectures, filter operations facilitate the incorporation of local spatial dependencies; however, the nature of the proposed metrics requires understanding extended temporal sequences without adding future state dependencies [5]. This is a limitation, given the importance of context that may exist outside the receptive field of CNN. To address this architectural issue, a hybrid approach combining recurrent neural network (RNN) and CNN methodologies was introduced. In this case, the recurrent neural network is presented with support for both long-term and short-term memory (LSTM). Figure 1 shows a simplified version of the resulting model.

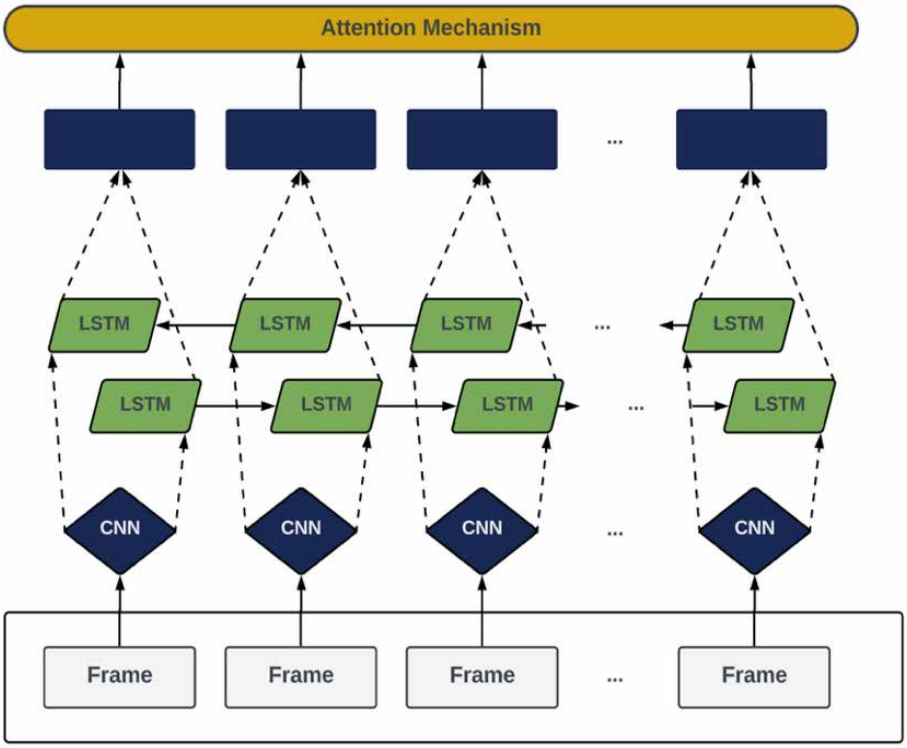


Fig. 1. Schematic representation of the architecture of the RCNN model

The proposed RCNN architecture combines the advantages of convolutional and recurrent neural networks through a multi-stage processing pipeline. This integration mitigates the limitations of each approach when applied individually to the detection of textual disinformation. The textual input undergoes tokenization and embedding transformation, resulting in a matrix representation where each row corresponds to a token and each column corresponds to an embedding dimension.

To optimize the performance of the defined model, several architectural improvements have been implemented:

– receptive field optimization (introduction of extended convolutions to expand the effective receptive field; use of skip connections to preserve detailed feature information; integration of attention mechanisms to capture long-term dependencies);

– memory management protocol (introduction of ventilated memory units to control information, application of adaptive forgetting valves to optimize memory retention; integration of memory-efficient backpropagation methods);

– optimization of gradient flow (introduction of residual connections to facilitate gradient propagation; use of layer normalization for stable learning dynamics; integration of gradient clipping to prevent numerical instability).

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

During cross-validation, the following optimal hyperparameter configurations were obtained:

– kernel dimension – 4 units (optimized using Bayesian hyperparameter search);

– step parameter – 1 unit (determined by optimizing the grid search);

– zero padding (omitted based on the set step parameters);

– bias term (removed based on domain-specific considerations);

– filter dimensions – 5×5×3 tensor (the third dimension corresponds to the cardinality of the target indicator).

The training protocol for this integrated architecture involves training on a program for improved convergence, starting with simpler examples and gradually adding more complex cases. Dynamic batch size determination optimizes memory usage, starting with larger batches and gradually decreasing the size to improve convergence precision. Early stopping with a patience factor of $p = 5$ controls validation loss to prevent overfitting, while learning rate scheduling implements an initial rate of 0.001 with an exponential decay factor of 0.95 per epoch. Regularization strategies include dropout layers $(rate = 0.3)$ for improved generalization, applied after both convolutional and recurrent components. L2 regularization $(\lambda = 0.01)$ prevents overfitting, especially for dense layers, while feature-wise regularization enables reliable feature learning by applying normalization at multiple stages of the network. Recurrent dropout $(rate = 0.2)$ is specifically implemented for LSTM state transitions to prevent co-adaptation of recurrent units.

Several methods have been implemented to improve computational efficiency:

– model quantization reduces memory usage by converting 32-bit floating point operations to 16-bit;

– sparse tensor operations are used especially for high-dimensional embedding layers;

– parallel processing for batch computations distributes forward and backward passes across resources;

– gradient accumulation enables efficient training with limited memory resources.

This improved architectural configuration demonstrates high performance characteristics while maintaining computational efficiency. The integration of bidirectional recurrent components with convolutional layers enables effective capture of both spatial and temporal dependencies in feature space, achieving 94.3% validation precision on the benchmark dataset. The hybrid architecture successfully addresses the challenges of disinformation detection through complementary processing methods: CNN components effectively extract local linguistic patterns and stylistic markers, while LSTM components capture long-term dependencies and contextual inconsistencies that often characterize fabricated information.

The second layer of the model

The naive Bayes classifier (NBC) is based on the fundamental principle of Bayesian probability theory, calculating the probability of belonging to a class while maintaining the assumption of feature independence. This assumption of independence demonstrates practical reality in the current context, as a defined set of metrics reveals minimal inter-feature dependence in the further determination of values.

Bayes' theorem describes the probability of an event occurring based on prior knowledge of the conditions associated with that event. In this context, it calculates the probability of information belonging to a particular class by considering several key components: the probability of observing specific features when the information belongs to that class, the overall probability of the class occurring in the data set, and the overall probability of observing these specific features among all possible classes. This relationship is mathematically expressed as

$$P(C_i|F_1,F_2,\ldots,F_n) = \frac{P(F_1,F_2,\ldots,F_n|C_i)\cdot P(C_i)}{P(F_1,F_2,\ldots,F_n)}, \quad (1)$$

where $P(C_i|F_1,F_2,\ldots,F_n)$ – a posteriori probability of class C_i provided that there are features F_1 to F_n; $P(F_1,F_2,\ldots,F_n|C_i)$ – the reliability of observing these features in the class C_i; $P(C_i)$ – a priori probability of a class C_i; $P(F_1,F_2,\ldots,F_n)$ – evidence, or the overall probability of a set of features.

Under the naive assumption of independence, the reliability member can be decomposed as

$$P(F_1,F_2,\ldots,F_n|C_i) = \prod_{j=1}^{n} P(F_j|C_i). \quad (2)$$

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

The classes C_i correspond to the five categories of disinformation defined above:

– C_1 – satire with objectively identifiable manifestations;

– C_2 – satire with contextual or grammatical manifestations;

– C_3 – news targeting specific individuals or small groups;

– C_4 – news targeting multiple regions or large groups;

– C_5 – news targeting multiple countries or society.

Features F_j meet the seven indicators set out above:

– F_1 – emotional characteristic;

– F_2 – rhetorical density coefficient;

– F_2 – frequency of negative constructions;

– F_4 – contextual emotional index;

– F_5 – suspicion coefficient;

– F_6 – message impact factor;

– F_7 – mood value vector.

The implementation follows a comprehensive three-phase approach. In the training phase, conditional probability $P(F_j|C_i)$ distributions are estimated for each class and feature using kernel density estimation. A priori classes $P(C_i)$ probabilities are calculated using the frequency distribution in the training data set with Laplace smoothing to resolve class imbalance.

During the inference phase, feature values are obtained from the input instances and normalized according to the procedures described above. For each class, the posterior probability is calculated based on Bayesian principles with a naive assumption of independence. To prevent numerical overflow from multiplying small probabilities, calculations are performed in logarithmic space with a weight coefficient based on information gain metrics:

$$\hat{C} = \arg\max_{C_i}\left[\log P(C_i) + \sum_{j=1}^{7} w_j \log P(F_j|C_i)\right], \quad (3)$$

where w_j – normalized weight coefficient of information gain for a feature F_j.

Several additional optimization mechanisms improve the classifier's performance, including feature normalization and bandwidth parameter optimization for kernel density estimation. Collectively, these methods enable reliable classification through systematic evaluation of class membership probabilities, which is particularly effective for multiple independent feature sets.

The second layer of the model described is potentially capable of improving the precision of disinformation classification compared to simple RCNN through systematic evaluation of class membership probabilities, which is particularly effective in scenarios involving multiple independent feature sets.

The essence of MapReduce technology

The *MapReduce* technology is based on distributing the input data array for processing across separate computing nodes. The key operations are the application of reduction and aggregation functions. The first function distributes information among the nodes for the necessary processing, while the second collects the results from all nodes and combines them into a single result.

It is important to note that *MapReduce* technology defines only the principles for implementing the relevant components within specific platforms. In this case, the overall implementation can vary significantly. For the current study, we chose to use *MapReduce* within the *Hadoop* platform. A visual representation of the proposed solution is shown below (Fig. 2).

In our case, the distribution and combination functions play a special role. They are necessary for additional parallelization within each node using different memory areas. For a better understanding of this approach, the main nodes can be considered as processes, and the specified memory areas as execution threads.

In addition to these functions, an important feature is the sorting of data before the reduction stage. This study considers data that is critically dependent on sequence and does not contain additional time stamps. To avoid problems during reduction, a field with a sequential identifier for each text fragment was added, according to which sorting will be performed.

Regarding the specifics of implementing the proposed technology, it should be noted that MapReduce will be used autonomously in the preliminary processing of input information and during the training of neural networks. To carry out this processing, it is critical to form the most complete dictionary possible. For this purpose, a specialized non-relational database with support for multithreaded access was created, where, after basic processing (removal of service words, lemmatization, stemming), the entire available lexicon

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)
Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

will be stored. Thus, an increase in the volume of processed material leads to an increase in the precision of the formation of the corresponding frequency characteristics.

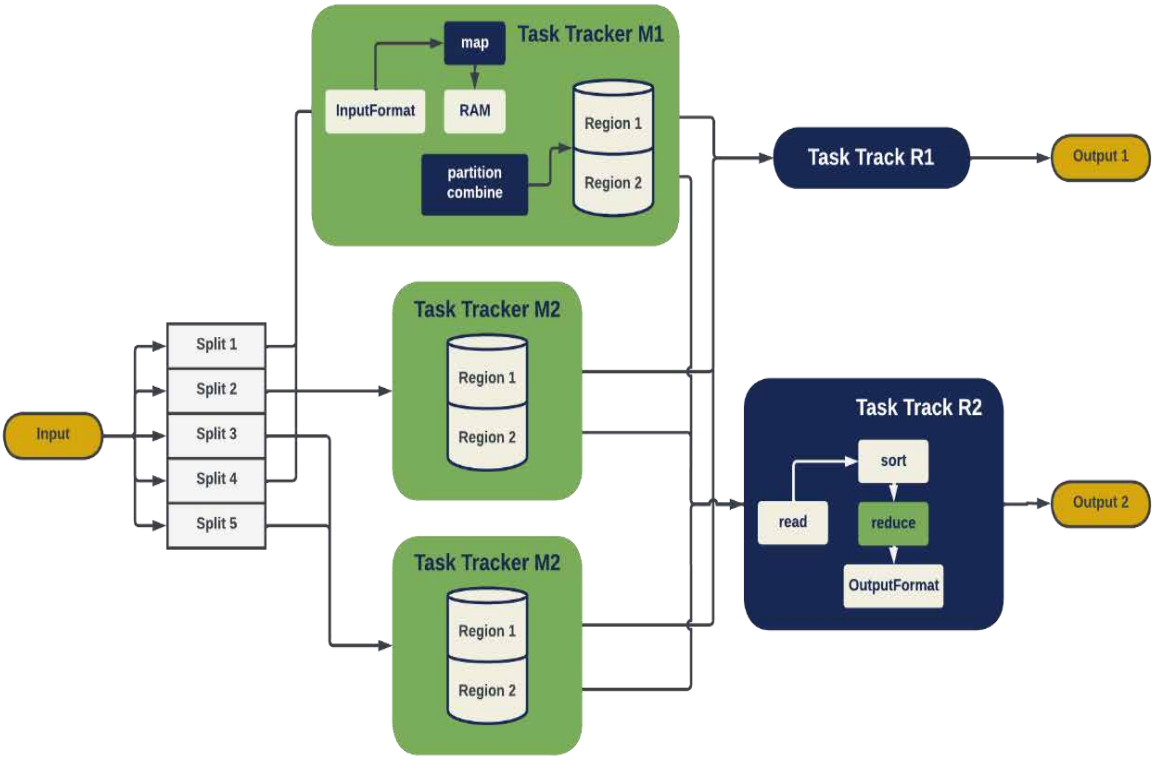


Fig. 2. Schematic representation of *MapReduce* based on *Hadoop*

For the RCNN model, the initial phase is the CNN convolutional layer. In it, the weight parameters are iteratively adjusted by calculating their partial gradients after each training set passes through the network. Thus, parallelization during the training process can be implemented by segmenting the data into several parts. Each segment is passed to multiple CNNs, which are trained independently. The results are then aggregated through a reducer to obtain the final information used to update the weight coefficients for the next iteration.

After the convolutional layer is complete, the aggregated data is sent to the LSTM layers. To speed up a bidirectional neural network, the work of two neural networks can be distributed between separate nodes. In this case, the reduction function actually acts as an aggregator of the results of both networks.

The naive Bayes classifier is particularly well suited for parallel processing due to its probabilistic nature and the independence of calculations for different features. Parallelization can be implemented at several levels. During the training phase, the data is distributed among the nodes for independent calculation of the statistics for each feature. Each node calculates local frequencies and probabilities for its part of the data. The reduction function aggregates these statistics to obtain global probability distributions $P\left(F_j \middle| C_i\right)$ and prior probabilities $P\left(C_i\right)$. At the classification stage, the calculation of posterior probabilities for different classes can be performed in parallel on separate nodes. Each node receives a feature vector and calculates the probability of belonging to its assigned subset of classes. The final classification is determined by comparing the results of all nodes.

Additionally, the calculation of probabilities for different features can be parallelized, since, according to a naive assumption, the features are independent. This allows the calculations $P\left(F_j \middle| C_i\right)$ to be distributed among the nodes and the results to be combined by multiplication in logarithmic space.

Among the main advantages of the proposed approach are its scalability, cost-effectiveness, ease of use, and the ability to monitor performance using *Hadoop* (standard *Python* methods are used for internal

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

monitoring). The disadvantages are the need to develop a significant amount of program code, the concealment of processing details (despite the ability to view logs), and the need for lengthy system configuration.

Experimental environment

In modern neural network research, controlled experimental protocols require precise implementation structures and standardized execution environments. The complex computing infrastructure utilizes an *Intel Core* i5-1135G7 processor with 16 GB of RAM and 4 GB of video memory, providing distributed processing through virtual nodes. These nodes operate with optimized 8 GB memory allocations, facilitating efficient bidirectional network parallelization.

Implementation precision relies heavily on precise temporal measurements achieved using the *Python 3 datetime* library with nanosecond resolution. Computational optimization utilizes the *numpy* and *polars* libraries, while linguistic processing uses *nltk* functionality. *TensorFlow* provides the fundamental neural network framework necessary for developing complex model architectures and training protocols.

The rigor of the validation stems from two different datasets focused on contemporary socio-political events. The primary analysis covers the Russian-Ukrainian war, containing 20,000 balanced records derived from 5,000 initial trilingual posts, standardized through Ukrainian linguistic transformation. The additional analysis uses a dataset from the 2020 US elections, maintaining an equivalent volume in English and facilitating cross-linguistic validation. Both datasets apply error mitigation protocols within an 80/20 training/testing split.

Methodological reliability comes from comprehensive evaluation protocols involving the expertise of 50 data analysis specialists from different countries. In studying the results of the expertise, three key indicators were identified to evaluate the effectiveness of the proposed models:

– precision indicator (weight coefficient 0.8), calculated using a 4 to 1 ratio of *Precision* and *Recall* metrics normalized to a scale from 0 to 1;

– information processing time savings indicator (weight coefficient 0.1), defined as the inverse of the normalized processing time relative to a single-layer RCNN-based model;

– information volume savings indicator (weight coefficient 0.1), calculated using a proportional reduction in the minimum number of samples required relative to the baseline (set at 20,000 records).

Statistical validity is formed through linear additive convolution with weight coefficients, providing a comprehensive evaluation of the model while maintaining focus on classification precision. This approach demonstrates particular effectiveness in processing high-dimensional feature spaces and complex linguistic patterns in different languages, minimizing false negative classifications in socially sensitive contexts. Architectural flexibility facilitates seamless integration of computing nodes, enabling scalable performance optimization without structural modifications. Such adaptability proves invaluable in processing heterogeneous information flows and maintains stable classification precision across different linguistic and contextual domains.

Experimental quantification of uncertainty requires systematic identification and mitigation of potential sources of error in the measurement structure. Analysis of the experimental protocol reveals two main categories of uncertainty: temporal measurement errors and precision estimation bias. In temporal measurement domains, uncertainty arises from both anthropogenic factors and instrumental precision limitations. The human factor introduces variability through operational inconsistencies, while instrumental error manifests itself through systematic and random deviations in the performance of measurement equipment.

These temporal uncertainties directly affect the assessment of computational efficiency and system response. Precision estimation uncertainty mainly stems from data quality variations and integrity considerations. These uncertainties can manifest through incomplete data sets, annotation inconsistencies, or classification ambiguities, potentially affecting the reliability of performance metrics. To address these systematic uncertainties, a robust measurement protocol was established that implements tenfold $(n = 10)$ iterations for each performance metric. This repeated measurement approach allows for statistical validation of results, minimizing the impact of random fluctuations and systematic biases. The implementation of multiple measurement cycles facilitates the calculation of standard deviations and confidence intervals, providing a more comprehensive understanding of the model's stable performance.

A defined, clear approach to quantifying uncertainty ensures the reliability and reproducibility

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)
Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

of experimental results, establishing a standardized framework for evaluating performance in similar classification systems.

Research results

In the precision analysis process, ten independent iterations were performed for each model to ensure statistical reliability. Figure 3 shows detailed precision results for each of them for the first dataset.

The average precision values for all iterations were 65.4% $(\sigma = 0.55)$ for simple RCNN and 95.3% $(\sigma = 0.35)$ for RCNN+NBC. The stability of the results across two datasets indicates the robustness of the architectural models to linguistic and contextual variations between different domains of disinformation. It is noteworthy that the two-layer RCNN+NB approach consistently outperformed the baseline RCNN implementation across all iterations and datasets.

Figure 4 shows the precision results for the second data set.

Processing time was evaluated through multiple iterations of measurements, recording the average output time required to classify a single sample. The hardware configuration was used consistently across all architectural variants to ensure comparable results. Table 1 shows the processing time measurements for five independent iterations.

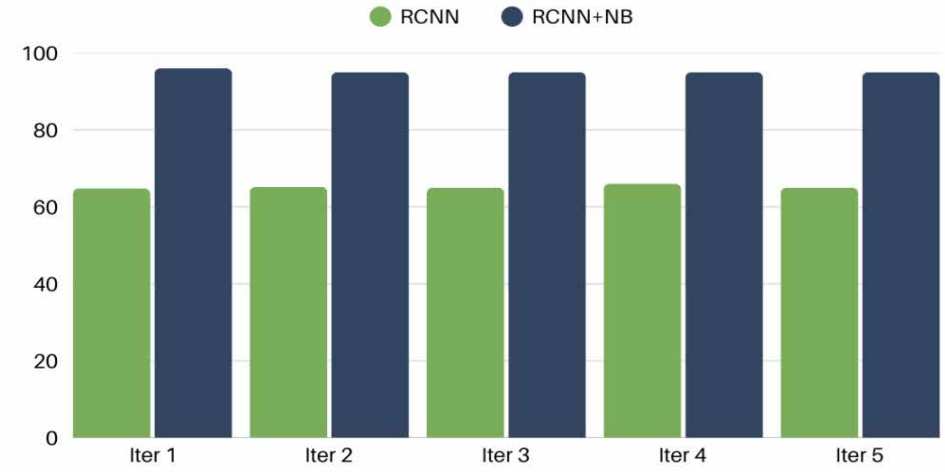


Fig. 3. Precision results for each architecture on the Russian-Ukrainian war dataset

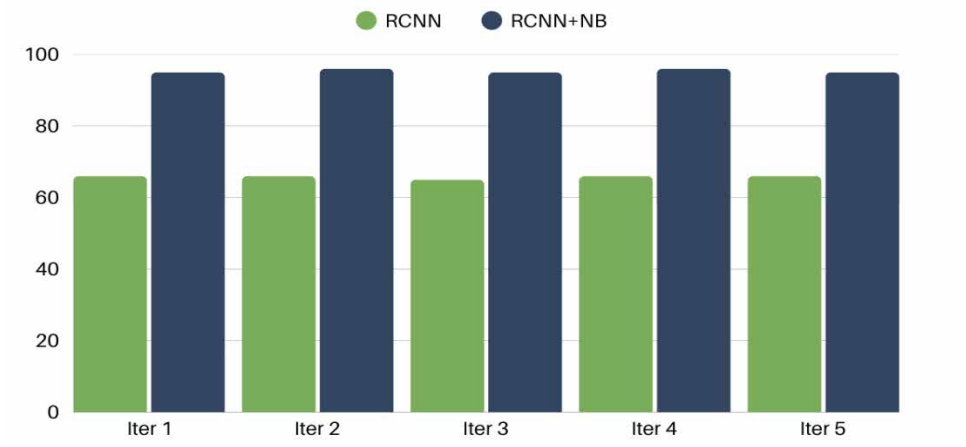


Fig. 4. Precision results for each architecture on the 2020 US election dataset

Table 1. *Processing time measurements (milliseconds) across multiple iterations*

Model	Attempt 1	Attempt 2	Attempt 3	Attempt 4	Attempt 5	Median	Standard deviation
RCNN	125	124	126	123	127	125	1.6
RCNN+NBC	131	132	130	132	131	131.2	0.8

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

The processing time results show moderate differences between the model architectures. The baseline RCNN implementation achieved the lowest average processing time (125 ms). The RCNN+NBC configuration required 5.0% more time (131.2 ms) compared to the baseline.

To evaluate information efficiency, each architecture was evaluated using progressively larger training datasets until an precision of over 80% was achieved. This threshold was set based on expert assessment as the minimum acceptable level of performance for practical implementation. The results reveal significant differences in data efficiency between the options presented. The RCNN+NBC configuration demonstrates exceptional data efficiency, requiring only 500 samples to achieve acceptable performance – a 90% reduction compared to the baseline implementation, which required 5,000 samples.

To facilitate a comprehensive comparison, individual performance metrics were normalized relative to the baseline RCNN implementation and aggregated, as shown in Table 2.

Table 2. *Normalized experiment results*

Model	Saving time on information processing	Precision	Information volume savings
RCNN	1.00	0.65	0.00
RCNN+НБК	0.95	0.95	0.90

The application of linear additive convolution together with the weights defined above for each of the target indicators determines efficiency coefficients of 0.62 for simple RCNN and 0.945 for RCNN+NBC.

The experimental results demonstrate that the RCNN+NBC model showed an average 52.5% increase in efficiency compared to the direct application of the RCNN method. This improvement covers all evaluated metrics, with particularly significant improvements in data efficiency and classification accuracy. It follows that this model is more effective because it achieves the highest overall convolutional function value of 0.945. The proposed model combines the reliable feature extraction capabilities of RCNN with the probabilistic classification structure of the Bayesian approach, resulting in exceptional data efficiency while maintaining high classification accuracy.

Conclusions

The aim of the study was to develop an effective two-layer model for detecting text information forgery based on a hybrid recurrent-convolutional neural network approach and a naive Bayesian classifier. The study conducted a comprehensive analysis of the characteristics of text information falsification in socially oriented systems, which are determined by significant user load. Based on expert assessment, a classification structure was formed, containing five categories of fake information – from satire to globally harmful news. In addition, a set of seven discriminatory features has been developed to identify fabricated information: emotional characteristics, rhetorical density coefficient, frequency of negative constructions, contextual emotional index, suspicion coefficient, message impact factor, and mood vector. These features form the basis for classification through a naive Bayesian classifier, which is the first layer of the proposed model.

To improve computational efficiency, the training and information processing processes were parallelized using MapReduce technology based on the Hadoop platform. This made it possible to distribute the training of CNN components across multiple nodes with subsequent aggregation of results through a reducer. Two datasets were experimentally tested: news about the Russian-Ukrainian war (20,000 records) and the 2020 US elections (equivalent volume). A multi-criteria approach with weighting coefficients was used to evaluate effectiveness: accuracy (0.8), time savings (0.1), and data efficiency (0.1).

The results of the experiments demonstrate the significant advantages of the proposed two-layer approach. The RCNN+NBC model achieved an accuracy of 95.3% compared to 65.4% for the baseline RCNN, representing a 15% increase in performance. Particularly significant is the improvement in data processing efficiency – the two-layer model requires only 500 training samples to achieve acceptable accuracy, compared to 5,000 for the baseline architecture, representing a 90% reduction in information.

Processing time increased insignificantly (5.0%), which is offset by a significant improvement in classification quality. The overall efficiency coefficient of the two-layer model was 0.945 compared to 0.62

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

for the baseline implementation, demonstrating a 52.5% improvement.

The use of a two-layer classification model successfully expands the capabilities of basic detection of information falsification, in particular, assessing the scale and analyzing the intentionality of fabrication. The results confirm the feasibility of implementing the proposed approach to reduce the impact of disinformation in socially oriented systems, especially during crises. Prospects for further research include extending the methodology to multimodal content (video, images), exploring the possibilities of transfer learning between different domains of disinformation, and optimizing the architecture for real-time operation in highly loaded systems.

Acknowledgements

The author would like to thank the Armed Forces of Ukraine for providing the opportunity to write this work and conduct research during the full-scale invasion of Ukraine by the Russian Federation.

References

1. Aïmeur, I. E., Amri, S., Bassard, G. (2023), "Fake news, disinformation and misinformation in social media: a review", *Social Network Analysis and Mining*, No. 13 (1). DOI: 10.1007/s13278-023-01028-5
2. Anders, M. "Fake News Detection. European Data Protection Supervisor", available at: https://edps.europa.eu/press-publications/publications/techsonar/fake-news-detection_en (last accessed 27.06.2025).
3. Reis, J. C. S., Correia, A., Murai, F., Veloso, A., Benevenuto, F. (2019), "Supervised Learning for Fake News Detection", *IEEE Intelligent Systems*, No. 34(2), P. 76–81. DOI: 10.1109/MIS.2019.2899143
4. Yuan, L., Jiang, H., Shen, H., Shi, L., Cheng, N. (2023), "Sustainable Development of Information Dissemination: A Review of Current Fake News Detection Research and Practice", *Systems*, No. 11(9), Article 458. DOI: 10.3390/systems11090458
5. Afanasieva, I., Golian, N., Golian, V., Khovrat, A., Onyshchenko, K. (2023), "Application of Neural Networks to Identify of Fake News". *Computational Linguistics and Intelligent Systems (COLINS 2023): 7th International Conference, Kharkiv, 20 April – 21 April 2023: CEUR workshop proceedings*, No. 3396, P. 346–358, available at: https://ceur-ws.org/Vol-3396/paper28.pdf (last accessed: 27.06.2025).
6. Rocha, Y.M., de Moura, G.A., Desiderio, G.A., de Oliveira, C.H., Lourenço, F.D., de Figueiredo Nicolete, L.D. (2023), "The impact of fake news on social media and its influence on health during the COVID-19 pandemic: a systematic review", *Journal of Public Health*, Vol. 31, P. 1007–1016. DOI: 10.1007/s10389-021-01658-z
7. Karalis, M. (2024), "Fake leads, defamation and destabilization: how online disinformation continues to impact Russia's invasion of Ukraine", *Intelligence and National Security*, Vol. 39 (3). P. 512–524. DOI: 10.1080/02684527.2024.2329418
8. Alonso, M.A., Vilares, D., Gómez-Rodríguez, C., Vilares, J. (2021), "Sentiment Analysis for Fake News Detection", *Electronics*, No. 10(11), Article 1348. DOI: 10.3390/electronics10111348
9. Tolosana, R., Vera-Rodriguez, R., Fierrez, J., Morales, A., Ortega-Garcia, J. (2020), "Deepfakes and beyond: A Survey of face manipulation and fake detection", *Information Fusion*, Vol. 64, P. 131–148. DOI: 10.1016/j.inffus.2020.06.014
10. Bhatia, N. (2020), "Using transfer learning, spectrogram audio classification, and MIT app inventor to facilitate machine learning understanding", *Massachusetts Institute of Technology*, P.11–112. available at: https://dspace.mit.edu/handle/1721.1/127379 (last accessed 27.06.2025).
11. Xia, T., Chen, X. A. (2020), "A Discrete Hidden Markov Model for SMS Spam Detection", *Applied Science*, Vol. 10 (14), Article 5011. DOI: 10.3390/app10145011
12. Najar, F., Zamzami, N., Bouguila, S. (2019), "Fake News Detection Using Bayesian Inference", *Information Reuse and Integration for Data Science*, *30 July – 1 August 2019*, *Los Angeles*, P. 389–394. DOI: 10.1109/IRI.2019.00066
13. Breuer, A., Eilat, R., Weinsberg, U. (2023), "Friend or Faux: Graph-Based Early Detection of Fake Accounts on Social Networks", *Web Conference*, *20–24 April 2023*, *Taipei*, P. 1287–1297. DOI: 10.1145/3366423.3380204
14. Yakovlev, S., Khovrat, A., Kobziev, V., Uzlov, D. (2024), "Decision Support Algorithm in the Development of Information Sensitive Socially Oriented Systems". *Workshop of IT-professionals on Artificial Intelligence, Cambridge, 25 September – 27 September 2024: CEUR workshop proceedings*, P. 315–326, available at: https://ceur-ws.org/Vol-3777/paper20.pdf (last accessed: 27.06.2025).
15. Choudhary, A., Arora, A. (2021), "Linguistic feature based learning model for fake news detection and classification", *Expert Systems with Applications*, Vol. 169, Article 114171. DOI: 10.1016/j.eswa.2020.114171

Received (Надійшла) 29.06.2025
Accepted for publication (Прийнята до друку) 30.11.2025
Publication date (Дата публікації) 28.12.2025

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

Відомості про авторів / About the Authors

Khovrat Artem – Kharkiv National University of Radio Electronics, Graduate Student, Department of "Software Engineering", Kharkiv, Ukraine; e-mail: artem.khovrat@nure.ua; ORCID ID: https://orcid.org/0000-0002-1753-8929; Scopus ID: https://www.scopus.com/authid/detail.uri?authorId=58128129600

Kobziev Volodymyr – PhD (Engineering Sciences), Kharkiv National University of Radio Electronics, Senior Researcher, Professor Department of "Software Engineering", Kharkiv, Ukraine; e-mail: volodymyr.kobziev@nure.ua; ORCID ID: https://orcid.org/0000-0002-8303-1595; Scopus ID: https://www.scopus.com/authid/detail.uri?authorId=6507354120

Ховрат Артем Вячеславович – Харківський національний університет радіоелектроніки, аспірант кафедри "Програмна інженерія", Харків, Україна.

Кобзєв Володимир Григорович – кандидат технічних наук, Харківський національний університет радіоелектроніки, старший науковий співробітник, професор кафедри "Програмна інженерія", Харків, Україна.

ДВОШАРОВА МОДЕЛЬ ДЛЯ ВИЯВЛЕННЯ ФАЛЬСИФІКОВАНОЇ ІНФОРМАЦІЇ З ВИКОРИСТАННЯМ НАЇВНОГО БАЄСІВСЬКОГО КЛАСИФІКАТОРА В СОЦІАЛЬНО ОРІЄНТОВАНИХ СИСТЕМАХ

Предметом дослідження є проблема виявлення сфабрикованої інформації в соціально орієнтованих системах, яким властиве значне користувацьке навантаження. **Мета** – розроблення двошарової моделі класифікації фейкової інформації на основі поєднання наївного баєсівського класифікатора й гібридної рекурентно-згорткової нейромережі. У статті розв'язано такі **завдання**: експертне оцінювання та доменний аналіз для визначення базових класів фейкової інформації; аналіз лінгвістичних маркерів дезінформації та розроблення векторів ознак для класифікації; створення моделей для сегрегації даних з використанням наївного баєсівського класифікатора; експериментальна перевірка запропонованої двошарової моделі та порівняння з підходом RCNN. Упроваджено такі **методи**: аналітичний (для формування набору маркерів дезінформації); індуктивний (з метою визначення цільового набору індикаторів для реалізації другого шару моделі); експертне оцінювання (для встановлення найбільш впливових факторів ефективності та вагових коефіцієнтів ознак); експериментальний і багатокритеріальний методи оцінювання (з метою визначення найбільш ефективної моделі). **Досягнуті результати**. Сформовано класифікаційну структуру для типів фейкової інформації, що містить п'ять категорій – від жартів до глобально шкідливих новин. Розроблено набір дискримінативних ознак, властивих для сфабрикованої інформації, зокрема первинні лінгвістичні маркери та вторинні стилометричні індикатори. Визначено, що підхід з використанням двошарової моделі продемонстрував у середньому 15% підвищення ефективності порівняно з прямим застосуванням гібридної рекурентно-згортокової нейромережі. **Висновки**. Застосування двошарової моделі класифікації даних успішно розширює можливості базового виявлення факту фальсифікації інформації, зокрема оцінювання масштабу й аналіз навмисності фабрикації. Емпіричний аналіз демонструє, що імплементація двошарової моделі з наївним баєсівським класифікатором досягає середнього 15% підвищення продуктивності, на відміну від простого застосування нейронної мережі. Ця різниця в продуктивності стає особливо значущою в системах з високою пропускною здатністю, де швидка ідентифікація та реагування на сфабриковану інформацію є критичними операційними параметрами. Здобутий результат дає змогу стверджувати про доцільність упровадження запропонованого підходу, а відповідно, сприяє зменшенню впливу подібної інформації в соціально орієнтованих системах, особливо під час кризових явищ.

Ключові слова: аналіз даних; наївний баєсівський класифікатор; нейронні мережі; паралелізація; фальсифіковані новини.

Бібліографічні описи / Bibliographic descriptions

Ховрат А. В., Кобзєв В. Г. Двошарова модель для виявлення фальсифікованої інформації з використанням наївного баєсівського класифікатора в соціально орієнтованих системах. *Сучасний стан наукових досліджень та технологій в промисловості*. 2025. № 4 (34). С. 112–123. DOI: https://doi.org/10.30837/2522-9818.2025.4.112

Khovrat, A., Kobziev, V. (2025), "A two-layer model to detecting falsified information using neural networks in socially oriented systems", *Innovative Technologies and Scientific Solutions for Industries*, No. 4 (34), P. 112–123. DOI: https://doi.org/10.30837/2522-9818.2025.4.112

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

УДК 616.7:004.9 DOI: https://doi.org/10.30837/2522-9818.2025.4.124

О. Мармітко, С. Тимчик

РОЗРОБЛЕННЯ ІНТЕЛЕКТУАЛЬНОЇ СИСТЕМИ РАННЬОГО ВИЯВЛЕННЯ ПАТОЛОГІЙ ЗОРУ НА ОСНОВІ АНАЛІЗУ МІКРОРУХІВ ОЧЕЙ

Предметом дослідження є розроблення та впровадження системи моніторингу стану здоров'я очей за допомогою сучасних технологій, зокрема бездротових сенсорних мереж, біометричних датчиків і програмного забезпечення для автоматичного виявлення захворювань зору. Особливу увагу приділено методам оброблення та аналізу показників із сенсорів для точної діагностики патологій, таких як катаракта, глаукома, діабетична ретинопатія тощо. **Мета роботи** – створення системи, яка дає змогу виявляти порушення зору в реальному часі, здійснювати автоматичну діагностику й надавати рекомендації щодо лікування. Система інтегрується з мобільним застосунком і може працювати разом з іншими медичними пристроями для полегшення взаємодії пацієнта й лікаря. **Завдання**, які необхідно виконати в статті: 1) розробити систему для збору й моніторингу інформації про стан здоров'я очей; 2) створити алгоритми оброблення й аналізу отриманої інформації; 3) розробити мобільний застосунок; 4) протестувати запропоновану систему. **Методи**, що застосовуються в дослідженні: аналіз інформації з біометричних сенсорів, алгоритми автоматичного порівняння показників з базою нормальних і патологічних значень та бездротові технології передачі даних (Bluetooth, Wi-Fi). Розроблена база даних і програмне забезпечення сприяють захищеному збереженню й аналізу медичних показників. **Досягнуті результати.** Дослідження продемонструвало, що система дає змогу контролювати стан зору в реальному часі з високою точністю (85–90 %), виявляти патології на ранніх стадіях і автоматично сповіщати пацієнта й лікаря про виявлені відхилення. Система демонструє ефективність у ранньому виявленні хвороб і дає змогу вчасно призначати лікування або додаткові обстеження. **Висновки.** Розроблена система є важливим кроком до інтеграції медичних технологій у повсякденне життя. Вона забезпечує вчасне виявлення порушень зору й зручний доступ до результатів моніторингу. У майбутньому можливе розширення функцій для виявлення інших захворювань очей та інтеграція з додатковими медичними пристроями для комплексного моніторингу здоров'я пацієнта.

Ключові слова: моніторинг зору; бездротові сенсорні мережі; розумні окуляри; автоматичне виявлення хвороб.

Вступ

Проблеми, пов'язані з органами зору, посідають одне з провідних місць серед хронічних захворювань і суттєво впливають на якість життя людини, обмежуючи її професійну, соціальну та повсякденну активність. Погіршення зору може призводити до зниження працездатності, обмеження можливостей для навчання, участі в соціальних і культурних заходах, а також до зростання ризику нещасних випадків у побуті й на роботі. За інформацією Всесвітньої організації охорони здоров'я (ВООЗ), понад 2,2 млрд людей у світі мають проблеми із зором, водночас приблизно 1 млрд випадків можна було б запобігти або успішно вилікувати за допомогою ранньої діагностики та відповідного лікування. Ця статистика свідчить про глобальну значущість проблеми й нагальну потребу в упровадженні ефективних методів моніторингу та профілактики захворювань очей.

Одними з найбільш поширених і серйозних очних захворювань є катаракта, глаукома, діабетична ретинопатія, а також інші патології сітківки ока, які без вчасного виявлення можуть призводити до значного погіршення зору або навіть повної сліпоти. Такі патології не лише знижують якість життя, але й створюють суттєве навантаження на систему охорони здоров'я та економіку держави через потребу в тривалому лікуванні, хірургічних втручаннях і соціальній підтримці пацієнтів із порушенням зору. Тому раннє виявлення зазначених патологій, регулярне спостереження й постійний моніторинг стану зору є критично важливими для збереження здоров'я очей, запобігання тяжким ускладненням та зниженням соціально-економічних витрат, пов'язаних із лікуванням і адаптацією людей із порушенням зору.

Традиційні методи діагностики зорових функцій зазвичай передбачають фізичний візит до медичного закладу й використання спеціалізованого обладнання, такого як офтальмоскопи, периметри та інші діагностичні пристрої. Проте такі підходи можуть бути обмежені в досяжності для населення, особливо для мешканців віддалених, сільських або важкодоступних регіонів. Крім того, обмежена частота

ISSN 2522-9818 (print)
Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34) ISSN 2524-2296 (online)

планових оглядів, очікування черг і логістичні труднощі призводять до запізнілого виявлення змін у стані зору, коли лікування стає більш складним, витратним і менш ефективним.

У зв'язку зі стрімким розвитком сучасних технологій і зростанням потреби в покращенні доступності медичних послуг з'являється нагальна необхідність у створенні інноваційних систем для моніторингу стану здоров'я очей. Такі системи мають забезпечувати дистанційний, безперервний і високоточний контроль за станом зору на основі передових технологічних рішень: бездротових сенсорних мереж, розумних пристроїв, алгоритмів оброблення інформації в реальному часі й методів штучного інтелекту. Впровадження подібних технологій дає змогу автоматично виявляти відхилення від нормальних показників, формувати попереджувальні повідомлення для користувача й медичного фахівця, а також вчасно реагувати на розвиток патологічних процесів.

Мета цього дослідження – розробити комплексну систему для моніторингу стану очей та раннього виявлення очних захворювань із застосуванням бездротових сенсорних мереж. Особливу увагу приділено використанню розумних окулярів із вбудованими сенсорами, здатними відстежувати зміни зору, фіксувати фізіологічні параметри користувача й автоматично передавати показники на мобільні пристрої. Інтеграція таких технологій із мобільними застосунками забезпечує можливість постійного моніторингу в режимі реального часу, автоматичний збір і аналіз даних, а також оперативне інформування пацієнта й лікаря про потенційні проблеми.

Запропонована система поєднує сучасні сенсорні технології, алгоритми оброблення й аналізу даних, інтуїтивний інтерфейс користувача та інтеграцію з мобільними пристроями, що допомагає підвищити ефективність ранньої діагностики, забезпечити доступність медичних послуг у будь-який час і в будь-якому місці, а також сприяє розвитку наукових досліджень і підвищенню практичної значущості моніторингу очних захворювань. Крім переліченого, така система відкриває нові перспективи для персоналізованої медицини, де спостереження за станом зору здійснюється індивідуально для кожного користувача з огляду на його фізіологічні особливості, стиль життя й умови навколишнього середовища.

Мета й завдання

Моніторинг здоров'я очей є одним із ключових складників сучасних медичних технологій, оскільки вчасне виявлення змін у стані зору дає змогу попередити розвиток серйозних офтальмологічних захворювань. До таких патологій належать катаракта, глаукома, діабетична ретинопатія, дегенерація сітківки та інші ураження очного апарату, які без ефективного контролю можуть призводити до значного погіршення зору або навіть сліпоти. Вчасна діагностика цих станів є критичною не лише для збереження гостроти зору, але й для підтримки високої якості життя пацієнтів, зменшення соціально-економічного навантаження на систему охорони здоров'я та підвищення загальної безпеки людини в повсякденній діяльності.

В умовах стрімкого розвитку бездротових сенсорних мереж, носимих пристроїв і мобільних технологій з'являється можливість створення інноваційних систем для моніторингу стану очей у режимі реального часу. Такі системи дають змогу автоматично відстежувати зміни зору, вчасно виявляти порушення на ранніх стадіях, а також надавати користувачам і медичним працівникам оперативні рекомендації щодо подальших дій. Використання сучасних алгоритмів оброблення та аналізу інформації допомагає зменшити вплив випадкових шумів, коригувати артефакти й порівнювати показники користувачів із базами даних нормальних і патологічних значень, що стрімко підвищує точність діагностики.

Мета дослідження – розроблення комплексної системи для бездротового моніторингу стану очей, яка використовує розумні окуляри із вбудованими сенсорами для автоматичного збору інформації про стан зорових функцій. Система призначена для виявлення таких змін, як зниження гостроти зору, поява ореолів, плям або затемнень у полі зору, які можуть свідчити про розвиток серйозних офтальмологічних захворювань. Важливим аспектом є здатність системи передавати інформацію в реальному часі на мобільні пристрої користувача, а також забезпечувати їх надійне збереження, оброблення та аналіз для подальшого застосування в консультаціях з лікарем.

Основною метою дослідження є створення технології бездротового моніторингу стану очей із використанням розумних окулярів, оснащених

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

сенсорами для фіксації змін зору й передачі показників за допомогою Bluetooth або Wi-Fi на мобільні пристрої. Ключовим аспектом є інтеграція цієї технології в мобільні застосунки, які забезпечують пацієнта й медичного фахівця зручним інтерфейсом для отримання результатів вимірювань, формування звітів і надання рекомендацій щодо подальших дій. Така інтеграція дає змогу здійснювати безперервний моніторинг стану зору, аналізувати динаміку змін і вчасно реагувати на потенційно небезпечні симптоми.

У межах цього дослідження передбачено розв'язання низки завдань:

– розроблення системи збору й моніторингу інформації про стан здоров'я очей із використанням сенсорів, вбудованих у розумні окуляри або інші носимі пристрої, для забезпечення безперервного й точного вимірювання показників зору;

– створення алгоритмів оброблення й аналізу отриманих показників для автоматичного виявлення змін у зорових функціях, які можуть свідчити про наявність патологій. Це передбачає фільтрацію шумів, нормалізацію сигналів, корекцію артефактів і порівняння з базою даних нормальних і патологічних показників;

– розроблення мобільного застосунку для зберігання історії вимірювань і доступу до результатів моніторингу в реальному часі. Застосунок дає змогу відстежувати динаміку змін стану зору користувачів, формувати аналітичні звіти, забезпечувати інтерфейс для взаємодії з лікарем і надсилати нагадування про необхідність перевірки зору;

– тестування розробленої системи для оцінювання її точності, надійності та ефективності в реальних умовах експлуатації. Це передбачає перевірку здатності системи точно виявляти зміни зору, правильно інтерпретувати результати й надавати коректні рекомендації користувачам і медичним працівникам.

Дослідження також передбачає проведення експериментальних випробувань для перевірки працездатності системи та підтвердження її ефективності в ранньому виявленні патологій. Очікується, що впровадження такої технології підвищить ефективність моніторингу стану очей, сприятиме вчасній медичній допомозі, зменшенню ризику прогресування серйозних захворювань органів зору, а також відкриє нові перспективи для розвитку персоналізованої офтальмологічної медицини, де спостереження за станом зору здійснюється індивідуально для кожного користувача з огляду на його фізіологічні особливості та спосіб життя.

Аналіз останніх публікацій

У сучасних системах моніторингу стану здоров'я очей особливу увагу приділяють не лише збору інформації, а і її високоточному та вчасному аналізу. Це є критично важливим, оскільки раннє виявлення змін у зорових функціях дає змогу запобігти розвитку серйозних офтальмологічних захворювань, таких як катаракта, глаукома, діабетична ретинопатія або дегенеративні зміни сітківки. Вчасна діагностика підвищує ефективність профілактичних заходів, допомагає призначити ефективне лікування та зменшити ризик ускладнень.

Сучасні дослідники пропонують комплексні підходи до створення таких систем, активно інтегруючи технології штучного інтелекту, мобільні платформи й сенсорні пристрої. Завдяки штучному інтелекту й методам машинного навчання можна автоматично аналізувати великі масиви даних про стан зору, виявляти приховані закономірності та прогнозувати розвиток патологій. Мобільні платформи допомагають реалізувати дистанційне спостереження, надаючи користувачам і медичним фахівцям актуальну інформацію в режимі реального часу.

Сенсорні пристрої, зокрема високоточні датчики руху очей, оптичні сенсори й біометричні модулі, забезпечують безперервний збір фізіологічних і оптичних параметрів ока. Поєднання таких технологій створює можливість для автоматизованого й персоналізованого контролю за станом зору, зменшуючи залежність від частих відвідувань клінік і підвищуючи доступність медичної допомоги. Завдяки цьому сучасні системи моніторингу зору стають ефективним інструментом ранньої діагностики, профілактики та підтримки здоров'я очей у повсякденному житті. Автори роботи [1] досліджують застосування глибоких нейронних мереж для автоматичного прогнозування рефракційних помилок за допомогою зображень очного дна. Упровадження цього підходу сприяє досягненню високої точності діагностики та може бути інтегроване в мобільні застосунки для моніторингу зору в реальному часі. Така технологія підвищує ефективність раннього виявлення патологій і забезпечує вчасну профілактику офтальмологічних захворювань.

ISSN 2522-9818 (print)
Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34) ISSN 2524-2296 (online)

У праці [2] запропоновано розроблення мобільного пристрою для офтальмологічного огляду, який допомагає здійснювати офтальмоскопію в польових умовах із використанням стандартних мобільних технологій. Дослідження демонструє ефективність дистанційного моніторингу зору й потенціал для впровадження подібних рішень у системи, що дають змогу користувачам контролювати стан очей за допомогою смартфонів.

Автори роботи [3] розглядають методи автоматичного стеження за рухом зіниць (*eye-tracking*), які можуть бути застосовані для дистанційного моніторингу зору. Використання алгоритмів аналізу руху зіниць дає змогу створювати інтерактивні системи для діагностики, що здатні виявляти порушення на ранніх етапах розвитку патологій, підвищуючи точність і оперативність моніторингу.

У публікації [4] описано систему, що інтегрує сенсори й мобільні платформи для вимірювання різних показників здоров'я, зокрема зорові функції. Така система може бути адаптована для дистанційного моніторингу зору із використанням мобільних застосунків для збору інформації та віддаленого контролю, що робить спостереження доступним у будь-який час і в різних умовах.

У роботі [5] проаналізовано методи застосування штучного інтелекту для скринінгу офтальмологічних захворювань, зокрема діабетичної ретинопатії, яка є одним із найбільш поширених патологій, що негативно впливають на зір. Результати цього дослідження демонструють потенціал інтеграції AI в системи моніторингу для раннього виявлення патологій очей, що значно підвищує ефективність профілактичних заходів і дає змогу лікарям оперативно реагувати на зміни в стані пацієнта.

У дослідженні [6] описано використання спеціалізованого програмного забезпечення для підвищення точності діагностики в офтальмології. Комплекс забезпечує автоматизований аналіз зображень очного дна й визначає наявність захворювань із високою точністю, що є важливим аспектом у створенні систем постійного моніторингу стану зору. Упровадження подібних технологій допомагає ефективно відстежувати динаміку змін у стані очей, вчасно виявляти розвиток захворювань і надавати рекомендації для лікування або додаткових обстежень.

У роботі [7] розглянуто застосування *eye-tracking* у поєднанні з алгоритмами глибокого навчання для раннього виявлення діабетичної ретинопатії. Використання цього підходу дає змогу визначати зміни зору на ранніх стадіях та інтегрувати технологію в мобільні або носимі системи моніторингу, підвищуючи точність і вчасність діагностики.

Автори публікації [8] пропонують інтелектуалізовану систему оцінювання динамічних змін біомедичних зображень, яка автоматично обробляє і аналізує медичні показники. Така технологія може бути застосована для дистанційного моніторингу стану очей і створення систем відстеження змін зорових функцій у реальному часі.

У роботі [9] описано багатофункціональні гнучкі контактні лінзи із вбудованими сенсорами для моніторингу стану очей. Ця розробка демонструє потенціал інтеграції біосенсорів у носимі пристрої для постійного відстеження фізіологічних параметрів і зорових функцій, що розширює можливості сучасних систем мобільного моніторингу.

Дослідження [10] присвячено медико-організаційним підходам до оптимізації спостереження за порушенням зору в школярів. Автори наголошують на важливості організаційних і технологічних рішень для систематичного збору інформації про стан очей у дітей, що може бути основою для впровадження моніторингових систем у навчальних закладах.

У праці [11] запропоновано гібридну архітектуру для аналізу контекстних параметрів руху очей з метою ефективної діагностики різних патологій. Використання такого підходу дає змогу підвищити точність класифікації захворювань та інтегрувати його в системи моніторингу для раннього виявлення офтальмологічних проблем.

Автори дослідження [12] пропонують метод аналізу електроретиносигналів для підвищення інформативності систем діагностики стану сітківки. Метод сприяє більш точному оцінюванню стану зорових функцій та підвищенню ефективності автоматизованих систем моніторингу очей.

Працю [13] присвячено використанню цифрових показників руху очей (DEMOs) як біомаркерів для виявлення неврологічних станів. Результати дослідження підтверджують потенціал застосування подібних показників у системах моніторингу для раннього виявлення зорових і нейрологічних порушень.

У дослідженні [14] описано системи візуального моніторингу в медицині та їх застосування для відстеження стану зору. Робота демонструє

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

ефективність інтеграції сенсорних технологій і програмного забезпечення для автоматичного збору інформації та підтримки медичної діагностики.

Автори праці [15] запропонували високощільну напівсферично вигнуту матрицю зображення на основі гетероструктури MoS_2-графен, яка надихається людським оком. Ця м'яка оптоелектронна система здатна детектувати оптичні сигнали та здійснювати електричну стимуляцію зорового нерва з мінімальними механічними навантаженнями на сітківку. Використання атомно тонких матеріалів і розрахунків методом скінченних елементів підтверджує ефективність запропонованої конструкції.

Опис системи моніторингу стану здоров'я очей

Запропонована система передбачає використання спеціальних розумних окулярів із вбудованими сенсорними датчиками, які забезпечують безперервний та автоматизований моніторинг стану очей користувача. Сенсорні модулі окулярів здатні збирати широкий спектр параметрів зору, зокрема рухи очей (*eye-tracking*), фокусна здатність, рівень освітленості, розпізнавання ореолів або затемнень у полі зору, а також інші фізіологічні показники, що можуть сигналізувати про ранні ознаки очних захворювань.

Окуляри оснащені модулем бездротової передачі інформації (Bluetooth або Wi-Fi), що дає змогу оперативно надсилати зібрані показники на мобільний пристрій користувача – смартфон, планшет або інший сумісний гаджет. На мобільному пристрої встановлений спеціалізований застосунок, що виконує автоматичне оброблення отриманих показників за допомогою алгоритмів штучного інтелекту й сучасних методів оброблення сигналів. Застосунок оцінює стан очей користувача, виявляє потенційні відхилення від нормальних показників і надає зрозумілі повідомлення з рекомендаціями, наприклад звернутися до лікаря або пройти додаткове обстеження.

Система забезпечує постійне відстеження змін у стані очей, що дає змогу вчасно виявляти ранні симптоми таких захворювань, як катаракта, глаукома, діабетична ретинопатія, дегенерація сітківки тощо. Крім того, вона веде детальну історію стану зору користувача, що допомагає здійснювати довгостроковий моніторинг, аналізувати динаміку змін і забезпечувати персоналізовані рекомендації на основі накопиченої інформації.

Схему функціональних модулів системи моніторингу стану здоров'я очей подано на рис. 1. Вона містить такі основні компоненти:

– датчики окулярів для збору параметрів зору й фізіологічних показників користувача;

– модуль бездротової передачі інформації, що забезпечує швидку й надійну синхронізацію інформації з мобільним застосунком;

– мобільний застосунок для автоматичного аналізу, оброблення та візуалізації показників, формування звітів і рекомендацій;

– система повідомлень користувача, яка інформує про виявлені відхилення, надає поради щодо профілактичних дій або необхідності консультації лікаря.

Така архітектура дає змогу створити інтегроване й зручне у використанні рішення для регулярного контролю за станом очей у будь-яких умовах – вдома, на роботі або під час подорожей. Упровадження подібної системи підвищує ефективність раннього виявлення офтальмологічних патологій, сприяє вчасній профілактиці захворювань і допомагає користувачу контролювати стан зору в реальному часі, не відвідуючи медичний заклад.

Рис. 1. Схема функціональних модулів системи моніторингу стану здоров'я очей

Система сенсорних окулярів

Датчики, інтегровані в розумні окуляри, призначені для виявлення широкого спектра змін у зорі, зокрема зниження його гостроти, поява ореолів, плям, затемнень у полі зору або порушення концентрації погляду. Такі симптоми можуть бути ранніми ознаками розвитку серйозних очних захворювань, серед яких катаракта, глаукома, діабетична ретинопатія, дегенерація сітківки,

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

запальні процеси або травматичні ушкодження ока. Для вчасного виявлення патологій система здійснює автоматичне зіставлення отриманих результатів із великою базою нормальних і патологічних параметрів, що дає змогу визначати відхилення від стандартних значень, оцінювати ризики розвитку захворювань і прогнозувати їх потенційний прогрес.

Основу системи становлять високоточні сенсори відстеження руху очей (*eye-tracking sensors*), які використовують інфрачервоні камери й спеціальні джерела світла для фіксації положення та руху зіниці щодо ока й голови користувача. Завдяки цьому можна оцінювати координацію рухів очей, концентрацію погляду, швидкість реакції на візуальні стимули, а також виявляти аномалії в роботі очних м'язів, що свідчать про нейрологічні або офтальмологічні проблеми.

Оптичні сенсори аналізують властивості світла, яке відбивається від сітківки, рогівки та кришталика. Це допомагає оцінювати фокусування зорового апарату, прозорість оптичних середовищ ока й виявляти непрозорі утворення, властиві для катаракти або інших захворювань. Застосування таких сенсорів забезпечує високоточне вимірювання параметрів зору без необхідності інвазивних процедур, що робить моніторинг комфортним і безпечним для користувача.

Біометричні сенсори додатково відстежують загальний фізіологічний стан пацієнта, а саме: пульс, температуру тіла, тиск і рухи голови. Це дає змогу системі брати до уваги зовнішні фактори й фізіологічні коливання, які можуть впливати на результати вимірювань, підвищуючи точність, надійність та індивідуалізацію діагностики.

Зібрані показники автоматично передаються на смартфон користувача за допомогою бездротових каналів зв'язку, такі як Bluetooth або Wi-Fi. Мобільний застосунок аналізує інформацію, порівнює її з базою даних нормальних і патологічних показників, формує детальні звіти про стан зору й надає пацієнтові зрозумілі рекомендації. Це може бути попередження про відхилення, поради щодо профілактичних заходів, а також рекомендація звернутися до офтальмолога для подальшого обстеження або лікування.

Поєднання сенсорних технологій, оптичного та біометричного контролю, а також алгоритмів оброблення інформації та штучного інтелекту забезпечує всебічний моніторинг стану очей у реальному часі. Така система не лише дає змогу на ранньому етапі виявити патологію, але й відстежити динаміку змін зору, проаналізувати ефективність лікування й профілактики, підвищуючи загальну результативність охорони здоров'я очей та якість життя користувачів.

Опис програмного забезпечення

Розроблений алгоритм оброблення інформації із сенсорних окулярів побудований на принципі порівняння отриманих показників із великою базою даних нормальних і патологічних значень. Після прийому інформації від сенсорів – *eye-tracking*, оптичних і біометричних – алгоритм виконує багаторівневе попереднє оброблення сигналів. Цей етап передбачає:

– фільтрацію шумів для усунення електронних і механічних перешкод, що виникають під час збору показників;

– нормалізацію сигналів, що забезпечує порівняність показників між різними користувачами й умовами вимірювання;

– усунення артефактів, викликаних рухами очей, морганням або зміною умов освітлення.

Такі процедури забезпечують високу точність аналізу й мінімізують вплив зовнішніх факторів на результати вимірювань, що критично для вчасного виявлення ранніх ознак офтальмологічних захворювань.

На наступному етапі система формує індивідуальний профіль стану зору користувача, який містить ключові параметри: гостроту зору, наявність ореолів, плям або затемнень, а також показники руху очей і фізіологічні параметри (пульс, температура, рухи голови). Алгоритм автоматично зіставляє отриману інформацію з патернами, що зберігаються в базі даних патологій. Це дає змогу:

– визначати наявність або відсутність ознак конкретних захворювань, таких як катаракта, глаукома, діабетична ретинопатія, дегенеративні зміни сітківки, запальні або травматичні ураження;

– попередньо прогнозувати розвиток патологій на основі аналізу динаміки показників;

– адаптувати порогові значення оцінки для конкретного користувача, беручи до уваги історію вимірювань та індивідуальні особливості зорового апарату.

Для підвищення точності діагностики алгоритм застосовує методи машинного навчання, зокрема класифікаційні та прогнозні моделі, які дають змогу

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

визначати тенденції змін зору та прогнозувати ймовірність розвитку захворювань у майбутньому. Це особливо важливо для раннього виявлення прогресуючих станів, таких як глаукома або діабетична ретинопатія, що на початкових стадіях можуть не виявлятися очевидними симптомами.

На виході алгоритм формує структуровані та зрозумілі повідомлення для користувача й лікаря, що передбачають:

– інформацію про виявлені відхилення;
– можливу патологію, що спричиняє відхилення;
– рекомендації щодо подальших дій, наприклад консультація офтальмолога, додаткові обстеження або повторні вимірювання після певного проміжку часу.

Завдяки такому комплексному підходу система забезпечує безперервне, автоматизоване і вчасне виявлення змін стану зору, спираючись на порівняння з достовірною інформацією про патології. Це підвищує як наукову значущість, так і практичну цінність запропонованої технології, робить її ефективним інструментом для ранньої діагностики, постійного моніторингу та профілактики офтальмологічних захворювань.

Мобільний застосунок

Мобільний застосунок для моніторингу стану очей створено для платформи *Android* із потенційною можливістю адаптації під *iOS* і з використанням середовища розроблення *Android Studio* та мови програмування *Kotlin*. Для забезпечення оброблення показників у реальному часі застосовано багатопотокові процеси, що дають змогу одночасно отримувати інформацію від сенсорів окулярів і виконувати її аналіз без затримок і переривань. Такий підхід гарантує плавну роботу застосунку навіть за великих обсягів інформації та високої частоти оновлення показників.

Застосунок отримує інформацію від сенсорних окулярів через бездротові інтерфейси Bluetooth та Wi-Fi, що забезпечує стабільну синхронізацію та надійну передачу показників, мінімізуючи втрати пакетів даних і затримки у відтворенні результатів. Для підвищення надійності застосовано механізми повторної передачі та буферизації інформації, що допомагає забезпечити безперервність моніторингу навіть у разі тимчасових перерв у з'єднанні.

Інтерфейс користувача розроблено з використанням XML для розмітки екранів та сучасних елементів *Material Design*. Це забезпечує інтуїтивне управління, зручну навігацію та швидкий доступ до основних функцій застосунку, таких як перегляд стану зору, історії вимірювань і повідомлень про відхилення. Для локального зберігання інформації використовується *SQLite*, що дає змогу вести архів вимірювань без доступу до інтернету. Для хмарного зберігання та синхронізації застосовано *Firebase Realtime Database*, що допомагає відстежувати динаміку змін стану зору користувача та надавати лікарю або медичному персоналу доступ до актуальної інформації в режимі реального часу.

Для оброблення медичних показників розроблено власні алгоритми на *Kotlin*, які порівнюють результати з базою нормальних і патологічних значень. Алгоритми аналізують зміни в зорових функціях, формують попередні висновки та готують повідомлення для користувача або лікаря. Для підвищення безпеки та конфіденційності інформації реалізовано механізми її шифрування, управління доступом і автентифікації користувачів, що забезпечує захист персональної медичної інформації відповідно до сучасних стандартів безпеки, разом з GDPR- та HIPAA-подібними вимогами.

Застосунок має модульну архітектуру, що дає змогу легко додавати нові функції та підтримувати сумісність із різними моделями сенсорних окулярів. Це забезпечує масштабованість системи й можливість інтеграції додаткових методів аналізу інформації в майбутньому, зокрема алгоритмів машинного навчання для прогнозування змін зору або раннього виявлення патологій на основі динаміки показників.

Схему взаємодії компонентів програмного забезпечення, що демонструє потік інформації від сенсорів окулярів до мобільного застосунку й хмарного сховища, подано на рис. 2. Вона містить такі модулі:

– збір інформації – прийом сигналів від *eye-tracking*, оптичних і біометричних сенсорів;
– оброблення сигналів – попередня фільтрація, нормалізація, корекція артефактів;
– формування профілю користувача – індивідуальні параметри стану зору й фізіологічні параметри;
– генерація повідомлень – формування зрозумілих результатів і рекомендацій для користувача й лікаря;
– зберігання та синхронізація – локальна база *SQLite* та хмарна база *Firebase* для архівування та віддаленого доступу.

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

Така архітектура забезпечує комплексний та інтегрований підхід до моніторингу стану очей, поєднуючи високоточний збір показників, надійне оброблення, зручний інтерфейс і безпечне зберігання медичної інформації, що робить систему ефективним інструментом для постійного й дистанційного контролю за здоров'ям очей.

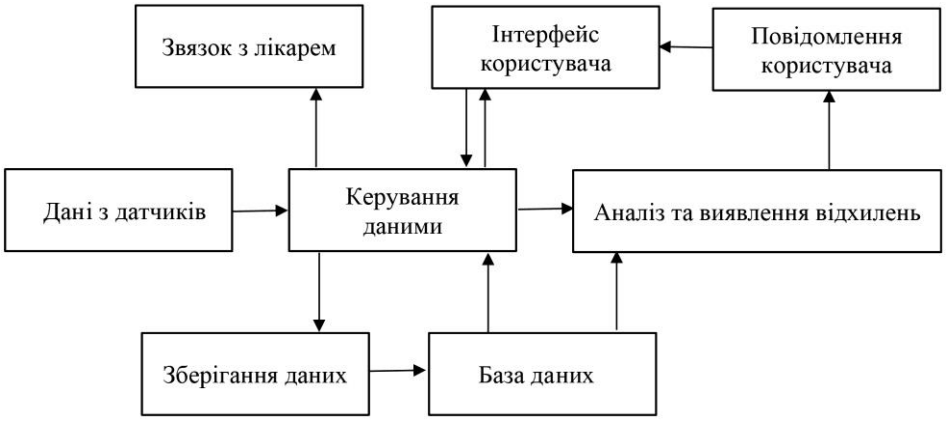


Рис. 2. Схема взаємодії компонентів програмного забезпечення

Результати

Основна мета проведеного експерименту – комплексно оцінити ефективність і точність розробленої системи моніторингу зору на основі сенсорних окулярів у виявленні найбільш поширених офтальмологічних захворювань, таких як катаракта, глаукома та діабетична ретинопатія. Експеримент також передбачав визначення продуктивності системи в реальних умовах використання, оцінювання стабільності роботи алгоритмів і здатності до вчасного виявлення змін у стані зору користувачів. Особливу увагу приділено можливості системи виявляти ранні симптоми захворювань, що дає змогу вчасно реагувати й проводити профілактичні або лікувальні заходи.

Під час експерименту система була протестована на групі пацієнтів різного віку, статі та з різним рівнем тяжкості зазначених захворювань, що допомогло всебічно оцінити роботу алгоритмів у різноманітних клінічних умовах. Такий підхід дав змогу брати до уваги індивідуальні особливості пацієнтів, зокрема анатомічні відмінності очей, наявність супутніх патологій та інші фізіологічні фактори, які можуть впливати на точність вимірювань.

Для аналізу результатів використовувалися стандартні медичні метрики, зокрема точність і чутливість, що допомагають порівняти автоматизовану діагностику з результатами класичних медичних обстежень. Зокрема як контрольні методи застосовували офтальмоскопію для огляду сітківки та рогівки, тонографію для оцінювання внутрішньоочного тиску, а також оптичну когерентну томографію (ОСТ), що забезпечує детальну візуалізацію структур ока. Упровадження таких порівняльних методів дало змогу отримати об'єктивну оцінку ефективності сенсорних окулярів і підтвердити достовірність автоматизованих результатів.

Після завершення тестування зібрано основні показники ефективності й точності (див. табл. 1).

Таблиця 1. *Результати дослідження*

Номер дослідження	Точність діагностики (катаракта), %	Точність діагностики (глаукома), %	Точність діагностики (ретинопатія), %
1	95	90	98
2	92%	85%	89%
3	93%	96%	90%
4	98%	92%	88%
5	89%	87%	95%
6	91%	94%	96%

Аналіз отриманої інформації показав, що система продемонструвала високий рівень точності діагностики для всіх трьох захворювань. Найвищу точність виявлення мала катаракта із середнім показником 94%, що свідчить про надійність алгоритмів у розпізнаванні цього захворювання. Дещо нижчий рівень точності спостерігався для глаукоми, середній показник якої становить 91%, а для діабетичної ретинопатії – 93%. Виявлені незначні відхилення в точності діагностики для окремих пацієнтів, особливо в разі глаукоми (85–96%), можуть бути пов'язані з індивідуальними

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)
Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

особливостями пацієнтів, такими як анатомічні відмінності очей, а також із впливом факторів зовнішнього середовища під час тестування. Ці результати свідчать про те, що, попри високий загальний рівень точності, існує потенціал для подальшого вдосконалення алгоритмів системи, зокрема завдяки оптимізації оброблення інформації та підвищення адаптивності моделей до різних умов використання. Отримані результати підтверджують, що розроблена система сенсорних окулярів є ефективним інструментом для автоматизованого моніторингу стану зору, що дає змогу вчасно виявляти зміни й потенційні патології, сприяючи підвищенню якості медичної діагностики та ранньому втручанню в разі очних захворювань.

Незважаючи на високий загальний рівень точності діагностики, результати експерименту вказують на наявність певних обмежень і потенційних недоліків системи моніторингу зору на основі сенсорних окулярів. Ці обмеження пов'язані як із характеристиками сенсорних технологій, так і з особливостями алгоритмів оброблення інформації та реальними умовами використання.

По-перше, спостерігаються коливання точності в разі виявлення глаукоми, де показники варіюють від 85% до 96%. Це свідчить про вплив індивідуальних особливостей пацієнтів, таких як анатомічні відмінності очей, форма зіниці, товщина рогівки та інші фізіологічні фактори, на ефективність алгоритмів. В окремих випадках ці фактори можуть знижувати точність автоматизованого виявлення патології, особливо на ранніх стадіях захворювання, коли симптоми не значні та їх важко зафіксувати сенсорами.

По-друге, зовнішні умови тестування, зокрема рівень освітленості, рухи голови, моргання пацієнта або інші непередбачувані фактори довкілля, можуть впливати на якість сигналів від сенсорів і, відповідно, на результати аналізу. Це наголошує на необхідності подальшого вдосконалення фільтрів оброблення інформації, алгоритмів компенсації артефактів і впровадження механізмів самокорекції для зменшення впливу зовнішніх факторів на точність діагностики.

По-третє, система демонструє певні обмеження в процесі діагностики діабетичної ретинопатії та катаракти в пацієнтів із нестандартними або складними клінічними випадками. Хоча середні показники точності залишаються високими (93–94%), окремі результати демонструють незначні відхилення, що свідчить про потенціал для інтеграції додаткової інформації або більш складних методів оброблення, таких як алгоритми глибокого навчання, мультипараметричний аналіз і поєднання різних сенсорних сигналів для підвищення точності в складних випадках.

По-четверте, на цьому етапі система орієнтована переважно на визначення конкретних захворювань і не забезпечує комплексну оцінку всіх можливих офтальмологічних патологій, таких як дегенеративні зміни сітківки, запальні процеси або травматичні ушкодження ока. Це обмежує її універсальність і вимагає подальшого розроблення, масштабування алгоритмів і розширення бази даних патологій для забезпечення більш комплексного моніторингу стану зору.

З огляду на зазначені обмеження доцільно впроваджувати адаптивні алгоритми, що беруть до уваги індивідуальні особливості користувачів, підвищувати стійкість сенсорів до зовнішніх факторів, розширювати базу даних патологій і застосовувати методи прогнозування змін зору на основі аналізу динаміки показників. Такі вдосконалення дають змогу підвищити ефективність і надійність системи в реальних умовах експлуатації та забезпечити більш точне раннє виявлення широкого спектра офтальмологічних захворювань.

Висновки

Розроблено комплексну систему для моніторингу стану здоров'я очей, основану на розумних окулярах із вбудованими сенсорами й бездротовою передачею інформації. Система забезпечує автоматичний і безперервний збір ключових показників про стан зору користувача: гострота, рухи очей, оптичні характеристики рогівки та сітківки, а також фізіологічні параметри, такі як пульс, температура тіла й рухи голови. Ця інформація передається на мобільний застосунок, що допомагає користувачам і лікарям оперативно отримувати результати моніторингу й приймати обґрунтовані рішення щодо необхідності медичного втручання.

У системі реалізовано алгоритми оброблення та аналізу інформації, які автоматично порівнюють показники зорових функцій із великою базою нормальних і патологічних значень. Такий підхід дає змогу вчасно виявляти ранні зміни в стані зору, що можуть свідчити про розвиток таких захворювань, як катаракта, глаукома, діабетична ретинопатія,

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

дегенеративні зміни сітківки, запальні процеси або травматичні ушкодження ока. Алгоритми працюють у режимі реального часу й демонструють високу точність діагностики – середня точність оцінки для різних захворювань коливається в межах 91–94%, що підтверджує надійність і практичну ефективність системи.

Для зберігання та оброблення результатів створено спеціалізовану базу медичної інформації з метою відстеження історії змін стану зору користувача. Це уможливлює аналіз динаміки розвитку патологій, допомагає лікарям отримувати актуальну інформацію для консультацій, прогнозувати прогресування захворювань і приймати обґрунтовані клінічні рішення. Для забезпечення конфіденційності та безпеки персональних показників реалізовано їх шифрування, контроль доступу й автентифікацію користувачів, що відповідає сучасним стандартам захисту медичної інформації, зокрема GDPR і HIPAA.

Систему експериментально оцінено в реальних умовах використання. Тестування показало високий рівень точності діагностики й здатність системи вчасно виявляти зміни в зорових функціях користувачів. Незначні відхилення в точності для окремих захворювань свідчать про потенціал подальшого вдосконалення алгоритмів, зокрема способом інтеграції адаптивних методів оброблення інформації та алгоритмів машинного навчання, що сприятиме підвищенню ефективності й надійності системи в процесі моніторингу широкого спектра офтальмологічних патологій.

Розроблена система також передбачає масштабованість та інтеграцію нових технологій: можливість під'єднання додаткових сенсорів, адаптацію алгоритмів для різних типів користувачів, зокрема дітей і літніх людей, а також інтеграцію з електронними медичними картками для централізованого обліку інформації.

Отже, система забезпечує комплексний та інтегрований підхід до автоматизованого моніторингу стану очей, поєднуючи високоточні сенсорні технології, алгоритми аналізу показників у реальному часі, безпечне зберігання інформації та можливість дистанційного спостереження. Це підвищує як практичну цінність, так і наукову значущість запропонованого рішення, створюючи ефективний інструмент для ранньої діагностики, моніторингу й профілактики очних захворювань.

References

1. Sorrentino, F. S., et al. (2024), "Novel Approaches for Early Detection of Retinal Diseases Using Artificial Intelligence", *Journal of Personalized Medicine*, 14(7), 690 p. DOI:10.3390/jpm14070690
2. Parmar, U. P. S., et al. (2024), "Artificial Intelligence (AI) for Early Diagnosis of Retinal Diseases", *Medicina*, 60(4), 527 p. DOI:10.3390/medicina60040527
3. Solomin, A. V., Kovalenko, M. V. (2020), "Automated system for pupil tracking" ["Avtomatyzovana systema stezhennia za zinytsyamy ochei"], *Biomedical Engineering and Technology,* (3), P. 25–29. DOI: 10.20535/2617-8974.2020.3.195134
4. Sedinkin O. A., Derkach M. V., Skarga-Bandurova I. S., Matyuk D. S. (2024), "Eye tracking system based on machine learning" ["Systema vidstezhennia pohliadu na osnovi mashynnoho navchannia"], *Computer-integrated technologies: education, science, production*, (55), P. 199–205. DOI: 10.36910/6775-2524-0560-2024-55-25
5. Ostrovska, K. Y., Porokhnyavyi, V. H. (2025), "Research of neural network models for gaze tracking and object fixation" ["Doslidzhennia neironetworkovykh modelei dlia vidstezhennia pohliadu ta fiksatsii obiektiv"], *Systemni tekhnolohii*, 2(157), P. 170–178. DOI: 10.34185/1562-9945-2-157-2025-17
6. Zaporozhets A. O., Verpeta V. O., Pluto I. V., Komisarenko Yu. I., Pavlenko Yu. R. (2012), "Development of DIOLAS software package for diagnostics in ophthalmoscopy" ["Rozrobka prohramnoho kompleksu DIOLAS dlia diahnostyky v oftalmoskopii"], *Science-Based Technologies*, 14(2), P. 114–118. DOI: 10.18372/2310-5461.14.5104
7. Jiang, H., et al. (2023), "Eye tracking based deep learning analysis for the early detection of diabetic retinopathy: A pilot study", *Biomedical Signal Processing and Control*, 84, 104830 p. DOI: 10.1016/j.bspc.2023.104830
8. Golikov M. S., Donets V. V., Strilets V. Ye. (2025), "Intelligent analysis of optical images based on a hybrid approach" ["Intelektualnyi analiz optychnykh zobrazhen na osnovi hibrydnoho pidkhodu"], *Bulletin of National Technical University "KhPI". Series: System Analysis, Control and Information Technologies*, 1(13), P. 83–87. DOI: 10.20998/2079-0023.2025.01.12
9. Xie, M., et al. (2022), "Multifunctional flexible contact lens for eye health monitoring using inorganic magnetic oxide nanosheets", *Journal of Nanobiotechnology*, 20(1). DOI: 10.1186/s12951-022-01415-8
10. Reshetnyak V. V., For E. E. (2024), "Eye tracking as a tool for researching user behavior" ["Vidstezhennia pohliadu yak instrument doslidzhennia povedinky korystuvachiv"], *Computer-integrated technologies: education, science, production*, 55, P. 181–190. DOI: 10.36910/6775-2524-0560-2024-55-23
11. El Hmimdi, A. E., Palpanas, T., Kapoula, Z. (2024), "Efficient diagnostic classification of diverse pathologies through contextual eye movement data analysis with a novel hybrid architecture", *Scientific Reports*, 14, 21461 p. DOI: 10.1038/s41598-024-68056-9

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

12. Andrushchenko M. A., Selivanova K. G. (2024), "Prospects of diagnosing movement disorders using computer vision methods on a mobile device" ["Perspektyvy diahnostyky rukhovykh rozladiv za dopomohoiu metodiv kompiuternoho zoru na bazi mobilnoho prystroiu"], *Opt.-el. inf.-enerh. tekhn.*, 48(2), P. 96–103. DOI: 10.31649/1681-7893-2024-48-2-96-103
13. Graham, L., et al. (2024), "Digital Eye-Movement Outcomes (DEMOs) as Biomarkers for Neurological Conditions: A Narrative Review", *Big Data and Cognitive Computing*, 8 (12), 198 p. DOI: 10.3390/bdcc8120198
14. Ivaskevych D. R., Popov A. A., Rizun V. S., Gavrylets Yu. P., Petrenko-Lysak A. O., Yachnyk Yu. O., Tukaev S. I. (2023), "Age-related differences in fixation gaze length while reading the news with negative text elements" ["Vikovi vidminnosti v tryvalosti fiksatsii pohliadu pid chas chytannia novyn z nehatyvnymy elementamy tekstu"], *East European Journal of Psycholinguistics*, 10(1), P. 36–47. DOI: 10.29038/eejpl.2023.10.1.iva
15. Choi, C., et al. (2017), "Human eye-inspired soft optoelectronic device using high-density MoS2-graphene curved image sensor array", *Nature Communications*, 8(1). DOI: 10.1038/s41467-017-01824-6

Received (Надійшла) 16.06.2025
Accepted for publication (Прийнята до друку) 30.11.2025
Publication date (Дата публікації) 28.12.2025

Відомості про авторів / About the Authors

Мормітко Олексій Михайлович – Вінницький національний технічний університет, аспірант кафедри біомедичної інженерії та оптико-електронних систем, Вінниця, Україна; e-mail: alexchrist1488@gmail.com; ORCID ID: https://orcid.org/0009-0004-0601-6238

Тимчик Сергій Васильович – кандидат технічних наук, Вінницький національний технічний університет, доцент кафедри біомедичної інженерії та оптико-електронних систем, Вінниця, Україна; e-mail: tymchyksv@ukr.net; ORCID ID: https://orcid.org/0000-0003-2977-1602; Scopus ID: https://www.scopus.com/authid/detail.uri?authorId=55225643900

Mormitko Oleksiy – Vinnytsia National Technical University, Postgraduate Student, Department of Biomedical Engineering and Opto-Electronic Systems, Vinnytsia, Ukraine.

Tymchyk Serhiy – PhD (Engineering Sciences), Vinnytsia National Technical University, Associate Professor, Department of Biomedical Engineering and Opto-Electronic Systems, Vinnytsia, Ukraine.

DEVELOPMENT OF AN INTELLIGENT SYSTEM FOR EARLY DETECTION OF VISION PATHOLOGIES BASED ON ANALYSIS OF EYE MICROMOVEMENT

The **subject** of the research is the development and implementation of an eye health monitoring system using modern technologies, in particular wireless sensor networks, biometric sensors and software for automatic detection of vision diseases. Special attention is paid to methods of processing and analyzing data from sensors for accurate diagnosis of pathologies such as cataracts, glaucoma, diabetic retinopathy and other eye diseases. **The aim** of the work is to create a system that allows detecting visual impairments in real time, performing automatic diagnostics and providing treatment recommendations. The system integrates with a mobile application and can work together with other medical devices to facilitate patient-doctor interaction. **The tasks** solved in the article: 1) develop a system for collecting and monitoring eye health data; 2) create algorithms for processing and analyzing the obtained data; 3) develop a mobile application; 4) test the developed system. **Methods** used in the study: data analysis from biometric sensors, algorithms for automatic comparison of indicators with a database of normal and pathological values, and wireless data transmission technologies (Bluetooth, Wi-Fi). The developed database and software provide secure storage and analysis of medical data. **Results**. The results of the study showed that the system allows monitoring the state of vision in real time with high accuracy (85–90%), detecting pathologies in the early stages and automatically notifying the patient and doctor about detected deviations. The system demonstrates effectiveness in early detection of diseases and allows for timely prescribing of treatment or additional examinations. **Conclusions**. The developed system is an important step towards integrating medical technologies into everyday life. It provides timely detection of vision disorders and convenient access to monitoring results. In the future, it is possible to expand the functions to detect other eye diseases and integrate with additional medical devices for comprehensive monitoring of the patient's health.

Keywords: vision monitoring; wireless sensor networks; smart glasses; automatic disease detection.

Бібліографічні описи / Bibliographic descriptions

Мормітко О. М., Тимчик С. В. Розроблення інтелектуальної системи раннього виявлення патологій зору на основі аналізу мікрорухів очей. *Сучасний стан наукових досліджень та технологій в промисловості*. 2025. № 4 (34). С. 124–134. DOI: https://doi.org/10.30837/2522-9818.2025.4.124

Mormitko, O., Tymchyk, S. (2025), "Development of an intelligent system for early detection of vision pathologies based on analysis of eye micromovement", *Innovative Technologies and Scientific Solutions for Industries*, No. 4 (34), P. 124–134. DOI: https://doi.org/10.30837/2522-9818.2025.4.124

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

UDC 621.865.8
DOI: https://doi.org/10.30837/2522-9818.2025.4.135

I. NEVLIUDOV, S. NOVOSELOV, O. SYCHOVA, S. ZYGIN

THE METHOD DEVELOPMENT FOR CONTROLLING THE MOBILE PLATFORM WITH FOUR STEERING WHEELS

The **subject matter** is a method for determining the robot trajectory with four steering wheels to reach a given point on a terrain map. **The research goal** is to develop a method for determining the orientation of the wheels depending on the trajectory of the mobile platform to increase the maneuverability of an autonomous robotic vehicle in a limited production space. **Tasks to be solved**: to analyze similar solutions, describe the proposed design of the steering unit mechanism for a mobile robotic cart, describe the kinematics of a mobile robot with four steerable wheels, develop an algorithm for the steering unit control module, propose a method for controlling a mobile platform with four steerable wheels, and perform experimental studies on the application of the proposed method. Scientific novelty: a method for determining the orientation of the wheels to reach a given point on the terrain plan has been proposed. An algorithm for performing calculations using a software tool has been developed. A mathematical justification for the method of controlling individual wheel blocks of a mobile platform has been provided. **Methods of the study**: modeling methods and automatic control theory, methods for describing linear dynamic systems, analytical modeling methods, computer modeling in the Matlab/Simulink environment. **Results and conclusions**: The mobile platform movement principle using four independent steering wheels is considered. A method for determining the orientation of the steering wheels depending on the trajectory of movement is proposed, which is based on the geometric analysis of the position of the platform and the target point, which allows calculating the angle of rotation of each wheel in such a way as to ensure movement to a given point without lateral slippage. A mathematical model of the control system is built, a structural and functional diagram is developed, an algorithm for processing commands, calculating the angles of rotation is described, and a three-level control system is implemented: linear speed, wheel orientation angle and angular speed of the entire platform. The developed mock-up sample of the mechatronic steering wheel assembly is described. The simulation conducted in the Simulink environment confirmed the operability of the proposed system.

Keywords: mobile robot; AGV; intelligent manufacturing; control method; automated system.

Introduction

In modern manufacturing environments, automated guided vehicles (AGVs) play a key role in intra-shop logistics, cargo transportation, and production line maintenance. However, the effective use of AGVs is often complicated by the specifics of the premises: limited space, narrow aisles, a large number of obstacles, and the need to perform precise maneuvers. Mobile transport robots operating in confined spaces, such as warehouses and factory floors, have strict requirements for omnidirectional mobility. Currently, various drive units have been developed that can move in all directions [1–5]. The two most common technologies that solve the problem of omnidirectional movement can be distinguished: omnidirectional Wheels [6] and rotary steering wheels [7].

The technology of using four steering wheels (4WS) opens up new opportunities for improving the maneuverability and precision of control of mobile platforms. The basic idea of 4WS is that each of the four wheels is not only driven by a separate motor, but can also change its angle of rotation independently of the others. This allows the AGV to perform complex maneuvers, including:

– circular movement with a minimum turning radius;

– "crab movement" (lateral movement of the entire platform);

– U-turn on the spot;

– precise positioning in the loading or service area.

In production facilities with narrow aisles, this allows you to reduce the area required for maneuvering, which, in turn, increases the density of equipment placement and optimizes logistics routes. In addition, by reducing the number of reverse movements, the load on the drive mechanisms is reduced, energy efficiency is increased and wear on the platform elements is reduced.

Intelligent control systems used for AGVs with four steering wheels are able to process sensor data, analyze trajectories and make decisions in real time. This ensures safe navigation in a dynamic environment, collision avoidance, and adaptation to environmental changes. This technology is especially effective for robots working in warehouse complexes, in electronics or pharmaceutical production, where space constraints are tight and movement accuracy is critical.

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)
Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

This study is relevant because mobile platforms with four steering wheels have increased maneuverability in a confined space compared to traditional wheeled and even tracked types of control. However, such platforms are not yet very widespread due to the complexity of their control methods. The complexity of the control methods is due to the peculiarities of the kinematic scheme of the mobile platform, in which each wheel turns a certain angle independently of each other.

The goal of the work is to develop a method for determining the orientation of the wheels depending on the trajectory of the mobile platform to increase the maneuverability of an autonomous robotic vehicle in a limited production space. The object of the study is an automated system for controlling the movement of a mobile platform. The subject of the study is a method for determining the trajectory of a robot with four steering wheels to reach a given point on the terrain map.

Research task statement: to achieve the goal, it is necessary to analyze similar solutions, describe the proposed design of the mechanism of the steering wheel block for a mobile robotic cart, describe the kinematics of a mobile robot with four steering wheels, describe the developed algorithm for the operation of the steering wheel block control module, describe the proposed method for controlling a mobile platform with four steering wheels, and perform experimental studies on the application of the proposed method.

Analysis of last achievements and publications

A wheeled platform with independent drives and steering capabilities has several key advantages that make it particularly effective for mobile robots in various applications. Independent control of each wheel allows the robot to perform complex maneuvers, such as turning on the spot or moving in a tight space. This is especially important for tasks where space is limited, such as in production and warehouse facilities with narrow aisles. Thanks to the independent wheel drive, the platform can easily adapt to different surface conditions and provide reliable traction with it. This reduces the risk of slipping, allowing precise control of movement even on uneven or slippery surfaces. Steering of each wheel allows you to configure different movement modes, such as parallel control to reduce the turning radius or crab movement, where all wheels turn at the same angle for lateral movement. This significantly expands the possibilities of using a mobile autonomous robot in various scenarios of its use.

In works [8-10] a comparative analysis of different types of kinematic schemes of mobile robots is presented (Fig. 1). Scheme 1 shows an example of the implementation of a three-wheeled mobile robot with a motorized rear wheel drive and a neutral steering front wheel, with which the direction of movement is controlled. Scheme 3 shows an example of the implementation of a three-wheeled mobile robot, in which a front drive steering wheel with a motor is used. The rear wheels in this scheme are neutral without motors.

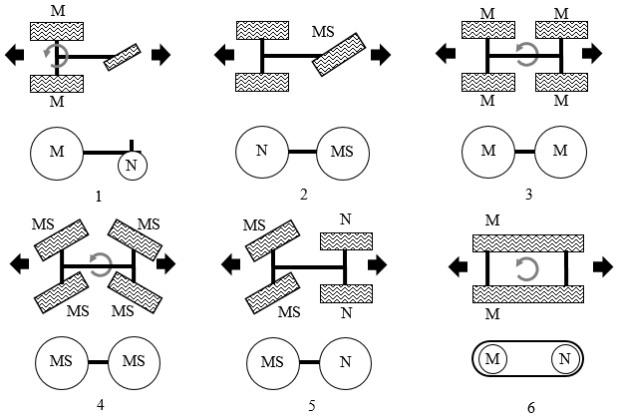


Fig. 1. The mobile robots main kinematic schemes classification:
M – Motor wheel; N – Neutral wheel;
MS – Motor and Steering wheel

Scheme 3 is a classic implementation of a mechanism with differential control (Differential Drive) [11]. In such a scheme, each wheel has its own separate drive (motor), but the direction of movement and turns of the robot are achieved due to the different speeds of rotation of the left and right wheels, without turning the wheels themselves. The advantage is a simple design, high maneuverability, the ability to turn in place, but the disadvantage is the complexity of control on uneven surfaces.

Scheme 4 uses four independent wheels, each of which can rotate. This is the so-called four-wheel scheme with independent wheel control (Four-Wheel Independent Steering and Drive) [12]. Each wheel has its own drive and can be independently controlled. This allows you to control the direction of movement of the robot without having to turn the entire platform. The advantage is high maneuverability, the ability to move sideways (crab movement), the ability to turn in place, but the

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

disadvantage is the complex design and the need for complex control.

In scheme 5, steering is achieved using Ackermann Steering [13]. This scheme is used in cars and has motorized front wheels that turn and rear wheels that move straight. Turning is achieved by changing the angle of the front wheels. The advantages are stability at high speeds, effective steering over large areas. The disadvantages include poor maneuverability in confined spaces and the inability to turn on the spot.

Scheme 6 is a tracked platform where tracks are used instead of wheels, which allows the robot to move over complex and uneven surfaces (Tracked Drive) [14]. Tracks can be controlled using the differential control principle. The advantage is high cross-country ability and ability to overcome obstacles. The disadvantages of such an implementation are limited maneuverability compared to wheeled platforms, high friction.

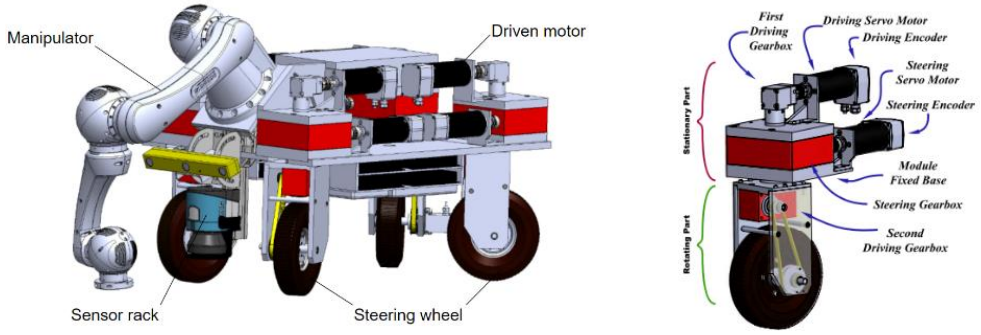
The studies conducted in [8] showed the effectiveness of a wheeled platform with independent drives and the ability to steer.

The article [15] discusses the development of a new control strategy for a four-wheeled mobile robot, which allows it to accurately follow the route in off-road conditions, while maintaining the desired orientation and avoiding lateral slippage. The main goal of the study is to develop a method that will not only provide the robot with the ability to follow the trajectory, but also help maintain the desired orientation of the robot, compensating for the effect of slippage, which is typical for off-road conditions. Also in this work a new approach to control the angles of rotation of the front and rear wheels using the strategy of "parallel control" is proposed. To increase the accuracy of control in sliding conditions, it was proposed to independently control the angles of rotation of the front and rear wheels. The rear wheels are responsible for the correction of lateral deviation, while the front ones control the orientation of the robot relative to the trajectory. To implement this strategy, two control laws are used, based on the backstepping method, which allows to achieve both the required trajectory and the desired orientation.

A feature of the proposed solution is the use of an observer to estimate the side slip angles. This allows to increase the control accuracy, taking into account the adhesion conditions with the surface and adjusting the algorithms based on the current slip conditions. The application of the developed strategy is promising, in particular, for solving tasks where mobile robots must perform tasks in confined spaces or on rough terrain. The results of the simulations confirm that the combination of the observer with the control laws based on the backward search method allows the robot to perform complex maneuvers with high accuracy.

In [16], the design of a mechatronic system of a four-wheeled robot with independent control of each wheel and a mechanism of fault-tolerant feedback by odometry is described. To perform tasks in a confined environment, the authors decided to use a platform with independent control of each wheel (4WS), which provides high maneuverability and efficiency of movement. One of the key design tasks was to synchronize eight electric drives to provide independent control of the movement and rotation of each wheel. This allows the robot to perform complex maneuvers in confined spaces. To implement these tasks, an embedded controller based on an embedded PC is used to provide the necessary computing power.

The article presents a 3D model and kinematic diagram of the mobile platform, and provides a description of the rotary unit for four independent wheels (Fig. 2).



a) b)

Fig. 2. Proposed design of a mobile robot with four independent swivel wheels [14]:
a) 3D model of the mobile robot; b) design of the swivel wheel mechanism

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)
Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

The kinematic models presented in the article are used both for calculating motion control signals and for estimating the robot speed. Special attention is paid to solving problems with odometry and fault-tolerant estimation of the robot speed and position. The authors note that the developed system provides high maneuverability and reliability in difficult conditions. Fault tolerance is achieved due to special kinematics estimation algorithms that allow ignoring incorrect sensor data and maintaining operability even in the event of partial equipment failures.

In [17], a motion control algorithm for a four-wheel autonomous steering system (4WS) is presented, which independently controls the steering angles of the front and rear wheels. This capability allows the use of three different driving modes: forward steering mode, crab steering mode, and symmetrical steering mode. The proposed algorithm effectively uses these modes to achieve precise maneuverability, facilitating accurate navigation to the target location. The authors conducted a kinematic analysis of a four-wheel steered vehicle, focusing on its motion characteristics.

A separate task should be to plan the trajectory of a mobile autonomous robot with four independent swivel wheels. In [18], a description of the trajectory planner method based on the adaptive control strategy for a four-wheel steering autonomous vehicle (4WS) is presented. The article describes two control strategies for following the trajectory of a mobile robot with four steering wheels based on the feedback control method. In [19], a method for controlling a four-wheeled vehicle with steering wheels is proposed based on the calculus of variations, and the corresponding optimal path is selected according to a predefined objective function. In [20], a new approach to local path planning for an independent mobile base with four wheels (I4WS) is presented. The proposed method adaptively steers and rotates the platform, continuing to follow the given path and avoiding obstacles. The implemented predictive control model (MPC) generates optimized trajectories to avoid collisions up to several meters ahead, taking into account a set of data on a predefined trajectory, using the method of laser reading of the surrounding space.

Method for determining the orientation of the wheels depending on the trajectory of the mobile platform

In the process of moving a mobile platform, the tasks of changing the rectilinear direction of movement arise. For this purpose, various principles of constructing steering mechanisms are used. In our work, the principle of control with four independent swivel wheels is considered. Fig. 3 shows a kinematic model of a mobile platform with four wheels that can change the angle of rotation.

This kinematic model is characterized by the fact that the perpendicular lines to each wheel meet at a single point, which is the center of rotation of the mobile platform. This condition guarantees a rotation without the effect of slipping. The intersection point $C(x, y)$ is the center of rotation or the instantaneous center of rotation of the robotic vehicle. It can change during the movement – for rectilinear movement, the radius from $C(x, y)$ to each wheel has an infinite length, while for motion in place it is equal to the distance from the center of the wheel mounting to the center of the platform $M(x_0, y_0)$.

The center of mass of the robotic vehicle $M(x_0, y_0)$ rotates along a circular trajectory with a radius R, as well as a linear velocity V and an angular velocity ω. The distance between the front and rear axles is the wheelbase l, the distance between the wheels of one axle is the track d. Each wheel has a linear velocity vector V_i and a rotation angle δ_i, which is measured between the longitudinal direction of the robotic vehicle and the direction of the steering wheel.

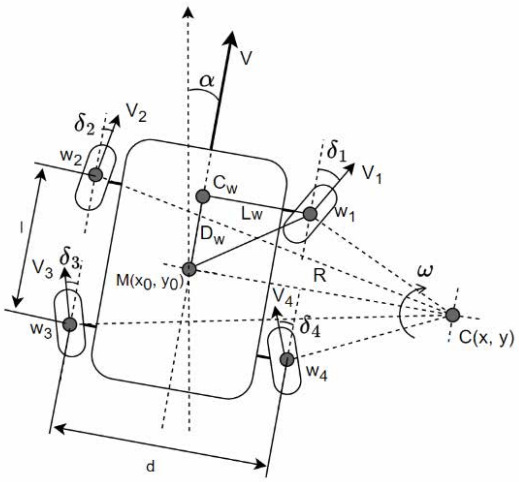


Fig. 3. Kinematic model of a moving platform with four wheels that can change the angle of rotation

The problem being solved has the following initial requirements. For a platform located at point $M(x_0, y_0)$, and currently having a velocity vector v (Fig. 4), it is

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

necessary to determine such a rotation angle for each wheel so that it reaches point $T(x_1, y_1)$.

The following method is proposed to solve the problem.

On the map of the area, the coordinates of the location of the mobile platform and the desired point to which it is necessary to move this platform are determined.

If the target point is not on the axis of the velocity vector v, then it is necessary to find such a position of each steering wheel to ensure smooth rotation of the platform during movement to this point.

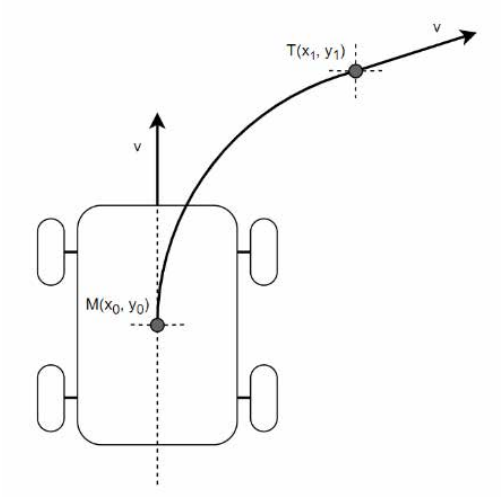


Fig. 4. Mobile platform movement trajectory

At the next stage, it is necessary to find the coordinates of the center of the circle $C(x, y)$, on which the trajectory of the platform movement lies (Fig. 5).

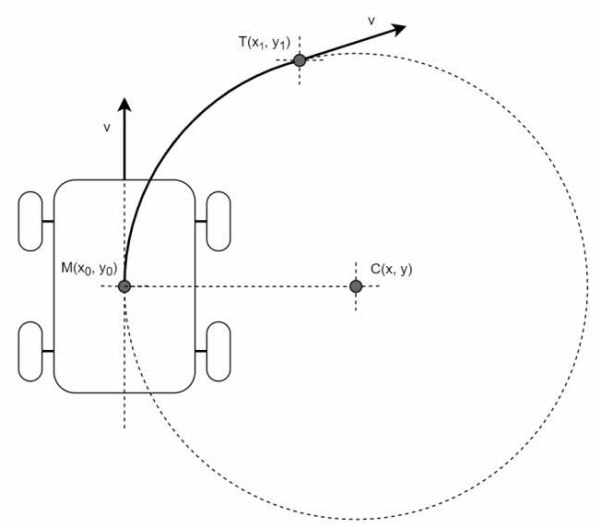


Fig. 5. Coordinates of the center of the circle $C(x, y)$, on which the trajectory of the platform movement lies

When determining the center of the circle, it is taken into account that its center is always located on the axis emanating from the geometric center of the platform (Fig. 5), perpendicular to the velocity vector v.

To calculate the value of the radius R, it is necessary to determine the angle of inclination of the trajectory of motion α (segment MT in Fig. 6).

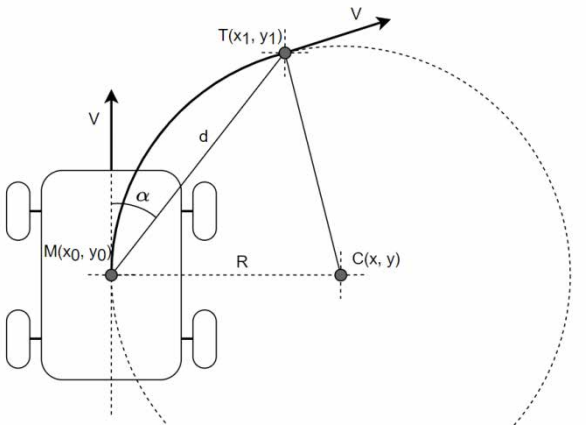


Fig. 6. Radius R calculation

By the law of cosines we get:

$$d^2 = 2R^2\left(1-\cos(2\alpha)\right), \qquad (1)$$

where d is distance between the center of the mobile platform and the target point, R is the radius of the circle on the boundary of which points M and T lie; α is the angle between the velocity vector v and the line d.

Let us rewrite (1) as follows:

$$d^2 = 2R^2\left(1-\cos(2\alpha)\right) = 4R^2\sin^2\alpha, \qquad (2)$$

Now from (2) we find R:

$$R = \frac{d}{2\sin\alpha}. \qquad (3)$$

In order for the platform to reach point T, knowing the radius of the circle that is the trajectory of its movement, it is necessary to determine the angle of rotation of each wheel.

Let us consider the procedure for determining the angle of rotation of each wheel. To do this, we will build a geometric model that describes the behavior of the wheel during its rotation. Such a model for one wheel is shown in Fig. 7.

The main task in this case is to find the coordinates of the center of the wheel W_1 relative to the coordinates of the center of the mobile platform M. When we know this coordinate, we can draw a segment W_1C from the center of the circle of the platform's trajectory.

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)
Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

The direction of rotation of the wheel will be normal to the segment W_1C.

In the general case, the platform will be in an arbitrary position relative to the central axis, which is the base for the terrain map and parallel to one of the x or y axes in the coordinate system. In this case, it is necessary to take into account the angle of rotation of the entire platform α (Fig. 7).

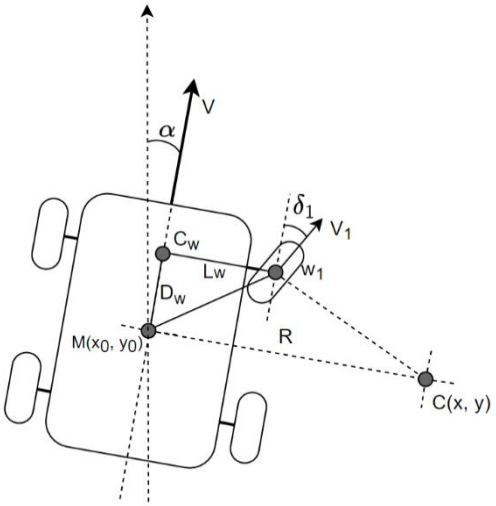


Fig. 7. Geometric model describing the behavior of the wheel

In Fig. 7 D_W is the distance from the center of the platform M to the line connecting the centers of the front wheels. L_W is the distance between the central axis C_W of the platform and the center of the wheel W_1.

The x coordinate of the point C_W, which is the center of the segment connecting the centers of wheels W_1 and W_2 (Fig. 3), is calculated by the formula:

$$C_{W,x} = C_x - D_W \cos\alpha, \quad (4)$$

where α is the angle of rotation of the mobile platform relative to the base axis; D_W is the distance from the center of platform M to the line connecting the centers of the front wheels.

The y coordinate of point C_W is calculated similarly, but instead of cosine we use sine:

$$C_{W,y} = C_y - D_W \sin\alpha. \quad (5)$$

Finally, the coordinate of the wheel center $W_{1,x}$ can be found using the:

$$W_{1,x} = C_x - L_W \cos(\alpha - 90), \quad (6)$$

where L_W is the distance between the central axis C_W of the platform and the center of the wheel W_1.

The coordinate of the wheel center $W_{1,y}$ is found by the formula:

$$W_{1,y} = C_y - L_W \sin(\alpha - 90), \quad (7)$$

The angle of rotation of the wheel is found as the normal to the radius extending from the center of the wheel $W_{1,(x,y)}$ to the center of the circle $C_{x,y}$.

To find the normal, we need to know the diameter of the wheel. Fig. 8 shows a schematic diagram of the main parameters of the mobile platform for calculating the angle of rotation of the wheel.

The coordinates of points D_{Wa} and D_{Wb} are calculated using the formulas:

$$D_{Wa} = W_{1,x} - D_W \frac{C_y - W_{1,y}}{\sqrt{(C_x - W_{1,x})^2 + (C_y - W_{1,y})^2}}, \quad (8)$$

$$D_{Wb} = W_{1,y} - D_W \frac{C_x - W_{1,x}}{\sqrt{(C_x - W_{1,x})^2 + (C_y - W_{1,y})^2}}, \quad (9)$$

where D_W is the wheel diameter; $W_{1,x}$ is x coordinate of the wheel center; $W_{1,y}$ is y coordinate of the wheel center; C_x is x coordinate of the trajectory circle center; C_y is y coordinate of the trajectory circle center.

The wheel rotation angle δ_1 is calculated as the angle between the central axis of the mobile platform, which is represented by points P_f and P_b, and the vector V_1 connecting points D_{Wa} and D_{Wb}.

$$\theta_1 = tg^{-1}(D_{Wa,y} - D_{Wb,y}, D_{Wa,x} - D_{Wb,x}), \quad (10)$$

$$\theta_2 = tg^{-1}(P_{f.y} - P_{b,y}, P_{f.x} - P_{b,x}), \quad (11)$$

where D_{Wa} and D_{Wb} are the points that define the diameter of the wheel; $P_{f.y}$ and $P_{f.x}$ are the points that define the length of the platform along the main axis.

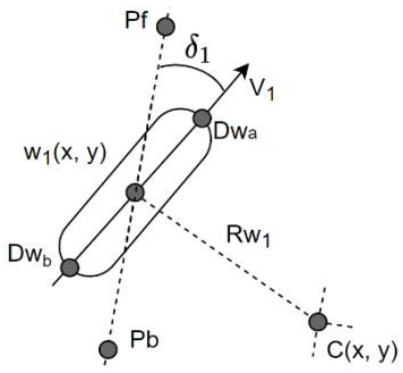


Fig. 8. Calculation of the wheel rotation angle

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

The modulus of the difference between the values of θ_1 and θ_2 is calculated by the formula:

$$d_\theta = |\theta_1 - \theta_2|, \quad (12)$$

The wheel rotation angle δ_1 is calculated by the formula:

$$\delta_1 = \min\{d_\theta, |180 - d_\theta|\}. \quad (13)$$

Design of the swivel wheel unit mechanism

To build our own mechanism for the swivel unit, existing designs of similar solutions were analyzed. The MK4 Swerve Module [21] is an example of an SDS Swerve drive. The MK4 is equipped with a 1.5-inch wide swivel wheel. The MK4 uses a set of SDS 2nd generation bevel gears to drive the swivel wheel (Fig. 9). The MK4 uses a centrally located steering encoder to directly measure the angle of rotation without the use of gears. This eliminates encoder backlash and reduces the complexity of the design.

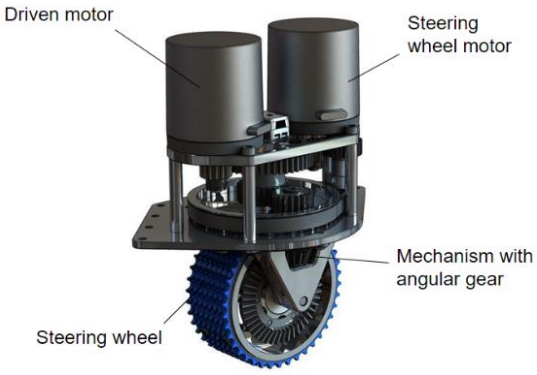


Fig. 9. SDS swivel wheel design

The compactness of the design is achieved by the vertical arrangement of the wheel drive motors and the rotation of the entire mechanism relative to the platform.

Fig. 10 shows an example of the implementation of the swivel wheel mechanism using the differential construction principle [22].

This configuration of the drive motors allows for a reduced height of the mechanism, which is suitable for the implementation of autonomous mobile robotic platforms for warehouse purposes. Such mechanisms can be located in the corners of the platform. The MK4i module inverts the motors on the MK4 module to a lower position, where they do not interfere with other components of the system. Thus, the overall height and center of gravity are lower. The MK4i module also moves the wheel further into the corner of the chassis for a wider wheelbase, which provides more stable operation of the mobile robot. Since the wheel is moved to the area usually occupied by the frame, the MK4i includes an auxiliary plate for proper placement in the chassis.

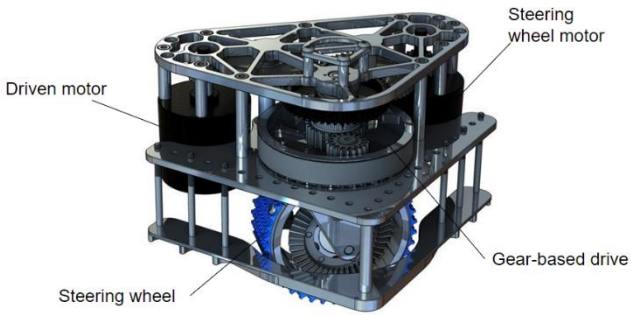


Fig. 10. Differential drive of the swivel wheels

Thus, the considered examples of the implementation of swivel wheel mechanisms showed that each design has its own advantages for a certain sector of application and can be changed depending on the purpose of the end device. This work proposes a universal design of a mechatronic module that uses the advantages of a parallel arrangement of motors to simplify the implementation of the wheel drive, and a differential one, which is characterized by a lower height.

Fig. 11 shows the proposed design of the swivel wheel unit.

This design uses a gear mechanism to rotate the wheel relative to the chassis (Fig. 11, a). Using a mounting plate, the wheel rotation unit is attached to the chassis with four screws. Thanks to the drive gear, which is mounted on the axis of the stepper motor, rotation is carried out to a given angle. The gear of the rotation angle sensor (Fig. 11, b) rotates synchronously with the drive gear, through the main gear. The rotation angle sensor is built on the basis of a multi-turn variable resistor. The rotating wheel is driven by its own motor. A belt drive is used to connect the wheel with the motor.

The entire mechanical part of the rotating wheel unit is mounted on a base plate (Fig. 11, a, b), which rotates relative to the axis on which the main gear is mounted, which is rigidly attached to the chassis of the mobile robot. An electronic control unit based on a microcontroller is also mounted on the base plate. This unit receives commands from the central control unit via the RS-485 interface and Modbus protocol.

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

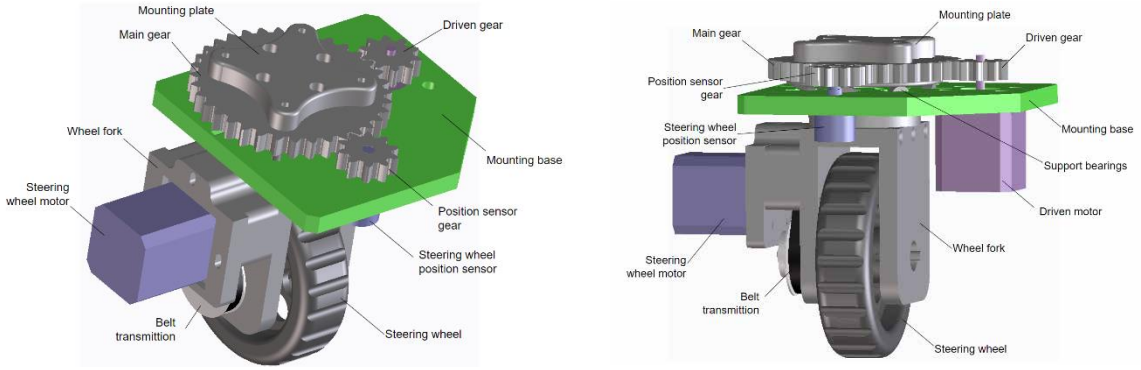


Fig. 11. Proposed design of the swivel wheel unit:
a) view of the gear mechanism for rotating the wheel; b) view of the location of the motors and the rotation sensor

Description of the operating principle of the automated system for controlling the position of the steering wheel unit

Fig. 12 shows a block diagram of the motion control system of a mobile transport robot with swivel wheels.

The central control unit performs the function of collecting information from the local navigation subsystem and transmitting commands to the executive modules. It coordinates the operation of the turning wheel units, based on commands from the remote control or the autonomous navigation subsystem. At each step, depending on the current location coordinate on the map of the production facility, the angle of rotation of each wheel is calculated. The obtained parameters are transmitted using the RS-485 interface and the Modbus protocol to each individual turning wheel unit.

The autonomous navigation unit is responsible for the automatic movement of the robot according to the programmed route or built-in navigation algorithms (for example, using GPS, lidar, ultrasonic sensors, etc.). It can determine the current trajectory of movement and the position of the robot in space and transmits this information to the central control unit for further processing.

Independent turning wheel control units directly control the drives of each wheel. They receive commands from the central control unit and autonomously regulate the rotation speed and angle of rotation of the wheels. Thanks to this architecture, each wheel can move independently, which allows the robot to perform complex maneuvers, such as turning on the spot or crab movement (sideways movement).

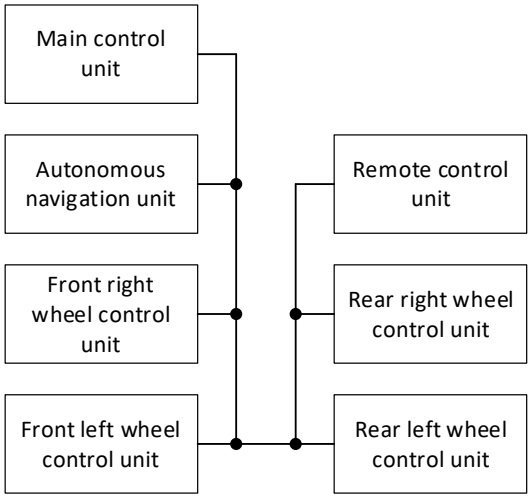


Fig. 12. Block diagram of the motion control system of a mobile transport robot with swivel wheels

Fig. 13 shows a diagram of the control algorithm for the rotary wheel module. At the beginning of the work, the wheel is set to the initial position. The initial position is considered to be the position when the limit sensor installed on the base is triggered. All other positions of the wheel are counted from this position.

After finding the initial position, the wheel is moved to the middle position. The middle position is considered to be the position in which the wheel is oriented clearly parallel to the direction of movement of the mobile platform.

When the module completes the calibration stage, the module controlling its operation goes into the waiting mode for commands from the main controller. Upon receipt of a command from the controller, it is processed and the specified angle of rotation of the wheel is determined.

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

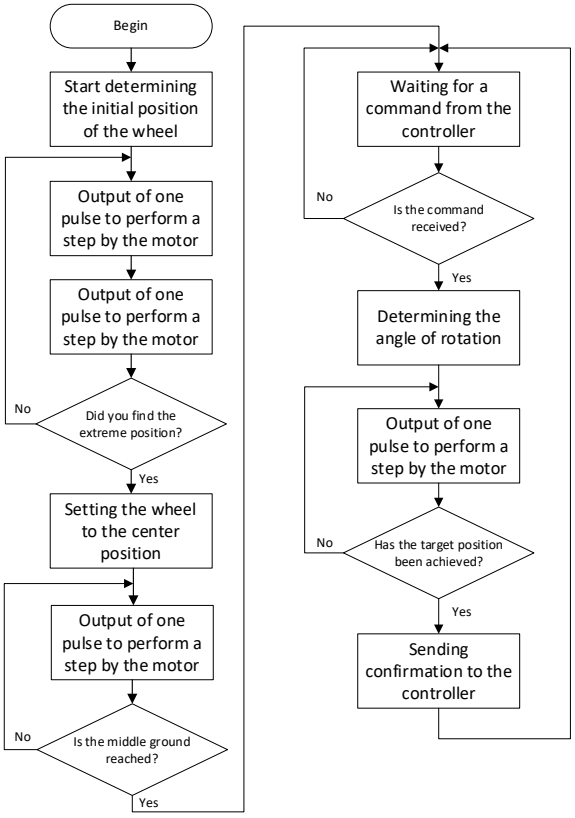


Fig. 13. Swivel wheel module control algorithm

In case of receiving a specified angle, the process of rotating the wheel to the specified position is started. This position is determined by the number of pulses to the stepper motor from the control module, and is also controlled by data from the rotation angle sensor. After reaching the specified position, the stepper motor stops and holds the wheel in the specified position, and the control module transmits a data packet to the main controller to confirm the operation of performing the wheel orientation.

Fig. 14 shows a diagram of the algorithm for synchronous control of the complex of rotating wheel modules.

At the beginning of operation, the initial position of the wheels is initialized. For this, a corresponding command is issued, which is recognized by each individual rotation module. Each module, independently of each other, executes the initialization command and reports this to the main module at its request.

The main module sequentially polls each rotation module and checks whether it has reached the initial position. Only after taking the specified position, the program proceeds to the stage of directly controlling the movement of the mobile platform. The determination of a certain angle of rotation of each wheel is based on data from the main control unit, which is responsible for laying the path. Typically, such a module is implemented on the basis of a Raspberry PI mini-computer.

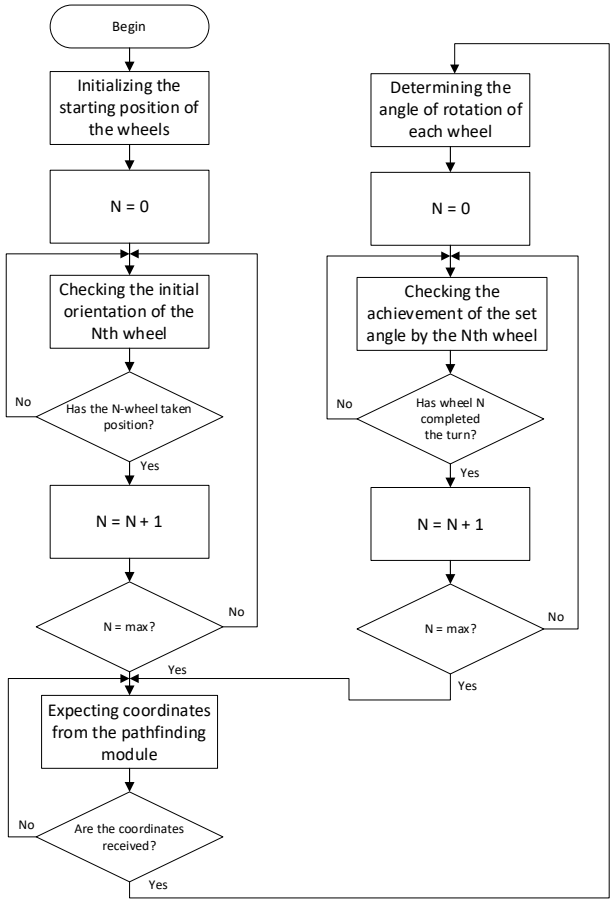


Fig. 14. Algorithm diagram for synchronous control of a complex of rotary wheel modules

Based on the received target coordinate and the current point in space, the method proposed in this work determines the angle of rotation of each wheel. Before starting the movement, the main controller sends a command to each control module with the corresponding data and monitors the completion of the process of acquiring the wheels of the given orientation. After completing the operation of turning the wheels, the module goes into the waiting mode for new coordinates.

Synthesis of the automatic control system block diagram

Fig. 15 shows a block diagram of the system for automatically adjusting the steering wheel angle to a given angle and linear speed of movement.

In this scheme, three loops can be distinguished. The first loop is responsible for regulating the linear speed and providing the setpoint V_{target} using

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)
Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

a PID controller. The speed of movement is controlled by an encoder mounted on the wheel axis. A speed sensor is included in the feedback loop. The difference between the current and desired speeds $e_{\omega,i}$ is used to control the motor, which provides the necessary acceleration or deceleration. The encoder provides feedback and measures the actual speed V_i, which allows the controller to adjust the motor operation to achieve the desired speed value V_{target}.

The second loop is responsible for controlling the rotation of each controlled wheel to the desired angle δ_{target}. An angle sensor – a multi-turn potentiometer – is used as a sensor to measure the current angle of rotation of the wheel. The error $e_{\delta,i}$ between the actual angle δ_{out} and the specified angle of rotation $\delta_{target,i}$ is processed by a PID controller that regulates the motor, which ensures smooth and accurate positioning of the wheel mechanism. The angle of rotation is monitored by a sensor based on a multi-turn resistor. The model takes into account a reducer based on a gear transmission to provide the required ratio.

The third loop regulates the angular velocity ω_i of the mobile robotic platform. Each wheel has its own angular velocity, which is determined by the radius of rotation for the i-th wheel R_i and the linear velocity of the i-th wheel V_i. Angular velocity

$$\omega_i = V_i / R_i, \quad (14)$$

is determined by the turning radius, which is calculated by the formula

$$R_i = \frac{l}{tg\left(\delta_{\omega,i}\right)}, \quad (15)$$

where $\delta_{\omega,i}$ is the angle of the wheel rotation; l is the distance between the front and rear axles.

The error between the actual angular velocity ω_i and the average angular velocity ω_{avg} determined by the other steered wheels is fed to the PID controller to correct the steering angle of the wheels.

The lateral component of the velocity for each wheel is defined as:

$$V_{y,i} = V_i \sin\left(\delta_i\right). \quad (16)$$

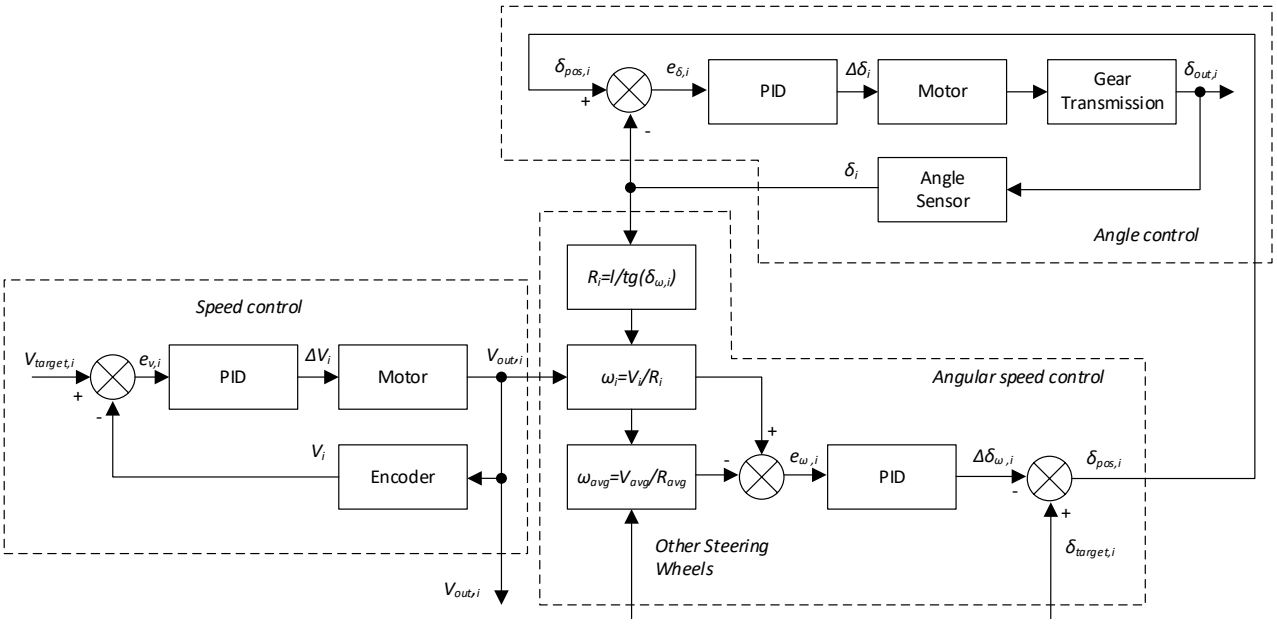


Fig. 15. Block diagram of the system for automatic adjustment of the steering wheel angle to a given angle and movement linear speed

The lateral component of velocity is the component of the velocity vector that determines the speed of an object in a direction perpendicular to the main direction of its movement. This speed is directed across the axis of motion of the vehicle or robot. The lateral component of velocity takes into account lateral displacements that occur during turns or other maneuvers.

The circuits work in conjunction to ensure smooth movement of the mobile robot. The first circuit controls the speed of the platform, the second is responsible for precise control of the angle of rotation of the wheels,

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

and the third ensures uniform execution of the turn by all wheels of the platform.

The total angular velocity of the mobile robot ω_{avg} is determined based on the average turning radius and the average linear velocity. The average turning radius R_{avg} will be calculated taking into account all four wheels, taking into account their individual radii R_i and linear velocities V_i:

$$R_{avg} = \sum_{i=1}^{4} R_i V_i \Big/ \sum_{i=1}^{4} V_i \,, \tag{17}$$

Then the angular velocity ω_{avg} of the robot can be calculated as:

$$\omega_{avg} = V_{avg} / R_{avg} \,, \tag{18}$$

where R_{avg} is the average turning radius, taking into account the different turning angles for each wheel; V_{avg} is the average linear velocity, calculated by the formula:

$$V_{avg} = \frac{1}{4} \sum_{i=1}^{4} V_i \,, \tag{19}$$

Given the independent control of each wheel and their different angles of rotation, the formula for angular velocity will look like this:

$$\omega_{avg} = \frac{1}{4} \left(\sum_{i=1}^{4} V_i \right)^2 \Big/ \sum_{i=1}^{4} \frac{l V_i}{tg(\delta_i)} \,. \tag{20}$$

Formula (20) takes into account the individual angles of rotation and speed of each wheel of the mobile robotic platform, which allows the robot to adapt to complex trajectories and turns with different angles for each steering wheel.

Thus, the proposed model allows you to control the movement of a mobile platform with independent control of each wheel, to ensure high maneuverability in difficult operating conditions.

Description of the research results

Let us consider separately the circuit responsible for precise control of the wheel rotation angle (Fig. 16).

This diagram corresponds to the manufactured mock-up sample of the swivel wheel module, which is shown in Fig. 17.

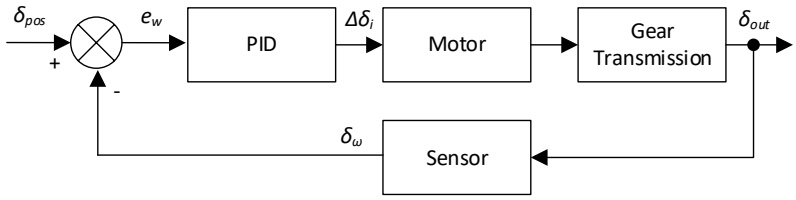


Fig. 16. Block diagram of the wheel steering angle control circuit

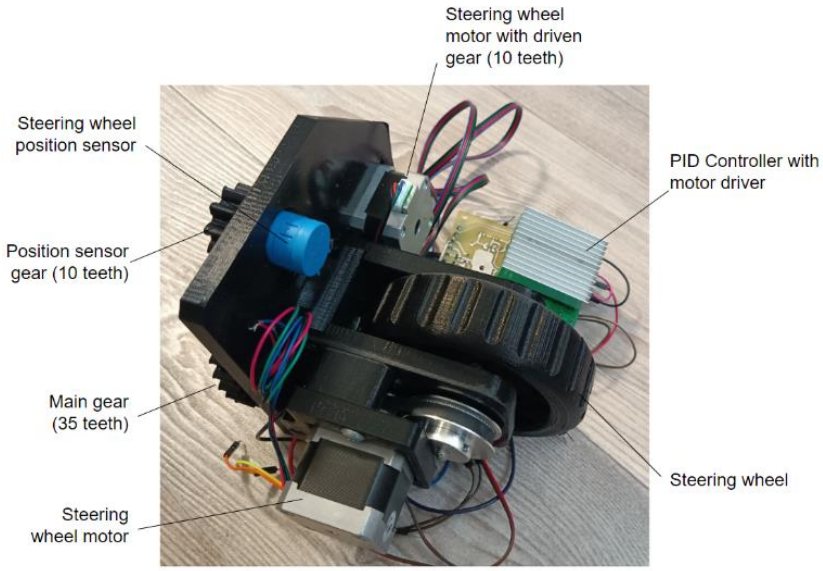


Fig. 17. Appearance of the mock-up sample of the swivel wheel module

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

The parameters of the components of the layout are as follows:

– number of teeth of the main gear – 35 pcs.;

– number of teeth of the driven gear – 10 pcs.;

– gain of the positioning sensor based on a variable resistor – 5;

– type of stepper motor – NEMA17;

– rated current of the stepper motor – 2 A;

– torque – 40 N·cm.

Let us describe each component of the structural diagram separately. The PID controller has three main components:

– proportional component (P) provides a proportional response to the error;

– integral component (I) provides correction of the constant error by integrating the error over time;

– differential component (D) takes into account the rate of change of the error, which allows to increase the stability of the system.

The transfer function of the PID controller is written as:

$$W_{PID}(s) = K_p + \frac{K_i}{s} + K_d s\,, \qquad (21)$$

where K_p is the proportional component coefficient; K_i is the the integral component coefficient; K_d is the differential component coefficient; s is the Laplace operator.

The transfer function of a stepper motor driving a rotary mechanism is written in the following form:

$$W_{motor}(s) = \frac{K_m}{T_m s + 1}, \qquad (22)$$

where K_m is the motor gain, which characterizes the conversion of electrical pulses into angular motion; T_m is the motor time constant, which determines the inertia of the system.

The design of the swivel wheel mechanism uses a transmission system in the form of two gears with a number of teeth of 35 and 10 – this determines the ratio between the angles of rotation of the stepper motor and the swivel wheel mechanism. If N_1 is the number of teeth on the driving gear (10 teeth), and N_2 is the number of teeth on the driven gear (35 teeth), then the gear ratio:

$$N_2 / N_1 = 35/10 = 3.5\,.$$

Thus, the transfer function of the transmission mechanism will be:

$$W_{gears}(s) = 1/3.5\,, \qquad (23)$$

A variable resistor is used to measure the angle of rotation through feedback. In this case, the transfer function of the sensor can be approximated by a linear relationship that converts the angle signal into an electrical voltage:

$$W_{sensor}(s) = K_{sensor}\,, \qquad (24)$$

where K_{sensor} is the coefficient of conversion of the angle of rotation into voltage, determined by the characteristics of the variable resistor. This coefficient is usually a constant value.

Let us substitute the determined transfer characteristics into the corresponding components of the steering wheel angle control system (Fig. 18).

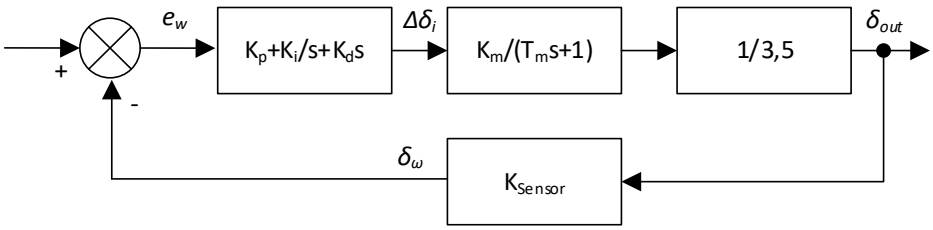


Fig. 18. Block diagram of the circuit for automatic steering angle adjustment with defined transfer functions

To synthesize the overall transfer function of the system, we need to combine the transfer functions of each element. The overall transfer function will look like this:

$$W_{sw}(s) = W_{PID}(s) + W_{motor}(s) + W_{gears}(s) + W_{sensor}(s)\,. \qquad (25)$$

We substitute the transfer functions of each element:

$$W_{sw}(s) = \left(K_p + \frac{K_i}{s} + K_d s\right) \cdot \frac{K_m}{T_m s + 1} \cdot \frac{1}{3.5} K_{sensor}\,. \qquad (26)$$

Simplifying this equation, we get:

$$W_{sw}(s) = \frac{K_m K_{sensor}\left(K_p s + K_i + K_d s^2\right)}{3.5 s\left(T_m s + 1\right)}\,. \qquad (27)$$

Let's simulate the developed automatic steering angle adjustment circuit using the Mathlab Simulink tool.

Let's define the coefficients of the transfer characteristics. For a stepper motor, the gain K_m depends

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

on the torque M developed by the motor at the rated current $I_0 = 2A$:

$$K_m = \frac{M}{I_0} = \frac{40}{2} = 20\left(\frac{N \cdot cm}{A}\right). \quad (28)$$

The system time constant T_m is determined by the inductance characteristics of the motor windings and can be calculated by the formula:

$$T_m = L/R, \quad (29)$$

where L is the winding inductance (for the selected motor type, 5 mH); R this is the winding resistance (for the selected motor type, 2 Ohms).

Thus, for the selected type of Nema 17 motor with an inductance of 5 mH and a resistance of 2 Ohms:

$$T_m = \frac{5 \cdot 10^{-3}}{2} = 0.0025(s).$$

Thus, after substituting the coefficients in (22), we obtain:

$$W_{motor}(s) = \frac{20}{0.0025s + 1}.$$

Fig. 19 shows a circuit diagram of the automatic steering angle adjustment built in Simulink.

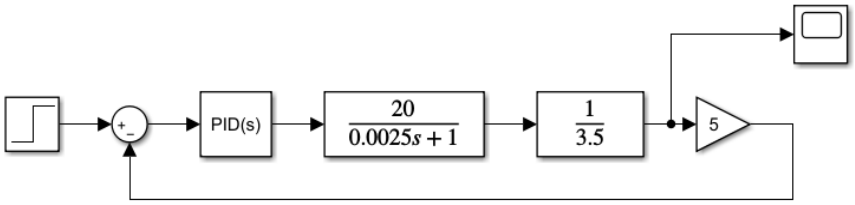


Fig. 19. Modeled circuit diagram of the automatic steering wheel angle adjustment

The PID controller parameters are as follows:

– proportional component coefficient $K_p = 2.5$;

– integral component coefficient $K_i = 25$;

– differential component coefficient $K_d = 0.0125$.

A graph of the transition process was constructed for the selected coefficients (Fig. 20).

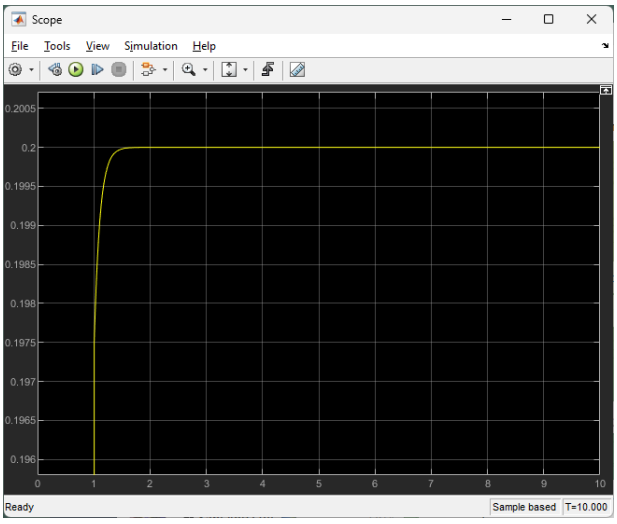


Fig. 20. Transient process graph for the modeled system

As can be seen from the graph above, the system reaches a stable state quite quickly, approximately 1–1.5 seconds after the start of the simulation. The value of the output parameter stabilizes at 0.2, which corresponds to a change in the wheel orientation by 1 degree. The gain coefficient of the rotation angle sensor is 5. It can also be seen that there is no significant excess of the set parameter, and the system quickly becomes stable. The output variable increases rapidly at the beginning of the simulation, which indicates that the system has good stability and fast response

Conclusions

This work studies a current topic related to solving the problem of increasing the degree of mobility of a mobile autonomous robot in a confined space. The principle of turning a mobile platform using four independent steering wheels is considered. A method for determining the orientation of the steering wheels depending on the trajectory of movement is proposed, which is based on the geometric analysis of the position of the platform and the target point, which allows calculating the angle of rotation of each wheel in such a way as to ensure movement to a given point without lateral slippage. A mathematical model of the control system was built, a structural and functional diagram was developed, an algorithm for processing commands, calculating the angles of rotation was described, and a three-level control system was implemented - linear speed, wheel orientation angle and angular velocity of the entire platform. The work

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

also describes the hardware implementation of the steering wheel module, which uses a stepper motor, gear transmission and a position sensor based on a variable resistor. The developed prototype of the mechatronic assembly of the steering wheels is described. The simulation conducted in the Simulink environment confirmed the operability of the proposed system: the system quickly reaches a stable state, demonstrates the absence of readjustments and high positioning accuracy. Thus, the developed control method provides high maneuverability and accurate tracking of the trajectory of the platform with four independently steered wheels, which is an important step towards the development of autonomous robotics.

References

1. Nevludov, I. et al. (2021), "Control System for Agricultural Robot Based on ROS", *IEEE International Conference on Modern Electrical and Energy Systems (MEES)*, P. 1–6, DOI: 10.1109/MEES52427.2021.9598560
2. Shi, Q., Wang, J. (2022), "Design and Realization of Mobile Robot Driven by Omnidirectional Wheels", *Journal of Physics: Conference Series*. 2402, 012035, DOI: 10.1088/1742-6596/2402/1/012035
3. Nevliudov,I., Novoselov, S., Sychova, O., Gopejenko, V., Kosenko, N. (2021), "Decentralized information systems in intelligent manufacturing management tasks", *Advanced Information Systems*, 8(3), P. 100–110. DOI: 10.20998/2522-9052.2024.3.12
4. Kuchuk, H., Husieva, Y., Novoselov, S., Lysytsia, D., Krykhovetskyi, H. (2025), "Load balancing of the layers IOT fog-cloud support network", *Advanced Information Systems*, 9(1), P. 91–98. DOI: 10.20998/2522-9052.2025.1.11
5. Bezkorovainyi, V., Binkovska, A., Noskov, V., Gopejenko, V., Kosenko, V. (2025), "Adaptation of logistics network structures in emergency situations", *Advanced Information Systems*, 9(4), P. 39–50. DOI: 10.20998/2522-9052.2025.4.06
6. Neamah, H., Akkad, H. (2023), "A Comparative Review of Omnidirectional Wheel Types for Mobile Robotics", *SSRN*, 20 p. DOI: 10.2139/ssrn.4488112
7. Artuso, P., Bocci, E. Ronci, M. (2011), "Four Independent Wheels Steering System Analysis", *SAE Technical Paper*, 01-0241, DOI: 10.4271/2011-01-0241
8. Nevliudov, I., Novoselov, S., Sychova, O. Tesliuk, S. (2021), "Development of the Architecture of the Base Platform Agricultural Robot for Determining the Trajectory Using the Method of Visual Odometry", *IEEE XVIIth International Conference on the Perspective Technologies and Methods in MEMS Design (MEMSTECH)*, Ukraine, P. 64–68, DOI: 10.1109/MEMSTECH53091.2021.9468008
9. Owen, B. et al. (2014), "A lightweight, modular robotic vehicle for the sustainable intensification of agriculture", *IEEE International Conference on Robotics and Automation*, P. 1–9, available at: https://eprints.qut.edu.au/82219/
10. Nevliudov, I., Yevsieiev, V., Maksymova, S., Gopejenko, V., Kosenko, V. (2025), "Development of mathematical support for adaptive control for the intelligent gripper of the collaborative robot manipulator", *Advanced Information Systems*, 9(3), P. 57–65. DOI: 10.20998/2522-9052.2025.3.07
11. Singh, R., Singh, G., Kumar, V. (2020), "Control of closed-loop differential drive mobile robot using forward and reverse Kinematics", *Third International Conference on Smart Systems and Inventive Technology (ICSSIT)*, Tirunelveli, India, P. 430–433, DOI: 10.1109/ICSSIT48917.2020.9214176
12. Zhao, Y., Feng, F., Zhang, R. (2015), "Development of a four wheel Independent Drive and Four Wheel Independent Steer Electric Vehicle", *Sixth International Conference on Intelligent Systems Design and Engineering Applications (ISDEA)*, Guiyang, China, P. 319–322, DOI: 10.1109/ISDEA.2015.85
13. Ge, P., Guo, L., Chen, J. (2021), "Electronic Differential Control for Distributed Electric Vehicles Based on Optimum Ackermann Steering Model", *5th CAA International Conference on Vehicular Control and Intelligence (CVCI)*, Tianjin, China, P. 1–6, DOI: 10.1109/CVCI54083.2021.9661256
14. Bae, J., Lee, D.-H. (2024), "Circular Path Tracking Control Scheme of the Self-Driving Robot Driven by Two BLDC Motors", *IEEE Symposium on Industrial Electronics & Applications (ISIEA)*, Kuala Lumpur, Malaysia, P. 1–5, DOI: 10.1109/ISIEA61920.2024.10607318
15. Deremetz, M. et al. (2017), "Path tracking of a four-wheel steering mobile robot: A robust off-road parallel steering strategy", *ECMR The European Conference on Mobile Robotics*, Paris, France. P. 62–68, available at: https://hal.inrae.fr/hal-02607237/file/pub00057050.pdf
16. Aref, M. M., Ghabcheloo, R., Mattila, J. (2013), "Mechatronic Design of a four wheel Steering Mobile Robot with Fault-Tolerant Odometry Feedback", *6th IFAC Symposium on Mechatronic Systems the International Federation of Automatic Control*, Hangzhou, China, 46(5), P. 663–669. DOI: 10.3182/20130410-3-CN-2034.00092

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

17. Cho, Y., Noh, K., Lee, E. Jeong, H. (2025), "Introduction to a motion control algorithm for the autonomous 4-wheel steering vehicle", *IEEE/SICE International Symposium on System Integration (SII)*, Munich, Germany, P. 1090–1094, DOI: 10.1109/SII59315.2025.10871133
18. Al-saedi, Mazin, I., Hiba Mohsin Abd Ali AL-bawi (2024), "Towards Sustainable Transportation: Adaptive Trajectory Tracking Control Strategies of a Four-Wheel-Steering Autonomous Vehicle for Improved Stability and Efficacy", *Processes 12*, No. 11: 2401. DOI: 10.3390/pr12112401
19. Tourajizadeh, H., Sarvari, M. Afshari, S. (2024), "Path planning and optimal control of a 4WS vehicle using calculus f variations", *Acta Mech. Sin*. 40(2), 523217, DOI: 10.1007/s10409-023-23217-x
20. Liu, R., Wei, M., Sang, N., Wei, J. (2021), "A trajectory-tracking controller for improving the safety and stability of four-wheel steering autonomous vehicles", *Transport*, 36(2), P. 1–17, DOI: 10.3846/transport.2021.14291
21. "SDS MK4 Swerve Modules. Educational Robotic Parts. AndyMark, Inc." available at: https://www.andymark.com/products/mk4-swerve-modules?via=Z2lkOi8vYW5keW1hcmsvV29ya2FyZWE6OkNhdGFsb2c6OkNhdGVnb3J5LzVhZjhkN2Y1YmM2ZjZkNWUzNmYyMzRkOA (date of access: 16.03.2025).
22. "SDS MK4i Swerve Modules. Educational Robotic Parts. AndyMark, Inc." available at: https://www.andymark.com/products/mk4i-swerve-modules (date of access: 16.03.2025).

Received (Надійшла) 01.10.2025
Accepted for publication (Прийнята до друку) 30.11.2025
Publication date (Дата публікації) 28.12.2025

Відомості про авторів / About the Authors

Nevlyudov Igor – Doctor of Sciences (Engineering), Professor, Kharkiv National University of Radio Electronics, Head at the Department of Computer-Integrated Technologies, Automation and Robotics, Kharkiv, Ukraine; e-mail: igor.nevliudov@nure.ua; ORCID ID: https://orcid.org/0000-0002-9837-2309; Scopus ID: 57216434058

Novoselov Sergiy – PhD (Engineering Sciences), Associate Professor, Kharkiv National University of Radio Electronics, Professor at the Department of Computer Integrated Technologies, Automation and Robotics, Kharkiv, Ukraine; e-mail: sergiy.novoselov@nure.ua; ORCID ID: https://orcid.org/0000-0002-3190-0592; Scopus ID: 57201604404

Sychova Oksana – PhD (Engineering Sciences), Kharkiv National University of Radio Electronics, Associate Professor at the Department of Computer-Integrated Technologies, Automation and Robotics, Kharkiv, Ukraine; e-mail: oksana.sychova@nure.ua; ORCID ID: https://orcid.org/0000-0002-0651-557X; Scopus ID: 16647283500

Zygin Sergii – Kharkiv National University of Radio Electronics, Graduate student Department of Design Automation, Kharkiv, Ukraine; e-mail: sergey.zygin@gmail.com; ORCID ID: https://orcid.org/0009-0003-3537-7713; Scopus ID: 59937709800

Невлюдов Ігор Шакирович – доктор технічних наук, професор, Харківський національний університет радіоелектроніки, завідувач кафедри комп'ютерно-інтегрованих технологій, автоматизації та робототехніки, Харків, Україна.

Новоселов Сергій Павлович – кандидат технічних наук, доцент, Харківський національний університет радіоелектроніки, професор кафедри комп'ютерно-інтегрованих технологій, автоматизації та робототехніки, Харків, Україна.

Сичова Оксана Володимирівна – кандидат технічних наук, Харківський національний університет радіоелектроніки, доцент кафедри комп'ютерно-інтегрованих технологій, автоматизації та робототехніки, Харків, Україна.

Зигін Сергій Євгенович – Харківський національний університет радіоелектроніки, аспірант кафедри автоматизації проєктування обчислювальної техніки, Харків, Україна.

РОЗРОБЛЕННЯ МЕТОДУ УПРАВЛІННЯ МОБІЛЬНОЮ ПЛАТФОРМОЮ З ЧОТИРМА КЕРОВАНИМИ КОЛЕСАМИ

Предметом вивчення є метод визначення траєкторії руху робота з чотирма керованими колесами для досягнення заданої точки на карті місцевості. **Мета дослідження** – розробити метод визначення орієнтації коліс відповідно до траєкторії руху мобільної платформи для підвищення маневреності автономного роботизованого транспортного засобу в обмеженому виробничому просторі. **Завдання**, які необхідно виконати: проаналізувати аналогічні рішення;

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) *Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)*

описати запропоновану конструкцію механізму рульового керування мобільного роботизованого візка; подати кінематику мобільного робота з чотирма керованими колесами; розробити алгоритм модуля управління рульовим керуванням; запропонувати метод управління мобільною платформою з чотирма керованими колесами; експериментально дослідити ефективність запропонованого методу. **Наукова новизна**: запропоновано метод визначення орієнтації коліс для досягнення заданої точки на плані місцевості; розроблено алгоритм розрахунків за допомогою програмного засобу; математично обґрунтовано метод керування окремими колісними блоками мобільної платформи. **Методи дослідження**: моделювання й теорія автоматичного керування, опис лінійних динамічних систем, аналітичні методи моделювання, комп'ютерне моделювання в середовищі *Matlab*/*Simulink*. **Досягнуті результати й висновки**. Розглянуто принцип руху мобільної платформи за допомогою чотирьох незалежних керованих коліс. Запропоновано метод визначення орієнтації керованих коліс відповідно до траєкторії руху, основаного на геометричному аналізі положення платформи й цільової точки, що дає змогу розрахувати кут повороту кожного колеса таким чином, щоб забезпечити рух до заданої точки без бокового прослизання. Побудовано математичну модель системи керування, розроблено структурно-функціональну схему, описано алгоритм оброблення команд, обчислення кутів повороту, а також реалізовано трирівневу систему керування: лінійною швидкістю, кутом орієнтації коліс і кутовою швидкістю всієї платформи. Описано розроблений макетний зразок мехатронного вузла керма. Моделювання, проведене в середовищі *Simulink*, підтвердило працездатність запропонованої системи.

Ключові слова: мобільний робот; AGV; інтелектуальне виробництво; метод керування; автоматизована система.

Бібліографічні описи / Bibliographic descriptions

Невлюдов І. Ш., Новоселов С. П., Сичова О. В., Зигін С. Є. Розроблення методу управління мобільною платформою з чотирма керованими колесами. *Сучасний стан наукових досліджень та технологій в промисловості*. 2025. № 4 (34). С. 135–150. DOI: https://doi.org/10.30837/2522-9818.2025.4.135

Nevlyudov, I., Novoselov, S., Sychova, O., Zygin, S. (2025), "The method development for controlling the mobile platform with four steering wheels", *Innovative Technologies and Scientific Solutions for Industries*, No. 4 (34), P. 135–150. DOI: https://doi.org/10.30837/2522-9818.2025.4.135

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

УДК 621.396
DOI: https://doi.org/10.30837/2522-9818.2025.4.151

О. Семенова, А. Джус, В. Мартинюк

ЗАСТОСУВАННЯ ТЕХНОЛОГІЙ ОБЧИСЛЮВАНОГО ІНТЕЛЕКТУ В ЗАДАЧАХ КЛАСТЕРИЗАЦІЇ БЕЗДРОТОВИХ СЕНСОРНИХ МЕРЕЖ

Предметом дослідження є процес вибору основного вузла кластера в бездротових сенсорних мережах (БСМ) із використанням інтелектуальних підходів, здатних адаптуватися до мінливих умов середовища. БСМ містять значну кількість сенсорних вузлів, що збирають, обробляють і передають інформацію. Ефективна кластеризація є одним із важливих механізмів оптимізації роботи БСМ, адже дає змогу зменшити енергоспоживання, підвищити надійність і масштабованість мережі. **Метою роботи** є аналіз особливостей застосування сучасних інструментів і методів обчислювального інтелекту для підвищення ефективності процесу кластеризації сенсорних вузлів, що допомагає брати до уваги множину факторів у прийнятті рішень про формування кластерів і обрання основних вузлів. Традиційні алгоритми кластеризації не завжди здатні адаптуватися до змін у параметрах мережі, особливо за наявності неоднорідних вузлів або змін у топології. У зв'язку з цим усе більшої актуальності набувають методи, основані на обчислюваному інтелекті, зокрема генетичні алгоритми, нейронні мережі, нечітка логіка, а також гібридні підходи. Ці методи дають змогу зважати на низку параметрів у процесі формування кластерів та вибору їх основних вузлів. **Завдання дослідження**: аналіз наявних підходів до кластеризації в БСМ; розроблення нечіткої системи кластеризації; побудова бази правил для прийняття оптимальних рішень; експериментальна перевірка запропонованої моделі. **Застосовані методи**: інструменти обчислювального інтелекту, зокрема нейромережеве навчання, генетична оптимізація та нечітке управління, комп'ютерне моделювання. У статті проаналізовано переваги використання кожного з наявних підходів. **Результати**: розроблено структуру нечіткої системи логічного висновку; визначено вхідні та вихідні змінні; сформовано базу нечітких правил і функцій належності; змодельовано роботу системи в середовищі MATLAB; оптимізовано розроблену систему й валідацію її роботи. **Висновки**: застосування гібридних інтелектуальних підходів має суттєві переваги для розв'язання задач кластеризації в БСМ, що може свідчити про перспективи подальшого розвитку систем, здатних ефективно функціювати в умовах обмежених ресурсів і високої складності середовища.

Ключові слова: бездротова сенсорна мережа; кластеризація; нечітка логіка; нейронна мережа; генетичний алгоритм.

Вступ

Наразі актуальною відповіддю на потреби в модернізації промисловості в умовах глобальної цифрової трансформації є впровадження технологій Індустрії 4.0. Вони забезпечують інтеграцію інформаційних технологій, Інтернету речей та систем автоматизації для створення інтелектуальних виробничих комплексів. Упровадження концепції Індустрії 4.0 дає змогу підприємствам залишатися конкурентоспроможними, підвищувати продуктивність і швидко адаптуватися до змін ринку й запитів споживачів.

Важливим технологічним складником Індустрії 4.0 є бездротові сенсорні мережі (БСМ), що забезпечують безперервний моніторинг фізичних процесів у режимі реального часу. Завдяки здатності до самоорганізації та автономної роботи БСМ сприяють цифровізації виробничих систем і підвищенню їх гнучкості. Інтеграція бездротових сенсорних мереж з кіберфізичними системами допомагає оперативно реагувати на зміни в середовищі, оптимізуючи витрати енергії, матеріалів і часу. Збір і передача показників сенсорами формують інформаційну основу для аналітики, прогнозування та адаптивного управління процесами. Отже, БСМ відіграють визначальну роль у створенні "розумних" фабрик, сприяючи підвищенню ефективності, безпеки та інноваційності промислового виробництва.

БСМ – це самоконфігуровані мережі, що містять значну кількість сенсорних вузлів, призначених для моніторингу фізичних умов, таких як температура, вологість, тиск, рух тощо [1]. Зазвичай цим вузлам властиві обмежені обчислювані та енергетичні ресурси, що робить проблему енергозбереження однією з найважливіших для забезпечення надійного функціювання БСМ. Через особливості структури бездротових сенсорних мереж, зокрема високу щільність вузлів, обмежений радіус дії та функціювання в складних або важкодоступних середовищах, часто виникає потреба у високоефективних механізмах організації структури мережі. Одним із таких механізмів є кластеризація. Це процес розподілу вузлів на окремі групи – кластери, – у кожній з яких

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)
Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

один сенсорний вузол призначається основним вузлом кластера (*cluster head*, CH) та координує комунікацію вузлів як всередині кластера між собою, так і з базовою станцією. Цей вузол відповідає за збір інформації від вузлів-членів свого кластера, їх агрегацію (об'єднання та оброблення) й передачу агрегованих даних до базової станції. Такий підхід дає змогу зменшити кількість передач на значні відстані, скоротити загальне навантаження на мережу, продовжити тривалість роботи вузлів, підвищити масштабованість мережі [2].

Практична реалізація ефективної кластеризації пов'язана з технічними складнощами. Так, динамічні зміни мережевої топології, зумовлені мобільністю або ж відмовами вузлів, ускладнюють підтримання стабільної кластерної структури. Крім того, виникають труднощі, пов'язані з ефективною маршрутизацією даних у кластерах, особливо в умовах шуму й перешкод. Також обмежені обчислювальні можливості сенсорних вузлів створюють додаткові перешкоди для реалізації складних алгоритмів кластеризації. Ще одним питанням є забезпечення енергоефективності, що пов'язано з обмеженими енергетичними ресурсами сенсорних вузлів. Тому сучасні алгоритми кластеризації мають мінімізувати витрати енергії на передачу інформації та оброблення сигналів. Крім того, особливу увагу варто приділяти правильному вибору основних вузлів кластерів, адже саме вони виконують додаткові функції агрегації та ретрансляції даних, що призводить до швидшого вичерпання їх енергії. Також надзвичайно важливо забезпечити баланс навантаження між основними вузлами для запобігання їх передчасному виходу з ладу [3–5].

Аналіз останніх досліджень і публікацій

Для початку розглянемо методи кластеризації [6]. Так, LEACH (*Low-Energy Adaptive Clustering Hierarchy*) є одним із перших і найвідоміших протоколів кластеризації. У ньому використовується ймовірнісний механізм вибору основних вузлів кластера. Ключовою його перевагою є простота реалізації, однак є і недолік – відсутність урахування залишкової енергії вузлів. Протокол TEEN (*Threshold-sensitive Energy Efficient sensor Network protocol*) використовує порогові значення для ініціації передачі даних. Він підходить для застосунків із високою чутливістю до змін, однак може втратити інформацію, якщо певна подія не перевищує поріг. У протоколі HEED (*Hybrid Energy-Efficient Distributed Clustering*) у процесі кластеризації береться до уваги початкова й залишкова енергія. Він забезпечує більш збалансоване енергоспоживання, проте має вищу складність обчислень. Протокол DEEC (*Distributed Energy-Efficient Clustering*) під час кластеризації зважає на початкову й залишкову енергію, забезпечуючи цим кращу стабільність у мережах з неоднорідною енергією. Він є ефективним у довгострокових сценаріях. Протокол PEGASIS (*Power-Efficient GAthering in Sensor Information System*) не кластеризує сенсорні вузли, а формує ланцюг послідовних передач. Він є більш енергоефективним, хоча й має вищу затримку.

Водночас у наявних методах кластеризації не беруться до уваги складні умови розгортання БСМ, до яких належать нерівномірне споживання енергії, мобільність вузлів, перешкоди під час передачі сигналу та інші фактори. Хоча таким методам властиві проста структура й швидка реалізація, вони обмежені в адаптації до змін середовища й не зважають на невизначеність або складні залежності між вузлами [7].

З огляду на ці обмеження значну увагу дослідників привертає застосування технологій обчислюваного інтелекту (англ. *Computational Intelligence*), які охоплюють такі підходи, як штучні нейронні мережі, генетичні алгоритми (ГА), нечіткі системи, алгоритми ройового інтелекту та інші еволюційні обчислення. Ці методи мають здатність до самонавчання, адаптації, оптимізації та ефективного оброблення нечітких, неповних або змінних даних, що робить їх особливо придатними для складних розподілених систем, таких як БСМ. Їх застосування дає змогу здійснювати динамічну кластеризацію, що бере до уваги такі параметри й властивості, як залишкова енергія вузлів, топологічна щільність, пріоритети передачі, попередні характеристики трафіку, збої в роботі вузлів або втрата вузлів, особливості навколишнього середовища.

Отже, підвищити ефективність розв'язання задачі кластеризації БСМ запропоновано за допомогою застосування технологій обчислювального інтелекту, а саме нечіткої логіки, штучних нейронних мереж і генетичних алгоритмів.

Загалом системи прийняття рішення на основі нечіткої логіки дають змогу моделювати ситуації, коли дані є нечіткими, неточними або частково визначеними. У разі кластеризації БСМ нечіткі системи логічного висновку використовуються для

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

оцінювання ступеня придатності деякого сенсорного вузла мережі до виконання ролі основного вузла кластера за певною множиною критеріїв [8]. Це забезпечує гнучкість у прийнятті рішень у складних умовах без потреби в створенні точних математичних моделей. Так, у роботі [9] запропоновано протокол, що поєднує нечітку кластеризацію *k*-середніх з модифікованим протоколом LEACH, до того ж результати моделювання свідчать про значне подовження часу життя мережі порівняно з класичним методом. Ще один енергоефективний розподілений алгоритм кластеризації на основі нечіткої логіки, використання якого дало змогу подовжити життєвий цикл мережі, запропоновано в праці [10].

У дослідженні [11] розглянуто модель, яка поєднує нечітку кластеризацію *k*-середніх із нейронною мережею, що забезпечує вищу енергоефективність завдяки навчанню нейронної мережі на визначених параметрах вузлів.

Окрім того, штучні нейронні мережі часто використовуються для моделювання складних залежностей між вхідними параметрами. У задачі кластеризації нейронні мережі можуть бути застосовані для виявлення патернів з метою більш ефективного формування кластерів. Оскільки вони здатні самонавчатися на основі статистичних показників, це може сприяти покращенню ефективності прийнятих рішень [12].

Завдяки своїй еволюційній природі генетичні алгоритми здатні знаходити наближено оптимальні рішення для задач кластеризації навіть у великих мережах [13]. Так, у праці [14] запропоновано оптимізований метод вибору основного вузла кластера на підставі генетичного алгоритму для подовження строку роботи гетерогенної бездротової сенсорної мережі.

Також у роботі [15] показано, що використання нейронної мережі та еволюційних алгоритмів у процесі кластеризації дає змогу підвищити ефективність передачі даних у БСМ.

Результати порівняння розглянутих підходів до кластеризації подано в табл. 1.

Таблиця 1. *Порівняння технологій і методів кластеризації*

Критерій	Традиційні методи	Нейронні мережі	Нечітка логіка	Генетичні алгоритми
Обчислювальна складність	низька	висока	середня	висока
Адаптивність до змін мережі	низька	висока	висока	висока
Потреба в навчальних даних	немає	висока	немає	помірна
Робота в умовах невизначеності	обмежена	середня	висока	середня
Можливість оптимізації параметрів	обмежена	висока	середня	висока
Придатність до режиму реального часу	висока	обмежена	середня	обмежена
Якість кластеризації	помірна	висока в умовах якісного навчання	висока	висока
Складність реалізації	низька	середня або висока	середня	висока

Отже, традиційні методи кластеризації є доцільними для простих мереж з обмеженими ресурсами, тоді як нечіткі системи, нейронні мережі та генетичні алгоритми мають значний потенціал для складних мереж. Методи на основі нечіткої логіки забезпечують баланс між гнучкістю та обчислювальною ефективністю, особливо в умовах нечітких або неповних даних.

Розроблення нечіткої системи

З огляду на обмеження окремих методів обчислювального інтелекту оптимальним рішенням для кластеризації в бездротових сенсорних мережах є гібридизація кількох підходів. Об’єднання генетичного алгоритму, нечіткого контролера та штучної нейронної мережі дає змогу ефективно поєднати їх властивості, а саме здатність генетичних алгоритмів до глобальної оптимізації, гнучкість нечітких контролерів, адаптивність і здатність до навчання нейронних мереж. Така гібридизація сприяє компенсації недоліків окремих технологій і досягненню більш високих показників ефективності в умовах динамічного середовища функціювання БСМ.

Отже, у межах цього дослідження розроблено структуру нечіткої системи логічного висновку для оцінювання придатності сенсорних вузлів до виконання ролі основного вузла кластера. Запропонована система бере до уваги ключові параметри сенсорних вузлів. Використання нечіткої логіки допомагає приймати рішення на основі неповної або неточної інформації, що властиво для умов функціювання бездротових сенсорних мереж.

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

На відміну від класичних підходів, що ґрунтуються на жорстких порогових значеннях параметрів, нечітка система дає змогу зважати на ступінь належності сенсорного вузла до певних категорій за показниками, такими як рівень залишкової енергії, щільність сусідніх вузлів, відстань до базової станції, навантаження на вузол, показник довіри тощо. Такий підхід забезпечує гнучкість, точність і адаптивність процесу вибору основного вузла.

Пропонована в межах цього дослідження процедура кластеризації в бездротових сенсорних мережах із використанням нечіткої логіки передбачає здійснення логічного висновку з огляду на три вхідні параметри, кожен з яких буде перетворено у відповідну нечітку множину за допомогою функцій належності, що відтворюють різні рівні значень параметра (наприклад, низький, середній, високий рівень). Такий підхід допомагає коректно описати реальні дані, що надходять від сенсорів і можуть бути неповними або неточними через складні умови середовища.

Подальший етап передбачає розроблення та застосування правил нечіткого логічного висновку, сформованих на основі евристичних співвідношень між параметрами. Ці правила дають змогу оцінити доцільність вибору конкретного вузла як основного вузла кластера, зважаючи на взаємозалежність вхідних параметрів. Наприклад, вузол із високим рівнем енергії, незначною відстанню до базової станції та середньою щільністю сусідів має вищу ймовірність бути обраним як основний порівняно з іншими вузлами.

Отже, формування кластерів здійснюватиметься на основі отриманих вихідних значень нечіткої системи логічного висновку (оцінок) способом об’єднання сенсорних вузлів навколо основних, які мають максимальний рівень ймовірності бути обраними. Такий підхід сприяє більш збалансованому розподілу навантаження та мінімізації енергетичних витрат на передачу даних усередині кластера.

Після етапу кластеризації структура БСМ періодично оновлюється, що забезпечує адаптацію до змін середовища, наприклад, вичерпання енергоресурсів або поява нових вузлів. Завдяки цьому підтримується енергетична ефективність, стійкість і тривалість життєвого циклу мережі.

Як вхідні величини нечіткої системи логічного висновку запропоновано використання таких параметрів БСМ: залишкова енергія вузла, щільність сусідніх вузлів і відстань до базової станції. Вхідні величини було обрано з огляду на такі міркування.

Залишкова енергія вузла в БСМ – це кількість енергетичних ресурсів (наприклад, заряд батареї), що залишилися в сенсорному вузлі після виконання ним комунікаційних і обчислювальних операцій, і яка визначає подальшу спроможність вузла брати участь у мережевій активності. Рівень залишкової енергії вузла є важливим фактором у процесі кластеризації, оскільки вибір вузлів з більш високим енергетичним резервом як основних дає змогу збалансувати енергоспоживання в мережі, подовжити її життєвий цикл та зменшити ймовірність передчасного виходу з ладу окремих компонентів. Вузли з недостатнім рівнем енергії не здатні ефективно виконувати функції основних, адже це може призвести до зниження надійності мережі. У багатьох алгоритмах кластеризації залишкова енергія є одним із критеріїв прийняття рішення. Для цієї вхідної величини лінгвістичні множини визначено відповідно до виразу

$$E \in \{\textit{Низька}, \textit{Середня}, \textit{Висока}\}. \quad (1)$$

Щільність сусідніх вузлів у БСМ відповідає кількості сенсорних вузлів у межах зони прямого радіозв’язку певного вузла, визначаючи таким чином локальну насиченість мережевої топології. Щільність сусідніх вузлів також впливає на процес кластеризації, оскільки вузли з високою щільністю мають більший потенціал для ефективного управління кластером, забезпечуючи зменшення кількості міжкластерних з’єднань та зниження енергетичних витрат. Водночас надмірна щільність може спричинити перевантаження окремих основних вузлів, що вимагатиме балансування навантаження між кластерами. Для цієї вхідної величини лінгвістичні множини визначено за виразом

$$N \in \{\textit{Мала}, \textit{Середня}, \textit{Велика}\}. \quad (2)$$

Відстань від вузла до базової станції у БСМ – це просторове розташування сенсорного вузла щодо центрального приймального пункту, що безпосередньо впливає на енергетичні витрати під час передачі даних. Відстань від вузла до базової станції також суттєво впливає на кластеризацію, оскільки вузли, розташовані ближче до базової станції, потребують менше енергії для передачі даних, що робить їх більш придатними для виконання ролі основних. Водночас надмірна концентрація таких вузлів у центрі мережі може призвести до дисбалансу навантаження, тому ефективні алгоритми кластеризації мають брати

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

до уваги не лише відстань, а й розподіл вузлів у просторі. Для цієї вхідної величини лінгвістичні множини визначено відповідно до виразу

$$D \in \{\textit{Близька}, \textit{Середня}, \textit{Далека}\}. \quad (3)$$

Для всіх трьох вхідних величин функції належності є трапецієподібні й розраховуються так:

$$\mu(x) = \begin{cases} 0 \ \textit{якщо}\ x \le p_1 \ x > q_2, \\ \dfrac{x - p_1}{p_2 - p_1} \ \textit{якщо}\ p_1 < x \le p_2, \\ 1 \ \textit{якщо}\ p_2 < x \le q_1, \\ \dfrac{q_2 - x}{q_2 - q_1} \ \textit{якщо}\ q_1 < x \le q_2, \end{cases} \quad (4)$$

де $\mu(x)$ – функція належності для вхідної змінної; x – значення вхідної змінної; p_1 – нижня межа, де функція належності починає зростати від 0; p_2 – нижня межа, де функція належності досягає 1; q_1 – верхня межа, де функція належності все ще рівна 1; q_2 – верхня межа, після якої функція належності спадає до 0.

Вибір саме цього типу функцій зумовлений їх простотою реалізації та здатністю адекватно відтворювати нечіткі межі між лінгвістичними термами. Трапецієподібна форма дає змогу гнучко описувати змінні, для яких існують чітко визначені ділянки "низьких" і "високих" значень, а також перехідні зони, де належність до певної категорії змінюється поступово. Такий підхід є доцільним для БСМ, де показники енергетичного стану або просторового розташування вузлів можуть варіюватися в широкому діапазоні. Крім того, трапецієподібні функції належності мають незначну кількість параметрів для налаштування, що спрощує процес їх подальшої оптимізації.

Вихідною величиною пропонованої нечіткої системи логічного висновку є ймовірність для конкретного вузла бути обраним основним вузлом кластера. Це значення відтворює ступінь придатності вузла до виконання функцій кластерного керування на основі його поточних характеристик, таких як залишкова енергія, відстань до базової станції та щільність сусідніх вузлів. Динамічне оновлення цього значення дасть змогу адаптувати процес кластеризації до змін у мережі, підвищуючи в такий спосіб її гнучкість і надійність. Для вихідної величини лінгвістичні множини визначено за виразом

$$P \in \{\textit{дуже мала}, \textit{мала}, \textit{середня}, \textit{велика}, \textit{дуже велика}\}. \quad (5)$$

Для вихідної величини в розробленій нечіткій системі функції належності задаються у вигляді констант (синглетонів). Такий вибір дає змогу надалі легко інтегрувати нечітку систему в структуру нейронечіткої системи. Отже, функції належності вихідної величини обчислено за формулою

$$\mu(y) = \begin{cases} 0 \ if\ y \ne m, \\ 1 \ if\ y = m, \end{cases} \quad (6)$$

де $\mu(y)$ – функція належності для вихідної змінної; y – значення вихідної змінної; m – еталонне значення.

Нечітка база знань є набором нечітких правил типу "якщо, тоді":

$$If\ E = A_i\ and\ N = B_j\ and\ D = C_k\ then\ P = p_{ijk}, \quad (7)$$

де A_i, B_j, C_k – відповідні нечіткі множини для кожної вхідної змінної; p_{ijk} – вихідне значення для *i, j, k*-го правила.

На вхід нечіткої системи логічного висновку подаються конкретні значення трьох вхідних величин: $E = e$, $N = n$, $D = d$. Далі для кожного нечіткого правила обчислено ступінь його активації за формулою (8) і його вихідне значення p_{ijk}.

$$w_{ijk} = \mu_{Ai}(e) \cdot \mu_{Bj}(n) \cdot \mu_{Ck}(d), \quad (8)$$

де w_{ijk} – ваговий коефіцієнт (ступінь активації) i, j, k-го правила; $\mu_{Ai}(e)$ – функція належності значення e вхідної змінної E до нечіткої множини A_i; $\mu_{Bj}(n)$ – функція належності значення n вхідної змінної N до нечіткої множини B_j; $\mu_{Ck}(d)$ – функція належності значення d вхідної змінної до D нечіткої множини C_k.

Вихідне значення, що формує нечітка система логічного висновку, обчислено за виразом

$$P = \frac{\sum_{i,j,k} w_{ijk} \cdot p_{ijk}}{\sum_{i,j,k} w_{ijk}}. \quad (9)$$

Оскільки в цьому дослідженні запропоновано гібридний підхід, то розроблена нечітка система логічного висновку оптимізується з використанням генетичного алгоритму для налаштування параметрів функцій належності. Такий підхід забезпечить наближення до глобального оптимуму в процесі вибору основних вузлів кластерів за багатьма критеріями.

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

Застосування генетичного підходу дає змогу здійснити глобальний пошук оптимальних рішень у багатовимірному просторі параметрів, уникаючи потрапляння в локальні мінімуми, що часто є проблемою традиційних методів оптимізації. Завдяки цьому забезпечується підвищення точності нечіткої системи й узгодженість між вхідними параметрами та результатами оцінювання доцільності вибору основних вузлів кластерів. Генетичний алгоритм виконує евристичний пошук способом еволюційного відбору, схрещування та мутації потенційних рішень – набором параметрів функцій належності. Критерієм оптимізації є так звана функція пристосованості. У генетичних алгоритмах розв'язки задачі оптимізації подаються у вигляді двійкових хромосом. У цьому випадку, оскільки використані функції належності є трапецієподібними, хромосома є набором параметрів трапецієподібних функцій належності й визначається так:

$$\chi = \left[a_1, b_1, c_1, d_1, a_2, b_2, c_2, d_2, ..., a_9, b_9, c_9, d_9\right] \in \mathbb{R}^{36}. \quad (10)$$

У цьому разі буде використано класичний генетичний алгоритм, що має такий вигляд [16]:

1) створюємо початкову популяцію випадкових хромосом;

2) оцінюємо кожну хромосому;

3) обираємо найкращі хромосоми;

4) схрещуємо пари хромосом-батьків, щоб створити нові хромосоми-нащадки;

5) мутуємо хромосоми;

6) замінюємо гірших, оновлюємо популяцію;

7) повторюємо кроки 2–6 протягом кількох поколінь;

8) обираємо найкращий результат – оптимальні параметри функцій належності.

На наступному етапі оптимізована нечітка система трансформується в нейронечітку систему типу ANFIS (*Adaptive Neuro-Fuzzy Inference System*), яка поєднує адаптивність і здатність до навчання нейронних мереж із гнучкістю та інтерпретованістю нечітких систем. У ANFIS параметри функцій належності та вагові коефіцієнти правил уточнюються за допомогою процедур навчання на основі змодельованих даних. Це дає змогу системі самостійно адаптуватися до змінних умов функціювання БСМ, підвищуючи точність прийняття рішень і знижуючи необхідність у ручному налаштуванні параметрів.

П'ятишарова нейромережа ANFIS на основі розробленої нечіткої системи формується за класичною структурною схемою, де кожен шар виконує певну функцію в процесі нечіткого логічного висновку та адаптивного навчання.

Перший шар виконує фазифікацію вхідних змінних, тобто перетворення чітких значень трьох вхідних параметрів у відповідні ступені належності до нечітких множин. Другий шар призначений для комбінування значень нечітких змінних і обчислення сили активації кожного правила в базі знань. Як наслідок, формується набір зважених значень, що відтворюють поточну активність правил у системі. Третій шар виконує нормалізацію отриманих значень сили активації, забезпечуючи їх приведення до відносних вагових коефіцієнтів. Четвертий шар відповідає за зважування та обчислення вихідних значень нечітких правил. На цьому етапі відбувається поєднання нормалізованих вагових коефіцієнтів із результатами лінійних комбінацій вхідних змінних, що відтворює вплив кожного правила на кінцеве значення. П'ятий шар реалізує агрегацію результатів усіх нечітких правил, утворюючи кінцеве значення вихідної змінної – оцінку придатності вузла до виконання ролі основного вузла кластера.

У процесі навчання ANFIS застосовує комбінований підхід, що поєднує метод найменших квадратів для коригування параметрів наслідків і зворотне поширення похибки для налаштування параметрів функцій належності. Завдяки цьому досягнуто високу точність апроксимації.

Комп'ютерне моделювання

Моделювання нечіткої системи логічного висновку й ANFIS-структури та оптимізацію генетичним алгоритмом реалізовано за допомогою середовища MATLAB, яке має розвинуті інструменти для моделювання нечітких систем (*Fuzzy Logic Designer*), оптимізаційних задач (*Global Optimization Toolbox*) та побудови нейронечітких моделей (ANFIS *Editor*). Такий вибір інструментарію забезпечує ефективну інтеграцію методів і візуалізацію результатів моделювання.

Загальний вигляд розробленої нечіткої системи логічного висновку, призначеної для розв'язання задач кластеризації (нечіткої системи кластеризації) в програмі MATLAB, подано на рис. 1. За основу обрано модель Сугено, яка має перевагу в обчислювальній ефективності (якщо порівнювати з моделлю Мамдані) завдяки використанню вихідних

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

функцій-констант, що спрощує етап дефазифікації. Це дає змогу простіше реалізовувати модель Сугено в системах реального часу з обмеженими ресурсами. Крім того, ця модель краще підходить для оптимізації та навчання за допомогою алгоритмів машинного навчання.

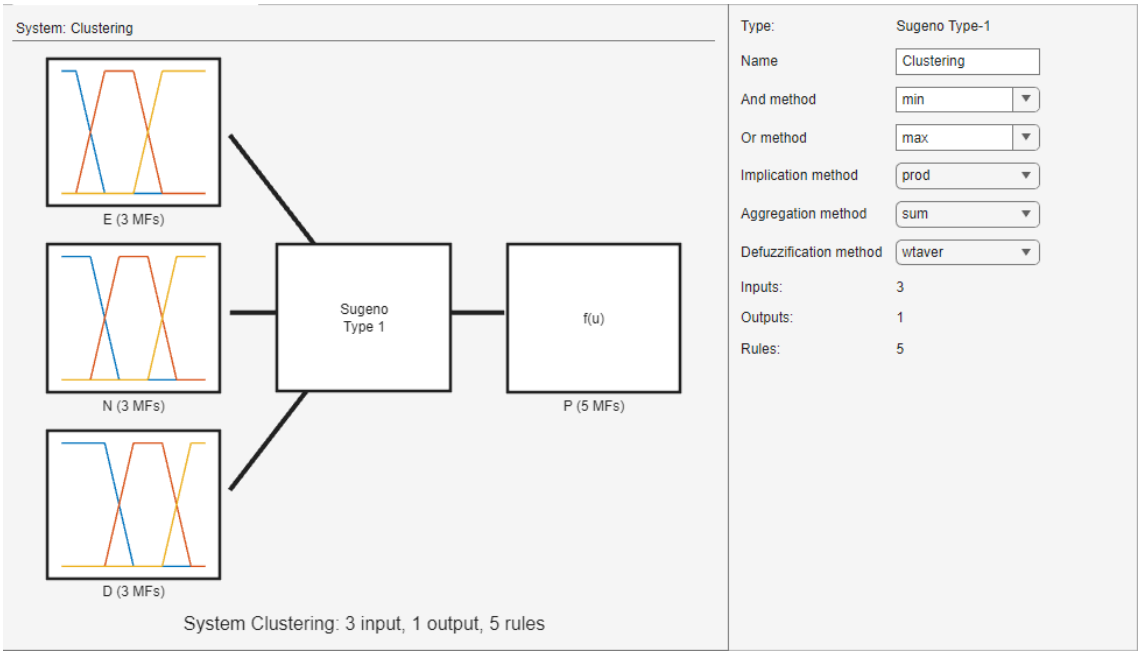
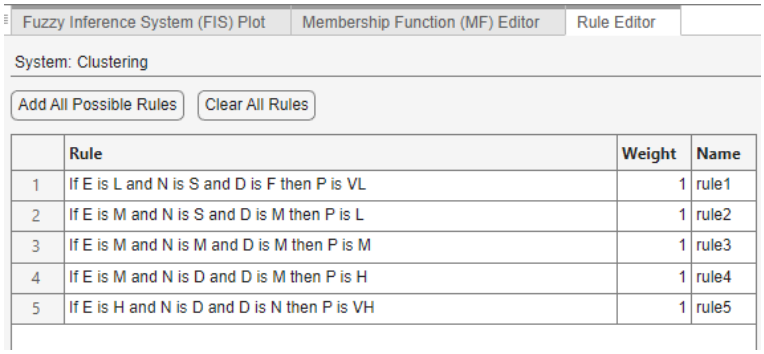


Рис. 1. Нечітка система кластеризації в програмі MATLAB

На першому етапі задано діапазони значень вхідних і вихідної змінних. Потім сформовано базу правил, яка встановлює логічні зв'язки між вхідними параметрами й вихідною реакцією системи. Нечітка система кластеризації працює на основі п'яти нечітких правил, що допомагає достатньо точно описати складні взаємодії параметрів і забезпечити ефективне функціювання моделі в межах поставленого завдання. Вигляд бази правил в інтерфейсі програми MATLAB подано на рис. 2.



	Rule	Weight	Name
1	If E is L and N is S and D is F then P is VL	1	rule1
2	If E is M and N is S and D is M then P is L	1	rule2
3	If E is M and N is M and D is M then P is M	1	rule3
4	If E is M and N is D and D is M then P is H	1	rule4
5	If E is H and N is D and D is N then P is VH	1	rule5

Рис. 2. База нечітких правил у програмі MATLAB

Моделювання роботи нечіткої системи кластеризації виконано в середовищі *Fuzzy Logic Designer* за допомогою вбудованих інструментів візуалізації та аналізу поведінки системи. Ці засоби дають змогу швидко перевірити коректність побудованої бази правил і функцій належності. Результат моделювання роботи нечіткої системи кластеризації продемонстровано на рис. 3.

Отже, якщо залишкова енергія сенсорного вузла становить 60%, щільність навколишніх вузлів – 70%, відстань від цього вузла до базової станції – 35% від максимально допустимої, тоді цей вузол може бути обраний основним вузлом кластера з імовірністю 80,7%.

Далі в розроблену систему інтегровано ще одну технологію обчислювального інтелекту – генетичний алгоритм. Результат оптимізації структури нечіткої

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

системи кластеризації за допомогою генетичного алгоритму подано на рис. 4. Оптимізація полягає в автоматичному налаштуванні параметрів нечіткої системи, зокрема форм функцій належності та вагових коефіцієнтів правил з метою мінімізації середньоквадратичної помилки.

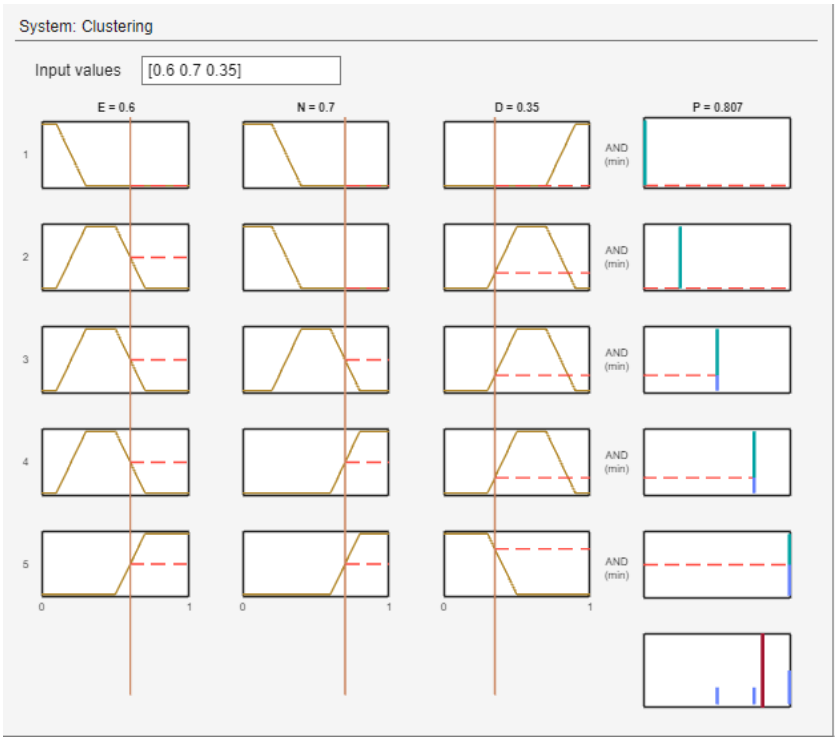


Рис. 3. Результат моделювання

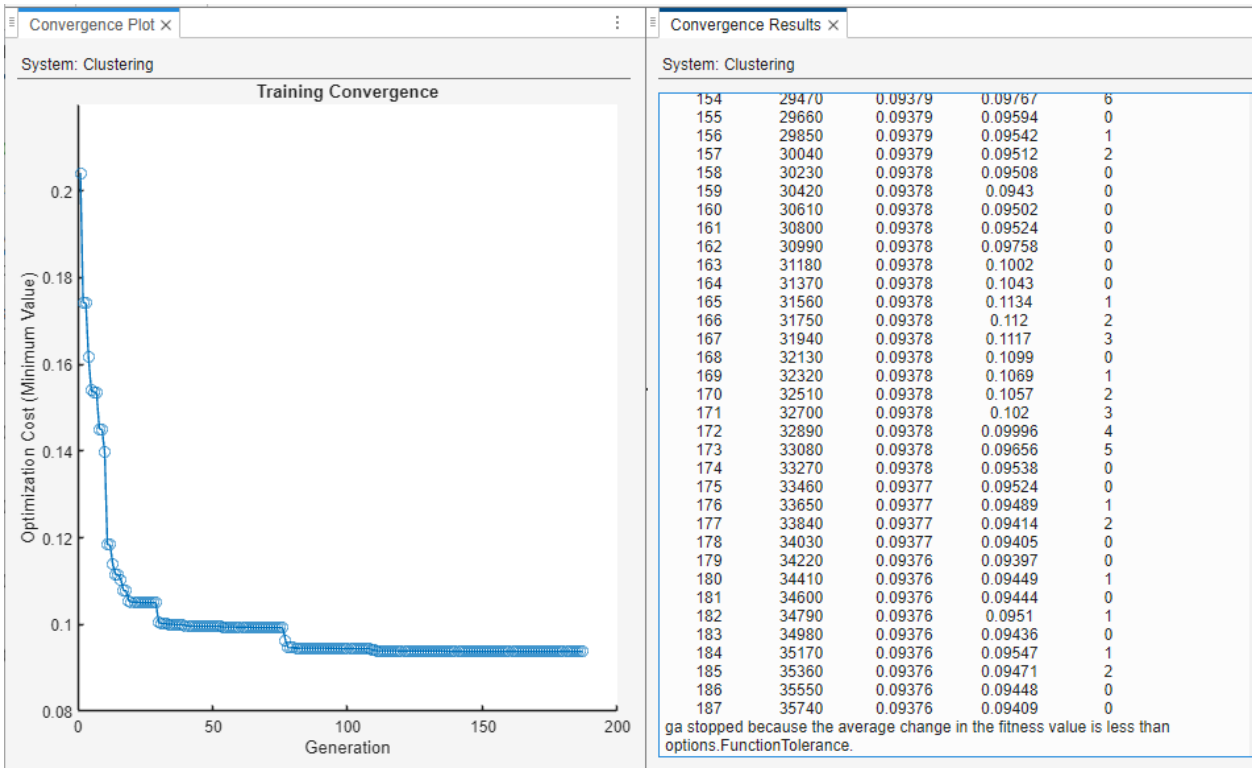


Рис. 4. Результат оптимізації

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

Далі додаємо третю технологію обчислювального інтелекту – штучну нейронну мережу. Для цього перетворюємо оптимізовану нечітку систему в адаптивну нейронечітку систему логічного висновку ANFIS. Це здійснюється з використанням інструменту ANFIS *Editor* GUI. Цей процес передбачав створення навчальної вибірки, на основі якої ANFIS адаптує параметри нечіткого виведення за допомогою нейромережевих методів. Результат перетворення нечіткої системи кластеризації в нейронечітку систему зображено на рис. 5.

Після налаштування параметрів навчання виконується тренування, унаслідок якого ANFIS автоматично коригує параметри функцій належності для забезпечення мінімальної похибки. ANFIS була навчена за методом зворотного поширення похибки. Для цього використано експериментальну вибірку вхідних і вихідних даних. Результат навчання на рис. 6 демонструє, що похибка роботи системи є допустимою, тобто система навчена правильно й може буди застосована для кластеризації в БСМ.

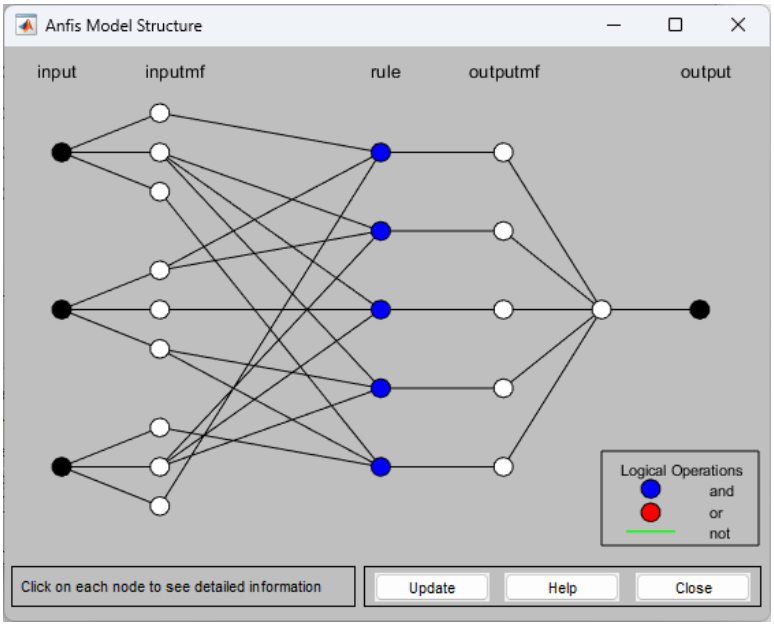


Рис. 5. Нейронечітка система кластеризації в програмі MATLAB

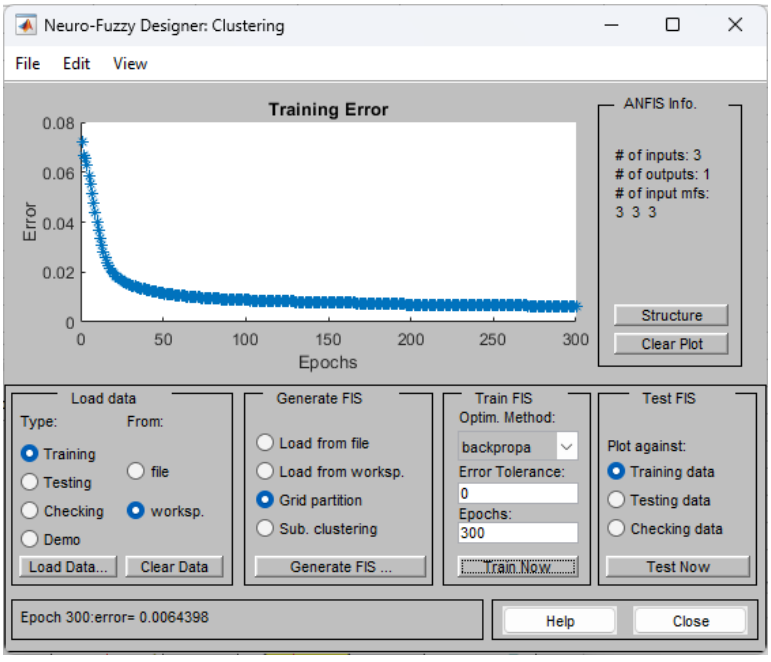


Рис. 6. Результат навчання адаптивної нейронечіткої системи кластеризації в програмі MATLAB

ISSN 2522-9818 (print)
ISSN 2524-2296 (online)
Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)

Наступний етап – валідація отриманої системи кластеризації. Валідація нечіткої системи виконується з метою оцінювання здатності системи узагальнювати знання на основі зовнішніх тестових показників. Після імпорту даних система порівнює результати з еталонною вихідною інформацією. Інтерфейс валідації надає графіки порівняння, статистику похибок, а також числові показники точності, зокрема середньоквадратичну помилку. Це дає змогу візуально й кількісно оцінити ефективність побудованої нечіткої системи логічного висновку. У разі незадовільної якості валідації можна внести зміни до бази правил, функцій належності або структури моделі та повторити перевірку. Як видно з рис. 7, значні розбіжності між еталонним і отриманими значеннями відсутні, відтак нечітка система працює правильно й не потребує додаткових налаштувань параметрів.

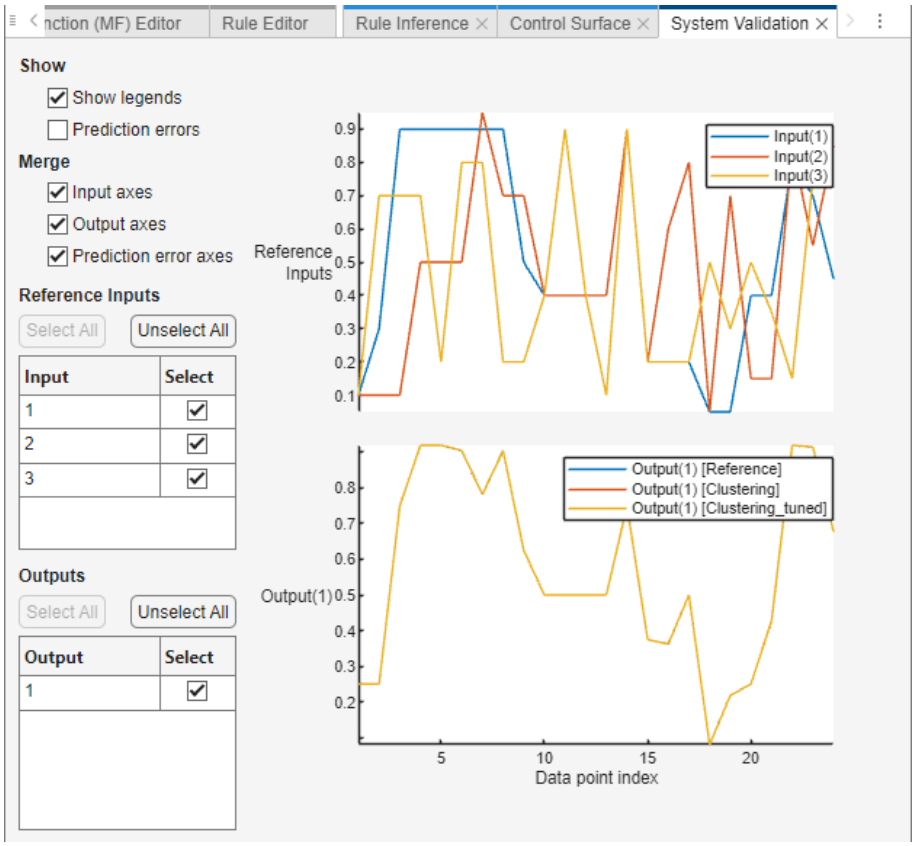


Рис. 7. Результати валідації нечіткої системи кластеризації в програмі MATLAB

Отже, результати дослідження демонструють переваги впровадження технологій обчислюваного інтелекту для кластеризації в БСМ та вказують на перспективи подальшого розвитку адаптивних систем, здатних ефективно функціювати в умовах обмежених ресурсів і високої складності середовища.

Висновки

У статті досліджено можливості й доцільність застосування технологій обчислювального інтелекту для розв'язання задачі кластеризації в бездротових сенсорних мережах. Аналіз основних питань, зокрема обмеженість енергетичних ресурсів, масштабованість, динамічні топології та балансування навантаження, дав змогу визначити ключові проблеми в побудові ефективних механізмів кластеризації.

Розглянуто три провідні технології обчислювального інтелекту: генетичний алгоритм, нечітку логіку та штучні нейронні мережі. Кожен з них продемонстрував окремі переваги. Водночас найкращим можна вважати гібридний підхід.

Аналіз довів, що гібридні підходи, які поєднують можливості генетичних алгоритмів, нечіткої логіки й нейронних мереж, забезпечують значно кращі результати порівняно з традиційними алгоритмами кластеризації.

У роботі розроблено нечітку систему логічного висновку, оптимізовану за допомогою генетичного

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

алгоритму та перетворену в нейронечітку систему. Комп'ютерне моделювання роботи нечіткої системи підтвердило правильність розрахунків.

Отримані результати створюють підґрунтя для подальших досліджень, спрямованих на розроблення самонавчальних кластеризаційних механізмів у БСМ на основі сучасних методів обчислювального інтелекту.

Отже, застосування обчислювального інтелекту, особливо у вигляді гібридних методів, є перспективним напрямом підвищення ефективності та надійності кластеризації в БСМ. Подальші дослідження будуть спрямовані на впровадження методів глибокого навчання, що дасть змогу ще краще адаптувати процес кластеризації до умов реального часу та високої мобільності вузлів.

References

1. Abdulwahid, A. and Salih, M. (2022), "Wireless sensor networks applications, challenges, and security requirements", in *Proceedings of 2nd International Multi-Disciplinary Conference Theme: Integrated Sciences and Technologies, IMDC-IST 2021*. DOI: 10.4108/eai.7-9-2021.2314823
2. Afroz, F. and Braun, R. (2020), "Energy-efficient MAC protocols for wireless sensor networks: a survey", *International journal of sensor networks*, 32(3), P. 150-173. DOI: https://doi.org/10.1504/ijsnet.2020.105563
3. Doddapaneni, K. *et al.* (2014), "Deployment challenges and developments in wireless sensor networks clustering", in *2014 28th International Conference on Advanced Information Networking and Applications Workshops*. IEEE. Victoria, BC, Canada. DOI: 10.1109/WAINA.2014.46
4. Jubair, A.M. *et al.* (2021), "Optimization of clustering in wireless sensor networks: Techniques and protocols", *Applied sciences*, 11(23), 11448 p. DOI: https://doi.org/10.3390/app112311448
5. Kalkha, H., Satori, H., Satori, K. (2017), "A dynamic clustering approach for maximizing scalability in wireless sensor networks", *Transactions on machine learning and artificial intelligence*, 5(4). DOI: https://doi.org/10.14738/tmlai.54.3328
6. Kaur, L., Kad, S. (2017), "Clustering techniques in wireless sensor network: A review", *International journal of computer applications*, 179(5), P. 30–34. DOI: https://doi.org/10.5120/ijca2017915952
7. Ali-Gburyi, K., Shah, A.F.M.S. (2023), "Performance comparison of PEGASIS, HEED and LEACH protocols in wireless sensor networks", *Celal Bayar University Journal of Science*, 19(1), P. 11–18. DOI https://doi.org/10.18466/cbayarfbe.1165816
8. Balakrishnan, B., Balachandran, S. (2017), "FLECH: Fuzzy Logic Based Energy Efficient Clustering Hierarchy for nonuniform wireless sensor networks", *Wireless communications and mobile computing*, 2017, P. 1–13. DOI: https://doi.org/10.1155/2017/1214720
9. Chit, T.A., Maung, N.A.M. and Zar, K.T. (2021), "Modified LEACH and fuzzy C-means based clustering protocol for wireless sensor networks", *International journal of advances in scientific research and engineering*, 07(01), P. 48–54. DOI: https://doi.org/10.31695/ijasre.2021.33963
10. Alkwai, L.M. and Yadav, K. (2025), "Energy-efficient clustering algorithm using distributed fuzzy-logic to prolong the survivability of wireless sensor networks", *Internet technology letters*, 8(2). DOI: https://doi.org/10.1002/itl2.549
11. Kaushik, A.K. (2016), "A Hybrid Approach of Fuzzy C-means Clustering and Neural network to make Energy-Efficient heterogeneous Wireless Sensor network", *International Journal of Electrical and Computer Engineering (IJECE)*, 6(2), 674 p. DOI: https://doi.org/10.11591/ijece.v6i2.pp674-681
12. Nayak, P. *et al.* (2024), "Artificial neural network-based clustering in Wireless sensor Networks to balance energy consumption", *Cogent engineering*, 11(1). DOI: https://doi.org/10.1080/23311916.2024.2384652
13. Al-Mousawi, A.J. (2020), "Evolutionary intelligence in wireless sensor network: routing, clustering, localization and coverage", *Wireless networks*, 26(8), P. 5595–5621. DOI: https://doi.org/10.1007/s11276-019-02008-4
14. Sahoo, B.M., Pandey, H.M. and Amgoth, T. (2022), "A genetic algorithm inspired optimized cluster head selection method in wireless sensor networks", *Swarm and evolutionary computation*, 75(101151), 101151 p. DOI: https://doi.org/10.1016/j.swevo.2022.101151
15. Balobaid, A.S. *et al.* (2023), "Neural network clustering and swarm intelligence-based routing protocol for wireless sensor networks: A machine learning perspective", *Computational intelligence and neuroscience*, 2023(1), 4758852 p. DOI: https://doi.org/10.1155/2023/4758852
16. Mohsin, R. razaq. (2022), "Genetic Algorithm: A study survey". *Iraqi Journal of Science*. Vol 63 No 3. P. 1215–1231. DOI: https://doi.org/10.24996/ijs.2022.63.3.27

Received (Надійшла) 02.07.2025
Accepted for publication (Прийнята до друку) 30.11.2025
Publication date (Дата публікації) 28.12.2025

ISSN 2522-9818 (print)
ISSN 2524-2296 (online) *Innovative technologies and scientific solutions for industries. 2025. No. 4 (34)*

Відомості про авторів / About the Authors

Семенова Олена Олександрівна – кандидат технічних наук, доцент, Вінницький національний технічний університет, доцент кафедри "Інфокомунікаційні системи і технології", Вінниця, Україна; e-mail: semenova.o.o@vntu.edu.ua; ORCID ID: https://orcid.org/0000-0001-5312-9148; Scopus ID: https://www.scopus.com/authid/detail.uri?authorId=36728178500

Джус Андрій Васильович – Вінницький національний технічний університет, аспірант, кафедра "Інфокомунікаційні системи і технології", Вінниця, Україна; e-mail: dzhuz1988@gmail.com; ORCID ID: https://orcid.org/0009-0005-3583-5766; Scopus ID: https://www.scopus.com/authid/detail.uri?authorId=59897049100

Мартинюк Володимир В'ячеславович – Вінницький національний технічний університет, аспірант, кафедра "Інфокомунікаційні системи і технології", Вінниця, Україна; e-mail: vm4ukr@gmail.com; ORCID ID: https://orcid.org/0009-0006-8421-0348

Semenova Olena – PhD (Engineering Sciences), Associate Professor, Vinnytsia National Technical University, Associate Professor at the Department of Infocommunication Systems and Technologies, Vinnytsia, Ukraine.

Dzhus Andrii – Vinnytsia National Technical University, Graduate Student at the Department of Infocommunication Systems and Technologies, Vinnytsia, Ukraine.

Martyniuk Volodymyr – Vinnytsia National Technical University, Graduate Student at the Department of Infocommunication Systems and Technologies, Vinnytsia, Ukraine.

APPLICATION OF COMPUTATIONAL INTELLIGENCE TECHNOLOGIES IN CLUSTERIZATION PROBLEMS OF WIRELESS SENSOR NETWORKS

The subject matter of the study is the process of selecting the head node of a cluster in wireless sensor networks (WSNs) using intelligent approaches that can adapt to changing environmental conditions. WSNs consist of a large number of sensor nodes with that collect, process and transmit data. Effective clustering is one of the main mechanisms for optimizing the operation of WSNs, as it allows reducing energy consumption, increasing network reliability and scalability. **The goal** of the study is to analyze the features of using modern computational intelligence tools and methods to increase the efficiency of the sensor node clustering process, which allow taking into account a variety of factors when making decisions about cluster formation and selecting head nodes. Traditional clustering algorithms are not always able to adapt to changes in network parameters, especially in the presence of heterogeneous nodes or changes in topology. In this regard, methods based on computational intelligence, in particular genetic algorithms, neural networks, fuzzy logic, as well as hybrid approaches, are becoming increasingly relevant. These methods allow taking into account a number of parameters when forming clusters and selecting cluster heads. **Tasks** of the study are analysis of existing approaches to clustering in BSM; development of a clustering fuzzy inference system; construction of a rule base for making optimal decisions; experimental verification of the proposed system. **Methods** of the study are tools of computational intelligence, in particular neural network learning, genetic optimization and fuzzy control, as well as computer modeling. The article analyzes the advantages of using each of the existing approaches. **Results** are: a structure of the fuzzy inference system was developed, input and output variables were determined, a database of fuzzy rules and membership functions was formed. The operation of the fuzzy system was simulated in the MATLAB environment. The developed system was also optimized and its operation validated. **Conclusions**: the use of hybrid intelligent approaches has significant advantages for solving clustering problems in BSM, which may indicate the prospects for further development of systems capable of functioning effectively in conditions of limited resources and high environmental complexity.

Keywords: wireless sensor network; clustering; fuzzy logic; neural network; genetic algorithm.

Бібліографічні описи / Bibliographic descriptions

Семенова О. О., Джус А. В., Мартинюк В. В. Застосування технологій обчислюваного інтелекту в задачах кластеризації бездротових сенсорних мереж. *Сучасний стан наукових досліджень та технологій в промисловості.* 2025. № 4 (34). С. 151–162. DOI: https://doi.org/10.30837/2522-9818.2025.4.151

Semenova, O., Dzhus, A., Martyniuk, V. (2025), "Application of computational intelligence technologies in clusterization problems of wireless sensor networks", *Innovative Technologies and Scientific Solutions for Industries*, No. 4 (34), P. 151–162. DOI: https://doi.org/10.30837/2522-9818.2025.4.151

Сучасний стан наукових досліджень та технологій в промисловості. 2025. № 4 (34)
ISSN 2522-9818 (print)
ISSN 2524-2296 (online)

АЛФАВІТНИЙ ПОКАЖЧИК

ALPHABETICAL INDEX

НАУКОВЕ ВИДАННЯ

Сучасний стан наукових досліджень та технологій в промисловості

Щоквартальний науковий журнал

№ 4 (34), 2025

Відповідальний секретар журналу *І.Г. Перова*
Відповідальний за випуск *А.А. Коваленко*
Відповідальний за ліцензування *В.В. Косенко*
Редактор *Л.В. Кузьміна*
Комп'ютерна верстка *Л.Ю. Свєтайло*

Формат 60×84/8. Умов. друк. арк. 23,25.
Тираж 150 прим.

Віддруковано з готових оригінал-макетів
в типографії ФОП Андреєв К.В.
Єдиний державний реєстр юридичних осіб
та фізичних осіб-підприємців.
Запис №24800170000045020 від 30.05.2003.

61157, Харків, вул. Акад. Богомольця, 9, кв. 50,
тел. +38 (063) 993-62-73
e-mail: ep.zakaz@gmail.com

SCIENTIFIC PUBLICATION

Innovative technologies and scientific solutions for industries

Quarterly scientific journal

No. 4 (34), 2025

Responsible secretary of journal *I. Perova*
Responsible for release *A. Kovalenko*
Responsible for licensing *V. Kosenko*
Editor *L. Kuzmina*
Computer layout *L. Svietailo*

Format 60×84/8. Conventional printed sheets 23,25.
Edition of 150 copies.

Printed from ready-made original layouts
in the printing house
of Individual Entrepreneur Andreev K.V.
Unified State Register of Legal Entities
and Individual Entrepreneurs.
Entry No. 24800170000045020 of 30.05.2003.

fl. 50, 9, Acad. Bogomolets Str., Kharkiv, 61157,
тел. +38 (063) 993-62-73
e-mail: ep.zakaz@gmail.com