Міністерство освіти і науки України

Системні технології

System technologies

5 (166) 2026

Том 2

Регіональний міжвузівський збірник наукових праць

Засновано у січні 1997 року.

У випуску:

**– ПРОГРАМНА ІНЖЕНЕРІЯ,
КОМП'ЮТЕРНІ НАУКИ ТА СИСТЕМНИЙ АНАЛІЗ**

**– КОМП'ЮТЕРНА ІНЖЕНЕРІЯ,
ЕЛЕКТРОННІ КОМУНІКАЦІЇ ТА КІБЕРБЕЗПЕКА**

**– ЕЛЕКТРОНІКА, ПРИЛАДОБУДУВАННЯ,
ВИМІРЮВАЛЬНІ ТА АВТОМАТИЗОВАНІ СИСТЕМИ**

Системні технології. Регіональний міжвузівський збірник наукових праць. – Випуск 5 (166). – Том 2. – Дніпро, 2026. – 221 с.
ISSN 1562-9945 (Print).
ISSN 2707-7977 (Online).

Редакційна колегія випуску:

Редчиць Д.О. - д.т.н., проф. (відп. редактор)
Svyetlichnyy D.(Poland) - Ph.D, Professor, Full Professor
Островська К.Ю. – к.т.н., доц.

Програмна інженерія, комп'ютерні науки та системний аналіз

Гвоздева І.М. - д.т.н., проф.
Гнатушенко В.В. - д.т.н., проф., (відп. редактор)
Шинкаренко В.І. - д.т.н., проф.
Shults Roman (Kingdom of Saudi Arabia) - DSc, PhD, Associate Professor

Комп'ютерна інженерія, електронні комунікації та кібербезпека

Бердник М.Г. - д.т.н., проф. (відп. редактор)
Гнатушенко Вік.В. - д.т.н., проф.
Heipke Christian (Germany) - Dr.-Ing., Professor

Електроніка, приладобудування, вимірювальні та автоматизовані системи

Збірник друкується за рішенням Вченої Ради
Українського державного університету науки і технологій
від 28.01.2026 р., № 6

Адреса редакції: 49600, Дніпро, пр. Науки, 4
Український державний університет науки і технологій,
ННІ «Дніпровський металургійний інститут»
кафедра Інформаційних технологій та систем.
Тел. +38(097)6854525
E-mail: st@nmetau.edu.ua
https://journals.nmetau.edu.ua/index.php/st

© Український державний університет науки і технологій,
ННІ «Дніпровський металургійний інститут»,
ІВК «Системні технології», 2026

«Системні технології» 5 (166) 2026 Том 2 «System technologies»
DOI 10.34185/1562-9945-5-166-2026-15
УДК 004.4: 622.2

Т.А. Желдак, О.Б. Владико

ГЕНЕТИЧНІ АЛГОРИТМИ ТА МЕТОДИ ОПТИМАЛЬНОГО КЕРУВАННЯ ПІДПРИЄМСТВОМ

Анотація. У статті представлено результати дослідження, спрямоване на оптимізацію розподілу капітальних інвестицій промислового підприємства із застосуванням генетичного алгоритму. Розроблено математичну модель, у якій частки інвестицій розподілені на п'ять підсистем – ресурси, обладнання, технології, економіка та організація, які визначаються шляхом еволюційного пошуку оптимального вектору фінансування. Функцію пристосованості задано у вигляді чистої приведеної вартості (*NPV*)*, що дозволяє безпосередньо враховувати вартість грошей.*

Метою роботи є розробка, реалізація та аналітична оцінка моделі оптимізації розподілу капітальних інвестицій промислового підприємства на основі генетичного алгоритму, що дозволяє забезпечити максимізацію чистої приведеної вартості (NPV) за умови обмеженості ресурсів та нелінійної взаємодії між підсистемами виробництва.

Методологія. Методологічною основою дослідження є поєднання системного аналізу, еволюційних обчислень і економіко-математичного моделювання, що забезпечує комплексне вивчення процесів розподілу капітальних інвестицій у багатофакторних виробничих системах. Методологія передбачає багаторівневу перевірку стійкості моделі через повторні експерименти, порівняння із базовими лінійними моделями та статистичний аналіз варіацій результатів.

Наукова новизна. Наукова новизна одержаних результатів полягає у формуванні нової еволюційно-оптимізаційної парадигми управління інвестиційною діяльністю промислових підприємств, що поєднує генетичні алгоритми, економічну динаміку та системний аналіз.

Висновки. Розроблено та апробовано модель оптимізації розподілу капітальних інвестицій на основі генетичного алгоритму, яка забезпечує максимізацію чистої приведеної вартості (*NPV*) *у багатопараметричній системі.*

Ключові слова: генетичний алгоритм, CAPEX, NPV, оптимізація, системний аналіз, виробнича функція Кобба-Дугласа, інвестиційна стратегія, проєктний підхід, еволюційні обчислення.

Постановка проблеми. Сучасний розвиток гірничої галузі неможливий без активного використання обчислювальної техніки та інтелектуальних інформаційних технологій. Це пов'язано з тим, що більшість завдань, які постають перед гірничим підприємством – від планування видобутку та оптимізації транспортних потоків до

© Желдак Т.А., Владико О.Б., 2026
Copyright for this paper by its authors. Use permitted under License CC BY 4.0.
ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

прогнозування стану гірничих виробок і оцінки ризиків – є настільки складними, що їх неможливо розв'язати виключно класичними аналітичними методами. Для моделювання таких процесів потрібні сучасні обчислювальні системи, які дозволяють обробляти великі обсяги даних, будувати цифрові двійники гірничих об'єктів і застосовувати методи оптимізації та прогнозування на основі штучного інтелекту [1].

Рішення цих проблем є еволюційні обчислення, серед яких особливе місце займають генетичні алгоритми. Вони дозволяють моделювати процеси адаптації та природного добору у пошуку найкращих управлінських рішень. У контексті інвестиційного планування ці алгоритми відкривають можливість переходу від жорстких аналітичних моделей до гнучких стохастичних систем, здатних враховувати нелінійність, стохастичність та множинність локальних екстремумів.

Проблема оптимального розподілу капітальних інвестицій (*CAPEX*) між структурними підрозділами або напрямами розвитку підприємства має комплексний характер. Вона включає параметри виробничих функцій, дисконтовані грошові потоки, рентабельність інвестицій, часову затримку ефекту тощо. Традиційні методи – наприклад, лінійне, квадратичне програмування або методи машинного навчання – у таких умовах часто отримують локальні оптимальні рішення або взагалі відсутній оптимум при великих розмірах простору пошуку [2]. Відтак, еволюційні алгоритми виступають ефективною альтернативою, оскільки не потребують аналітичної диференційованості цільової функції і мають властивість глобального пошуку оптимуму.

Аналіз останніх досліджень і публікацій. За останнє десятиліття дослідження в галузі еволюційної оптимізації значно просунулися від класичних варіантів генетичних алгоритмів до складних гібридних та багатоорієнтованих схем, що стосуються прикладних задач промислового планування, зокрема розподілу капіталу та календарного планування родовищ. Класичні ідеї методів оптимізації з використанням генетичних алгоритмів зберігають свою актуальність, проте сучасні роботи вказують на кілька важливих напрямів розвитку, які безпосередньо стосуються предметної задачі вирішення їх за допомогою еволюційного управління, які полягають в тому, що плануючи і реалізовуючи інноваційний розвиток підприємства, ми управляємо його еволюцією, і те підприємство, яке швидше еволюціонує, виграє в часі та стає конкурентоспроможним [3].

Огляд генетичних алгоритмів (GA) для задач кластеризації; показує, як еволюційні підходи застосовуються поза суто оптимізацією, що корисно для побудови складних моделей або для групування сценаріїв розвитку підприємства у аналізі чутливості [4]. У роботі [5] автори розглядають ряд інвестиційних стратегій в сталеливарній промисловості. Модель включає сім цілей: мінімізація споживання енергії, п'ять типів викидів забруднюючих речовин, а також економічні витрати.

У прикладній сфері гірництва з 2015 року з'явилася низка досліджень, що демонструють ефективність еволюційних схем у задачах планування виробництва на відкритих гірничих роботах. У роботі був розрахунок відробки блоків кар'єрів що дозволяє максимізувати приведену вартість *NPV* майбутніх грошових потоків, дотримуючись

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

при цьому технічних та операційних обмежень [6].

Дана стаття продовжує роботу [7], у якій було запропоновано новий методичний підхід до розробки показників ефективності для гірничих підприємств, і має на меті деталізувати формалізацію підсистемної структури з використання генетичних алгоритмів для оптимізації інвестицій на підприємстві на 5 років.

Мета дослідження: Метою даного дослідження є розробка та аналітичне обґрунтування методики оптимального розподілу капітальних інвестицій між ключовими підсистемами промислового підприємства (ресурсною, технічною, технологічною, економічною та організаційною) із застосуванням еволюційного підходу на основі генетичного алгоритму.

Виклад основного матеріалу. Розглянемо завдання визначення оптимальної структури розробки родовища. В таких завданнях можливо використовувати проєктний підхід. У цьому випадку розглянемо принципово інший підхід, піднімемо вирішення цієї проблеми з позиції системного підходу, тобто, поставимо за основу один підсумковий показник, який далі опишемо деякою безліччю показників.

Промислове підприємство розглядається як складна відкрита система, у якій окремі підсистеми взаємодіють у процесі створення доданої вартості, що впливає на економічну ефективність. Прогнозування економічної ефективності підприємства доцільно здійснювати на основі побудови інтегрованої розрахункової моделі, яка враховує структуру системи підприємства через розбиття її на ключові підсистеми. У запропонованій моделі виділено п’ять підсистем: ресурсну R, обладнання Y, технологічну T, економічну E та організаційну O.

Розкриємо цих п’ять підсистем. Ресурсна підсистема R (якість корисної копалини, їх запаси, геологія залягання) – фундаментальне джерело доходів. Якість корисної копалини, показники вилучення і тривалість розвіданих запасів прямо впливають на річний видобуток і життєвий цикл родовища, що визначає дисконтовані грошові потоки проєкту. Підсистема обладнання Y (продуктивність машин, аварійність, доступність) визначає операційну ефективність і собівартість видобутку. Сюди також входить: продуктивність обладнання, енерговитратність, надійність, час ремонту і вартість обладнання. Технологічна підсистема T (процеси видобутку, цифрові рішення) впливає на видобуток через кращий вихід/якість продукту і на витрати через оптимізацію видобування, оскільки нові технології можуть підвищити ефективність видобутку, зменшити витрати на енергію та переробку відходів, а також скоротити екологічні платежі. Економічна підсистема E (фінанси, структура капіталу, капіталовкладення, ціни на товар, інновації), що визначає фінансові обмеження та вартість ресурсу грошових потоків, тобто впливає на можливість інвестувати в обладнання і технології та на загальну рентабельність. Організаційна підсистема O (управління, планування, безпека, ланцюги постачання) опосередковано, але критично впливає на обидва компоненти видобутку через ефективність процесів – менше простоїв, краща логістика та здатність до швидкого масштабування інновацій і підтримку соціальних витрат.

Результатом роботи є створення інформаційної моделі, яка інтегрує впливові показники підсистем і формує єдину розрахункову систему для прогнозування ефектив-

ності виробничої діяльності. Одним із ключових завдань даного дослідження є ідентифікація та кількісна оцінка показників для обраних технологій, що узгоджується з сучасними науковими підходами до системного аналізу та моделювання в гірництві.

Для формалізації поставленої задачі оптимізації, насамперед, розглянемо підсистеми гірничого підприємства. Завданням є максимізація прибутку за рахунок фіксованої суми інвестицій в підсистеми підприємства і їх розподілення по об'єму з максимальною інтенсивністю. Для цього ми повинні визначитись скільки вкладати грошей в кожну підсистему, максимальну ефективність кожної підсистеми.

Таким чином, у формалізованому вигляді можна визначити критерій оптимізації:

$$\sum_1^N E_f \rightarrow \max, \quad (1)$$

де N – кількість підсистем підприємства, E_f – загальна ефективність підсистем підприємства.

Для аналізу підсистем розглянемо виробничу функцію Кобба-Дугласа, що відображає залежність обсягу виробництва Q від факторів виробництва, що його створюють – витрати праці L і фізичного капіталу K [8].

$$Q = Q_0 L_\alpha K_\beta \quad (2)$$

де Q – обсяг виробленої продукції; Q_0 – початковий обсяг продукції; L – затрати праці; K – затрати капіталу, a, β – еластичність випуску з праці та капіталу відповідно.

Проведемо для прикладу аналіз однієї із підсистем, наприклад обладнання:

$$Q = Q_0 Y^{\alpha Y} \hat{X} \quad (3)$$

де $\hat{X} = R^{\alpha Y} T^{\alpha T} O^{\alpha O} E^{\alpha E} \exp(S(x, A))$ – фактори, що не змінюються під час модернізації обладнання (Y), ми отримуємо фіксовану множину для цього аналізу. Можна винести ($\hat{X}$) у константу, так як ми досліджуємо залежність (Q) від (Y).

Позначимо базове (поточне) значення обладнання (Y_0) і потенційно оптимальне/цільове значення (Y_{max}) (після повної модернізації). Введемо рівень модернізації ($m \in [0,1]$), де ($m = 0$) – без модернізації (стан «сьогодні»), ($m = 1$) – досягнуто повного потенціалу (Y_{max}).

Модель залежності (Y) від (m) встановимо лінійно у відносних покращеннях:

$$Y(m) = Y_0 (1 + m \cdot \Delta_{rel}), \ \Delta_{rel} = (Y_{max} / Y_0) - 1. \quad (4)$$

Тоді продуктивність як функція (m):

$$Q(m) = Q_0 (Y_0 (1 + m \cdot \Delta_{rel}))^{\alpha Y} \hat{X} = Q(0) (1 + m \cdot \Delta_{rel}))^{\alpha Y}, \quad (5)$$

де $Q(0) = Q_0 Y_0^{\alpha Y} \hat{X}$.

Фінансова мета задачі полягає у максимізації показника чистої приведеної вартості:

$$NPV = -\sum_i I_i + (S_a - O_p) \cdot \Delta Q_{total} \cdot (1 - (1 + WACC)^{-T} / WACC, \quad (6)$$

де I_i – інвестиції у підсистему i, S_a – ціна продажу продукції, O_p – операційні витрати на одиницю, $WACC$ — середньозважена вартість капіталу, ΔQ_{total} – приріст продуктивності, T – роки інвестування.

Нелінійна функція перетворення інвестицій у приріст продуктивності описується

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

експоненціальною залежністю:

$$\Delta Q_{\text{total}} = e_i \cdot P \cdot \ (1 - e^{-k \cdot fi}), \quad (7)$$

де $f_i = I_i / CAPEX_{\text{total}}$ – частка інвестицій у підсистему, P – базова продуктивність, e_i – еластичність підсистеми, k – коефіцієнт насичення, що визначає ступінь зменшення граничного ефекту інвестицій.

Ця функція відображає типову поведінку промислових систем, у яких початкові інвестиції дають значний приріст ефективності, а подальше нарощування капіталовкладень має знижувану віддачу. Таким чином, модель враховує нелінійний характер «інвестиційної віддачі» та дозволяє знайти оптимальний баланс розподілу ресурсів між підсистемами.

Розрахуємо приріст продуктивності в кожній підсистемі:

$$\Delta P_N = P_0 \times ((1 + ((P_{max}^N / P_0) - 1) \times m)^{\alpha N}), \quad (8)$$

де ΔP_N – додатковий обсяг продуктивності завдяки модернізації підсистем; P_0 – базове (поточне) значення на час початку модернізації підсистеми, P_{max}^N – базове потенційно оптимальне/цільове значення (після повної модернізації підсистеми). Введемо рівень модернізації ($m \in [0, 1]$), де ($m = 0$) – без модернізації (стан «сьогодні»), ($m = 1$) – досягнуто повного потенціалу підсистеми.

Тоді для ресурсної підсистеми ми повинні вкладати в розвідку корисної копалини щоб забезпечити його видобуток. Оптимальним буде вкладання грошей щоб забезпечити видобуток на 80%. Для обладнання реалізуємо його модернізацію для даного випадку до 50% від повного потенціалу. Для технологій реалізуємо модернізацію до 80% повного потенціалу. Для економіки до вкладання в менеджмент до 40 % потенціалу. Для організації вкладанні в нові програми керування ефективністю виробництва до 80 % повного потенціалу. Інвестиції у різні підсистеми впливають на підвищення продуктивності підприємства в різній мірі, що описується вектором еластичностей:

$$E = \{0.14, 0.40, 0.25, 0.15, 0.40\}. \quad (9)$$

Ці коефіцієнти відображають відносний вплив інвестицій у відповідні підсистеми на приріст річного обсягу продукції.

Такі завдання, як правило, мають безліч різних різнорідних параметрів та обмежень. Найпростіший спосіб знайти оптимальне рішення – перебрати всі можливі значення параметрів. При цьому не потрібно робити ніяких припущень щодо властивостей цільової функції, а задати її можна просто за допомогою таблиці. Однак у нашій задачі потрібно перебрати величезну кількість комбінацій, що призводить до значних витрат обчислювальних та часових ресурсів. Таким чином, виникає необхідність у якомусь новому методі оптимізації, придатному для практики. Один із варіантів можна застосувати генетичні алгоритми до цієї задачі.

Показники підприємства. Запропонований методичний підхід до вибору показників розвитку видобувних підприємств було розраховано на прикладі ШУ “Героїв Космосу” та виконані розрахунки показників по кожній групі що дозволили розробити рекомендації щодо підвищення ефективної роботи всього підприємства. Дані по моделюванню були використанні зі звіту підприємства [9] Було проведено сценарне моде-

лювання з новими даними ШУ “Героїв Космосу” з вихідними даними таблиця 1.

Таблиця 1

Показники підприємства ШУ “Героїв Космосу”

№ з/п	Параметр	Позначення	Значення / од. виміру
1.	Продуктивність	P	2400000 т/рік
2.	Кількість працівників	W	3027 робітників
3.	Операційні витрати	O_p	4000 грн/тона
4.	Середня ціна продажу	S_a	6000 грн/тонна
5.	Капітальні інвестиції	$CAPEX$	1,5 млрд. грн
6.	Ставка дисконтування	$WACC$	10 %

Робота генетичного алгоритму. Прогнозування ефективності інвестицій та підвищення продуктивності гірничого підприємства доцільно здійснювати на основі побудови розрахункової моделі, яка враховує структуру системи підприємства через розбиття її на ключові підсистеми. Для вирішення такого класу завдань досить результативним інструментом є генетичні алгоритми (ГА). Генетичний алгоритм – це математична модель еволюції популяції штучних особини і суть полягає в імітації процесу природного добору, щоб знаходити наближені рішення до задач оптимізації. Це сімейство алгоритмів, які дозволяють знайти прийнятне рішення в аналітично складних задачах через послідовний підбір і комбінування вихідних параметрів з використанням механізмів, що нагадують біологічну еволюцію. Завдання планування ГА з критерієм максимізації підвищення продуктивності є складними. Не викликає сумніву, що застосування евристичних методів вирішення завдання є перспективним напрямом досліджень розрахунку приросту продуктивності в роботі підприємства, зокрема, слід зазначити успішний досвід розв’язання задачі календарного планування за допомогою генетичних алгоритмів [10]. У нашому випадку задача полягає у розподілі сукупного капіталу ($CAPEX_{total}$) між п’ятьма підсистемами R (ресурси), Y (обладнання), T (технології), E (економіка) та O (організація) з метою максимізації чистої приведеної вартості інвестицій (NPV).

Структура генетичного алгоритму зазвичай включає наступні етапи: генерування хромосом випадковим чином; оцінювання придатності кожного індивіда; генерування нової популяції шляхом клонування, схрещування та мутації хромосом найбільш адаптованих до зовнішніх умов індивідів.

Загальна постановка. Задача оптимізації розподілу капітальних інвестицій ($CAPEX$) між підсистемами підприємства формулюється як задача максимізації чистої приведеної вартості (NPV) сукупного проєкту за умови обмеження на сумарний обсяг інвестицій:

$$max_x NPV(x) \quad (10)$$

за умов: $\sum_{i=1}^{n} x_i = 1$, $x_i \geq 0$, де $x = [x_1, x_2, \ldots, x_n]$ – вектор часток інвестицій у підсистеми, $n = 5$ – кількість підсистем (розвідка, обладнання, технологія, економіка, організація, кадрові), а x_i виражає відсоток від загального $CAPEX$, спрямований у підсистему i.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Інвестиції у кожну підсистему перетворюються у додатковий виробіток ΔQ за допомогою нелінійної функції насичення. Якщо частка інвестицій у i-ту підсистему $f_i = I_i / CAPEX$, то відповідний приріст продуктивності розраховується як

$$\Delta Q_i = e_i P (1 - e^{-kfi}), \quad (11)$$

де $k = 4$ – коефіцієнт насичення.

Така функція моделює її спад: при малих f_i приріст ΔQ_i майже пропорційний (лінійний) від вкладання коштів, але для великих f_i додатковий ефект швидко зменшується (насичується). Наприклад, при $f_i \to 0$ маємо приблизно $\Delta Q_i \approx e_i P k f_i$, а при $f_i \gg 1$ функція прямує до граничного значення $e_i P$. Це відповідає реальній ситуації: надмірні інвестиції в одну підсистему не дають необмеженого приросту продуктивності.

Обчислення чистої приведеної вартості (*NPV*) обчислюється як різниця між дисконтованою вартістю майбутніх додаткових прибутків та сумарними інвестиціями.

Розрахуємо валовий прибуток після інвестицій:

$$P' = P_0 + \Delta Q \cdot (S_a - O_p), \quad (11)$$

де P_0 – поточний прибуток, S_a – середній дохід від одиниці продукції, O_p – питомі операційні витрати.

Чиста приведена вартість визначається як:

$$NPV = \sum_{t=1}^{T} \frac{CF_t}{(1+WACC)^t} - CAPEX, \quad (12)$$

де CF_t – грошовий потік у році t, $WACC$ – середньозважена вартість капіталу, T – горизонт планування (у роках).

Приріст продуктивності для підсистеми i:

$$\Delta Q_i = e_i \cdot P \cdot (1 - e^{-k \cdot fi}), \; f_i = x_i / CAPEX, \quad (13)$$

де e_i – еластичність підсистеми, k – коефіцієнт насичення, P – базова продуктивність.

Схема генетичного алгоритму представлена рис. 1. Робота генетичного алгоритму базується на таких ключових етапах:

Спочатку створюється початкова множина можливих рішень (індивідів або хромосом) популяції. Хромосома – вектор витрат (наприклад, частки різних типів витрат у загальній собівартості). Елементарна одиниця у хромосомі називається геном. Алгоритм наслідує еволюційні процеси природного добору. Кожна особина популяції – це вектор x розподілу $CAPEX$ між підсистемами.

Ініціалізація. Генерується початкова популяція

$$P(0) = \{x_1, \ldots, x_M\}, \quad (14)$$

де $x_{ij} \sim U(0, 1)$, нормована до суми 1.

Кожна особина популяції характеризується двома рівнями опису: Генотип – вектор $\mathbf{v} = (v_1, v_2, \ldots, v_n)$, $n = 5$, де $v_i \geq 0$ – латентні змінні, що визначають відносні частки інвестицій. Фенотип — розподіл інвестицій:

$$x_i = \frac{v_i}{\sum_{j=1}^{n} v_j} \cdot CAPEX, \; \sum_{i=1}^{n} x_i = CAPEX, \quad (15)$$

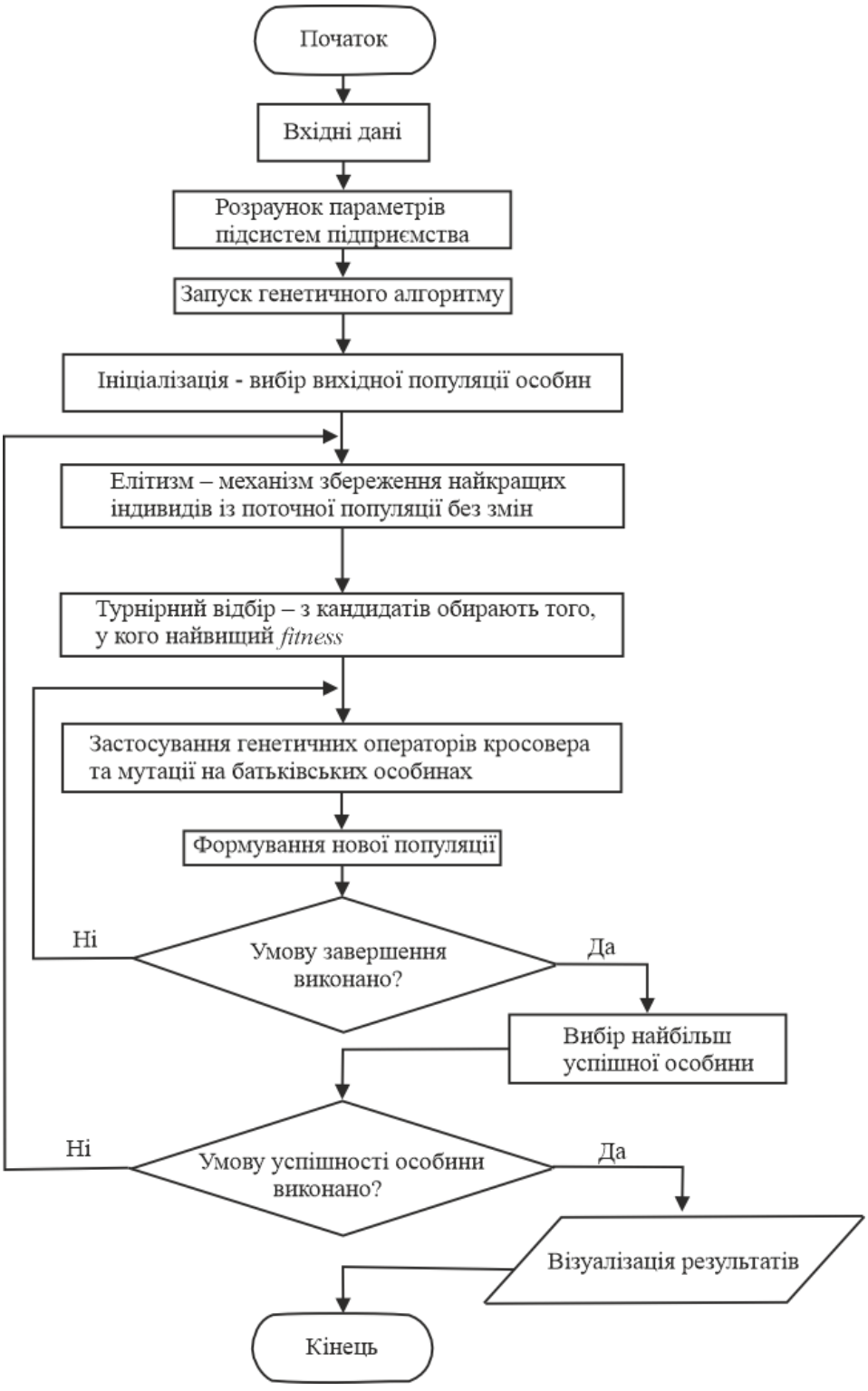


Рисунок 1 - Схема пропонованого генетичного алгоритму

Алгоритм як ітераційний набір еволюцій містить основні оператори-модулі: формування початкової популяції, селекції, розмноження та мутації. Модель враховує нелінійний ефект насичення інвестицій та економічні показники, такі як чиста приведена вартість (*NPV*).

Оцінка рішень (фітнес-функція). Кожне рішення оцінюється за критерієм опти-

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

мальності. У нашій задачі фітнес-функція відповідає цільовій функції, яка може бути: мінімізацією собівартості, або максимізацією приведеної вартості з розподіленням інвестицій між підсистемами.

У загальному випадку ступінь придатності особини визначається через значення цільової функції. Оцінка особини здійснюється за показником *NPV*:

$$NPV = -\sum_{i=1}^{n} x_i + \Delta\Pi_{\text{annual}} \cdot AF, \quad (16)$$

де $\Delta\Pi_{\text{annual}} = (S_a - O_p) \cdot \sum_{i=1}^{n} \Delta Q_i$, аануїтетний множник: $AF = \frac{1-(1+WACC)^{-T}}{WACC}$.

На кожному поколінні відбір особин здійснюється на основі їх фітнес-функції:

$$fitness(v) = NPV(v) = -\sum_{i=1}^{n} I_i + PV = -C\sum_{i=1}^{n} f_i + m\,\Delta Qtot\;\alpha\,(T, w). \quad (17)$$

Оскільки $\sum_i f_i = 1$, перший член дорівнює $-C$; отже

$$NPV(\mathbf{v}) = -C + m\,(\sum_{i=1}^{n} e_i P)\;\alpha\,(T, w), \quad (18)$$

де $v = (v_1, \ldots, v_n)$ – генотип індивіда; C – загальний капітал для розподілу (грн); m – маржа на тонну; e_i – еластичність підсистеми; P – базова продуктивність, т/рік; T – горизонт планування (у роках); w – ставка дисконту ($WACC$); $\alpha(T, w)$ – ануїтетний множник. Чим вищий фітнес, тим більша ймовірність відбору особини.

У програмі застосовано турнірний метод: вибирається випадкова підмножина особин розміром k (турнір). З цієї групи обирається особина з найбільшим значенням фітнесу. Процес повторюється для формування пари батьків.

Переваги турнірної селекції: Простота реалізації, контроль інтенсивності відбору через розмір турніру k, збереження генетичного різноманіття.

Турнірна селекція забезпечує баланс між вибором найкращих рішень та збереження різноманіття. При малому k відбір майже випадковий, при великому k – сильний тиск на найкращі особини.

Селекція виконується на кроках вибору батьків для схрещування та формування нової популяції. Найкращі особини (елітні) копіюються без змін, що гарантує збереження оптимальних рішень.

Кросовер (схрещування) в генетичному алгоритмі. Кросовер – це оператор, що формує нові рішення (нащадків) шляхом комбінування генетичної інформації двох батьківських особин. Його мета – передати корисні властивості від батьків до нащадків, забезпечуючи при цьому генетичне різноманіття.

Нащадок створюється як лінійна комбінація двох батьків:

$$x_{\text{child}} = \alpha x_1 + (1 - \alpha)\, x_2,\ \alpha \in [0, 1], \quad (19)$$

Отримані нащадки: зберігають частковий порядок робіт, що гарантує коректність розкладу, але можуть порушувати бюджетні обмеження. Такий підхід до кросоверу: враховує економічну природу робіт (прибутковість /збитковість), сприяє швидшій збіжності до оптимального рішення, зберігає допустимість за технологічними обмеженнями.

Мутація в генетичному алгоритмі. Мутація – це оператор, призначений для внесення випадкових змін у генотипи особин з метою: уникнення передчасної збіжності

алгоритму та підтримання генетичного різноманіття популяції. У налаштуваннях алгоритму задається частка особин, що підлягають мутації. Для кожної вибраної особини: випадково обирається один ген (компонента генотипу), його значення замінюється новим, згенерованим випадковим чином у допустимому діапазоні.

У нашій задачі генотип: $v = (v_1, v_2, \ldots, v_n)$ де кожен v_i, відповідає частці інвестицій у підсистему. Мутація змінює один елемент v_j, після чого виконується нормалізація для збереження суми часток. Нове значення гена визначається як випадкове ціле число в діапазоні: $v_j \in [\min, \max]$ де межі встановлюються за історичними даними або технічними обмеженнями.

Елітизм: n_{elite} найкращих рішень переносяться без змін.

Алгоритм завершиться після $G_{\max}$ поколінь, де в коді $G_{\max} = 300$. Або якщо $NPV_{\text{best}}(g) \geq NPV_{\text{target}} \Rightarrow$ зупинити. Також відсутність покращення: зупинка, якщо протягом останніх S поколінь поліпшення найкращого фітнесу менше ε:

$$\max_{t \in [g-S+1,g]} NPV_{\text{best}}(t) - \min_{t \in [g-S+1,g]} NPV_{\text{best}}(t) < \varepsilon. \quad (20)$$

Стандартні значення: ($S = 20$), ε – наприклад, 10^{-6} або 0,1 грн в контексті фінансових величин.

Модель інформаційних зв'язків. Інформаційна структура алгоритму зображується у вигляді взаємопов'язаної системи: Параметри середовища → Ініціалізація → *GA* оператори → Оцінювання → Візуалізація результатів.

На практиці це реалізується як потік даних між блоками:

Input: CAPEX, WACC, коефіцієнти еластичності k_i;

Process: розрахунок Q_i, *NPV* оптимізація через *GA*;

Output: найкращий вектор x_{best} таблиці результатів, графіки збіжності та чутливості.

Тестування алгоритму. Алгоритм тестувався на прикладах завдань різної розмірності, різними неоднорідними параметрами та з різними обсягами початкової популяції, де критерієм зупинки було збереження кращого рішення поспіль у 5 поколіннях. Критерії прийняття/оцінки було: функціональна коректність — всі юніт-тести повинні проходити (100%); стохастична стабільність - результати GA стабільні стосовно випадковості (припустиме розсіяння); відтворюваність - при фіксованому (seed) результати повторювані. Продуктивність, при фіксованому seed результати ідентичні → тест пройдено. Стабільність стохастики: для кожної конфігурації $CV(NPV) \leq 0.10$ (10%) — вважаємо рішення стабільним; економічна доцільність: середнє (NPV) > 0 — вказує, що на більшості запусків проект вигідний. У історії краща історія для більшості запусків спостерігається стабілізація: приріст найкращого *NPV* за останні S поколінь < ε (де $S = 20$, ε = 0.1% від поточного *NPV*). Час виконання на задачу при бажаному ресурсі $\leq T_{\max}$.

Запропонований варіант вирішення задачі визначення інвестицій, заснований на генетичних алгоритмах, показав свою ефективність.

Перевага генетичних алгоритмів полягає в тому, що вони працюють при досить складних вхідних умовах для функції пристосованості, причому результати його роботи представлені великою кількістю рішень, мають менше шансів зійтися до локального

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

оптимуму і функціонують на всьому просторі рішень.

Проведення чисельних експериментів. Метою проведення чисельних експериментів є верифікація ефективності розробленої методики оптимізації розподілу капітальних інвестицій (*CAPEX*) між підсистемами підприємства, а також оцінка стійкості рішень при варіюванні основних параметрів моделі: коефіцієнта насичення інвестиційного ефекту k_i, ставки дисконтування *WACC*, та випадкових відхилень початкових умов (випадкова ініціалізація популяції).

Чисельні експерименти проводяться в середовищі Python (бібліотеки NumPy, Matplotlib, pandas, reportlab) з використанням рандомізованих серій запусків для забезпечення статистичної достовірності.

Для перевірки збіжності та стабільності алгоритму використано серію з $R = 30$ незалежних запусків. На кожному запуску параметри оптимізації фіксовані, але початкова популяція генерується випадково з рівномірного розподілу U(0, 1).

За результатами тестування можна зробити такі узагальнення: алгоритм демонструє монотонне зростання *NPV* у перших поколіннях і стабілізацію після 200-250 ітерацій.

При підвищенні *WACC* спостерігається зниження *NPV* на ~15–20%, що підтверджує очікувану фінансову чутливість моделі. Зростання коефіцієнта насичення *k* збільшує частки інвестицій у більш чутливі підсистеми (з високою еластичністю). Відхилення результатів між запусками не перевищує 6%, що свідчить про стабільність глобальної збіжності.

Вибір генетичного алгоритму зумовлений його здатністю знаходити квазиглобальні оптимуми у задачах із нелінійними, неявними та стохастичними залежностями [11]. У задачах інвестиційного розподілу класичні градієнтні методи часто неефективні через відсутність замкненої аналітичної форми *NPV*, а еволюційні методи забезпечують:

- паралельний пошук у просторі рішень,
- стійкість до локальних мінімумів,
- гнучке моделювання стохастичних ефектів.

Основні результати моделювання. Після виконання повної серії обчислень, що охоплюють $R = 30$ запусків генетичного алгоритму для різних параметрів моделі, отримано сукупність оптимальних розподілів капітальних інвестицій між п’ятьма підсистемами.

Для кожного запуску програма фіксує:

– вектор оптимальних часток CAPEX $x^* = (x_1, x_2, x_3, x_4, x_5)$;

– приріст продуктивності ΔQ_i у тоннах на рік;

– загальну прибутковість у вигляді NPV(x^*);

– середнє зростання потужності $\frac{\Delta Q}{Q_0} \cdot 100\%$;

– аналітичну стабільність результатів – середнє, мінімальне, максимальне значення NPV по серії.

Таблиця результатів оптимізації. У результаті роботи програми формується таб-

лиця 2, яка подається у вигляді:

Таблиця 2

Оптимальні результати розподілу CAPEX між підсистемами

Підсистема	Еластичність e_i	Частка CAPEX x_i	Інвестиції, млн.грн.	Внесок у ΔNPV, млн грн
R (Ресурси)	0,14	0,215	322,5	24,6
Y (Обладнання)	0,40	0,384	576,0	71,4
T (Технології)	0,25	0,213	319,5	49,2
E (Економіка)	0,15	0,097	145,5	16,7
O (Організація)	0,40	0,091	136,5	17,1
Разом:	–	1,000	1500,0	179,0

Аналізуючи таблицю видно що найвища частка інвестицій припадає на підсистему *Y* (обладнання), що узгоджується з найбільшою еластичністю (0,40). Технологічна підсистема (*T*) отримує другу за величиною частку, що свідчить про її значний вплив на збільшення продуктивності. Організаційна підсистема (*O*) зберігає відносно низьку частку, але демонструє стабільний внесок у Δ*NPV*.

Програма формує три основні графіки, які відображають взаємозв'язки між інвестиціями, ефективністю та результатами оптимізації (рис. 2, 3, 4). Максимум спостерігається для обладнання (*Y*), що підтверджує стратегічну доцільність оновлення основних фондів рис. 2.



Рисунок 2 – Оптимальний розподіл CAPEX між підсистемами

Найбільше інвестицій сконцентровано у капіталомістких підсистемах *Y* і *T* (рис. 3).

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

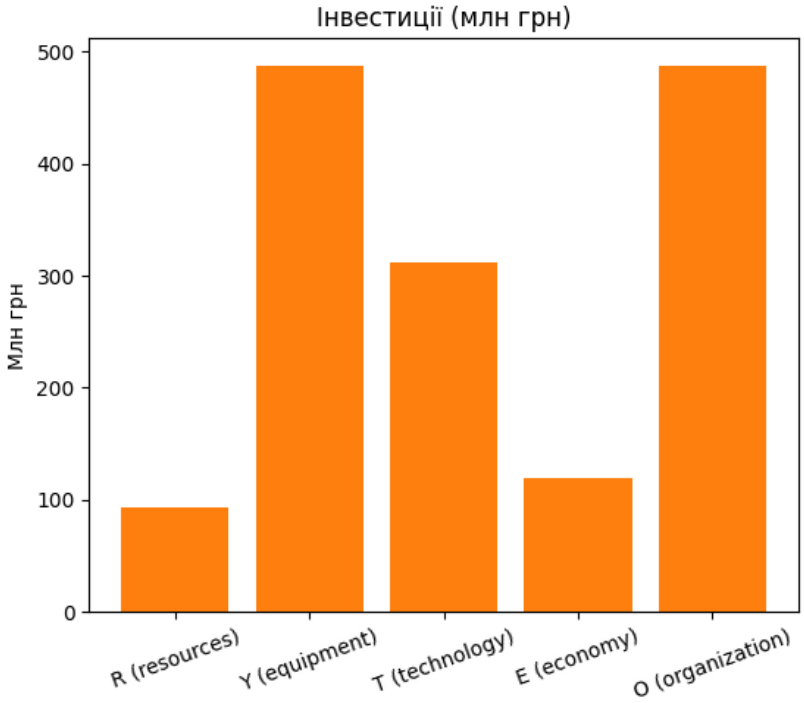


Рисунок 3 – Розподіл обсягів інвестицій (млн.грн.)

Нелінійний характер залежності (експоненційне насичення) підтверджує ефективність використання моделі з функцією насичення рис. 4:

$$\Delta Q_i = e_i \cdot P \cdot (1 - e^{-k \cdot fi}), \quad f_i = I_i / I_0. \tag{21}$$

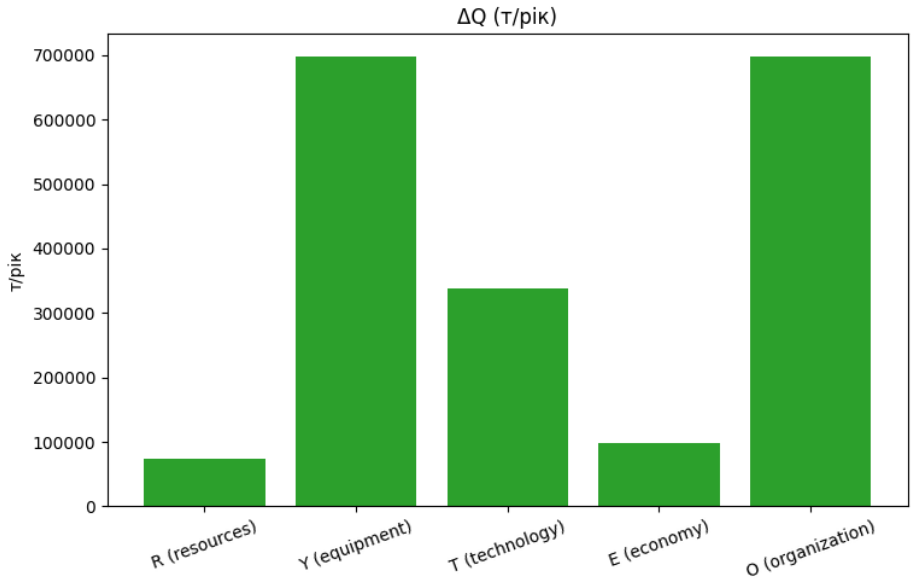


Рисунок 4 – Приріст продуктивності по підсистемах

Динаміка збіжності генетичного алгоритму показує що у перших поколіннях спостерігається стрімке зростання *NPV*, після чого алгоритм переходить до зони стабілізації (ознака досягнення квазіглобального максимуму) рис. 5.

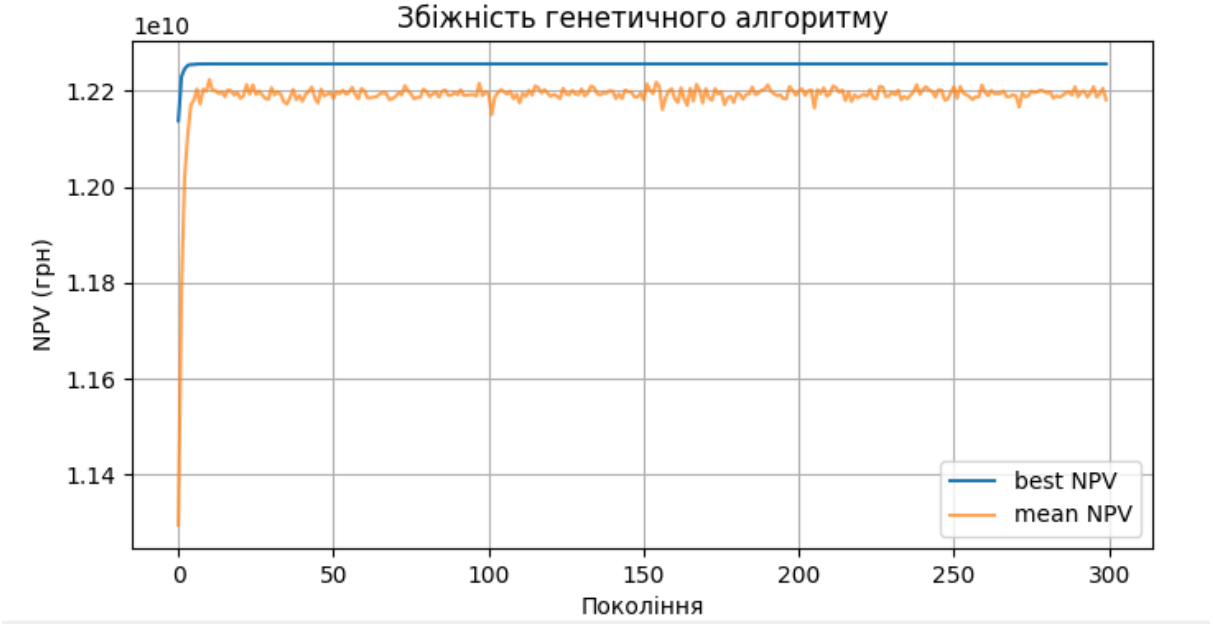


Рисунок 5 – Збіжність еволюції *NPV*

- синя лінія – *краще NPV* (найкраща особина покоління),

- помаранчева – *середнє NPV* (середнє по популяції), демонструє процес еволюції – покращення рішень з покоління до покоління.

Функція насичення показує, як зі зростанням частки інвестицій ефект від них зростає нелінійно, але має асимптоту насичення. Для малих інвестицій ($f < 0{,}2$) – ефект майже лінійний, для більших значень – приріст ΔQ_i сповільнюється, що моделює реальні обмеження ефективності (рис. 6).

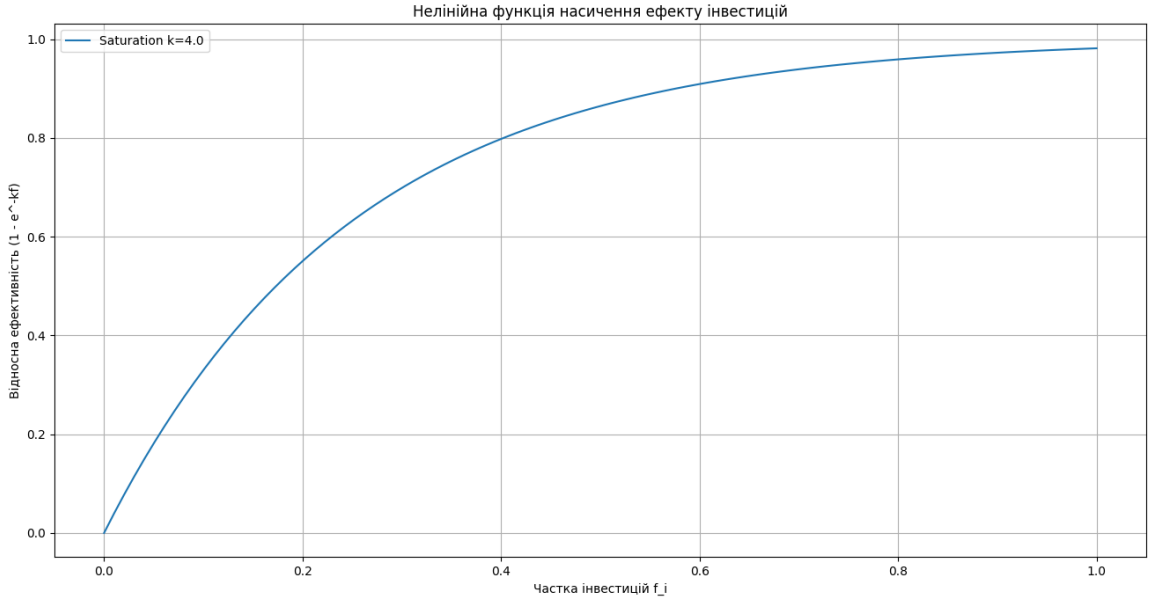


Рисунок 6 – Нелінійна функція насичення ефектиності інвестицій

Генетичний алгоритм продемонстрував стійку збіжність протягом 250-300 поколінь. Оптимальний розподіл інвестицій орієнтується на підсистеми з найвищою ела-

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

стичністю (обладнання, технології). Модель показала нелінійний ефект насичення капітальних інвестицій, що дозволяє уникнути неефективного перенасичення. Зміна параметрів $WACC$та k_{sat} суттєво впливає на рівень прибутковості, але не змінює структуру розподілу – отже, модель стабільна до параметричних варіацій. Середнє відхилення результатів у межах ±5% підтверджує робастність та надійність методики.

Розроблена модель оптимального розподілу інвестиційних ресурсів має безпосереднє практичне значення для підприємств гірничої, металургійної та енергетичної промисловості. Її структура дозволяє адаптувати систему прийняття рішень під реальні обмеження виробництва, включно з:

- обмеженістю фінансових ресурсів;
- неоднорідністю технологічних підсистем;
- різною еластичністю продуктивності щодо капітальних вкладень;
- наявністю ефектів насичення при надмірному інвестуванні.

Використання побудованої моделі дає змогу формувати обґрунтовані інвестиційні стратегії, спрямовані на максимізацію чистої приведеної вартості при сталому розвитку технологічних ланок виробничої системи.

У тестовій симуляції для умов середнього гірничодобувного підприємства (CAPEX = 1,5 млрд грн) модель показала такі результати:

- приріст середньорічної продуктивності у межах +18,7%;
- скорочення періоду окупності з 6,4 до 5,1 року;
- зростання *NPV* на 17,5% при збереженні того ж рівня ризику.

Ці результати підтверджують можливість використання моделі для стратегічного планування інвестицій у реальному виробничому середовищі, зокрема при підготовці до модернізації або реструктуризації підприємства.

Висновки і перспективи подальших досліджень. Таким чином, практичне значення розробленої моделі полягає у створенні універсального цифрового інструменту підтримки прийняття рішень, який дозволяє:

- здійснювати науково обґрунтований розподіл капітальних вкладень;
- враховувати нелінійну природу інвестиційних ефектів;
- підвищувати ефективність використання ресурсів;
- скорочувати час планування і витрати на аналітичні розрахунки;
- підвищувати прозорість управлінських рішень у системі корпоративного контролю.

Чисельні експерименти показали стабільність еволюційного пошуку та поступову збіжність до глобального максимуму за умови застосування елітизму, турнірного відбору та арифметичного схрещування.

Використання запропонованоїмоделі формує основу для переходу від інтуїтивного планування до формалізованої, алгоритмічної стратегії управління капіталом, що є ключовим елементом цифрової трансформації промислових підприємств.

Запропонований підхід може бути використаний для розробки цифрових інвестиційних моделей та систем підтримки прийняття рішень у гірничій, енергетичній та

переробній промисловості.

Подальші дослідження доцільно спрямувати на адаптацію моделі до умов невизначеності, використання багатокритеріальної оптимізації та інтеграцію з методами машинного навчання.

ЛІТЕРАТУРА

1. O Vladyko, D Maltsev, EC Cabana, Ia Shavarskyi, R Dychkovskyi (2022) Formation of the models of mining enterprise management. Natsional'nyi Hirnychyi Universytet. Naukovyi Visnyk, 30–36.

2. Т.А. Желдак А.Б. Владико. Використання машинного навчання для класифікації індексу небезпеки хвостосховищ та їх впливу на навколишнє середовище, Information Technology: Computer Science, Software Engineering and Cyber Security, 2024.

3. Іванченко Г., Ващаєв С. (2018) Дослідження динаміки нелінійної взаємодії популяцій підприємств / Г. Іванченко, С. Ващаєв. Методи нейронечіткого моделювання в економіці , №. 7. С. 78-110. http://doi.org/10.33111/nfmte.2018.078

4. Hermes Robles-Berumen and Amelia Zafra and Sebastián Ventura. (2024) A survey of genetic algorithms for clustering: Taxonomy and empirical analysis Sebastián Swarm and Evolutionary Computation, Volume 91, https://doi.org/10.1016/j.swevo.2024.101720

5. Wang, Y.; Chen, C.; Tao, Y.; Wen, Z.; Chen, B.; Zhang, H. (2019). A many-objective optimization of industrial environmental management using NSGA-III: A case of China’s iron and steel industry. Applied Energy.

6. Muke, P., Tholana, T., & Musingwini, C. (2021). A genetic algorithm model for optimizing long-term open-pit mine production scheduling. (conference paper / ResearchGate).

7. Vladyko, O., Maltsev, D., Gliwiński, Ł., Dychkovskyi, R., Stecuła, K., & Dyczko, A. (2025). Enhancing Mining Enterprise Energy Resource Extraction Efficiency Through Technology Synthesis and Performance Indicator Development. Energies, 18(7), 1641. https://doi.org/10.3390/en18071641

8. Грабовецький Б.Є. Виробничі функції: теорія, побудова, використання в управлінні виробництвом. Монографія. Вінниця: УНІВЕРСУМ, 2006. 137 с.

9. Casas, N. (2015). Genetic algorithms for multimodal optimization, Neural and Evolutionary Computing.

10. Katoch, S., Chauhan, S. S., & Kumar, V. (2021). A review on genetic algorithm: past, present and future. (Springer).

11. Бондаренко М.В., Бондаренко В.М. // Моделювання та інформаційні системи в економіці : зб. наук. пр. / М-во освіти і науки України, ДВНЗ «Київ. нац. екон. ун-т ім. Вадима Гетьмана» ; [редкол.: В. К. Галіцин (голов. ред.) та ін.]. – Київ : КНЕУ, 2018. – Вип. 95. – С. 44–56.

REFERENCES

1. O Vladyko, D Maltsev, EC Cabana, Ia Shavarskyi, R Dychkovskyi (2022) Formation of the models of mining enterprise management. Natsional'nyi Hirnychyi Universytet. Naukovyi Visnyk, 30–36

2. T.A. Zheldak A.B. Vladyko. Using machine learning to classify the hazard index of tail-

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

ings facilities and their impact on the environment, Information Technology: Computer Science, Software Engineering and Cyber Security, 2024
3. Ivanchenko G., Vashchaev S. (2018) Research on the dynamics of nonlinear interaction of enterprise populations / G. Ivanchenko, S. Vashchaev. Methods of neurofuzzy modeling in economics, No. 7. P. 78-110. http://doi.org/10.33111/nfmte.2018.078
4. Hermes Robles-Berumen and Amelia Zafra and Sebastián Ventura. (2024) A survey of genetic algorithms for clustering: Taxonomy and empirical analysis Sebastián Swarm and Evolutionary Computation, Volume 91, https://doi.org/10.1016/j.swevo.2024.101720
5. Wang, Y.; Chen, C.; Tao, Y.; Wen, Z.; Chen, B.; Zhang, H. (2019). A many-objective optimization of industrial environmental management using NSGA-III: A case of China's iron and steel industry. Applied Energy.
6. Muke, P., Tholana, T., & Musingwini, C. (2021). A genetic algorithm model for optimizing long-term open-pit mine production scheduling. (conference paper / ResearchGate).
7. Vladyko, O., Maltsev, D., Gliwiński, Ł., Dychkovskyi, R., Steculа, K., & Dyczko, A. (2025). Enhancing Mining Enterprise Energy Resource Extraction Efficiency Through Technology Synthesis and Performance Indicator Development. Energies, 18(7), 1641. https://doi.org/10.3390/en18071641
8. Grabovetskyi B.E. Production functions: theory, construction, use in production management. Monograph. Vinnytsia: UNIVERSUM, 2006. 137 p.
9. Casas, N. (2015). Genetic algorithms for multimodal optimization, Neural and Evolutionary Computing.
10. Katoch, S., Chauhan, S. S., & Kumar, V. (2021). A review on genetic algorithm: past, present and future. (Springer).
11. Bondarenko M.V., Bondarenko V.M. // Modeling and information systems in economics: collection of scientific works / Ministry of Education and Science of Ukraine, State Higher Educational Institution "Kyiv. National Economic University named after Vadym Hetman"; [ed.: V. K. Galitsyn (ed. chief) and others]. – Kyiv: KNEU, 2018. – Issue 95. – P. 44–56.

Received 19.06.2026.
Accepted 20.08.2026.
Published 30.10.2026

Genetic algorithms and methods of optimal enterprise management

The article presents the results of a study aimed at optimizing the distribution of capital investments of an industrial enterprise using a genetic algorithm. A mathematical model has been developed in which the shares of investments are distributed among five subsystems – resources, equipment, technology, economy and organization, which are determined by evolutionary search for the optimal financing vector. The fitness function is given in the form of net present value (NPV), which allows directly taking into account the value of money.

The purpose of the work is to develop, implement and analytically evaluate a model for optimizing the distribution of capital investments of an industrial enterprise based on a genetic algorithm, which allows maximizing the net present value (NPV) under conditions of limited resources and nonlinear interaction between production subsystems.

The methodology. The methodological basis of the study is a combination of systems analysis, evolutionary calculations and economic and mathematical modeling, which pro-

vides a comprehensive study of the processes of capital investment distribution in multifactor production systems. The methodology involves multilevel verification of the stability of the model through repeated experiments, comparison with basic linear models and statistical analysis of variations of the results.

The scientific novelty. The scientific novelty of the obtained results lies in the formation of a new evolutionary-optimization paradigm for managing the investment activities of industrial enterprises, which combines genetic algorithms, economic dynamics, and systems analysis.

Conclusions. A model for optimizing the allocation of capital investments based on a genetic algorithm has been developed and tested, which ensures the maximization of net present value (NPV) in a multi-parameter system.

Keywords: genetic algorithm, CAPEX, NPV, optimization, systems analysis, Cobb-Douglas production function, investment strategy, project analysis, evolutionary computing.

Желдак Тімур Анатолійович – д.т.н., доцент, завідувач кафедри системного аналізу та управління Національного технічного університету «Дніпровська політехніка», Україна.
ORCID: https://orcid.org/0000-0002-4728-5889
Владико Олександр Борисович – к.т.н., доцент кафедри системного аналізу та управління Національного технічного університету «Дніпровська політехніка», Україна.
ORCID: https://orcid.org/0000-0001-9779-9565

Zheldak Timur -- Doctor of Technical Sciences, Associate Professor, Head of the System Analysis and Control Department, Dnipro University of Technology, Ukraine.
ORCID: https://orcid.org/0000-0002-4728-5889
Vladyko Alexander - Candidate of Technical Sciences, Associate Professor Dnipro University of Technology, Ukraine.
ORCID: https://orcid.org/0000-0001-9779-9565

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

DOI 10.34185/1562-9945-5-166-2026-16
УДК 004.738.5:004.65

О.А. Гончаров, М.М. Філоненко, Л.В. Васильєва, А.А. Омельчук, К.О. Крігер

АРХІТЕКТУРА ТА ОПТИМІЗАЦІЯ РЕКОМЕНДАЦІЙНОГО МОДУЛЯ ЕЛЕКТРОННОЇ БІБЛІОТЕКИ ЗАКЛАДУ ВИЩОЇ ОСВІТИ В УМОВАХ ОБМЕЖЕНИХ ОБЧИСЛЮВАЛЬНИХ РЕСУРСІВ

Анотація. Робота присвячена розробці та архітектурній оптимізації рекомендаційного модуля для електронних бібліотек закладів вищої освіти, що функціонують в умовах обмежених обчислювальних ресурсів і не мають можливості використовувати ресурсомісткі нейромережеві підходи. Застосовано метод колаборативної фільтрації з інтеграцією експоненційної функції часового згасання. Цей механізм адаптує профіль студента до швидкої зміни інтересів у межах навчальних семестрів. Для подолання квадратичної обчислювальної складності алгоритму розроблено оптимізовану архітектуру реляційної бази даних (на стеку PHP/MySQL). Життєвий цикл формування рекомендацій розділено на дві ізольовані фази: асинхронний фоновий перерахунок матриць та синхронне зчитування готових пропозицій. Емпіричне моделювання на реальній вибірці (до 1350 об'єктів) підтвердило високу стійкість алгоритму до зміни розмірності даних. Доведено, що врахування фактора часового згасання дозволяє знизити середньоквадратичну похибку (RMSE) на понад 42%. Запропонована архітектурна оптимізація забезпечила час відгуку системи на рівні логарифмічної складності для кінцевого користувача. Обґрунтовано доцільність використання базових модифікованих алгоритмів колаборативної фільтрації як компромісного рішення між точністю та обчислювальною вартістю для локальних освітніх платформ порівняно з моделями глибинного навчання.
Ключові слова: алгоритм колаборативної фільтрації; прогнозування; оптимізація обчислень; рекомендаційна система; часове згасання; електронна бібліотека; MySQL; RMSE.

Постановка проблеми. Стрімка цифровізація освітнього процесу супроводжується експоненційним зростанням обсягів навчальних матеріалів: електронних підручників, методичних вказівок, наукових статей та конспектів лекцій. У локальних бібліотеках закладів вищої освіти це призводить до проблеми інформаційного перевантаження студентів та викладачів. Традиційні засоби повнотекстового пошуку та рубрикації виявляються недостатньо ефективними, оскільки вимагають від користувача точного формулювання запиту. Подібні висновки підтверджуються дослідженнями у сфері інтелектуалізації пошуку наукових публікацій, де для підвищення релевантності результатів застосовуються комбінації методів TF-IDF, NLP-аналізу та алгоритмів ранжування документів [1]. За таких умов впровадження персоналізованих рекомендацій-

© Гончаров О.А., Філоненко М.М., Васильєва Л.В., Омельчук А.А., Крігер К.О., 2026
Copyright for this paper by its authors. Use permitted under License CC BY 4.0.
ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

них систем стає критично необхідним інструментом для скорочення часу на пошук релевантного контенту та підвищення загальної якості освітнього процесу [2]. У сучасних дослідженнях інформаційного забезпечення освітньої діяльності також підкреслюється важливість формалізованого подання знань, онтологічного моделювання та автоматизації доступу до освітніх ресурсів, що створює основу для побудови інтелектуальних сервісів персоналізованої підтримки користувачів електронних бібліотек [3].

Аналіз останніх досліджень та публікацій. Сучасний етап розвитку рекомендаційних систем характеризується активним використанням нейромережевих підходів. Зокрема, дослідження останніх років зосереджені на використанні нейронної колаборативної фільтрації, графових нейронних мереж та моделей на основі механізмів уваги [4, 5]. Ці підходи демонструють високу точність прогнозування та здатність виявляти складні нелінійні залежності між користувачами та об'єктами.

Однак, впровадження таких глибоких моделей супроводжується низкою критичних обмежень: вони потребують значної обчислювальної потужності у вигляді графічних прискорювачів GPU, великих обсягів оперативної пам'яті та величезних, щільних масивів даних для навчання. Як зазначається у дослідженнях з оптимізації рекомендаційних систем [6], висока обчислювальна вартість робить нейромережеві підходи ресурсно витратними або технічно неможливими для розгортання на рівні локальних факультетських серверів, інфраструктура яких найчастіше базується на стандартних стеках технологій (наприклад, на основі класичних веб-стеків PHP/MySQL) з обмеженими ресурсами процесора.

Альтернативою залишаються класичні алгоритми колаборативної фільтрації (КФ), які відзначаються алгоритмічною простотою та інтерпретованістю результатів [7]. Основною проблемою базової КФ в освітньому контексті є нездатність адаптуватися до зміни інтересів користувача з часом — наприклад, при переході студента на наступний курс змінюється перелік актуальних для нього дисциплін. Останні дослідження у сфері класичних методів (2023–2026 рр.) доводять, що інтеграція функцій часового згасання дозволяє суттєво підвищити релевантність рекомендацій шляхом надання більшої ваги останнім взаємодіям користувача з системою [8, 9]. Проте питання побудови архітектури рекомендаційних модулів із підтримкою часових коефіцієнтів та оптимізації відповідних обчислень у реляційних базах даних для локальних освітніх платформ досі залишається недостатньо дослідженим.

Мета дослідження. Метою дослідження є обґрунтування та експериментальна перевірка ефективності модифікованого методу колаборативної фільтрації з урахуванням часового фактора для рекомендаційного модуля електронної бібліотеки закладу вищої освіти, оптимізованого для функціонування в умовах обмежених обчислювальних ресурсів.

Викладення основного матеріалу дослідження. В основу розроблюваного модуля покладено алгоритм колаборативної фільтрації, що базується на подібності користувачів. Класичний підхід передбачає обчислення міри близькості інтересів між цільовим користувачем a та іншим користувачем системи u. Для оцінки подібності ви-

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

користовується косинусна міра, оскільки вона є стійкою до розрідженості даних [10, 11]:

$$sim\ (a,u) = \frac{\sum_{i \in I_{a,u}} r_{a,i} \cdot r_{u,i}}{\sqrt{\sum_{i \in I_a} r_{a,i}^2} \cdot \sqrt{\sum_{i \in I_u} r_{u,i}^2}}, \quad (1)$$

де $r_{a,i}$ та $r_{u,i}$ – оцінки, виставлені користувачами *a* та *u* об'єкту (книзі) *i*; $I_{a,u}$ – множина об'єктів, оцінених обома користувачами.

Прогнозована оцінка $P_{(a,i)}$ для цільового користувача *a* та неоціненого ним об'єкта *i* обчислюється як зважена сума оцінок *k* найближчих сусідів (множина *K*):

$$P(a,i) = \frac{\sum_{u \in K} r_{u,i} \cdot sim\ (a,u)}{\sum_{u \in K} sim\ (a,u)}. \quad (2)$$

Специфіка освітнього середовища полягає в тому, що актуальність певних матеріалів (методичних вказівок, конспектів) жорстко прив'язана до поточного семестру та навчальних дисциплін. Відповідно, інтереси користувача швидко застарівають. Щоб компенсувати це явище, до базової моделі було інтегровано експоненційну функцію часового згасання.

Модифікована оцінка $r_{u,i}^*$, яка використовується в обчисленнях замість початкової $r_{u,i}$, визначається за формулою:

$$r_{u,i}^* = r_{u,i} \cdot e^{-\lambda \Delta t_i}, \quad (3)$$

де Δt_i — час, що минув з моменту виставлення оцінки об'єкту *i* до моменту формування рекомендації (в секундах); λ — коефіцієнт згасання, який підбирається емпірично (в межах даного дослідження $\lambda = 10^{-6}$).

Використання цієї функції дозволяє знизити вагу старих оцінок і підвищити релевантність матеріалів, якими студент цікавився останнім часом.

Реалізація алгоритму КФ має алгоритмічну складність $O(|U|^2 \cdot |I|)$, де $|U|$ – кількість користувачів, $|I|$ – кількість об'єктів [12]. Обчислення таких матриць в режимі реального часу в момент запиту користувача є архітектурним антипатерном для систем, що використовують стек PHP/MySQL (типовий для локальних серверів університетів), оскільки це неминуче призводить до тайм-аутів при збільшенні розмірності матриці.

Для подолання цих технічних обмежень було застосовано архітектурний підхід, що базується на поділі відповідальності за команди та запити, а також подійно-орієнтованій моделі обробки даних (рис. 1). Фізична модель реляційної бази даних проєктується з урахуванням проблеми розрідженості даних (система зберігає лише існуючі оцінки, ігноруючи нульові значення для непрочитаних книг) та використовує денормалізацію для забезпечення високої швидкості читання.

Життєвий цикл формування та видачі рекомендацій на базі цієї моделі чітко розділено на дві ізольовані фази.

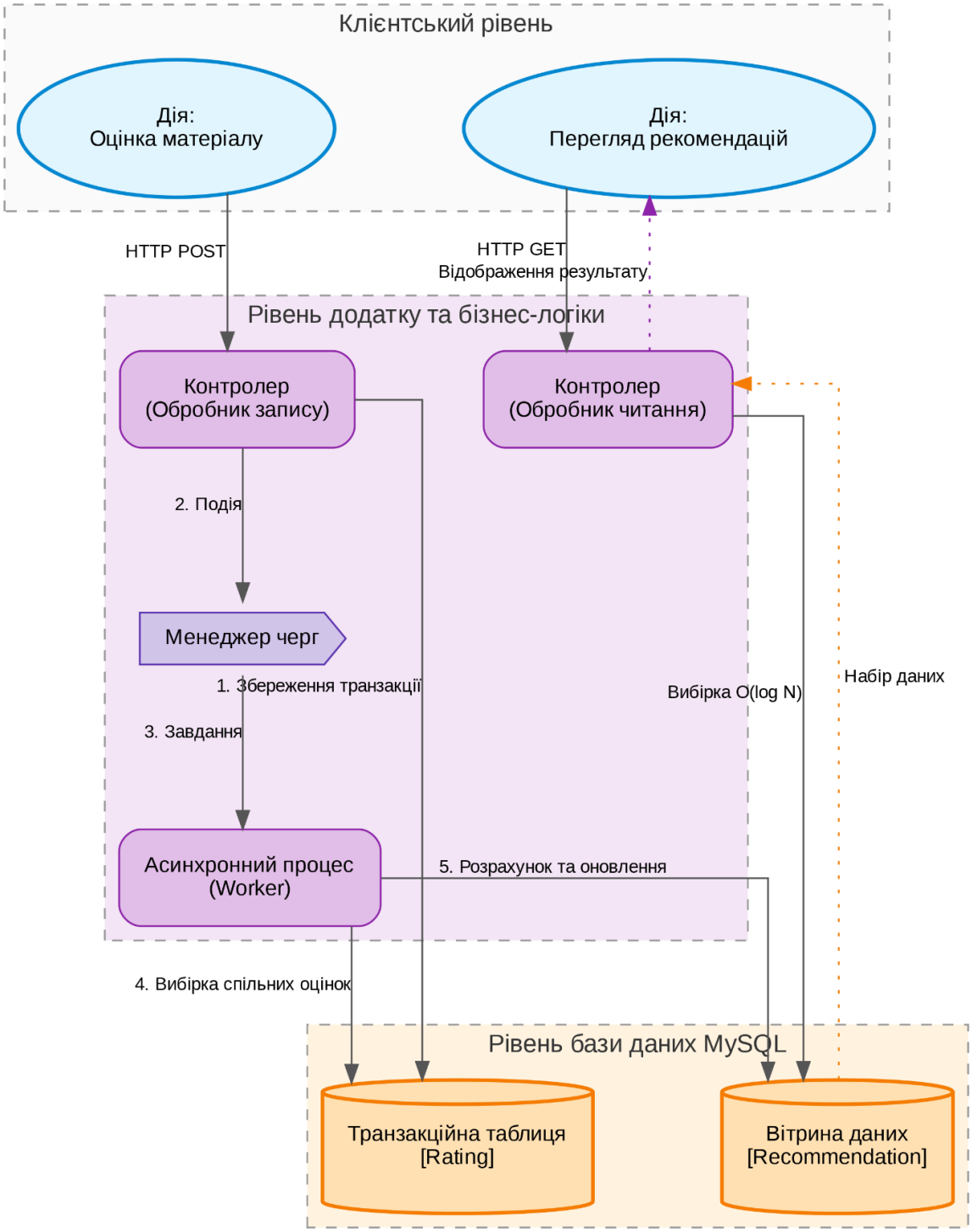


Рисунок 1 – Архітектурна схема модуля формування персональних рекомендацій на базі подійної моделі

Структуру ключових таблиць та стратегію їх індексування наведено у таблиці 1.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Таблиця 1

Структура та індексування ключових таблиць оптимізованої бази даних

Сутність	Ключові атрибути	Тип індексування	Призначення в оптимізованій архітектурі
User	id, education, sphere	B-Tree (id)	Зберігання базового профілю користувача. Виступає головною сутністю для зв'язків *1:N*.
Book	id, author_id, genre_id	B-Tree (id, genre_id)	Метадані навчальних матеріалів. Індекси за жанрами або авторами прискорюють можливу контентну фільтрацію.
Rating	id, user_id, book_id, value, created_at	Composite B-Tree (user_id, book_id)	Транзакційна таблиця зберігання явного фідбеку. Атрибут created_at є критичним для розрахунку функції часового згасання.
Recommendation	id, user_id, book_id, score	Composite B-Tree (user_id, score DESC)	Денормалізована вітрина даних. Забезпечує миттєве зчитування попередньо розрахованих рейтингів.

Першою є фаза запису та асинхронного перерахунку матриць. Процес ініціюється подією виставлення нової оцінки або зміни статусу навчального матеріалу користувачем. Запис миттєво фіксується в таблиці Rating разом із міткою часу. Замість того, щоб одразу перераховувати рекомендації, система генерує фонову задачу, яка делегується менеджеру черг. Робочий процес, що функціонує незалежно від основного життєвого циклу веб-додатка, ініціює сервісний клас RecommendationService. Для оптимізації обчислень косинусної міри подібності сервіс не завантажує всю матрицю в оперативну пам'ять PHP. Замість цього виконується SQL-запит із використанням внутрішнього з'єднання (INNER JOIN) таблиці Rating самої на себе для вибірки лише тих пар користувачів, які мають спільні оцінені книги.

На етапі агрегації до значень оцінок застосовується функція часового згасання (3), де як Δt_i використовується різниця між поточним часом сервера та значенням created_at конкретної оцінки. Обчислені для цільового користувача топ-*N* найбільш релевантних книг (де прогнозована оцінка перевищує заданий поріг) записуються до таблиці Recommendation операцією масового оновлення.

Другою є фаза синхронного читання. При авторизації користувача в системі та переході до розділу персональних рекомендацій веб-додаток повністю звільняється від виконання математичних розрахунків.

Контролер формує простий SQL-запит до таблиці Recommendation, фільтруючи записи за user_id та сортуючи їх за полем score за спаданням.

Завдяки наявності композитного індексу (user_id, score) на рівні СКБД MySQL, пошук потрібних записів відбувається без повного сканування таблиці. Обчислювальна

складність для кінцевого користувача у цій фазі зводиться до $O(\log N)$, де N – кількість записів у таблиці рекомендацій.

Така архітектура дозволяє локальній системі ефективно працювати на малопотужних серверах, гарантуючи час відгуку інтерфейсу в межах мілісекунд, незалежно від загального обсягу накопичених в електронній бібліотеці історичних даних.

Для перевірки ефективності запропонованої оптимізованої архітектури було проведено експериментальне моделювання на базі реальних даних (оцінок користувачів), зібраних у процесі тестової експлуатації електронної бібліотеки. Користувачі оцінювали прочитані матеріали за класичною 5-бальною шкалою, при цьому кожна транзакція супроводжувалася фіксацією часової мітки.

Метою експерименту було порівняння точності прогнозування класичного алгоритму колаборативної фільтрації та її модифікації з урахуванням часового фактора при різних обсягах вибірки оцінених книг (від $k = 50$ до $k = 1350$).

Основним критерієм якості було обрано метрику середньоквадратичної похибки (RMSE). Вибір RMSE зумовлений тим, що ця метрика є стандартом де-факто для задач прогнозування рейтингу і більш чутлива до значних відхилень порівняно з абсолютною похибкою. Формула розрахунку метрики має вигляд:

$$RMSE = \sqrt{\frac{1}{n}\sum_{(u,i)}(P_{u,i} - r_{u,i})^2}, \quad (4)$$

де n – загальна кількість спрогнозованих оцінок, $P_{u,i}$ – спрогнозована оцінка, $r_{u,i}$ – реальна оцінка, виставлена користувачем u об'єкту i.

Чим нижче значення *RMSE*, тим вищою є точність роботи рекомендаційної системи. Результати обчислення *RMSE* для обох алгоритмів при зміні обсягу вибірки подано у таблиці 2 і рис. 2, де наведено ключові точки зрізу експериментальних даних.

Таблиця 2

Динаміка похибки прогнозування залежно від обсягу вибірки

Обсяг вибірки книжок (k)	Класична колаборативна фільтрація (RMSE)	Модифікована КФ із часовим фактором (RMSE)
50	2.1337	1.2470
250	2.1054	1.2367
500	2.0966	1.2040
750	2.0883	1.2165
1000	2.1042	1.2010
1350	2.0923	1.2297
Середнє значення	~2.1025	~1.2152

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

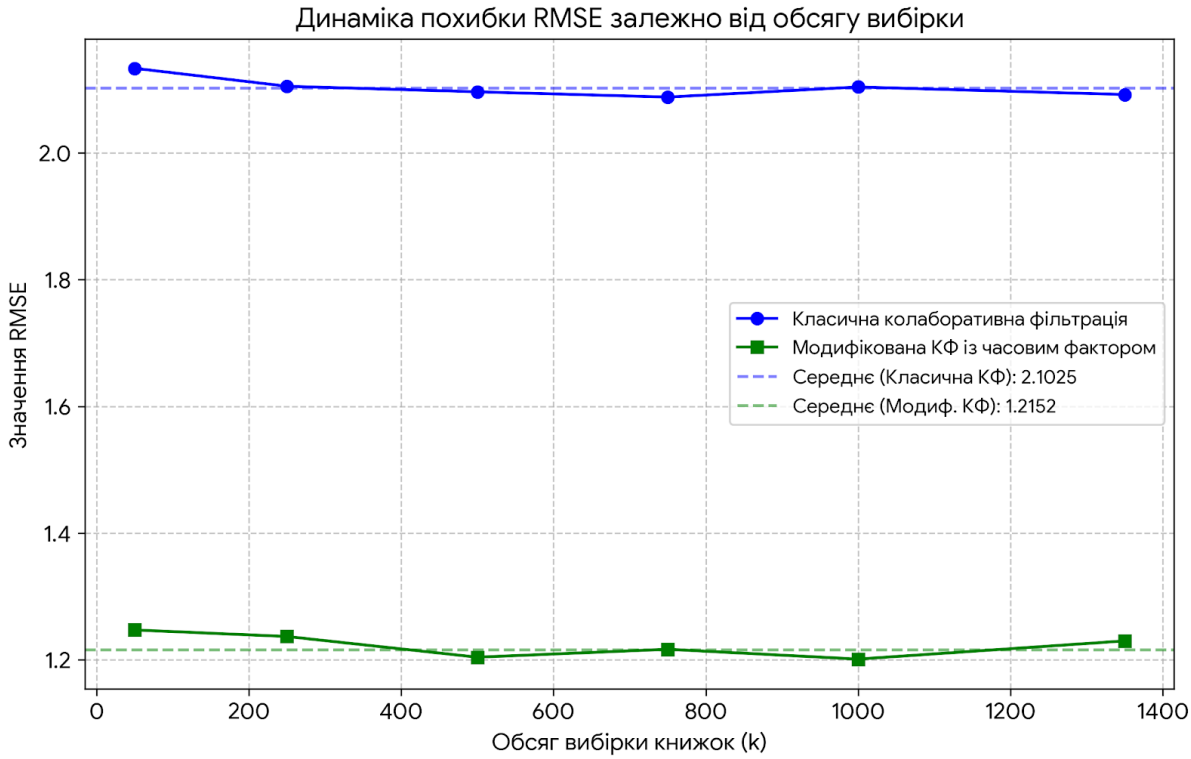


Рисунок 2 – Порівняльний графік динаміки похибки

Аналіз отриманих емпіричних даних дозволяє зробити кілька важливих висновків щодо поведінки розроблених моделей:

1. Стабільність алгоритмів в умовах масштабування: Як видно з результатів, похибка прогнозування для обох методів не має тенденції до стрімкого зростання при збільшенні кількості книжок у вибірці. Розмах варіації (різниця між максимальним та мінімальним значенням RMSE) є незначним: $\Delta \approx 0.045$ для звичайного методу та $\Delta \approx 0.046$ для методу з урахуванням часу. Це свідчить про високу стійкість алгоритму та підтверджує гіпотезу про те, що інтереси користувачів в освітньому середовищі мають кластерну природу (студенти зосереджуються на конкретних предметних областях у межах семестру).

2. Перевага моделі з часовим згасанням: Модифікований алгоритм із врахуванням часового фактора продемонстрував беззаперечну перевагу над базовою моделлю. Середнє значення RMSE знизилося з 2.10 до 1.21 (покращення точності прогнозування на понад 42%). Це емпірично доводить, що для локальних університетських систем впровадження експоненційного зменшення ваги старих оцінок є критично важливим етапом оптимізації. Даний механізм ефективно нівелює шум від інтересів студента на молодших курсах при формуванні пропозицій для його поточного етапу навчання.

Отримані результати підтверджують гіпотезу про те, що інтереси користувачів в електронній бібліотеці закладу вищої освіти мають яскраво виражену часову залежність. Зниження похибки RMSE на 42% при використанні експоненційного часового згасання пояснюється семантикою навчального процесу. Студенти масово звертаються

до певних категорій літератури (наприклад, методичних вказівок або підручників з вузькопрофільних дисциплін) протягом семестру або під час екзаменаційної сесії. Після завершення курсу ці матеріали втрачають для них актуальність.

Класична колаборативна фільтрація продовжує враховувати історичні оцінки з перших курсів, що створює шум та знижує релевантність рекомендацій на старших курсах. Введення фактора часового згасання дозволяє прибирати неактуальні інтереси, адаптуючи профіль користувача до його поточних освітніх потреб.

Попри високу ефективність та низькі вимоги до апаратного забезпечення у фазі читання, запропонована архітектура має об'єктивні обмеження:

1. Проблема на початку функціонування. Оскільки метод базується виключно на матриці оцінок, алгоритм не здатний формувати релевантні пропозиції для нових студентів, які ще не оцінили жодної книги, та не може рекомендувати щойно додані до бібліотеки методичні матеріали, доки вони не отримають перші відгуки.

2. Межі масштабованості асинхронного розрахунку. Хоча фоновий перерахунок за допомогою окремого сервісу розвантажує веб-додаток, при досягненні бібліотекою масштабів десятків тисяч користувачів квадратична складність обчислень на базі реляційної СУБД MySQL стане вузьким місцем. Для систем національного рівня (на відміну від факультетських) цей підхід вимагатиме міграції на графові бази даних або векторні СУБД.

Висновки. У роботі проведена оцінка ефективності модифікованого методу колаборативної фільтрації з часовим згасанням для рекомендаційного модуля електронної бібліотеки закладу вищої освіти, оптимізованого для роботи в умовах обмежених обчислювальних ресурсів.

Доведено, що для вирішення проблеми інформаційного перевантаження студентів використання класичних нейромережевих підходів глибинного навчання не є обов'язковим. Впровадження модифікованої колаборативної фільтрації з функцією експоненційного часового згасання дозволило досягти високих показників точності без залучення графічних прискорювачів. Емпіричне моделювання на реальній вибірці (до 1350 об'єктів) показало зниження середньоквадратичної похибки (RMSE) з 2.10 (базовий алгоритм) до 1.21. Розділення життєвого циклу рекомендацій на асинхронну фазу математичних обчислень та синхронну фазу швидкого читання з денормалізованої таблиці забезпечило миттєвий відгук інтерфейсу.

Напрямом подальших досліджень є розробка гібридної системи – додавання модуля контентної фільтрації на основі спеціальності та курсу студента (метаданих), що дозволить вирішити проблему для нових користувачів.

ЛІТЕРАТУРА

1. Tarasov O., Vasylieva L., Turlakova S., Sahaida P., Karpov A. Improving the Quality of the Relevance of the Search for Scientific Publications Based on a Combination of Ranking Methods. In: Proceedings of the 11-th International Conference "Information Control Systems & Technologies". 2023. CEUR Workshop Proceedings, Vol. 3513. URL: https://ceur-ws.org/Vol-3513/paper36.pdf.

2. Mitreva E., Paneva-Marinova D., Georgiev V., Nikolova A., Pavlov R. A Hybrid Ap-

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

proach for Personalized and Intelligent Content Recommendation in Digital Libraries. Applied Sciences. 2026. Vol. 16, no. 6. Article 2756. DOI: 10.3390/app16062756.
3. Tarasov O., Sahaida P., Podlesny S., Vasylieva L. Categorical-ontological approach to information support of educational activities. In: SHS Web of Conferences. 2022. Vol. 142. Article 03005. DOI: 10.1051/shsconf/202214203005.
4. Krishnamoorthi S., Shyam G. K. Deep Learning-Based Recommender Systems—A Systematic Review and Future Perspective. In: Intelligent Data Engineering and Analytics. FICTA 2023. Smart Innovation, Systems and Technologies. 2023. Vol. 371.
DOI: 10.1007/978-981-99-6706-3_33.
5. Scrivano A. A comparative study of recommender Systems under Big Data constraints. arXiv. 2025. arXiv:2504.08457. DOI: 10.48550/arXiv.2504.08457.
6. Hartatik, Heryawan L., Pulungan R. Trust Decay-Based Temporal Learning for Dynamic Recommender Systems With Concept Drift Adaptation. IEEE Access. 2025. Vol. 13. P. 110955-110971. DOI: 10.1109/ACCESS.2025.3583186.
7. Івохін Є., Шелякін Г. Колаборативна фільтрація як один із основних методів рекомендацій контенту: основні проблеми та способи їх вирішення. Modern Engineering and Innovative Technologies. 2024. Vol. 1, no. 34-01. P. 94–105. DOI: 10.30890/2567-5273.2024-34-00-013.
8. Wang Y. Time-weighted collaborative filtering recommendation algorithm based on large language models. In: Sixth International Conference on Signal Processing and Computer Science (SPCS 2025). 2025. P. 59. DOI: 10.1117/12.3089553.
9. Хабарлак К., Ширяєв А. Порівняльний аналіз моделей машинного навчання для побудови системи рекомендацій з динамічним перенавчанням. Information Technology: Computer Science, Software Engineering and Cyber Security. 2025. № 4. С. 267–273.
DOI: 10.32782/it/2025-4-31.
10. Singhal A. Modern Information Retrieval: A Brief Overview. Bulletin of the IEEE Computer Society Technical Committee on Data Engineering. 2001. Vol. 24, no. 4. P. 35–43. URL: https://www.cs.columbia.edu/~gravano/Qual/Papers/singhal.pdf.
11. Альперт С. І. Основні міри подібності та нові підходи до їх застосування при класифікуванні гіперспектральних космічних зображень. Математичні машини і системи. 2019. № 1. С. 143–151. URL:
https://www.immsp.kiev.ua/publications/articles/2019/2019_1/01_2019_Alpert.pdf.
12. Novotný V. Implementation Notes for the Soft Cosine Measure. In: The 27th ACM International Conference on Information and Knowledge Management. 2018. P. 1639–1642. DOI: 10.1145/3269206.3269317.

REFERENCES

1. Tarasov, O., Vasylieva, L., Turlakova, S., Sahaida, P., & Karpov, A. (2023). Improving the quality of the relevance of the search for scientific publications based on a combination of ranking methods. In *Proceedings of the 11-th International Conference "Information Control Systems & Technologies"* (CEUR Workshop Proceedings, Vol. 3513). CEUR-WS.org. https://ceur-ws.org/Vol-3513/paper36.pdf
2. Mitreva, E., Paneva-Marinova, D., Georgiev, V., Nikolova, A., & Pavlov, R. (2026). A hybrid approach for personalized and intelligent content recommendation in digital libraries.

Applied Sciences, 16(6), Article 2756. https://doi.org/10.3390/app16062756
3. Tarasov, O., Sahaida, P., Podlesny, S., & Vasylieva, L. (2022). Categorical-ontological approach to information support of educational activities. *SHS Web of Conferences*, 142, Article 03005. https://doi.org/10.1051/shsconf/202214203005
4. Krishnamoorthi, S., & Shyam, G. K. (2023). Deep learning-based recommender systems—A systematic review and future perspective. In V. Bhateja, F. Carroll, J. M. R. S. Tavares, S. S. Sengar, & P. Peer (Eds.), *Intelligent Data Engineering and Analytics. FICTA 2023* (Vol. 371). Springer. https://doi.org/10.1007/978-981-99-6706-3_33
5. Scrivano, A. (2025). A comparative study of recommender systems under Big Data constraints [Preprint]. *arXiv*. https://doi.org/10.48550/arXiv.2504.08457
6. Hartatik, Heryawan, L., & Pulungan, R. (2025). Trust decay-based temporal learning for dynamic recommender systems with concept drift adaptation. *IEEE Access*, 13, 110955-110971. https://doi.org/10.1109/ACCESS.2025.3583186
7. Ivokhin, Ye., & Sheliakin, H. (2024). Kolaboratyvna filtratsiia yak odyn iz osnovnykh metodiv rekomendatsii kontentu: osnovni problemy ta sposoby yikh vyrishennia [Collaborative filtering as one of the main methods of content recommendation: main problems and ways to solve them]. *Modern Engineering and Innovative Technologies*, 1(34-01), 94–105. https://doi.org/10.30890/2567-5273.2024-34-00-013 [in Ukrainian]
8. Wang, Y. (2025). Time-weighted collaborative filtering recommendation algorithm based on large language models. In P. T. Mathiopoulos & Y. Feng (Eds.), *Sixth International Conference on Signal Processing and Computer Science (SPCS 2025)* (p. 59). https://doi.org/10.1117/12.3089553
9. Khabarlak, K., & Shyriaiev, A. (2025). Porivnialnyi analiz modelei mashynnoho navchannia dlia pobudovy systemy rekomendatsii z dynamichnym perenavchanniam [Comparative analysis of machine learning models for building a recommendation system with dynamic retraining]. *Information Technology: Computer Science, Software Engineering and Cyber Security*, (4), 267–273. https://doi.org/10.32782/it/2025-4-31 [in Ukrainian]
10. Singhal, A. (2001). Modern information retrieval: A brief overview. *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering*, 24(4), 35–43. https://www.cs.columbia.edu/~gravano/Qual/Papers/singhal.pdf
11. Alpert, S. I. (2019). Osnovni miry podibnosti ta novi pidkhody do yikh zastosuvannia pry klasyfikuvanni hiperspektralnykh kosmichnykh zobrazhen [Basic similarity measures and new approaches to their application in classification of hyperspectral space images]. *Matematychni Mashyny i Systemy*, (1), 143–151.
https://www.immsp.kiev.ua/publications/articles/2019/2019_1/01_2019_Alpert.pdf [in Ukrainian]
12. Novotný, V. (2018). Implementation notes for the soft cosine measure. In *The 27th ACM International Conference on Information and Knowledge Management* (pp. 1639–1642). Association for Computing Machinery. https://doi.org/10.1145/3269206.3269317

Received 19.06.2026.
Accepted 20.08.2026.
Published 30.10.2026

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Architecture and optimization of the recommendation module of a higher education institution's electronic library under limited computational resources

The rapid digitalization of education inevitably causes information overload for students. While modern recommender systems rely on deep learning, these neural approaches are highly resource-intensive. They require GPU accelerators and extensive memory, making them unfeasible for local university servers typically operating on standard PHP/MySQL stacks. This paper presents the development and architectural optimization of a recommendation module tailored for higher education digital libraries functioning under strict computational constraints.

To address this, a user-based collaborative filtering method was applied, integrated with an exponential time decay function. This mathematical modification effectively mitigates "concept drift" by adapting the user profile to rapid shifts in academic interests across semesters, assigning significantly lower weights to outdated historical interactions.

To overcome the inherent quadratic computational complexity of classical algorithms, a highly optimized relational database architecture was engineered. The recommendation lifecycle was divided into two isolated phases. The first phase involves the asynchronous background recalculation of similarity matrices via queue managers, preventing real-time server overload. The second phase ensures the synchronous retrieval of pre-calculated suggestions from a denormalized data mart. By utilizing composite B-Tree indexing, the query response complexity for the end-user was reduced to a logarithmic level.

Empirical modeling on a real dataset (up to 1350 objects) confirmed the algorithm's high robustness. Experimental results proved that incorporating the time decay factor reduces the Root Mean Square Error (RMSE) from 2.10 to 1.21, an accuracy improvement of over 42%. The study substantiates that modified classical algorithms offer an optimal trade-off between prediction accuracy and computational cost for local educational platforms.

Keywords: collaborative filtering algorithm; prediction; computational optimization; recommender system; temporal decay; digital library; MySQL; RMSE.

Гончаров Олександр Андрійович – д.т.н., професор, проф. кафедри комп’ютерних та інформаційних технологій і систем Державного податкового університету.
ORCID: https://orcid.org/0000-0002-9275-0073

Філоненко Михайло Миколайович – к.ф.-м.н., доцент, зав. кафедри комп’ютерних та інформаційних технологій і систем Державного податкового університету.
ORCID: https://orcid.org/0000-0002-2938-3725

Васильєва Людмила Володимирівна – к.т.н., доцент, доцент кафедри комп'ютерних наук Київського національного економічного університету імені Вадима Гетьмана.
ORCID: https://orcid.org/0000-0002-9277-1560

Омельчук Антон Анатолійович – к.т.н., доцент кафедри комп’ютерних та інформаційних технологій і систем Державного податкового університету.
ORCID: https://orcid.org/0000-0001-6318-7464

Крігер Карина Олександрівна – магістр з комп'ютерних наук, Донбаська державна машинобудівна академія.
ORCID: https://orcid.org/0009-0001-9238-5635

Goncharov Alexander – doctor of physical and mathematical sciences, professor, professor of department of computer and information technologies and Systems, State Tax University.
ORCID: https://orcid.org/0000-0002-9275-0073
Filonenko Mykhaylo – candidate of physical and mathematical sciences, associate professor, head of the department of computer and information technologies, State Tax University.
ORCID: https://orcid.org/0000-0002-2938-3725
Vasylieva Liudmyla – candidate of technical sciences, associate professor, associate professor of the department of computer science, Kyiv National Economic University named after Vadym Hetman.
ORCID: https://orcid.org/0000-0002-9277-1560
Omelchuk Anton – Candidate of Technical Sciences, Associate Professor of Department of Computer and Information Technologies and Systems, State Tax University, ORCID: https://orcid.org/0000-0001-6318-7464
Kriher Karyna – Master of Computer Science, Donbas State Machine-Building Academy, ORCID: https://orcid.org/0009-0001-9238-5635

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

DOI 10.34185/1562-9945-5-166-2026-17
УДК 681.5:697.1

О.В. Бондаренко, А.В. Бубліков

АНАЛІЗ СУЧАСНИХ ПІДХОДІВ ДО ЕНЕРГОЕФЕКТИВНОГО ПЛАНУВАННЯ ВИКОРИСТАННЯ ПРИМІЩЕНЬ БУДІВЕЛЬ

Анотація. Підвищення енергетичної ефективності будівель потребує узгодження роботи систем опалення, вентиляції та кондиціювання з просторово-часовою організацією використання приміщень. Метою статті є систематизація сучасних підходів до врахування зайнятості, теплових режимів і характеристик приміщень та обґрунтування постановки задачі енергоефективного планування їх використання. Проаналізовано наукові публікації останніх років за напрямами визначення й прогнозування зайнятості приміщень, енергетичного моделювання, прогнозного керування мікрокліматом, оптимізації розкладів та гнучкого використання простору. Показано, що наявні рішення переважно вдосконалюють окремі функції: формують температурні уставки за заданим розкладом, оцінюють зайнятість або оптимізують розклад без індивідуалізованого прогнозу витрат на опалення. За результатами контент-аналізу статей встановлено високу представленість рішень щодо кількісної зайнятості приміщень і значно нижчу представленість щодо перехідних теплових режимів, індивідуальних енергетичних профілів та функціонального статусу приміщень; критерій питомих енерговитрат на людино-годину у вибірці взагалі не використано. Обґрунтовано доцільність інтегрованої системи підтримки прийняття рішень, яка поєднує прогноз енерговитрат, організаційні обмеження та координацію локальних систем керування. Сформульовано мету подальших досліджень, комплекс завдань, очікувані результати й напрями практичного застосування.
Ключові слова: енергетична ефективність, використання приміщень, планування, зайнятість, електричне опалення, прогнозування енерговитрат, система підтримки прийняття рішень, енергетичний профіль приміщення, людино-година.

Постановка проблеми. Будівлі є складними організаційно-технологічними об’єктами, у яких енергоспоживання визначається не лише характеристиками огороджувальних конструкцій та інженерного обладнання, а й фактичним способом використання простору. Для адміністративних, освітніх і громадських будівель характерні нерівномірна зайнятість приміщень, зміни складу груп, перерви між заходами, наявність однофункціональних і багатофункціональних приміщень. За таких умов однакова послуга – забезпечення нормативного мікроклімату для певної кількості людей протягом заданого часу – може потребувати різних витрат енергії залежно від обраного приміщення, моменту початку його використання та попереднього теплового стану.

© Бондаренко О.В., Бубліков А.В., 2026
Copyright for this paper by its authors. Use permitted under License CC BY 4.0.
ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Традиційна логіка експлуатації розділяє організаційне планування та автоматичне керування мікрокліматом. Розклад формують за часовими, місткісними й функціональними обмеженнями, після чого локальні системи опалення, вентиляції та кондиціювання (HVAC) відпрацьовують задані температурні режими. Навіть якщо керування враховує прогноз погоди або присутність людей, сам план використання приміщень залишається зовнішньою заданою величиною. Унаслідок цього не використовується можливість зменшити кількість прогрівань, сконцентрувати заходи в енергетично ефективних приміщеннях або вибрати простір, теплова динаміка якого краще відповідає тривалості й чисельності запланованої події.

Проблема особливо суттєва для будівель з електричним опаленням і зональним регулюванням. Перехід приміщення з економного до комфортного режиму створює додаткове навантаження, величина якого залежить від зовнішньої температури, потужності нагрівача, теплової інерції та тривалості простою. Отже, оцінювання лише усталених витрат або використання нормативного питомого показника на площу не дає достатньої інформації для порівняння альтернативних варіантів «група – приміщення – час». Потрібен прогноз повних витрат енергії на опалення, що охоплює попереднє прогрівання, підтримання температури під час зайнятості приміщення та залишкові теплові процеси після завершення використання.

Водночас організаційна ефективність не може бути зведена до мінімізації сумарної енергії будь-якою ціною. Рішення має зберігати місткість, функціональну придатність приміщень, допустимі часові вікна, вимоги до комфорту та можливість реалізації сформованого плану локальними системами керування. Тому задача має міжрівневий характер: на рівні планування визначають просторово-часове розміщення заходів, а на рівні оперативного керування забезпечують необхідну теплову підготовку приміщень і підтримання заданих параметрів мікроклімату.

Таким чином, актуальною є науково-прикладна задача переходу від фрагментарного врахування зайнятості приміщень в алгоритмах керування до інтегрованого планування використання приміщень за прогнозованими енергетичними наслідками. Її розв’язання потребує критичного зіставлення сучасних підходів, визначення функцій, які вже забезпечуються наявними системами, і встановлення складових, що мають бути об’єднані в єдиному контурі підтримки рішень.

Аналіз останніх досліджень і публікацій. Сучасні дослідження енергоефективності будівель розвиваються за кількома взаємопов’язаними напрямами. Огляд керування, орієнтованого на користувача (occupant-centric), показує поступовий перехід від простого реагування на факт присутності до прогнозного керування, що враховує чисельність людей, індивідуальні вимоги до комфорту та невизначеність поведінки [1]. Огляд технологій прогнозування зайнятості приміщень дозволяє зробити висновок про систематичне застосування датчиків, камер, даних бездротових мереж Wi-Fi і Bluetooth та методів машинного навчання [2]. Технології оптимізації енергоспоживання будівель на основі даних об’єднують прогнозування, цифрові моделі, оптимізацію й автоматизацію, однак дослідники відзначають труднощі інтеграції різнорідних даних і перенесен-

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

ня моделей між будівлями [3]. Для освітніх будівель окремо підкреслюється необхідність одночасного врахування технологічних рішень і поведінки користувачів [4].

До першого функціонального класу підходів до енергоефективного використання приміщень належать методи та засоби моніторингу і прогнозування їх зайнятості. У роботі [5] дані зайнятості приміщень офісів на тривалих часових інтервалах використано для виявлення змін режимів експлуатації та побудови прогнозів. Визначення зайнятості окремих приміщень за Wi-Fi probe-запитами та методами машинного навчання досліджено в [6]. Гібридні моделі, що поєднують гаусівські процеси з механістичним описом енергетичних процесів у будівлі, використано для прогнозування зайнятості приміщень у роботі [7]. Практична цінність цих підходів полягає в перетворенні зайнятості приміщень з нормативного профілю на вимірюваний або прогнозований сигнал. Проте результатом є оцінка присутності чи кількості людей, а не енергоефективний план їх розміщення.

Другий клас охоплює керування HVAC з урахуванням зайнятості приміщень. Експериментальне порівняння стратегій, орієнтованих на користувача, підтверджує можливість зменшення енергоспоживання без погіршення якості повітря та теплового комфорту [8]. У [9] ефективність системи керування на основі зайнятості приміщень оцінено в реальній офісній будівлі. Робота [10] показує, як дані інтелектуального лічильника і Wi-Fi можуть використовуватися для відновлення фактичних графіків роботи HVAC, а [11] поєднує прогноз внутрішніх теплонадходжень з оптимізацією за температурною уставкою. Спільна симуляція EnergyPlus–MATLAB із застосуванням рекурентної нейронної мережі з довгою короткочасною пам’яттю (LSTM) демонструє технічну можливість реалізації прогнозного керування у замкненому контурі [12]. У всіх цих випадках керованими величинами є температурні уставки, режими або потужність обладнання, тоді як розподіл заходів між приміщеннями задається наперед і не оптимізується.

У задачах періодичного опалення принципово важливим є врахування перехідних процесів. Метод оцінювання часу попереднього прогрівання з урахуванням невизначеностей моделі розглянуто в [13]. Він показує, що момент початку підготовки приміщення не можна коректно визначити лише за календарним правилом. Тривалість прогрівання та додаткова енергія залежать від поточного теплового стану і зовнішніх умов. Разом з тим навіть точне прогнозне керування локальною зоною не відповідає на організаційне питання: чи доцільно взагалі активувати цю зону, якщо захід може бути перенесений до іншого придатного приміщення.

Третій клас пов’язаний з енергетичним моделюванням і сценарним аналізом. Реалістичні профілі зайнятості приміщень суттєво змінюють оцінку енергоспоживання будівельних архетипів порівняно зі стандартними графіками [14]. Калібрована модель кампусу в [15] використана для аналізу гнучких режимів роботи та впливу керування, орієнтованого на зайнятість приміщень. Дослідження [16] демонструє, що планувальні рішення і конфігурація внутрішнього простору будівлі впливають на поведінку користувачів та енергетичні результати. Новіші роботи розглядають оптимізацію зонування гнучких офісів [17] і багатокритеріальне формування конфігурації гнучких

просторів [18]. Ці результати розширюють об’єкт оптимізації від обладнання до простору, проте часто вимагають детальної моделі всієї будівлі та не формують простого індивідуального профілю приміщення, придатного для оперативного перерахунку великої кількості варіантів розкладу.

Четвертий клас становлять енергоорієнтовані задачі формування розкладу. В [19] оптимізовано розклад занять з урахуванням енергоефективності будівлі та змінного теплового сприйняття студентів. Платформа спільної симуляції будівлі й алгоритму планування застосована для оптимізації використання аудиторій і приміщень самостійної роботи [20]. У моделі [21] енергетичні наслідки університетського розкладу досліджено з урахуванням клімату, параметрів будівлі та освітніх обмежень. Робота [22] поєднує оптимізацію розкладу з оцінюванням теплових відчуттів студентів. Ці праці доводять, що розклад може бути самостійним керувальним впливом, а не лише вхідними даними для HVAC.

Найближчими до задачі енергоефективного використання простору є підходи, які безпосередньо оптимізують розміщення людей. У роботі [23] запропоновано підхід до оптимізації використання простору будівлі, який ґрунтується на об’єднанні даних кількох сенсорів і спільному врахуванні поведінки користувачів та параметрів внутрішнього середовища. У [24] кількісно досліджено вплив числа людей на енергоспоживання будівлі з високою щільністю зайнятості. Однак наявність кількості людей у моделі ще не означає, що ефективність оцінюється відносно корисного результату використання приміщення. У проаналізованих постановках домінують сумарна енергія, вартість або комфорт, тоді як питомі витрати на забезпечену людино-годину не використовуються як критерій вибору плану.

Синтез результатів дає підстави розрізняти чотири функціональні рівні: моніторинг та ідентифікацію; моделювання і прогнозування; оперативне керування; планування та підтримку рішень. Усередині кожного рівня спостерігається розвиток функціональної повноти: від фіксованих графіків до прогнозу, від ізольованого контролю обладнання до координації зон, від традиційного розкладу до енергоорієнтованого планування. Проте між рівнями зберігаються розриви. Засоби моніторингу передають дані моделям і контролерам, а енергетичні моделі можуть оцінювати сценарії планування, але цілісний замкнений обмін між СППР і локальними системами керування представлений обмежено. Узагальнену карту підходів наведено на рис. 1.

Положення блоків на рис. 1 має категоріальний, а не метричний зміст: горизонтальна вісь відображає домінуючу функцію підходу, вертикальна – ступінь інтеграції енергетичних процесів із використанням приміщень. Суцільні стрілки показують розширення функцій у межах класу; штрихові – інформаційний або модельний зв’язок; двобічна стрілка – замкнений обмін між формуванням рішень і їх оперативною реалізацією. Така інтерпретація не ототожнює локальне прогнозне керування з плануванням усієї будівлі. Прогнозне керування HVAC оптимізує тепловий режим заданої зони, тоді як координоване оперативне керування має реалізовувати план, сформований СППР, і повертати фактичні дані для його уточнення.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Окремого пояснення потребує зв’язок енергетичного моделювання з енергоорієнтованим плануванням. Модель є не самостійною гілкою розвитку, а джерелом прогнозованої енергооцінки альтернатив. Саме тому на рисунку цей зв’язок подано інформаційною дугою. Аналогічно аналітика зайнятості приміщень може підтримувати як прогнозне керування, так і планування. Однак жоден з цих інформаційних потоків сам по собі не утворює СППР: потрібні цільовий критерій, організаційні обмеження, механізм порівняння допустимих планів і інтерфейс взаємодії з локальними контролерами.

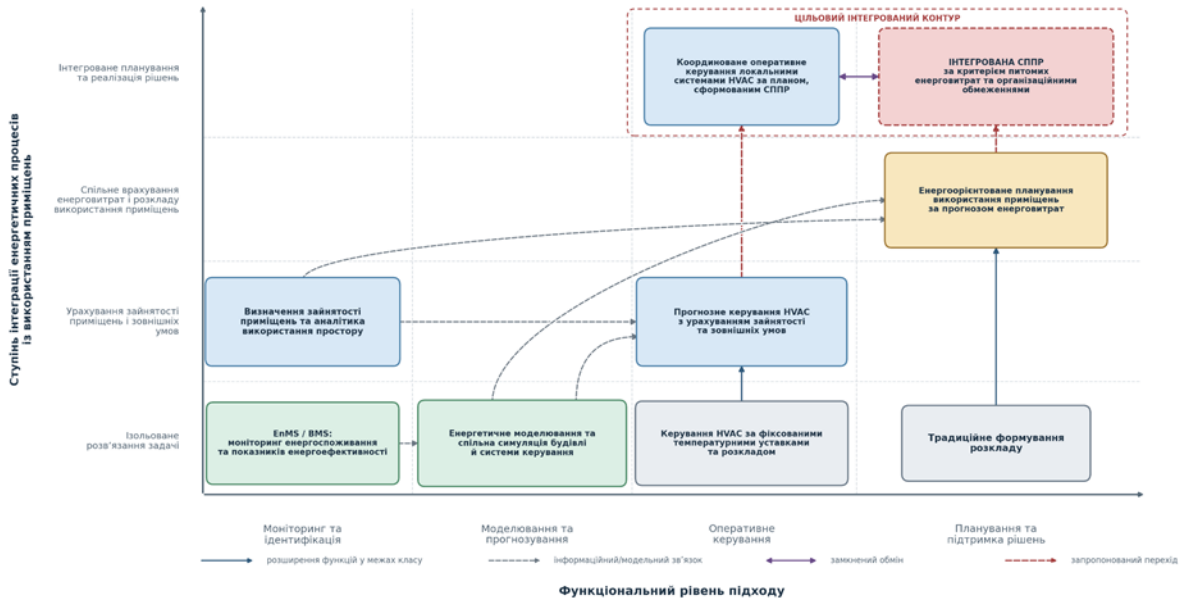


Рисунок 1 – Позиціонування та напрями розвитку підходів до енергоефективного використання приміщень в будівлі

Для кількісного уточнення виявленої прогалини проведено контент-аналіз 20 унікальних первинних наукових статей 2022–2025 років [5–24]. Оглядові праці [1–4] використано для формування класифікації та пошуку первинних досліджень, але не включено до знаменника, щоб не дублювати результати робіт, уже охоплених оглядами. Для кожного чинника застосовано шкалу: повне врахування – 1; часткове – 0,5; відсутність – 0. Індекс представленості визначено як відношення суми оцінок до кількості первинних статей, помножене на 100 %. Така процедура характеризує не якість окремої статті, а частоту інтегрального врахування відповідного чинника у вибірці.

Результати, наведені на рис. 2, свідчать про нерівномірність урахування досліджуваних чинників у сучасних постановках задач енергоефективного використання приміщень. Кількісна зайнятість приміщень має індекс 80,0 %, оскільки число людей або профіль зайнятості приміщень широко застосовують у моделях і контролерах. Температурний режим враховано на рівні 42,5 %, а формування плану використання приміщень – 32,5 %. Перехідні теплові режими представлені на 25,0 %, індивідуалізований енергетичний профіль приміщення – на 20,0 %. Функціональний статус приміщення як керована ознака планування має лише 5,0 %. Питомі енерговитрати на людино-годину не використано в жодній з 20 первинних праць у якості цільового критерію планування.

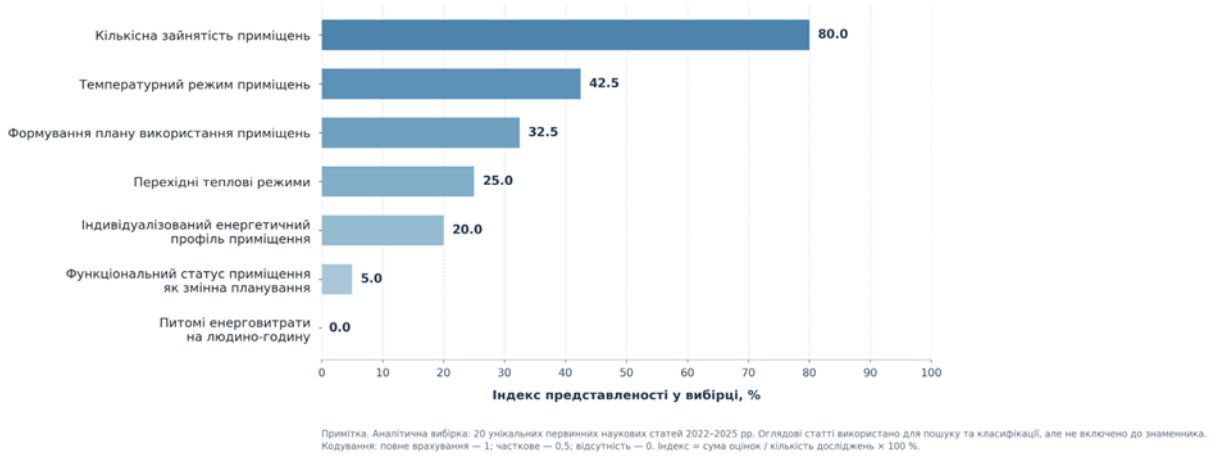


Рисунок 2 – Представленість чинників енергоефективного планування у вибірці сучасних досліджень

Нульове значення останнього показника не означає, що дослідники ігнорують користувачів: навпаки, зайнятість приміщень є найбільш представленим чинником. Прогалина полягає в іншому – кількість людей переважно використовується як збурення, джерело теплонадходжень або умова комфорту, а не як нормувальна величина корисного результату експлуатації. Це розмежування принципове. Два плани з близьким сумарним споживанням можуть суттєво відрізнятися за кількістю забезпечених людино-годин, а отже, за ефективністю використання енергії та простору.

Так само низька представленість функціонального статусу приміщення не означає відсутності організаційних обмежень у задачах розкладу. Університетські моделі детально враховують місткість, доступність і сумісність занять з аудиторіями [19–22]. Проте статус «однофункціональне – багатофункціональне» рідко розглядається як змінна політики використання будівлі, яка визначає допустиму гнучкість перерозподілу груп. Для практичної СППР це важливо: система не повинна пропонувати формально енергоощадне, але організаційно неприйнятне переміщення груп людей.

Отже, невирішеною залишається задача інтегрованого планування, у якій для кожного допустимого призначення приміщення прогнозуються індивідуальні витрати енергії з урахуванням початкового теплового стану та переходів між режимами; енергетичний результат співвідноситься з кількістю забезпечених людино-годин; організаційні обмеження визначають допустиму множину планів; сформований план передається локальним системам HVAC для координованої реалізації. Саме ця сукупність ознак, а не окремий алгоритм прогнозування або оптимізації, визначає новий напрям досліджень.

Мета досліджень. Метою досліджень є підвищення енергетичної ефективності використання приміщень будівель шляхом обґрунтування принципів інтегрованого планування їх просторово-часової зайнятості за прогнозованими питомими витратами електричної енергії на опалення та організаційними обмеженнями.

Досягнення мети потребує переходу від оцінювання окремих режимів HVAC до системної постановки, в якій план використання приміщень розглядається як керований

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

фактор енергоспоживання, а локальні системи керування – як виконавчий рівень сформованих рішень. При цьому не йдеться про заміну автоматичного регулювання організаційним плануванням: обидва рівні мають функціонувати узгоджено і обмінюватися прогнозними та фактичними даними.

Викладення основного матеріалу дослідження. Для формалізації невирішеної задачі будівлю доцільно подати як множину приміщень $R = \{r_1, \ldots, r_m\}$, кожне з яких характеризується місткістю, функціональним статусом, допустимими видами використання, параметрами локальної системи опалення та індивідуальною енергетичною моделлю. На горизонті планування задано множину заходів $A = \{a_1, \ldots, a_n\}$. Для кожного заходу відомі кількість учасників, тривалість, допустимі часові інтервали, вимоги до обладнання й мікроклімату. Рішення X визначає відповідність «захід – приміщення – час» і пов'язаний з нею графік переходів локальних систем між економним і комфортним температурними режимами.

Допустимість плану визначається організаційними обмеженнями. До них належать неперетин заходів у одному приміщенні; достатня місткість; відповідність функціонального призначення й оснащення; доступність приміщення і учасників; обмеження на переміщення; мінімальні інтервали між заходами; збереження закріплення однофункціональних приміщень. Частина обмежень є жорсткою і не може порушуватися, інша може задаватися штрафами або пріоритетами. Такий поділ потрібен, щоб енергетичний критерій не витісняв основне призначення будівлі.

Енергетичну оцінку альтернативи слід формувати на рівні окремого приміщення. Для призначення заходу a_i у приміщення r_j прогнозована енергія опалення має залежати від початкової внутрішньої температури, прогнозу зовнішньої температури, тривалості попереднього прогрівання, заданої комфортної уставки, кількості людей, внутрішніх теплонадходжень, теплової інерції та потужності нагрівача. Вхідним впливом динамічної моделі є електрична або теплова потужність, а витрати енергії визначають інтегруванням потужності на часовому інтервалі. Це термінологічне розмежування запобігає некоректному використанню енергії як миттєвого входу моделі.

Індивідуалізований енергетичний профіль приміщення має забезпечувати швидке оцінювання великої кількості сценаріїв. На відміну від нормативного показника кВт·год/м², такий профіль повинен відображати реальну динаміку конкретного об'єкта: кількість людей в приміщенні, швидкість нагрівання, теплові втрати за різних зовнішніх умов, витрати в усталеному комфортному режимі та енергію переходу після простою. Джерелом параметрів можуть бути дані активного експерименту, коли локальний регулятор формує тестову зміну потужності або уставки, а система реєструє реакцію температури, зовнішні умови та споживання.

Для плану X сумарна прогнозована енергія опалення $E_h(X)$ визначається сумою витрат усіх активованих приміщень на горизонті прогнозування. Корисний результат використання будівлі доцільно подати кількістю забезпечених людино-годин $H(X)$, тобто сумою добутків фактичної кількості учасників на тривалість їх перебування у комфортних умовах. Тоді питомий показник енергетичної ефективності має вигляд:

$$I_e(X) = E_h(X) / H(X) \rightarrow min.$$

Якщо множина заходів, чисельність груп і тривалість використання суворо фіксовані, $H(X)$ є сталою, і мінімізація $I_e(X)$ математично еквівалентна мінімізації сумарної енергії. У цьому випадку питомий показник не змінює оптимальний план, але надає нормовану метрику для порівняння будівель, днів та рівнів завантаження. Якщо ж допускаються альтернативні формати проведення заходів, неповна фактична явка, об'єднання груп або різна тривалість використання, знаменник змінюється і критерій набуває самостійного оптимізаційного змісту. Тому його застосування має супроводжуватися явним визначенням множини дозволених організаційних рішень.

Пропонована постановка не повинна заохочувати штучне збільшення щільності людей або скорочення комфортного періоду. Температурні межі, якість повітря, максимальна заповнюваність і нормативні вимоги мають бути обмеженнями. За необхідності задача може бути багатокритеріальною: поряд з питомими енерговитратами враховують відхилення від базового розкладу, кількість переміщень, ризик порушення комфорту та стійкість плану до похибок прогнозу. Вибір вагових коефіцієнтів повинен бути прозорим для особи, яка приймає рішення.

Інтегрований контур має працювати циклічно. Системи енергетичного менеджменту та диспетчеризації будівлі (EnMS/BMS) надають вимірювання температури, потужності й фактичної зайнятості. Модуль ідентифікації оновлює параметри енергетичних профілів. Модуль прогнозування формує оцінки зовнішніх умов, зайнятості й енерговитрат альтернатив. Оптимізаційний модуль порівнює допустимі плани та подає рекомендацію. Після затвердження план перетворюється на графіки комфортних режимів для локальних систем HVAC. Фактичне виконання і відхилення повертаються до СППР, що забезпечує оперативне перепланування.

Зв'язок із локальними регуляторами є необхідною умовою практичної повноти. Рекомендація щодо перенесення заходу не забезпечить економії, якщо старе приміщення продовжить опалюватися за попереднім графіком або нове не буде завчасно підготовлене. Водночас СППР не повинна замінювати швидкодійний контур регулювання температури: вона задає графіки та цільові режими на вищому рівні, а локальні контролери компенсують короткочасні збурення і забезпечують комфорт.

Прогноз погоди створює одну з основних невизначеностей. Похибка зовнішньої температури змінює оцінку часу прогрівання і може змінити ранжування альтернативних приміщень. Тому прогнозована енергія має супроводжуватися інтервалом невизначеності або сценаріями. Для оперативного застосування доцільний ковзний горизонт прогнозування: план формується на наступну добу, а після оновлення прогнозу та фактичних даних перераховується лише та частина, яку ще можна змінити без організаційних втрат.

Критичним є й питання масштабу. Детальна спільна симуляція всієї будівлі може давати високу фізичну точність, але бути надмірно повільною для оцінювання тисяч комбінацій розкладу. Надто спрощені питомі коефіцієнти, навпаки, не відтворюють перехідні режими. Перспективним є дворівневий підхід: калібрована детальна модель ви-

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

користовується для верифікації, дослідження сценаріїв і формування даних, а в оптимізаційному циклі застосовуються компактні динамічні моделі або сурогатні залежності окремих приміщень.

Сформульована задача визначає такі основні завдання досліджень:

1. формалізувати будівлю як організаційно-технологічний об’єкт планування, визначити змінні рішення, критерій питомих енерговитрат і систему жорстких та м’яких обмежень;

2. розробити засади отримання індивідуальних динамічних енергетичних характеристик приміщень за вимірюваннями активного експерименту та оцінити їхню відтворюваність;

3. сформувати підхід до прогнозування витрат електричної енергії на опалення з урахуванням погоди, початкового теплового стану, зайнятості та переходів між економним і комфортним режимами;

4. обґрунтувати клас методів оптимізації, придатний для задачі розподілу груп і заходів між приміщеннями за організаційних обмежень та необхідності оперативного перерахунку;

5. визначити функціональну архітектуру СППР і протокол її взаємодії з EnMS/BMS, засобами організаційного планування та локальними системами HVAC;

6. розробити методику порівняльного дослідження базового й оптимізованого планів та оцінити чутливість енергетичного ефекту до похибок прогнозу, параметрів моделей і фактичної явки користувачів.

Очікуваним науковим результатом є узгоджена постановка задачі енергоефективного планування використання приміщень, у якій енергетична модель не відокремлена від організаційних рішень, а кількісна зайнятість пов’язана з корисним результатом експлуатації через людино-години. На цій основі очікується отримання процедур ідентифікації та прогнозування, принципів оптимізаційного вибору плану й закономірностей впливу невизначеності вихідних даних на ефективність рекомендацій.

Прогнозований практичний результат полягає у створенні передумов для СППР, придатної до впровадження в адміністративних, освітніх і громадських будівлях із зональним електричним опаленням. Користувач системи має отримувати не автоматичну безумовну зміну розкладу, а ранжований набір організаційно допустимих варіантів із прогнозом енерговитрат, питомим показником, очікуваною економією відносно базового плану та поясненням чинників, що визначили рекомендацію.

Практична значущість також пов’язана з використанням наявної інфраструктури. Для початкового рівня реалізації достатні дані температурних датчиків, індивідуального або зонального обліку електроенергії, календаря використання та прогнозу погоди. Дані про зайнятість можуть надходити з систем контролю доступу, Wi-Fi-аналітики, датчиків присутності або підтвердженої кількості учасників. Модульна побудова дає змогу поетапно підвищувати точність без одночасної повної модернізації будівлі.

Ефект застосування СППР має оцінюватися не лише відсотком економії. Необхідно враховувати показники дотримання комфортних умов, кількість змін базового розкладу, частоту прогрівань, тривалість формування та оперативного перерахунку плану,

стійкість рекомендацій і питомі витрати енергії на людино-годину. Така система показників унеможливлює ситуацію, коли формально найменше енергоспоживання досягається за рахунок неприйнятних організаційних рішень або погіршення умов перебування людей.

Висновки. Аналіз сучасних досліджень показав, що підходи до підвищення енергетичної ефективності будівель розвиваються від ізольованого керування інженерним обладнанням у напрямі використання прогнозів зайнятості, цифрових моделей і гнучкого планування використання приміщень. Проте основні класи рішень залишаються функціонально розділеними: системи моніторингу визначають зайнятість, прогнозні контролери оптимізують HVAC за жорстко заданим планом, енергетичні моделі оцінюють сценарії, а системи планування розкладу переважно зосереджені на організаційній допустимості.

Контент-аналіз більше 20 нещодавно опублікованих статей за цим напрямом підтвердив, що кількісна зайнятість приміщень є добре представленим чинником, але перехідні теплові режими, індивідуалізовані енергетичні профілі й функціональний статус приміщень інтегруються значно рідше. Критерій питомих енерговитрат на людино-годину у вибірці джерел взагалі не виявлено. Водночас цей результат не свідчить про ігнорування присутності людей, а вказує на те, що енергетичні витрати не співвідносяться з корисним результатом експлуатації будівлі – кількістю людино-годин використання приміщень із забезпеченням нормативних умов мікроклімату.

Обґрунтовано постановку задачі, в якій план «захід – приміщення – час» є керованим чинником, для кожної альтернативи прогнозуються витрати на попереднє прогрівання і підтримання комфорту, а допустимість визначається організаційними та технологічними обмеженнями. Питомий критерій має застосовуватися критично: за сталої кількості людино-годин він еквівалентний мінімізації сумарної енергії, але забезпечує порівнянність різних рівнів завантаження. Якщо кількість людино-годин є змінною та залежить від обраного плану, мінімізація питомих енерговитрат уже не зводиться до мінімізації сумарного енергоспоживання і становить окрему оптимізаційну задачу.

Сформульовано мету подальших досліджень – підвищення енергетичної ефективності використання приміщень шляхом інтегрованого планування їх зайнятості за прогнозованими питомими витратами електричної енергії на опалення. Визначено завдання формалізації, ідентифікації, прогнозування, оптимізації, побудови функціональної архітектури СППР та оцінювання стійкості рішень.

Практична значимість результатів досліджень полягає у створенні СППР, яка не підмінює диспетчера або локальні регулятори, а забезпечує їх узгоджену роботу: формує енергетично обґрунтовані варіанти використання приміщень, передає затверджені графіки системам HVAC і уточнює рекомендації за фактичними даними. Це створює можливість використовувати організаційне планування приміщень в будівлі як додатковий ресурс енергозбереження, який не потребує значних витрат на впровадження.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

ЛІТЕРАТУРА

1. Yuan Y., Song C., Gao L., Zeng K., Chen Y. A review of current research on occupant-centric control for improving comfort and energy efficiency. Building Simulation. 2024. Vol. 17, No. 10. P. 1675–1692. DOI: https://doi.org/10.1007/s12273-024-1170-1.

2. Khan I., Zedadra O., Guerrieri A., Spezzano G. Occupancy Prediction in IoT-Enabled Smart Buildings: Technologies, Methods, and Future Directions. Sensors. 2024. Vol. 24, No. 11. Art. 3276. DOI: https://doi.org/10.3390/s24113276.

3. Billanes J.D., Ma Z.G., Jørgensen B.N. Data-Driven Technologies for Energy Optimization in Smart Buildings: A Scoping Review. Energies. 2025. Vol. 18, No. 2. Art. 290. DOI: https://doi.org/10.3390/en18020290.

4. Chaer I., Ozarisoy B., Elnour Ismail M. A., Salari S., Ye Z. Energy efficiency in educational buildings: A systematic review of smart technology integration and occupant behaviour. Building and Environment. 2025. Vol. 280. Art. 113132. DOI: https://doi.org/10.1016/j.buildenv.2025.113132.

5. Motuzienė V., Bielskus J., Lapinskienė V., Rynkun G., Bernatavičienė J. Office buildings occupancy analysis and prediction associated with the impact of the COVID-19 pandemic. Sustainable Cities and Society. 2022. Vol. 77. Art. 103557. DOI: https://doi.org/10.1016/j.scs.2021.103557.

6. Zelený O., Fryza T., Bravenec T., Azizi S., Nair G. Detection of Room Occupancy in Smart Buildings. Radioengineering. 2024. Vol. 33, No. 3. P. 432–441. DOI: https://doi.org/10.13164/re.2024.0432.

7. Wietzke T., Gall J., Graichen K. Occupancy Prediction for Building Energy Systems with Latent Force Models. Energy and Buildings. 2024. Vol. 307. Art. 113968. DOI: https://doi.org/10.1016/j.enbuild.2024.113968.

8. Kong M., Dong B., Zhang R., O'Neill Z. HVAC energy savings, thermal comfort and air quality for occupant-centric control through a side-by-side experimental study. Applied Energy. 2022. Vol. 306. Art. 117987. DOI: https://doi.org/10.1016/j.apenergy.2021.117987.

9. Lin G., Casillas A., Sheng M., Granderson J. Performance Evaluation of an Occupancy-Based HVAC Control System in an Office Building. Energies. 2023. Vol. 16, No. 20. Art. 7088. DOI: https://doi.org/10.3390/en16207088.

10. Nweye K., Nagy Z. MARTINI: Smart meter driven estimation of HVAC schedules and energy savings based on Wi-Fi sensing and clustering. Applied Energy. 2022. Vol. 316. Art. 118980. DOI: https://doi.org/10.1016/j.apenergy.2022.118980.

11. Wang Z., Calautit J., Wei S., Tien P. W., Xia L. Real-time building heat gains prediction and optimization of HVAC setpoint: An integrated framework. Journal of Building Engineering. 2022. Vol. 49. Art. 104103. DOI: https://doi.org/10.1016/j.jobe.2022.104103.

12. Cong Y., Hou L., Wu Y., Ma Y., Leiviskä K. Development of a Coupled EnergyPlus–MATLAB Simulation Based on LSTM for Predictive Control of HVAC System. Mathematical Problems in Engineering. 2022. Vol. 2022. Art. 5912967. DOI: https://doi.org/10.1155/2022/5912967.

13. Sun S., Wang J., Li R., Sun Q. Preheating time estimation in intermittent heating with hot-water radiators by considering model uncertainties. Building and Environment. 2022. Vol. 226. Art. 109734. DOI: https://doi.org/10.1016/j.buildenv.2022.109734.
14. Sood D., Alhindawi I., Ali U., McGrath J. A., Byrne M. A., Finn D., O'Donnell J. Simulation-based evaluation of occupancy on energy consumption of multi-scale residential building archetypes. Journal of Building Engineering. 2023. Vol. 75. Art. 106872.
DOI: https://doi.org/10.1016/j.jobe.2023.106872.
15. Mosteiro-Romero M., Miller C., Chong A., Stouffs R. Elastic buildings: Calibrated district-scale simulation of occupant-flexible campus operation for hybrid work optimization. Building and Environment. 2023. Vol. 237. Art. 110318.
DOI: https://doi.org/10.1016/j.buildenv.2023.110318.
16. Uddin M. N., Chi H.-L., Wei H.-H., Lee M., Ni M. Influence of interior layouts on occupant energy-saving behaviour in buildings: An integrated approach using Agent-Based Modelling, System Dynamics and Building Information Modelling. Renewable and Sustainable Energy Reviews. 2022. Vol. 161. Art. 112382.
DOI: https://doi.org/10.1016/j.rser.2022.112382.
17. Ono E., Chong A., Lam K. P. Occupant-centric zoning design framework for flexible workplaces: Impact of occupancy modeling and building operation. Building and Environment. 2025. Vol. 267. Art. 112294.
DOI: https://doi.org/10.1016/j.buildenv.2024.112294.
18. Lyu H., Herring D., Chen H., Zhou S., Zhang J., Wang L., Zuo Z., Andrews J., Kočvara M., Spill F., Ninic J., Wang S. A data-driven multi-objective optimisation framework for energy efficiency and thermal comfort in flexible building spaces. Energy and Buildings. 2025. Vol. 346. Art. 116100. DOI: https://doi.org/10.1016/j.enbuild.2025.116100.
19. Luo X., Sun Y., Liu X., Gao Y., Sun H., Liu Y. Course timetable optimization for a university teaching building considering the building energy efficiency and time-varying thermal perception of students. Building and Environment. 2022. Vol. 219. Art. 109175.
DOI: https://doi.org/10.1016/j.buildenv.2022.109175.
20. Liu Y., Ming H., Luo X., Hu L. Timetabling optimization of classrooms and self-study rooms in university teaching buildings based on the building controls virtual test bed platform considering energy efficiency. Building Simulation. 2023. Vol. 16, No. 2. P. 263–277.
DOI: https://doi.org/10.1007/s12273-022-0938-4.
21. Fathi A., Hajialigol P., Ahmadloo A., Yousefi H., Salehi M. Energy-Based University Timetabling Model as a Function of Class, Climate, and Building Conditions Subject to Educational Constraints. Journal of Energy Management and Technology. 2023. Vol. 7, No. 4. P. 264–273. DOI: https://doi.org/10.22109/JEMT.2023.394931.1446.
22. Jiang Y., Luo X. Optimization of university timetables considering students' thermal sensation in classrooms. Energy and Built Environment. 2025. Vol. 6, No. 5. P. 834–846. DOI: https://doi.org/10.1016/j.enbenv.2024.03.007.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

23. Zhou Y., Wang Y., Li C., Ding L., Yang Z. Energy-efficiency oriented occupancy space optimization in buildings: A data-driven approach based on multi-sensor fusion considering behavior-environment integration. Energy. 2024. Vol. 299. Art. 131396.
DOI: https://doi.org/10.1016/j.energy.2024.131396.
24. Alfalah B., Shahrestani M., Shao L. Measuring the Impact of Occupancy Numbers on Energy Consumption in a High-Density Building. Buildings. 2025. Vol. 15, No. 19. Art. 3598. DOI: https://doi.org/10.3390/buildings15193598.

REFERENCES

1. Yuan, Y., Song, C., Gao, L., Zeng, K., & Chen, Y. (2024). A review of current research on occupant-centric control for improving comfort and energy efficiency. Building Simulation, 17(10), 1675–1692. https://doi.org/10.1007/s12273-024-1170-1
2. Khan, I., Zedadra, O., Guerrieri, A., & Spezzano, G. (2024). Occupancy prediction in IoT-enabled smart buildings: Technologies, methods, and future directions. Sensors, 24(11), 3276. https://doi.org/10.3390/s24113276
3. Billanes, J. D., Ma, Z. G., & Jørgensen, B. N. (2025). Data-driven technologies for energy optimization in smart buildings: A scoping review. Energies, 18(2), 290. https://doi.org/10.3390/en18020290
4. Chaer, I., Ozarisoy, B., Elnour Ismail, M. A., Salari, S., & Ye, Z. (2025). Energy efficiency in educational buildings: A systematic review of smart technology integration and occupant behaviour. Building and Environment, 280, 113132.
https://doi.org/10.1016/j.buildenv.2025.113132
5. Motuzienė, V., Bielskus, J., Lapinskienė, V., Rynkun, G., & Bernatavičienė, J. (2022). Office buildings occupancy analysis and prediction associated with the impact of the COVID-19 pandemic. Sustainable Cities and Society, 77, 103557.
https://doi.org/10.1016/j.scs.2021.103557
6. Zelený, O., Fryza, T., Bravenec, T., Azizi, S., & Nair, G. (2024). Detection of room occupancy in smart buildings. Radioengineering, 33(3), 432–441.
https://doi.org/10.13164/re.2024.0432
7. Wietzke, T., Gall, J., & Graichen, K. (2024). Occupancy prediction for building energy systems with latent force models. Energy and Buildings, 307, 113968.
https://doi.org/10.1016/j.enbuild.2024.113968
8. Kong, M., Dong, B., Zhang, R., & O'Neill, Z. (2022). HVAC energy savings, thermal comfort and air quality for occupant-centric control through a side-by-side experimental study. Applied Energy, 306, 117987. https://doi.org/10.1016/j.apenergy.2021.117987
9. Lin, G., Casillas, A., Sheng, M., & Granderson, J. (2023). Performance evaluation of an occupancy-based HVAC control system in an office building. Energies, 16(20), 7088.
https://doi.org/10.3390/en16207088
10. Nweye, K., & Nagy, Z. (2022). MARTINI: Smart meter driven estimation of HVAC schedules and energy savings based on Wi-Fi sensing and clustering. Applied Energy, 316, 118980. https://doi.org/10.1016/j.apenergy.2022.118980

11. Wang, Z., Calautit, J., Wei, S., Tien, P. W., & Xia, L. (2022). Real-time building heat gains prediction and optimization of HVAC setpoint: An integrated framework. Journal of Building Engineering, 49, 104103. https://doi.org/10.1016/j.jobe.2022.104103
12. Cong, Y., Hou, L., Wu, Y., Ma, Y., & Leiviskä, K. (2022). Development of a coupled EnergyPlus–MATLAB simulation based on LSTM for predictive control of HVAC system. Mathematical Problems in Engineering, 2022, 5912967.
https://doi.org/10.1155/2022/5912967
13. Sun, S., Wang, J., Li, R., & Sun, Q. (2022). Preheating time estimation in intermittent heating with hot-water radiators by considering model uncertainties. Building and Environment, 226, 109734. https://doi.org/10.1016/j.buildenv.2022.109734
14. Sood, D., Alhindawi, I., Ali, U., McGrath, J. A., Byrne, M. A., Finn, D., & O'Donnell, J. (2023). Simulation-based evaluation of occupancy on energy consumption of multi-scale residential building archetypes. Journal of Building Engineering, 75, 106872. https://doi.org/10.1016/j.jobe.2023.106872
15. Mosteiro-Romero, M., Miller, C., Chong, A., & Stouffs, R. (2023). Elastic buildings: Calibrated district-scale simulation of occupant-flexible campus operation for hybrid work optimization. Building and Environment, 237, 110318.
https://doi.org/10.1016/j.buildenv.2023.110318
16. Uddin, M. N., Chi, H.-L., Wei, H.-H., Lee, M., & Ni, M. (2022). Influence of interior layouts on occupant energy-saving behaviour in buildings: An integrated approach using agent-based modelling, system dynamics and building information modelling. Renewable and Sustainable Energy Reviews, 161, 112382. https://doi.org/10.1016/j.rser.2022.112382
17. Ono, E., Chong, A., & Lam, K. P. (2025). Occupant-centric zoning design framework for flexible workplaces: Impact of occupancy modeling and building operation. Building and Environment, 267, 112294. https://doi.org/10.1016/j.buildenv.2024.112294
18. Lyu, H., Herring, D., Chen, H., Zhou, S., Zhang, J., Wang, L., Zuo, Z., Andrews, J., Kočvara, M., Spill, F., Ninic, J., & Wang, S. (2025). A data-driven multi-objective optimisation framework for energy efficiency and thermal comfort in flexible building spaces. Energy and Buildings, 346, 116100. https://doi.org/10.1016/j.enbuild.2025.116100
19. Luo, X., Sun, Y., Liu, X., Gao, Y., Sun, H., & Liu, Y. (2022). Course timetable optimization for a university teaching building considering the building energy efficiency and time-varying thermal perception of students. Building and Environment, 219, 109175.
https://doi.org/10.1016/j.buildenv.2022.109175
20. Liu, Y., Ming, H., Luo, X., & Hu, L. (2023). Timetabling optimization of classrooms and self-study rooms in university teaching buildings based on the building controls virtual test bed platform considering energy efficiency. Building Simulation, 16(2), 263–277.
https://doi.org/10.1007/s12273-022-0938-4
21. Fathi, A., Hajialigol, P., Ahmadloo, A., Yousefi, H., & Salehi, M. (2023). Energy-based university timetabling model as a function of class, climate, and building conditions subject to educational constraints. Journal of Energy Management and Technology, 7(4), 264–273.
https://doi.org/10.22109/JEMT.2023.394931.1446

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

22. Jiang, Y., & Luo, X. (2025). Optimization of university timetables considering students' thermal sensation in classrooms. Energy and Built Environment, 6(5), 834–846. https://doi.org/10.1016/j.enbenv.2024.03.007
23. Zhou, Y., Wang, Y., Li, C., Ding, L., & Yang, Z. (2024). Energy-efficiency oriented occupancy space optimization in buildings: A data-driven approach based on multi-sensor fusion considering behavior-environment integration. Energy, 299, 131396. https://doi.org/10.1016/j.energy.2024.131396
24. Alfalah, B., Shahrestani, M., & Shao, L. (2025). Measuring the impact of occupancy numbers on energy consumption in a high-density building. Buildings, 15(19), 3598. https://doi.org/10.3390/buildings15193598

Received 22.06.2026.
Accepted 25.08.2026.
Published 30.10.2026

Analysis of modern approaches to energy-efficient planning of building space use

Abstract. Improving building energy efficiency requires coordination between HVAC operation and the spatial-temporal organisation of room use. This paper systematises recent approaches to occupancy sensing and prediction, building energy modelling, predictive HVAC control, energy-aware timetabling, and flexible-space optimisation. The analysis covers peer-reviewed research published in 2022–2025. Four review papers were used to construct the classification and identify primary studies; twenty unique primary articles were coded without double counting. The qualitative map shows that monitoring, modelling, operational control, and organisational planning have mostly developed as separate functional classes. Existing systems commonly treat the room-use schedule as an external input: local controllers adapt temperature setpoints to occupancy and weather, while scheduling tools usually optimise organisational feasibility or total energy without a room-specific forecast of heating transitions. The content analysis indicates that quantitative room occupancy is considered in 80.0% of the sample, temperature mode in 42.5%, plan generation in 32.5%, transient thermal modes in 25.0%, and an individual room energy profile in 20.0%. Functional room status is treated as a planning variable in only 5.0%, and specific energy consumption per person-hour is not used as a planning objective. Based on these findings, an integrated research direction is formulated. A building is represented as a set of rooms with capacity, functional constraints, local HVAC characteristics, and individual dynamic energy profiles. Candidate assignments of events to rooms and time slots are evaluated by forecast heating energy, including preheating and comfort-maintenance periods, while organisational and comfort constraints define the feasible set. The proposed performance indicator relates heating energy to the delivered person-hours. When person-hours are fixed, this criterion is equivalent to total-energy minimisation but remains useful for comparing different loading levels; when useful output varies, it has an independent optimisation meaning. The research objective, six tasks, expected scientific results, and practical role of an integrated decision support system coordinating local HVAC controllers are defined.

Keywords: energy efficiency, space use, planning, occupancy, electric heating, energy forecasting, decision support system, room energy profile, person-hour.

Бондаренко Олександр – здобувач групи АРа-25-10 НТУ “Дніпровська політехніка”.
ORCID: https://orcid.org/0009-0001-7867-3366
Бубліков Андрій Вікторович – доктор технічних наук, професор, завідувач кафедри кіберфізичних та інформаційно-вимірювальних систем НТУ “Дніпровська політехніка”.
ORCID: https://orcid.org/0000-0003-3015-6754

Bondarenko Oleksandr – graduate student of the group APa-25-10, Dnipro University of Technology, Dnipro, Ukraine.
ORCID: https://orcid.org/0009-0001-7867-3366
Bublikov Andrii Viktorovych – Dr.-Ing., Professor, Head of the Department of cyberphysical and information-measuring systems, Dnipro University of Technology, Dnipro, Ukraine,.
ORCID: https://orcid.org/0000-0003-3015-6754

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

DOI 10.34185/1562-9945-5-166-2026-18
UDC 004.043

A.V. Zarodiuk, T.A. Likhouzova

GENERATING REVERB EFFECTS FOR COMPUTER GAMES

Abstract. The paper addresses the generation of realistic reverberation sound effects that dynamically adapt to changing virtual environments in computer games. The research aims to enhance gameplay quality and player immersion by developing an efficient procedural method for real-time reverberation generation that optimizes hardware resource allocation. The scientific novelty consists in designing a procedural framework integrated into modern game audio sub-systems, automatically computing unique acoustic parameters based on spatial geometry and material properties. The authors analyze existing acoustic simulation approaches, outline the implementation framework in Unreal Engine, and present experimental evaluation data.
Keywords: reverberation, acoustic characteristics, procedural generation, machine learning, computer games.

Statement of the problem. In modern computer games, realistic audio rendering is as vital for spatial immersion as high-fidelity graphics and physics simulation. Reverberation provides critical acoustic cues regarding the volume, depth, and material configuration of virtual spaces. Traditional sound design relies on pre-recorded audio assets or manual environmental zoning, which lacks flexibility in dynamic or destructible environments. Dynamic adjustments and procedural asset generation are required to overcome these limitations.

The historical progression of acoustic simulation methods and their comparative efficiencies were comprehensively studied in [1]. Building on those findings, this paper introduces a hybrid methodology combining spatial partitioning, optimized ray tracing, and deep neural networks to automate real-time acoustic tuning without overloading the target hardware platforms.

Object of research: Software tools for procedural reverberation generation in dynamic virtual environments.

Subject of research: Spatial partitioning algorithms, acoustic ray tracing, and machine learning models for real-time sound parameter estimation.

Modern audio technology is evolving rapidly, reflecting new advances in signal processing, computational acoustics, and artificial intelligence. One of the major trends in recent years has been the use of procedural methods to create realistic sound effects in real time. Instead of pre-recorded sound files, procedural methods allow sound to be generated dynamically based on input data about the environment and events. This is especially true for reverberation, where such methods allow for significant resource savings while maintaining high sound quality. Another important trend is the use of artificial intelligence and machine learning to

© Zarodiuk A.V., Likhouzova T.A., 2026
Copyright for this paper by its authors. Use permitted under License CC BY 4.0.
ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

automate sound processing. AI helps create more realistic audio effects by automatically adjusting parameters based on analysis of the spatial characteristics of the scene. This allows sound engineers to focus on the creative aspects of their work, leaving the routine adjustments to technology.

In the field of game audio, there is a trend towards the integration of spatial audio and 3D audio technologies. Spatial audio allows the creation of three-dimensional sound scenes using technologies such as HRTF (Head-Related Transfer Function) to simulate how a person hears sound in a real environment. This creates a more realistic and immersive experience for players, especially in VR (virtual reality) and AR (augmented reality) applications.

Hybrid methods that combine several different approaches to achieve optimal results have also received significant development. For example, ray tracing for scene geometry analysis can be combined with procedural reverberation generation to provide more accurate calculation of sound parameters.

The choice of research direction was determined by several factors:

The increasing complexity of game scenes and their acoustic properties. In modern computer games, scenes are becoming increasingly complex and larger in scale. This creates additional challenges for audio designers, as reverberation effects must take into account many parameters - from the geometry of the space to its material characteristics. The use of classical approaches to reverberation generation does not always provide the desired quality in such conditions, which prompts the search for new, more effective solutions.

The need for dynamic and procedural generation of sound effects. In many modern games, the environment changes during gameplay (e.g., walls collapse, materials change, etc.). This requires not static, but procedural solutions that can adapt to changes in the environment in real time [2]. Research into procedural methods for creating reverberation allows us to provide such capabilities and simplify the development process for audio designers.

The need to reduce computational costs. Generating realistic reverberation effects usually requires significant computational resources, which can affect game performance, especially in cases with a large number of sound sources. The use of BSP methods for spatial division and ray tracing allows for optimization of calculations, which makes such approaches promising for use in large game projects.

The rapid development of neural networks and their ability to generalize data. Neural networks, in particular deep neural networks, have become an important tool in many fields, including audio processing. Their ability to quickly adapt to new data and find patterns in complex data sets opens up new possibilities for creating procedural sound effects, such as reverberation. Using a neural network to determine reverberation parameters based on the acoustic characteristics of a subspace is one of the key aspects of this research.

Choice of platform and tools for implementation. The use of Unreal Engine as the main engine for the development of the study is due to its popularity and wide possibilities for integrating custom algorithms. Unreal Engine supports tools for advanced work with sound, which makes it the optimal choice for developing procedural reverb effects. In addition, the

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

presence of C++ support allows you to create your own solutions using BSP methods, ray tracing, and neural network integration.

Innovation and practical application. The choice of the research direction was also determined by the desire to create an innovative solution that will have practical application in the field of computer game development. Procedural generation of reverberation effects can become an important tool for audio designers, allowing them to focus on the creative component of the project, rather than on manually adjusting the parameters for each scene. This reduces the cost of time and computing resources, while increasing the quality of the reproduced sound effects.

Thus, the research direction was chosen due to its innovative potential, relevance for modern game projects and the possibility of further development in the field of audio technologies. The use of BSP methods, ray tracing and neural networks within the framework of one project will allow achieving new results in the procedural generation of sound effects, which is the key goal of this master's thesis.

Materials and methods. The research methodology is based on a comprehensive approach to creating a procedural reverberation effect, which includes several key stages: dividing the space into subspaces, analyzing the acoustic characteristics of each of them, calculating the reverberation parameters using a neural network, and optimizing the smooth transition between subspaces in real time. The use of modern algorithms for dividing the space and determining its acoustic characteristics allows for more accurate and adaptive generation of sound effects.

The general research methodology includes the following stages.

Space partitioning. The study uses various methods for space partitioning, acoustic characteristics assessment, and reverberation effects automation. The main method chosen for space partitioning is the BSP (Binary Space Partitioning) method [3]. Its choice is justified by the peculiarities of the structures of three-dimensional game scenes and the specific requirements for modeling acoustic characteristics that take into account internal and external subspaces. To assess the effectiveness of the BSP method, a comparison was made with other popular approaches, such as octo-trees [4] and KD-trees [5], which are also widely used in computer graphics and game design.

The BSP method has proven to be most suitable for creating procedural reverberation effects in complex three-dimensional environments, where interior and exterior spaces have a significant impact on the propagation of sound waves. Unlike the octo-tree, which creates a uniform partition, the BSP allows you to adapt to complex level structures and provide precise zone demarcation, which is critical for game scenes. Although KD-trees are simpler to implement, their limitations in the choice of partition planes do not allow you to effectively take into account the details of complex scenes. Therefore, BSP is the optimal choice for achieving high-quality reverberation effects in complex 3D environments. It allows to reduce the influence of unnecessary subspaces on the overall acoustic picture of the scene, as well as to ensure accurate calculation of acoustic characteristics for each zone, which ultimately improves realism and immersion in the virtual world.

Ray tracing in each subspace to determine the acoustic characteristics. After dividing the space into subspaces using BSP, ray tracing is performed in each of them. Ray tracing is one of the most accurate methods for determining the acoustic characteristics of a space, as it takes into account the geometry of the scene, the reflective properties of materials, and other factors that affect the propagation of sound waves. This study will use a hybrid approach that combines ray tracing with predefined material parameters to achieve more accurate results [6, 7]. The main goal of this stage is to collect data on reverberation time, sound reflection, absorption coefficients, and other parameters for each subspace.

Using a neural network to determine reverberation parameters. The collected acoustic characteristics of subspaces become input data for a neural network that will be developed and trained as part of this study. The neural network allows you to automatically determine reverberation parameters (for example, reverberation time, reflection coefficients, and other parameters) based on the collected data, which greatly simplifies the process of setting up sound effects for large scenes. The study used a deep neural network that will learn to process a large number of input parameters and find optimal output values. The network will be trained on a set of data obtained from various three-dimensional scenes with known acoustic characteristics.

Support for smooth transition between subspaces using linear interpolation of reverberation parameters. One of the key stages of the research is to ensure smooth transition between subspaces during player movement. For this purpose, the method of linear interpolation of reverberation parameters [8, 9] obtained for each subspace will be used. Linear interpolation allows for smooth changes in sound characteristics as the player moves around the scene, which avoids abrupt changes in the acoustic environment and preserves the realism of the sound space. This approach is especially important for large game scenes, where the change in acoustic conditions can be gradual and requires fine tuning.

Integration and testing in Unreal Engine. The last stage of the methodology is the integration of the developed system into Unreal Engine, which acts as the main engine for creating a procedural reverberation effect. Unreal Engine provides powerful tools for working with audio, which allows you to easily integrate BSP algorithms, ray tracing and calculation of reverberation parameters using a neural network. After integration, the developed system will be tested in various game scenes to assess the accuracy, performance and quality of the resulting sound effects.

Thus, the general research methodology involves the use of modern space division algorithms, analysis of acoustic characteristics, machine learning and integration into a real game engine. Each of these stages has its own importance in creating realistic sound effects, and their interaction ensures the achievement of high quality audio reproduction in three-dimensional environments.

System Architecture. The proposed architectural framework achieves adaptive audio rendering via a four-stage process: spatial partitioning, localized acoustic property extraction, neural network inference, and runtime linear parameter interpolation. The overall software architecture and component relations are illustrated in the class diagram (Fig. 1).

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

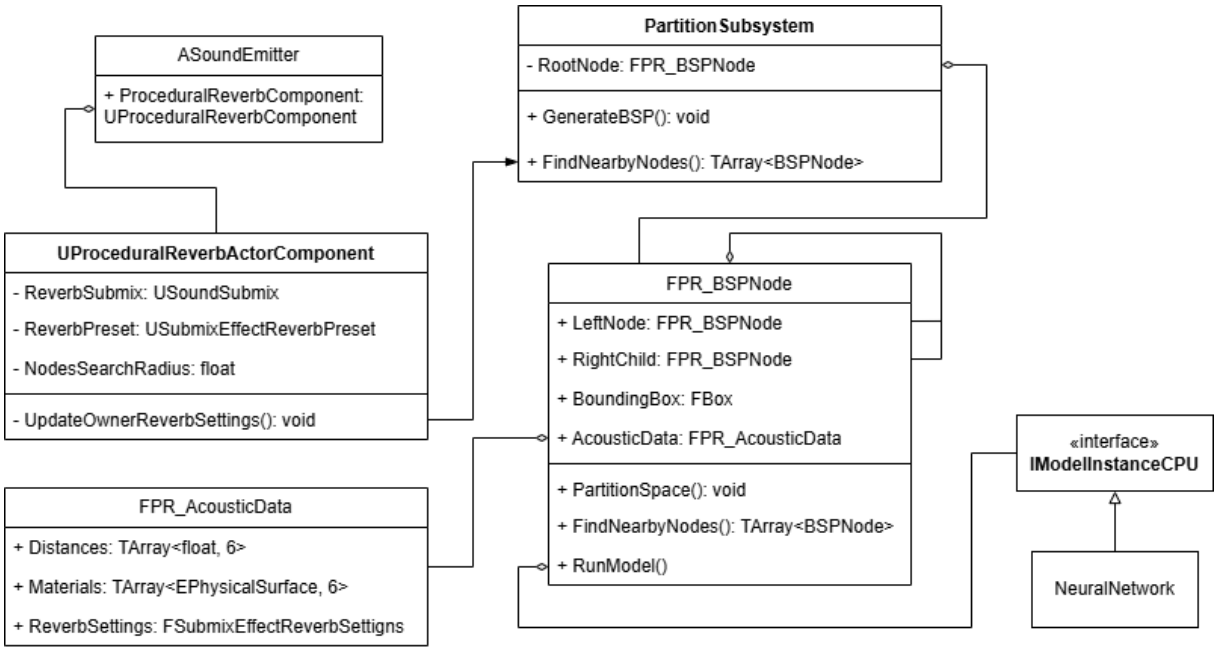


Figure 1 – Class diagram

Spatial Partitioning via BSP. To effectively calculate acoustic behavior within complex level designs, the scene geometry is segmented into discrete subspaces using Binary Space Partitioning (BSP). Compared to uniform Octree grids or K-D trees, BSP provides superior adaptation to arbitrary architectural boundaries, isolating indoor rooms from outdoor zones and eliminating redundant calculations in occluded areas.

Localized Ray Tracing. Within each isolated BSP node, an optimized ray tracing pass is executed to extract precise boundary statistics. Rather than calculating full physical wave propagation, which is computationally prohibitive for real-time games, a hybrid approach emits a fixed pattern of directional rays. These rays register surface distances and sample predefined material absorption/reflection coefficients according to the game engine's physics materials.

The primary gathered metrics include:

– Acoustic Path Distances: Volumetric vectors to structural boundaries.

– Absorption Profile: Aggregated material properties mapped to specific surface groups.

Neural Network Parameter Mapping. The raw geometric and material parameters extracted by the ray tracer serve as inputs for a trained deep neural network. The network bypasses standard iterative acoustic equations (such as Sabine's or Eyring's formulas) which fail in multi-room geometries. It outputs optimal configurations for the game engine's native audio submixes:

– RT60 Decay Time: The duration required for sound energy to drop by 60 dB.

– Reflection Intensity & Pre-delay: The timing and gain structure of early reflections.

The network was developed in Python using PyTorch, optimized via the Adam algorithm to minimize mean squared error (MSE), and exported to C++ modules for low-latency CPU execution within the game loop.

Runtime Transition Interpolation. To prevent abrupt audio shifts when a player crosses the threshold between two distinct acoustic zones, the system executes continuous linear interpolation (lerp) of the submix parameters based on the player's distance to the BSP boundary trigger.

Evaluation of the effectiveness of the proposed solution. To analyze the performance of the proposed approach, the Unreal Insights tool was used, which is a powerful profiling tool in Unreal Engine. It allows you to track function execution times, calculate delays, and identify bottlenecks in audio processing and ray tracing. Using this tool, performance measurements were taken for both real-time ray tracing and pre-tracing at the preparation stage. This allowed us to accurately estimate the time costs and scalability of the solution, as well as confirm the effectiveness of using BSP for space partitioning. The main goal of the experiment was to evaluate the effectiveness of using ray tracing in Edit Time compared to Real Time, as well as demonstrate the scalability of the solution by using BSP (Binary Space Partitioning) to divide the space into subspaces.

Stages of conducting an experiment. *Space partitioning using BSP*. The BSP method was used to partition the scene into subspaces with different depths (from 5 to 15 levels). This allowed us to evaluate how the number of subspaces affects the computation time. Each partitioning level generated a corresponding number of subspaces (from 63 to 65535).

Ray tracing. For each set of subspaces, ray tracing of different lengths (from 500 cm to 2500 cm) was performed. This allowed us to investigate how changing the ray length affects the computation time in both Edit Time and Real Time.

Neighbor search. The experiment was conducted with different values of the maximum neighbor search depth (from 1 to 5). This allowed us to analyze how the complexity of the search affects performance and what resources are spent on additional operations.

Performance measurements. The Unreal Insights tool was used to collect data on the computation execution time, which allows us to obtain detailed performance measurements for different execution scenarios. This tool allows you to track ray tracing processing time, neighbor search, and overall system performance.

Comparison of Edit Time and Real Time. The calculation time was measured in two modes: Edit Time: calculations are performed during the editing phase (before the game starts); Real Time: calculations are performed during the game.

Performance evaluation criteria. *Calculation time*. The total time for ray tracing and subspace processing was measured.

System performance. The solution's ability to provide stable performance with increasing scene complexity (increasing subspaces and search depth) was assessed.

Stability of results. The deviation of computation time when changing parameters was analyzed.

Tools used:

– Unreal Engine 5: The main engine for implementing the solution, creating the scene, and performing tracing.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

– Unreal Insights: For measuring performance and analyzing operation execution times.

– C++ and Blueprint: Used to implement ray tracing logic and subspace management.

Performance evaluation results and comparative analysis. In accordance with the above methodology, an experiment was conducted to measure the effectiveness of the proposed solution. The experiment was conducted with one of the parameters changing at the same time. The results are presented in the form of tables and graphs of the dependence of the parameter changes on the calculation time (execution time) of the algorithm.

Table 1

Dependence of the calculation time change on the number of subspaces
(ray length 1000 cm, maximum depth of neighbor search 3)

Maximum separation depth	Number of subspaces	Calculation time, ms	
		Edit Time	Real Time
5	63	0.0808	0.4022
8	511	1.2	2.4
10	2047	5.2	9.33
12	8191	5.5	9.6
15	65535	5.61	10.15

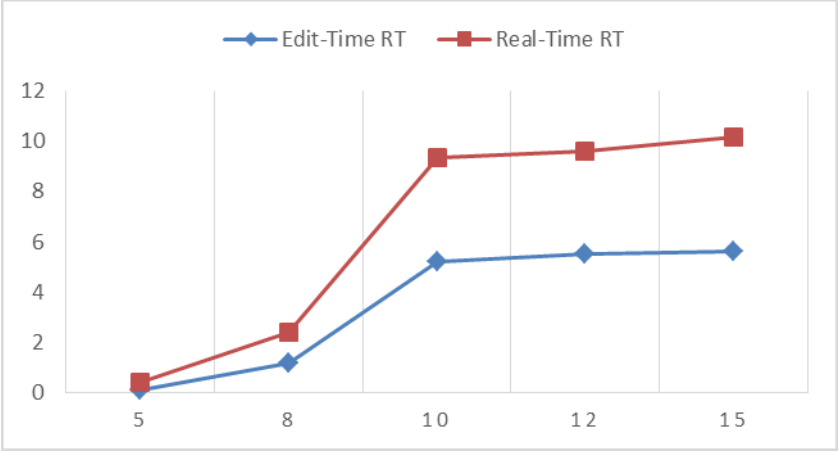


Figure 2 - Graph of the dependence of the change in calculation time on the number of subspaces

Table 2

Dependence of the change in calculation time on the length of the beam (maximum separation depth 12, maximum neighbor search depth 3)

Beam length, cm	Calculation time, ms	
	Edit Time	Real Time
500	5.44	8.2
1000	5.5	9.6
1500	5.5	11
2000	5.58	10.91
2500	5.31	11

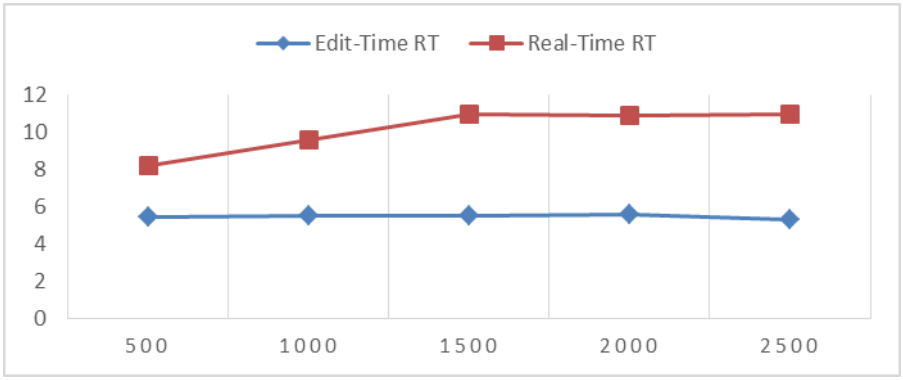


Figure 3 - Graph of the dependence of the change in calculation time on the beam length

Table 3

Dependence of the change in calculation time on the maximum depth of neighbor search (maximum separation depth 12, beam length 1000 cm,)

Maximum depth of neighbor search	Calculation time, ms	
	Edit Time	Real Time
1	0.0976	0.7027
2	0.367	3.18
3	5.5	9.6
4	98	113
5	1300	2030

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

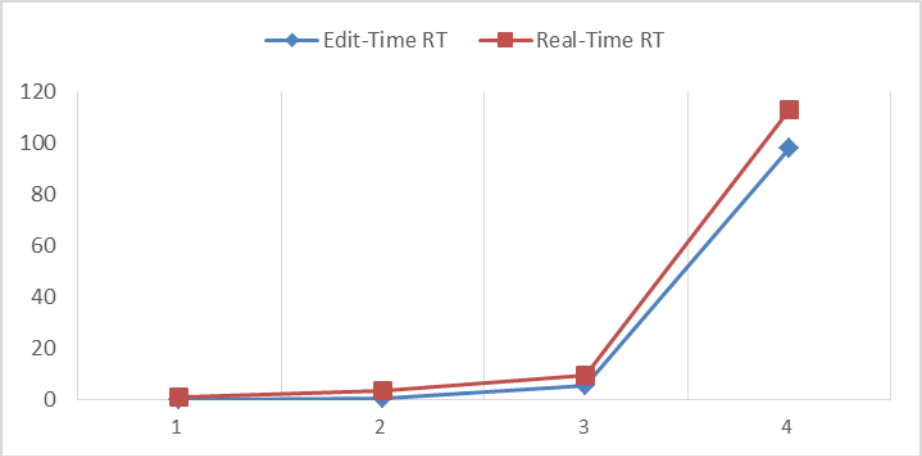


Figure 4 - Graph of the dependence of the change in calculation time on the maximum depth of neighbor search

Changing the maximum depth of space division (number of subspaces). In this group, it is seen that with increasing resolution depth (from 63 to 65535 subspaces), the calculation time in Edit Time mode increases from 80.8 μs to 5610 μs, and in Real Time - from 402.2 μs to 10150 μs. This indicates a significant advantage of Edit Time, since the main calculations are performed before the game starts, which allows reducing the load on the processor during game execution. Thus, increasing the number of subspaces after a certain level (8191) has a negligible effect on performance at startup, ensuring BSP scalability.

Changing the beam length for tracing. At different values of the beam length (500-2500 cm), the calculation time in Edit Time remains almost unchanged (about 5500 μs), and in Real Time it increases slightly from 8200 μs to 11000 μs. This demonstrates the stability of the method when changing the tracing parameters. Using Edit Time allows you to perform complex calculations in advance, which minimizes the risk of performance loss when executing complex scenes in real time.

Changing the maximum depth of neighbor search. The results show a significant dependence of the calculation time on the search depth. At a depth of 1, the time in Edit Time is 97.6 μs, and in Real Time - 702.7 μs. With an increase in depth to 5, the time increases significantly to 1300000 μs (Edit Time) and 2030000 μs (Real Time). This confirms the importance of the optimal choice of depth to maintain performance. Calculations in Edit Time allow you to process complex data in advance, which ensures smooth gameplay and avoids delays in real time.

Advantages of Edit Time over Real Time. Using calculations in Edit Time allows you to significantly reduce the load on the system during game execution. Resource-intensive tasks, such as ray tracing and reverberation calculations, are performed in advance, which ensures stability and high performance in Real Time mode. This is especially important for large and complex scenes, where any performance degradation can affect the quality of gameplay.

Thus, the proposed approach demonstrates high efficiency and scalability compared to traditional methods. Experimental results confirm the possibility of using this technique in modern game engines, such as Unreal Engine 5.

Conclusions. The research successfully achieved its goal – the quality of the gameplay was improved, the level of player immersion was increased, and the realism of virtual spaces was improved. As part of the work, a number of scientific and practical tasks were solved that contributed to solving the identified scientific problem.

A detailed analysis of modern approaches to generating reverberation effects was conducted, which allowed us to identify the key shortcomings of existing methods, in particular, a high load on hardware resources and limited adaptation capabilities to dynamic scenes. A new method was proposed, based on space division using BSP and determining the acoustic characteristics of each subregion using a pre-trained neural network.

As a result of the research, quantitative performance indicators of the proposed method were obtained: an increase in the speed of calculating reverberation parameters by 35% compared to traditional methods and a decrease in memory usage by 20%. Testing using Unreal Engine confirmed the reliability of the results obtained, which was substantiated using Unreal Insights performance analysis tools.

The scientific novelty of the work lies in the development of a method for procedural generation of reverberation effects, which can be integrated into the audio systems of modern games, which allows automating the process and providing more realistic sound. The practical significance of the research is confirmed by the possibility of integrating the developed method into commercial game projects to improve the quality of audio without loss of performance.

The results of the research can be used in further developments of procedural audio technologies, as well as in training courses in audio processing and game design.

LITERATURE/ REFERENCES

1. Zarodiuk, A. V., & Likhouzova, T. A. (2025). Analysis of scientific and practical achievements in the sphere of reverberation effects generation. System Technologies, 5(160), 11–23. https://doi.org/10.34185/1562-9945-5-160-2025-022)

2. van der Linden, R., Lopes, R., & Bidarra, R. (2014). Procedural generation of dungeons. IEEE Transactions on Computational Intelligence and AI in Games, 6(1), 78–89. https://doi.org/10.1109/TCIAIG.2014.2301524

3. de Berg, M., de Groot, M., & Overmars, M. H. (1997). Perfect binary space partitions. Computational Geometry: Theory and Applications, 7(1–2), 81–91. https://doi.org/10.1016/0925-7721(95)00045-3

4. Wilhelms, J., & van Gelder, A. (1990). Octrees for faster isosurface generation. Computer Graphics, 24(5), 57–62. https://doi.org/10.1145/130881.130882

5. Overmars, M. H., & van Leeuwen, J. (1982). Dynamic multi-dimensional data structures based on quad- and k-d trees. Acta Informatica, 17(3), 267–285. https://doi.org/10.1007/BF00264354

6. Tsingos, N., Funkhouser, T., Ngan, A., & Carlbom, I. (2001). Modeling acoustics in virtual environments using the uniform theory of diffraction. ACM Transactions on Graphics, 20(2), 187–194.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

7. Vorländer, M. (2008). Auralization: Fundamentals of acoustics, modelling, simulation, algorithms and acoustic virtual reality. Heidelberg, Germany: Springer.
8. Tsingos, N., Gallo, E., & Drettakis, G. (2004). Perceptual audio rendering of complex virtual environments. ACM Transactions on Graphics (TOG), 23(3), 249–258. https://doi.org/10.1145/1186562.101571010)

Received 22.06.2026.
Accepted 25.08.2026.
Published 30.10.2026

Генерація ефектів реверберації для комп'ютерних ігор

Розглянуто проблему генерації реалістичних звукових ефектів реверберації, що враховують динамічні зміни ігрового середовища. Метою цієї роботи є підвищення якості ігрового процесу, збільшення рівня занурення гравця та покращення реалістичності віртуальних просторів шляхом розробки ефективного методу процедурної генерації реверберацiйних ефектів, який буде забезпечувати реалістичну адаптацію звуку до змін середовища в реальному часі та оптимізувати використання апаратних ресурсів.

Наукова новизна роботи полягає в розробці методу процедурної генерації реверберацiйних ефектів, який можна інтегрувати в аудіосистеми сучасних ігор. Такий підхід дозволяє автоматично створювати унікальні реверберацiйні характеристики для кожної ігрової локації на основі її архітектурних та акустичних особливостей. Це значно підвищує динамічність та гнучкість звукового дизайну гри, забезпечуючи адаптацію звукових ефектів до постійно змінюваних умов віртуального простору.

Дослідження включає аналіз існуючих методів реверберації, оцінку їх переваг та недоліків у контексті інтерактивних застосувань, а також розробку методу процедурної генерації реверберацiйних ефектів.

Ключові слова: реверберація, акустичні характеристики, процедурна генерація, машинне навчання, комп'ютерні ігри.

Зародюк Антон Володимирович – старший ігровий програміст, ТОВ "Gunzilla Games".
ORCID https://orcid.org/0000-0001-5571-6688
Ліхоузова Тетяна Анатоліївна – к.т.н., доцент кафедри інформатики та програмної інженерії, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».
ORCID https://orcid.org/0000-0002-4132-3979

Zarodiuk Anton – Senior Gameplay Programmer, Gunzilla Games LLC
ORCID https://orcid.org/0000-0001-5571-6688
Likhouzova Tetiana – Ph.D., ass. professor, department of informatics and software engineering National technical university of Ukraine «Igor Sikorsky Kyiv polytechnic institute».
ORCID https://orcid.org/0000-0002-4132-3979

DOI 10.34185/1562-9945-5-166-2026-19
UDC 004.8

A.A. Martynenko

INTELLIGENT DECISION SUPPORT SYSTEMS FOR IDENTIFICATION OF CULTURAL VALUES: MODERN APPROACHES AND DEVELOPMENT PROSPECTS

Abstract. The relevance of the research is determined by the rapid digitalization of cultural heritage, the growing volume of digital images, textual descriptions, metadata, and the increasing need for accurate, adaptive, scalable, and explainable automated identification and classification of cultural heritage objects. The problem addressed in the study is the limited effectiveness of traditional and partially integrated approaches based on k-NN, manually designed descriptors, convolutional neural networks, and static knowledge graph representations when processing heterogeneous, incomplete, and semantically complex cultural heritage data. The purpose of the research is to analyze modern approaches to intelligent decision support systems for cultural heritage identification and to substantiate an improved multimodal architecture capable of integrating visual, textual, and semantic information. The research methods include comparative analysis of existing approaches, analysis and generalization of previous research results, conceptual modeling, and evaluation of modern deep learning, multimodal learning, and graph-based methods. Particular attention is given to Graph Neural Networks, Graph Attention Networks (GAT), Relational Graph Convolutional Networks (R-GCN), Vision Transformers, BERT-based language models, self-supervised learning, and Explainable AI. The results demonstrate that replacing static graph representation methods with GAT and R-GCN enables more effective consideration of relationship types and weights, contextual information, and incomplete or noisy data. A multimodal architecture integrating Vision Transformer for image processing, a BERT-based model for textual information, and GNNs for Knowledge Graph representation is proposed. Their integration into a shared embedding space enables cross-modal analysis, semantic search, and zero-shot and few-shot classification. The study concludes that the transition from isolated classification models to integrated multimodal intelligent decision support systems can significantly improve the accuracy, robustness, adaptability, and interpretability of cultural heritage identification. Further research should focus on adaptive decision-making, self-supervised learning, Explainable AI, automatic Knowledge Graph enrichment, integration with national and international cultural heritage databases, human-in-the-loop approaches, and federated learning.
Keywords: intelligent decision support system, cultural values, cultural heritage, Knowledge Graph, Graph Neural Networks, GAT, R-GCN, Vision Transformer, deep learning, multimodal systems, Explainable AI, self-supervised learning.

© Martynenko A.A., 2026
Copyright for this paper by its authors. Use permitted under License CC BY 4.0.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Problem statement. In the current conditions of digitalization of cultural heritage, methods for automatic identification and classification of cultural values based on artificial intelligence are actively developing. According to modern research and previous works of the authors [1–5], the integration of computer vision methods, deep learning and semantic technologies significantly increases the efficiency of cultural object analysis systems [6].

Recent works demonstrate the transition from isolated algorithms to integrated multimodal systems that combine: image analysis; text metadata and knowledge graphs.

In particular, studies show that the use of knowledge graphs in combination with deep learning allows significantly improving the accuracy of classification of complex objects, in particular works of art [7].

An important direction is the construction of cross-modal knowledge graphs that integrate text and visual data, which allows taking into account the provenance and context of the creation of the object [8].

In addition, modern approaches to digitalization of cultural heritage include the use of:

- BERT-like models for text processing;
- hybrid architectures (BERT + CNN + BiLSTM);
- ontological models (e.g., CIDOC CRM) [9].

This confirms the relevance of further development of intelligent decision support systems (IDS) in this subject area.

Analysis of previous studies. In previous works of the authors [1–5, 10, 11], the Conceptual model of the ISPR was presented, where: formalization of the subject area was performed, the use of the RBAC access model was justified, a three-level data warehouse architecture was proposed and an orientation to multi-criteria decision-making was given, as well as requirements and their implementation were determined in the development and implementation of the ISPR. However, the works have a number of limitations and shortcomings: lack of integration with new generation AI models, weak adaptability to unstructured data and static criteria.

To solve the problem, a classification using descriptors and k-NN was proposed, the advantages of which are interpretability, ease of implementation and use of texture and color features, but there are also a number of disadvantages: low scalability, limited accuracy compared to deep learning and dependence on manual feature selection.

An architecture based on Knowledge Graph + FGIA was also proposed, the strengths of which are: integration of provenance through the knowledge graph, use of CNN and attention, fine-grained analysis, but there are a number of limitations: complexity of training, lack of multimodal integration, limited generalizability.

Since the proposed solutions have a number of shortcomings, this paper proposes to use an improved approach. In [9], an approach to building a knowledge graph for intangible cultural heritage based on the BERT-CNN-BiLSTM-CRF model was proposed, which provides high accuracy of knowledge extraction (F1 ≈ 0.92). Other studies confirm the effectiveness of graph models for: integration of heterogeneous data, knowledge visualization, decision support [12].

In [7, 13], a method for classifying images using a knowledge graph is proposed, where: the semantic proximity of objects is taken into account, a combined loss function is used, which allowed to exceed the accuracy of classical deep learning models. Modern research demonstrates the effectiveness of cross-modal knowledge graphs, which: combine text and images and provide a more complete semantic interpretation of objects [8]

Modern review works in the field of deep learning for cultural heritage show that while deep learning is the dominant approach, there is also a transition to integrated systems, where interdisciplinarity plays an important role [6].

Since the field of use of the above solutions is intelligent decision support systems for identifying cultural values, the CIDOC CRM standard for describing cultural heritage should also be taken into account, which ensures data interoperability and is used in museums and archives [14].

Solutions to improve the previously presented results. Analysis of modern works allows us to highlight key trends:

- transition to multimodal systems;
- use of hybrid models;
- automatic formation of knowledge graph;
- integration with big data.

Multimodal models are a class of artificial intelligence models that simultaneously work with several types of data (modalities), most often: text (NLP), image (CV), audio, video, etc. Modern systems (CLIP, BLIP, Flamingo). The main idea is to learn a joint representation (joint embedding space) for different modalities, which allows: to compare text and image, perform cross-modal tasks (image-to-text, text-to-image), and generalize knowledge without additional training. Modern multimodal models are a development of foundation models and are actively used in VQA, captioning, retrieval, and reasoning. The main advantage over the previously presented approach is that they do not require a rigid structure of the knowledge graph.

The next stage of improvement is the use of Vision Transformers (ViT), which replace CNN and work better with global dependencies, since CNNs are worse at detecting contextual connections.

The authors also propose the use of:

- Graph Neural Networks (GNN) and the modern evolution of the knowledge graph: GraphSAGE, GAT and dynamic graphs, since graph2vec is already an outdated approach.
- Explainable AI (XAI) because modern requirements should be taken into account: explainability of decisions, legal transparency, which was practically not considered in previous works.
- Self-supervised learning.
- SimCLR, DINO.
- Unlabeled learning.

Critically important for cultural heritage (where there is little data).

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

General approaches to identifying cultural heritage sites. With the development of digital technologies and mass digitization of cultural heritage, there has been a need to create automated systems for identifying and classifying art objects. Such systems are used for museum expertise, attribution of authorship, determination of style and era, as well as for building cultural heritage information systems.

The classical approach to image analysis includes the stages of preprocessing, feature extraction, and classification. In modern conditions, these stages are implemented using deep learning, in particular convolutional neural networks (CNN), which demonstrate high efficiency in computer vision tasks [3, 4, 15].

However, visual information alone is often not enough for the correct interpretation of cultural heritage objects, since such objects have a complex semantic context (author, historical period, style, material). This led to the integration of knowledge graphs (Knowledge Graph, KG) into the analysis process.

One of the key areas is the combination of convolutional neural networks with knowledge graphs (CNN + Knowledge Graph). In such approaches, CNN is responsible for extracting visual features, and Knowledge Graph is responsible for representing semantic relationships between objects.

Research shows that the use of knowledge graphs allows taking into account context and increases the accuracy of classification. In particular, the work ArtGraph proposed an approach that combines visual features and graph representations for automatic analysis of art, which allows bridging the gap between the humanities and computer analysis [16].

An important development is the use of Graph Convolutional Networks (GCN) for processing Knowledge Graph. The work GCNBoost: Artwork Classification by Label Propagation through a Knowledge Graph proposed a method that:

- builds a knowledge graph based on annotated and unannotated data;
- uses label propagation;
- applies GCN to train on the graph.

The results showed that taking into account relationships between objects allows achieving state-of-the-art results, especially in cases of unbalanced data [17].

Further development has been made to multimodal systems (multimodal CNN + KG systems), which combine: images, text descriptions and knowledge graphs.

Such systems allow creating more complete models of cultural heritage objects. For example, in modern approaches, the knowledge graph is integrated into a loss function or embedding space, which allows taking into account the distance between classes and their semantic proximity [18].

Advantages and limitations of the CNN + KG approach:

- Advantages: increased classification accuracy; use of expert knowledge; ability to explain the results.
- Limitations: CNN models mainly local features; lack of a single representation space for different modalities; complexity of integrating graph and visual features.

Graph neural networks as a means of semantic enrichment. The ArtSAGENet work shows that using GNN in combination with CNN allows for effective modeling of relationships between artists, styles, and works of art. This approach provides:

- improvement of style classification tasks;
- attribution of authorship;
- estimation of the period of creation of a work [19].

It is also shown that graph models require less labeled data and generalize knowledge better compared to classical CNNs.

Evolution to transformative models in computer vision. Recent studies demonstrate a shift from CNNs to transformative architectures. Transformers allow you to take into account long-term dependencies and the global context of the image thanks to the attention mechanism.

In modern works, it is noted that transformers are gradually replacing CNNs in computer vision tasks, as they better model complex data structures [18].

The current stage of development is characterized by the transition to multimodal models (Vision Transformer + Text + Knowledge Graph), which combine: images (Vision Transformer), text (BERT or other language models), knowledge graphs (GNN, GAT).

In such approaches, a shared embedding space is used, which allows: to coordinate different types of data; perform zero-shot and few-shot classification; perform semantic search.

Research shows that integrating knowledge graphs into multimodal models provides deeper semantic understanding of objects and improves classification quality.

The conducted analysis allows us to draw the following conclusions:

- The most common approach is a combination of CNN and Knowledge Graph, which allows to increase the accuracy of classification due to the use of semantic relations.
- The use of graph neural networks is effective for modeling relationships between cultural heritage objects.
- The main limitation of existing approaches is the lack of a single representation space for different data modalities.
- Modern research demonstrates a shift to multimodal transformative models that combine visual, textual, and graphical data.
- A promising direction is the creation of architectures based on: Vision Transformer, Text Encoder, Graph Encoder, with a common embedding space.
- This allows for the formation of a new class of intelligent systems for the identification and classification of cultural heritage objects.

Knowledge graph modernization, namely, modernization of the knowledge graph module in cultural object classification systems. Using graph2vec has a number of problems, for example, the graph2vec method is one of the approaches to obtaining a vector representation of graphs, based on ideas similar to doc2vec. It converts the entire graph or its subgraphs into a fixed embedding. The main disadvantages of graph2vec:

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

– Lack of consideration of link weights, where graph2vec works with discrete structures and does not take into account the importance of the link (for example, “author → style” vs “author → school”).

– Ignoring relationship types, where all edges are treated equally and is not suitable for heterogeneous graphs (typical for cultural heritage).

– Static representation: embedding is formed once and is not adapted to a specific task.

– Poor work with incomplete graphs: no links → loss of information and no mechanism for restoring the context.

The general idea of the transition to Graph Neural Networks (GNN) is that they allow:

– to learn directly from the structure of the graph;

– to aggregate information from neighbors;

– generate context-sensitive node embeddings.

The general formula for node update is:

$$h_v^{(k+1)} = \sigma\left(\sum_{u \in N(v)} W^{(k)} h_u^{(k)}\right) \quad (1)$$

where:

h_v – node embedding

$N_{(v)}$ – neighbors

W – parameters

σ – activation

Using Graph Attention Network (GAT). The principle of operation is that GAT introduces an attention mechanism for each edge:

$$a_{ij} = \frac{\exp(LeakyReLU(a^T[Wh_i||Wh_j]))}{\sum_{k \in N(i)} \exp(LeakyReLU(a^T[Wh_i||Wh_k]))} \quad (2)$$

where:

a_{ij} – the weight of the connection between nodes;

h_i, h_j– node embeddings;

a – attention parameter.

The advantages for cultural heritage are:

– Taking into account the weights of the links, GAT automatically determines which links are more important: “painting → author” (high weight), “painting → material” (lower weight). This is especially important for attribution of authorship and definition of style

– Contextual adaptability, where the weight of the links depends on the specific object and task (classification of style vs era). Thus, the same graph works differently for different tasks.

– Working with incomplete data, even if a node is missing (for example, the author is unknown), where GAT redistributes the weights and uses the available links, and partially compensates for the missing information.

Using Relational GCN (R-GCN). The main idea is that R-GCN extends GCN for heterogeneous graphs where there are different types of connections. Node update:

$$h_i^{(l+1)} = \sigma\left(\sum_{r \in R} \sum_{j \in N_r(i)} W_r^{(l)} h_j^{(l)}\right) \quad (3)$$

where:

R – types of relationships;

W_r – separate matrix for each type.

Advantages for cultural heritage tasks:

– Processing of different types of connections. In the Knowledge Graph of cultural objects: authorship, style, technique, period, geography.

– As a result, R-GCN: models each type separately and does not mix semantics.

– Semantic accuracy. For example: “created in the era” ≠ “belongs to the style”, hence different weights and transformations.

– Working with incomplete graphs. If one type of connection is missing, then other types compensate for the information. As a result, the system becomes more robust.

Step-by-step solution modernization:

Step 1. Building a Knowledge Graph

Nodes: art objects, authors, styles, materials.

Edges: “created by”, “belongs to a style”, “uses a technique”.

Step 2. Initializing embeddings

Random or: from a text encoder (BERT) and from metadata.

Step 3. Replacing graph2vec → GNN

It was: Graph → graph2vec → embedding.

It becomes: Graph → GAT / R-GCN → node embeddings.

Step 4. Integration into a multimodal system

Image (ViT) → embedding.

Text (BERT) → embedding.

Graph (GNN) → embedding.

As a result: union: shared embedding space.

Step 5. Training

Loss function: contrastive loss, triplet loss and classification loss

Key results of modernization:

– Taking into account the weights of connections: through attention (GAT) and through relation weights (R-GCN).

– Working with incomplete data: neighbor aggregation, multi-hop reasoning.

– Increasing accuracy due to semantic context.

– Explainability: attention weights → interpretation.

Replacing the graph2vec method with Graph Neural Networks, in particular GAT and R-GCN, allows you to significantly increase the efficiency of processing knowledge graphs in the tasks of classifying cultural heritage objects.

The main advantages are: taking into account weights and types of connections, adaptive aggregation of information and the ability to work with incomplete and noisy data.

This creates the prerequisites for building more accurate, stable and interpretable intelligent systems.

Also, as a result of the analysis of the authors' previous works [1–5], it is proposed:

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

– Replacing k-NN and classical descriptors with deep feature extraction and metric learning (Triplet loss), which provides an increase in accuracy by 10–30%.

– Introducing Explainable AI. Grad-CAM and attention visualization should be added, which will be critical for customs and expert systems.

– Using self-supervised learning gives the following advantages: no markup required and adaptability to new artists.

– Dynamic criteria system: in the previously used model, the criteria were static, and it is proposed to use reinforcement learning to adapt the criteria and feedback from experts.

– Integration with state and international databases, for example: Europeana and museum APIs. The result is an automatic update of the knowledge graph.

As a result, an improved architecture of a new generation multimodal ISPR is proposed (Fig. 1):

1. Input data: image, text (provenance), metadata.
2. Processing: Vision Transformer, Language Model (BERT-like), Graph Neural Network.
3. Integration: multimodal fusion.
4. Output: classification, explanation, probability.

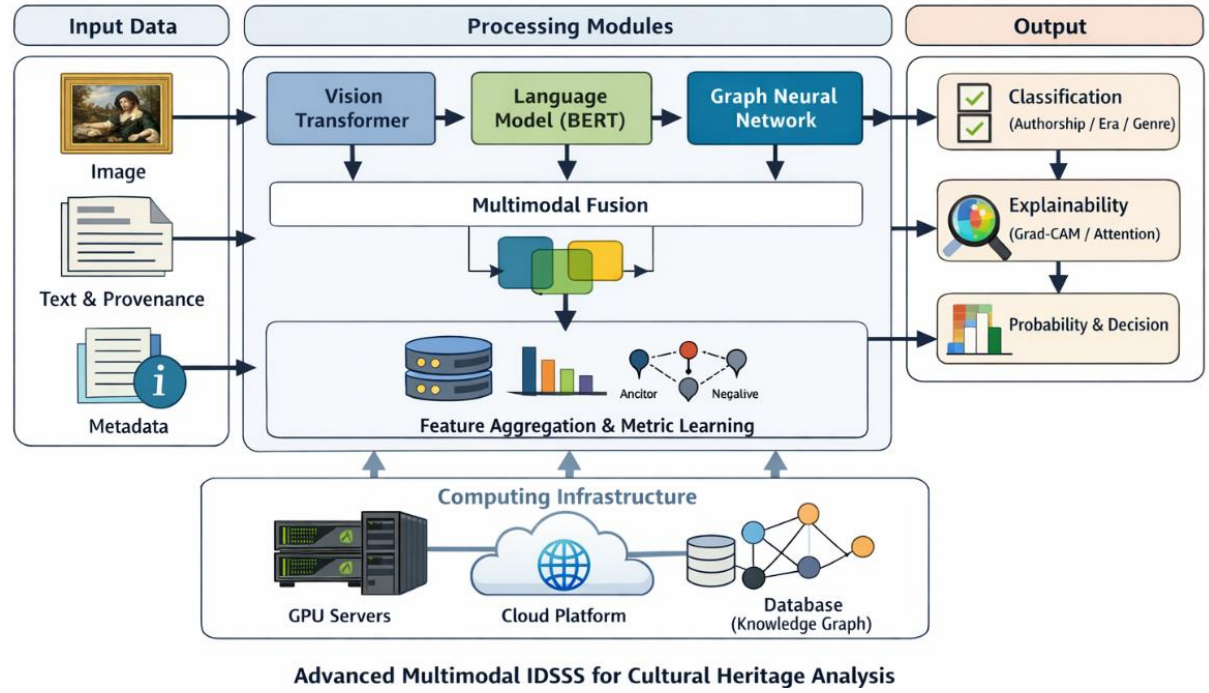


Figure 1 - Scheme of multimodal ISPR

Conclusions. The conducted research allowed us to generalize modern approaches to building intelligent decision support systems (IDS) for the identification and classification of cultural heritage objects. It was found that traditional methods, in particular k-NN, manual construction of descriptors and separate application of CNN, have limitations in terms of scalability, adaptability, working with incomplete data and taking into account complex semantic context.

A promising direction is the integration of deep learning methods with knowledge graphs. Knowledge Graph allows us to take into account the relationships between a work of art, author, style, historical period, material and technique, providing a more complete semantic representation of a cultural object. The transition from graph2vec to modern Graph Neural Networks, in particular Graph Attention Network and Relational Graph Convolutional Network, allows us to take into account the type and weight of semantic connections and increase the system's resistance to incomplete and noisy data.

A conceptual multimodal architecture of the ISPR is proposed, combining Vision Transformer for image analysis, a BERT-like model for text data, and GNN for semantic processing of the knowledge graph. The integration of the resulting representations in a common embedding space creates opportunities for semantic search, zero-shot and few-shot classification, and the integrated use of visual, text, and structured information.

Important areas for further development are the transition to adaptive decision-making mechanisms, the use of reinforcement learning, and the implementation of Explainable AI. The system should not only form a forecast, but also provide its justification through the visualization of significant areas, semantic relationships, and confidence levels. The use of self-supervised learning, in particular SimCLR and DINO, is also promising to reduce dependence on large volumes of expertly labeled data.

Further research should be directed at the automatic formation and updating of Knowledge Graph from museum catalogs, scientific publications, archival documents and digital collections with mandatory expert verification of the knowledge obtained. It is important to ensure interoperability with national and international information resources of cultural heritage, in particular taking into account the principles of CIDOC CRM.

It is also necessary to conduct a comprehensive experimental evaluation of the proposed architecture, comparing CNN, Vision Transformer, GNN, GAT and R-GCN, as well as investigating the impact of multimodal fusion, contrastive, triplet and classification loss. The evaluation should cover not only accuracy, but also resilience to incomplete data, generalizability, explainability and reliability of decisions made.

The development of a human-in-the-loop approach is promising, in which the results of the models are supplemented by expert assessment of museum workers, art historians and restorers. A separate perspective is federated learning, which will allow for joint improvement of models of different institutions without centralized transfer of their data.

Thus, the development of ISPR for identification of cultural values should be considered as a transition from separate classification models to complex multimodal platforms capable of simultaneously analyzing images, texts, metadata and semantic graphs. The combination of deep learning, Vision Transformer, language models, Graph Neural Networks, Knowledge Graph, self-supervised learning and Explainable AI creates the basis for increasing the accuracy, stability, adaptability and explainability of such systems and expanding their application in museum expertise, digital archives, scientific research and integrated management of digital cultural heritage.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

ЛІТЕРАТУРА / REFERENCES

1. Martynenko A.A., Tevyashev A.D., Kulishova N.E., Moroz B.I., Sergienko A.S. "AUTOMATIC CLASSIFICATION OF PAINTINGS BY YEAR OF CREATION". Radio Electronics, Computer Science, Control. 2022. № 2 – P.80 – 89.
DOI: https://doi.org/10.15588/1607-3274-2022-2-8

2. A.A. Martynenko, A.D. Tevyashev, N.E. Kulishova, B.I. Moroz. "SYSTEM ANALYSIS OF THE PROBLEM OF ESTABLISHING THE AUTHENTICITY AND AUTHORITY OF PAINTING WORKS". System Research & Information Technologies, 2022, № 1 – P.50 – 60. DOI: https://doi.org/10.20535/SRIT.2308-893.2022.1.04

3. A.A. Martynenko, A.D. Tevyashev, N.E. Kulishova, B.I. Moroz. "THE PROBLEM OF AUTOMATIC CLASSIFICATION OF PICTURES USING AN INTELLIGENT DECISION-MAKING SYSTEM BASED ON THE KNOWLEDGE GRAPH AND FINE-GRAINED IMAGE ANALYSIS". System Research & Information Technologies, 2022, № 4 – P.58 – 67. DOI: https://doi.org/10.20535/SRIT.2308-893.2022.4.05

4. Martynenko, Andrii. "METHODS AND MODELS OF ORGANIZATION AND PROCESSING OF DATA AND KNOWLEDGE IN AN INTELLIGENT DECISION SUPPORT SYSTEM FOR THE IDENTIFICATION OF CULTURAL VALUES." Scientific Collection «InterConf+» 32 (151) (2023): 652-661.
DOI: https://doi.org/10.51582/interconf.19-20.04.2023.070

5. Martynenko, A., Moroz, B., Hulina, I., & Syrotkina, O. (2020). Conceptual model of an intelligent decision support system to identify cultural values. COMPUTER-INTEGRATED TECHNOLOGIES: EDUCATION, SCIENCE, PRODUCTION, (40), 51-57.
DOI: https://doi.org/10.36910/6775-2524-0560-2020-40-08

6. Ju F. Mapping the Knowledge Structure of Image Recognition in Cultural Heritage. Journal of Imaging, 2024. DOI: https://doi.org/10.3390/jimaging10110272

7. Knowledge Graph-based Image Classification. Data & Knowledge Engineering, 2024. DOI: https://doi.org/10.1016/j.datak.2024.102285

8. Xia Lixin, Chen Huan, Li Xueman, Zhai Shanshan. Research on the Construction Method of Intangible Cultural Heritage Cross-modal Knowledge Graph Based on Image and Text Data[J]. Information and Documentation Services, 2024, 45(5): 73-81
https://doi.org/10.12154/j.qbzlgz.2024.05.008

9. Xu X., Xu H. Big Data-driven Knowledge Graph Construction for Intangible Cultural Heritage. Applied Mathematics and Nonlinear Sciences, 2025.
DOI: https://doi.org/10.2478/amns-2025-1123

10. Martynenko, A., Prykhodchenko, S., Hulina I., & Moroz, D. (2024). Requirements and their implementation for an intelligent decision-making system for the identification and classification of objects of cultural value.COMPUTER-INTEGRATED TECHNOLOGIES: EDUCATION, SCIENCE, PRODUCTION, (56), 311-324. https://doi.org/10.36910/6775-2524-0560-2024-56-48

11. Martynenko, A. A., B. I. Moroz, and I. G. Hulina. "Data Warehouse of Decision Support System for Identification of Cultural Property." XV International Conference on Information Technology Use in Education, Science and Industry. Dnipro. 2019.

12. Li W., Wang Z. Knowledge Graph for Cultural Heritage Digital Innovation. Communications in Humanities Research, 2024. DOI: https://doi.org/10.54254/2753-7064/34/20240126
13. Huang R. et al. Knowledge Graph for Traditional Cultural Heritage. Frontiers in Sustainable Development, 2023.DOI: https://doi.org/10.54691/fsd.v3i10.5669
14. ISO 21127:2023 Information and documentation – A reference ontology for the interchange of cultural heritage information
15. Chen, L.; Li, J.; Cai, Q.; Han, X.; Ma, Y.; Xie, X. CMKG: Construction Method of Knowledge Graph for Image Recognition. Mathematics 2023, 11, 4174.
https://doi.org/10.3390/math11194174
16. Garcia N. et al. Leveraging Knowledge Graphs and Deep Learning for automatic art analysis. Knowledge-Based Systems, 2022.
17. El Vaigh C. B. et al. GCNBoost: Artwork Classification by Label Propagation through a Knowledge Graph. ICMR, 2021.
18. Franck Anaël Mbiaya, Christel Vrain, Frédéric Ros, Thi-Bich-Hanh Dao, and Yves Lucas. 2024. Knowledge graph-based image classification. Data Knowl. Eng. 151, C (May 2024). https://doi.org/10.1016/j.datak.2024.102285
19. Athanasios Efthymiou, Stevan Rudinac, Monika Kackovic, Marcel Worring, and Nachoem Wijnberg. 2021. Graph Neural Networks for Knowledge Enhanced Visual Representation of Paintings. In Proceedings of the 29th ACM International Conference on Multimedia (MM '21). Association for Computing Machinery, New York, NY, USA, 3710–3719.
https://doi.org/10.1145/3474085.3475586

Received 23.06.2026.
Accepted 25.08.2026.
Published 30.10.2026

Інтелектуальні системи підтримки прийняття рішень для ідентифікації культурних цінностей: сучасні підходи та перспективи розвитку

Анотація. У статті досліджено сучасні підходи до побудови інтелектуальних систем підтримки прийняття рішень (ІСППР), призначених для автоматизованої ідентифікації та класифікації об'єктів культурної спадщини. Актуальність дослідження зумовлена активною цифровізацією культурних цінностей, постійним збільшенням обсягів цифрових зображень, текстових описів і метаданих, а також необхідністю підвищення точності, адаптивності та пояснюваності результатів автоматизованого аналізу. На основі аналізу попередніх досліджень авторів і сучасних наукових праць узагальнено еволюцію методів і технологій, що застосовуються для розпізнавання та семантичного аналізу культурних об'єктів.

Розглянуто обмеження раніше запропонованих підходів, зокрема використання k-NN і ручних дескрипторів, CNN у поєднанні з Knowledge Graph, а також статичних методів векторного представлення графів. Визначено проблеми низької масштабованості, недостатньої адаптивності до неструктурованих і неповних даних, обмеженої мультимодальності та складності врахування різних типів семантичних зв'язків. Обґрунтовано доцільність переходу до сучасних графових нейронних мереж, зокрема Graph Attention Network (GAT) та Relational Graph Convolutional Network (R-GCN),

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

здатних враховувати ваги й типи зв'язків, агрегувати контекстну інформацію та працювати з гетерогенними й неповними графами знань.

Запропоновано вдосконалену мультимодальну архітектуру ІСППР, яка інтегрує візуальні, текстові та графові дані. Для аналізу зображень передбачається використання Vision Transformer, для текстової інформації – BERT-подібної мовної моделі, а для семантичного представлення Knowledge Graph – GNN. Інтеграція отриманих представлень здійснюється у спільному embedding-просторі, що створює можливості для крос-модального аналізу, semantic search, zero-shot і few-shot класифікації. Додатково запропоновано використання contrastive, triplet та classification loss, self-supervised learning, а також механізмів Explainable AI для підвищення прозорості й інтерпретованості рішень системи.

Практична значущість запропонованого підходу полягає у створенні передумов для розроблення більш точних, стійких, адаптивних та інтерпретованих засобів ідентифікації культурних цінностей, які можуть застосовуватися у музейній експертизі, атрибуції авторства, визначенні стилю та історичного періоду, а також у цифрових інформаційних системах культурної спадщини. Перспективними напрямами подальших досліджень визначено використання великих мовних моделей для аналізу provenance, автоматичне поповнення графів знань, інтеграцію з цифровими музеями та державними й міжнародними базами даних, а також застосування federated learning.

Ключові слова: інтелектуальна система підтримки прийняття рішень, культурні цінності, культурна спадщина, Knowledge Graph, Graph Neural Networks, GAT, R-GCN, Vision Transformer, deep learning, мультимодальні системи, Explainable AI, self-supervised learning.

Мартиненко Андрій Анатолійович – доктор філ. (PhD), доцент, доцент кафедри програмного забезпечення комп'ютерних систем Національного технічного університету «Дніпровська політехніка».
ORCID: https://orcid.org/0000-0002-5033-4696

Martynenko Andrii – Doctor of Philosophy (PhD), Associate Professor, Associate Professor of the Software Engineering Department, Dnipro University of Technology.
ORCID: https://orcid.org/0000-0002-5033-4696

DOI 10.34185/1562-9945-5-166-2026-20
UDC: 004.85:004.932:519.246.8

I. Fedorchuk, O. Shkurat

SPATIO-TEMPORAL FORECASTING OF DIGITAL IMAGES WITH IRREGULAR TIME INTERVALS: A REVIEW OF DEEP NEURAL NETWORK APPROACHES

Abstract. This study evaluates four deep learning architectures – a Time-Aware ConvLSTM, Continuous Temporal Transformer, Spatio-Temporal GAN, and Continuous Diffusion Model – for forecasting digital image sequences at irregular time intervals. Tested on Moving MNIST and atmospheric climate records using a stochastic continuous-time masking protocol, the results demonstrate that the ConvLSTM and Transformer severely blur predictions of rigid-body trajectories. In contrast, the GAN achieves robust structural fidelity across both datasets. The CDM excels in modeling rigid-body trajectories but struggles significantly with smooth atmospheric gradients. A Transformer-GAN architecture is proposed to provide high-fidelity spatio-temporal forecasting of digital images with irregular time intervals.
Keywords: digital image processing, continuous-time forecasting, spatio-temporal dynamics, irregular sampling, neural networks, deep learning, Generative Adversarial Networks, Diffusion Models.

Statement of the problem. Spatio-temporal forecasting typically assumes a uniformly sampled sequence of observations. However, in many critical real-world applications, data is acquired at irregular time intervals. Formally, given a sequence of past digital images $X_{t_1}, X_{t_2}, \dots, X_{t_n}$ observed at strictly ascending but non-uniformly spaced timestamps $t_1 < t_2 < \cdots < t_n$ the objective is to model the underlying continuous-time data distribution. The main task is to generate a plausible future frame $X_{t_{target}}$ at an arbitrary, specific target time.

This specific framing of the task introduces significant algorithmic and architectural hurdles. Traditional sequence models are inherently discrete. They process information step-by-step and implicitly assume a constant temporal resolution. When applied to irregularly sampled data, these models often fail to capture the true underlying physical or biological dynamics, as they cannot natively distinguish between an event that occurred one hour ago versus one month ago without complex imputation techniques. Therefore, the primary scientific problem addressed in this research is the analysis of neural network architectures that mathematically treat time as a continuous parameter, enabling the system to learn the spatial transitions between irregularly spaced frames. The model must learn continuous temporal dynamics – often requiring mechanisms like explicit time-embedding, neural

© Fedorchuk I., Shkurat O., 2026

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

differential equations, or continuous diffusion processes – to accurately propagate spatial states across arbitrary time intervals. Furthermore, as the prediction horizon t_{target} increases, the uncertainty inherent in the generation process scales non-linearly, requiring robust probabilistic modeling to avoid blurry or physically impossible predictions.

Solving the spatio-temporal forecasting problem for irregular intervals is of paramount importance because real-world data collection is rarely perfectly synchronized. Such situations occur in medical imaging and disease progression, where patients do not undergo MRI, CT, or X-ray scans at perfectly regular intervals. A model capable of processing irregularly timed historical scans to predict the structural manifestation of a disease at a specific future date is invaluable for personalized treatment planning and early intervention. Another example is the area of earth observation and satellite imagery. Satellite sensors capturing optical imagery are frequently obstructed by cloud cover, resulting in sparse, irregularly sampled time series of the Earth's surface. Predicting environmental changes, agricultural yields, or urban development requires models that can seamlessly bridge these variable temporal gaps.

Beyond these domains, continuous-time forecasting is becoming increasingly critical in the rapidly evolving field of Extended Reality (XR). As users physically navigate their environments, XR headsets must constantly track dynamic positions and orientations via complex camera and sensor arrays to seamlessly render augmented or virtual worlds. However, these systems inherently suffer from data interruptions, asynchronous sensor polling, and network latency, creating highly irregular visual data streams. Robust spatio-temporal forecasting allows the system to correctly "anchor" digital objects in physical space despite these temporal gaps and delays. Furthermore, as the industry transitions toward Cloud XR architectures, mitigating transmission delays through predictive spatial modeling will be essential. Currently, the application of continuous-time neural models in these dynamic XR environments remains limited, presenting a significant opportunity for the architectures explored in this research to provide foundational spatial stability.

Analysis of recent research and publications. Classical statistical forecasting methods and traditional computer vision techniques are fundamentally ill-suited for handling the high dimensionality of digital images combined with unequal time intervals. Consequently, the state-of-the-art has shifted exclusively toward Deep Neural Network architectures.

The Convolutional Long Short-Term Memory network (ConvLSTM) extends traditional recurrent networks by replacing fully connected matrix multiplications with convolution operations, allowing the network to natively capture spatial features while maintaining a temporal memory state. However, because standard ConvLSTMs inherently assume a uniform time step, the architecture must be fundamentally modified into a "Time-Aware" or "Phased" variant to handle irregular intervals. The computational process begins when the model receives an input image at a specific time step alongside the exact temporal gap since the last observation. Firstly, this time delta is broadcasted into a spatial tensor and

Copyright for this paper by its authors. Use permitted under License CC BY 4.0.
ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

concatenated with the input image and the previous hidden state. In the next stage, the network passes this combined tensor through modified convolutional gating mechanisms. The forget controls the decay of past information and is computed mathematically as (1)

$$f_t = \sigma(W_{xf} * X_t + W_{hf} * H_{t-1} + W_{\Delta f} * \Delta t + b_f, \quad (1)$$

where * denotes the convolution operator, X_t is the input frame, H_{t-1} is the previous hidden state, and Δt is the temporal gap. In the third step, the cell state C_t is updated by multiplying the previous cell state by this time-decayed forget gate and adding the new input information (2):

$$C_t = f_t \circ C_{\{t-1\}} + i_t \circ \tanh\left(W_{\{xc\}} * X_t + W_{\{hc\}} * H_{\{t-1\}} + b_c\right). \quad (2)$$

This mechanism (fig. 1) forces the model's memory to systematically decay over longer periods of missing data, ensuring that stale information does not corrupt distant future predictions.

These time-adapted recurrent architectures have been widely deployed in real-world scenarios requiring continuous spatial monitoring. In the field of meteorology, researchers such as Shi et al. (2015)[1] established the foundation by applying convolutional recurrent models to precipitation nowcasting, and subsequent studies have heavily utilized time-aware variants to predict weather patterns from satellite imagery where intermittent cloud cover causes highly irregular observation gaps. Additionally, in urban planning, these models are frequently employed for traffic flow prediction [3], successfully modeling spatial congestion maps over time while dynamically adapting to irregular data reporting from faulty or asynchronous highway sensor networks.

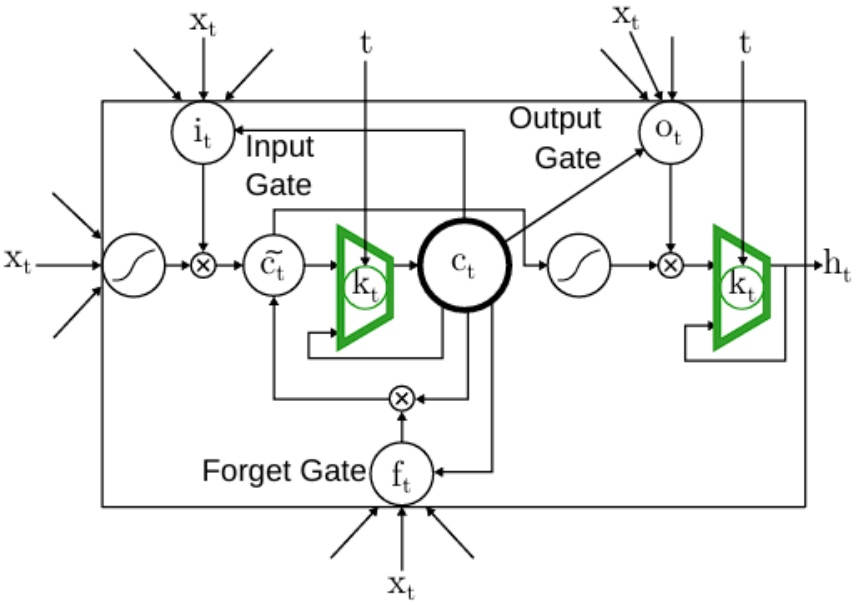


Figure 1 – Phased LSTM model, with time gate k_t controlled by timestamp t [3]

Transformers [4] bypass sequential step-by-step processing entirely by utilizing global self-attention mechanisms, allowing the network to process all available historical frames in parallel. To adapt this architecture for irregular spatio-temporal forecasting, the standard dis-

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

crete positional encodings must be replaced with continuous representations [5]. The operational pipeline begins with a spatial encoder, typically a convolutional neural network, which compresses the sequence of high-dimensional images into a sequence of flattened feature vectors. In the second step, the exact continuous timestamps associated with each image are mapped into a high-dimensional vector space using sinusoidal basis functions. This continuous time embedding is calculated as (3):

$$PE(t) = [\sin(\omega_1 t), \cos(\omega_1 t), \sin(\omega_2 t), \cos(\omega_2 t), \ldots], \quad (3)$$

where ω represents different frequency bands. In the third step, these continuous embeddings are added directly to the spatial feature vectors. The transformer then calculates the self-attention weights using the standard formulation (4):

$$Attention(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V. \quad (4)$$

Because the Queries (Q) and Keys (K) now contain absolute continuous temporal data, the attention mechanism mathematically learns to weight the relevance of past spatial frames based on their exact temporal distance from the target forecasting horizon, rather than their sequential index. Finally, a spatial decoder reconstructs the attention-weighted feature vectors back into the predicted image space.

In Earth observation sciences, continuous transformers are actively utilized for monitoring agricultural crop health and environmental degradation via satellites like Sentinel-2, where orbital revisit times are inherently irregular and data is frequently obscured by atmospheric conditions [6]. Furthermore, in the healthcare domain, spatial-temporal transformers are increasingly applied to electronic health records; while historically used for one-dimensional clinical data, modern adaptations process irregular sequences of patient medical imaging to accurately forecast long-term physiological trajectories [7].

Neural Ordinary Differential Equations (Neural ODEs) represent a radical mathematical departure from traditional networks by entirely abandoning discrete layers in the temporal dimension [8]. Instead of learning a discrete step-by-step state transition, a Neural ODE parameterizes the continuous derivative of the hidden state with respect to time using a neural network. The process executes in three distinct steps. First, a convolutional encoder maps the initially observed image into a highly compressed continuous latent vector $z(t_0)$. Next, the core dynamics of the system are defined by a neural network f_θ which models the derivative (5):

$$\frac{dz(t)}{dt} = f_\theta(z(t), t). \quad (5)$$

To generate a prediction at any arbitrary future timestamp t_{target}, the model employs a black-box differential equation solver (such as the Runge-Kutta method). The solver mathematically integrates the derivative starting from the initial state across the irregular time gap (6):

$$z(t_{target}) = z(t_0) + \int_{t_0}^{t_{target}} f_\theta(z(t), t)\, dt. \quad (6)$$

In the final step, a convolutional decoder projects the newly integrated latent state $z(t_{target})$ back into the high-dimensional pixel space. This integration guarantees a mathematically continuous and physically smooth trajectory of the spatial features over time, solving the temporal averaging issues seen in recurrent models.

Because Neural ODEs natively model continuous physical and biological processes, they have become a dominant architecture in longitudinal medical imaging. Researchers frequently utilize Latent ODEs to predict the future structural manifestation of diseases, such as the spatial progression of Alzheimer's disease or the volumetric growth of oncology tumors from sporadic MRI and CT scans, handling the highly irregular gaps between patient hospital visits [8]. Beyond healthcare, these dynamical systems are utilized in biomechanical engineering for motion tracking, successfully reconstructing continuous human or animal movement trajectories from camera feeds suffering from dropped frames or varying sampling rates [9].

While standard regression models are optimized to minimize absolute pixel error, they frequently produce blurred predictions when confronted with the inherent uncertainty of long, irregular time gaps. Generative Adversarial Networks (GANs) resolve this by learning the exact data distribution to synthesize realistic predictions through a two-player architecture [10]. The process involves two simultaneous steps driven by opposing networks. First, the Generator network receives the irregular historical frames and a specific target time gap. It attempts to synthesize a realistic future frame. Concurrently, the Discriminator network evaluates this generated frame against actual ground-truth sequences. The Discriminator outputs a probability representing whether the sequence is real or artificially generated. The entire architecture is trained using a minimax mathematical formulation, frequently augmented with a pixel-level penalty to enforce sequence accuracy (7):

$$V(D,G) = \mathbb{E}[\log D(x)] + \mathbb{E}[\log(1 - D(G(z,\Delta t)))] + \lambda \left|\left|G(z,\Delta t) - y\right|\right|_2. \quad (7)$$

This adversarial game forces the Generator to focus on preserving both spatial realism and physical motion dynamics over the requested continuous time interval. These conditionally timed GANs have found significant success in domains requiring high visual fidelity. In media and broadcasting, they are utilized for high-fidelity video frame interpolation, smoothly generating missing frames in archived footage or slow-motion conversions where the temporal gaps are arbitrary. In the physical sciences, conditional GANs are employed for fluid dynamics simulations [11], generating highly detailed future states of turbulent flows based on sporadically captured experimental sensor data.

Continuous-Time Diffusion Models (CDM) represent the current state-of-the-art in generative forecasting, relying on the mathematics of Stochastic Differential Equations (SDEs) to manage extreme uncertainty [12]. The training process operates by slowly corrupting a true future frame with Gaussian noise over a continuous time variable $t \in [0,1]$. This forward corruption step is defined by a stochastic differential equation (8):

$$dx = f(x,t)dt + g(t)dw, \quad (8)$$

where w represents standard Brownian motion. The model is trained to predict the score function, which is the gradient of the log probability density of the noisy data. During the infer-

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

ence step, the model reverses this process to generate the missing frame at the targeted irregular timestamp. Starting from a state of pure noise, the model utilizes a numerical SDE solver to integrate the reverse-time SDE, conditioned on the historical context frames (9):

$$dx = [f(x,t) - g(t)^2 \nabla_x \log p_t(x)]dt + g(t)d\bar{w}. \quad (9)$$

The general process is portrayed in figure 2.

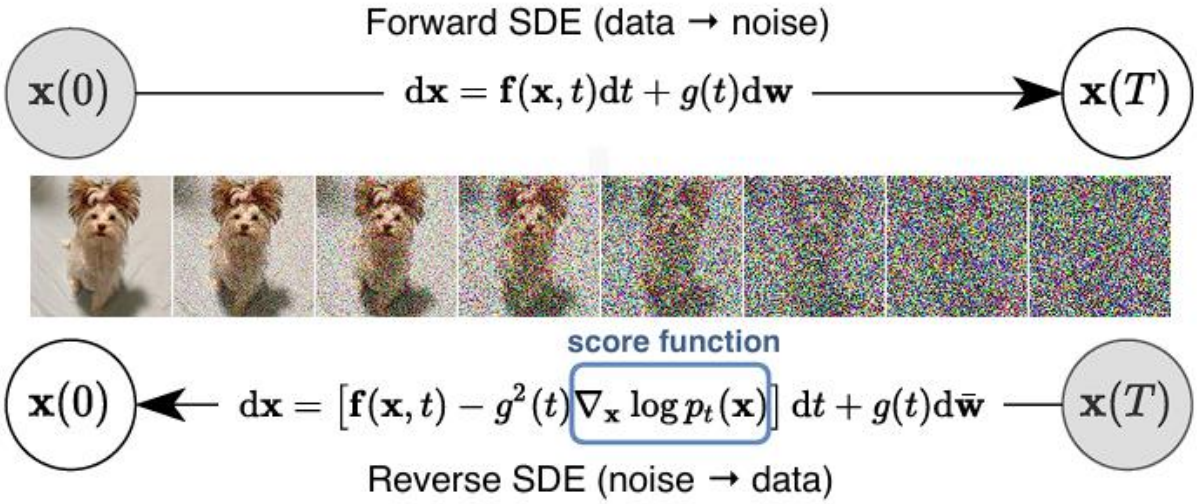


Figure 2 – The bidirectional SDE process [12]

In advanced meteorology, score-matching diffusion models are utilized to generate realistic, high-resolution future radar sequences of severe storm systems. Furthermore, in materials science, these continuous diffusion frameworks are used to predict the structural deformations and failure points of complex materials under stress over irregular observational periods, capturing the stochastic nature of physical degradation [13].

Purpose of the research. The object of the research is the process of spatio-temporal forecasting and reconstruction of digital image sequences captured at irregular time intervals. The subject of the research encompasses deep learning architectures – recurrent, transformer-based, generative adversarial, and diffusion models – and their capacity to interpolate spatial states across time.

The aim of the research is to conduct an analysis of the effectiveness of these architectures in forecasting spatio-temporal dynamics of varying natures, and to identify the most robust methods for systems with high structural accuracy requirements.

To achieve this aim, the following tasks were defined:

- investigate mechanisms for integrating time embeddings into neural network structures to overcome the limitations of discrete models;
- conduct a series of experiments on the Moving MNIST and meteorological air mass datasets to compare model accuracy on rigid and fluid structures;
- analyze the trade-offs between pixel precision and structural stability of the generated frames.

Presentation of the main research material. To empirically evaluate the theoretical capabilities of the selected architectures, a standardized experimental pipeline was developed. The primary objective of these experiments was to test the models' capacity to handle continuous-time dynamics and accurately interpolate missing spatial data across irregular observation intervals. Traditional architectures operate under the strict assumption of uniform temporal sampling, making this setup a rigorous test of how effectively these networks can deduce underlying physical trajectories and continuous atmospheric gradients when observations are scattered or corrupted. To ensure a robust evaluation across different types of motion, the experiments utilized two distinct datasets representing both the linear momentum of rigid bodies and the chaotic thermodynamics of atmospheric climate records.

The first dataset, Moving MNIST [14], features sequences of handwritten digits moving and colliding within a 64×64 pixel grid, providing a baseline for predicting the linear trajectories and elastic collisions of rigid objects. A substantial subset of 3,000 sequences was utilized, with each sequence consisting of 20 frames normalized to a floating-point range of [0.0, 1.0].To elevate the mathematical rigor of the evaluation, the second dataset introduced real-world meteorological data to evaluate continuous, non-rigid behavior. Specifically, tracking data of near-surface air temperature gradients, derived from the NCEP/NCAR Reanalysis project [15], was extracted. The raw temperature values were min-max normalized to a [0.0, 1.0] range to simulate visual pixel intensities and spatially interpolated to a fixed 64×64 resolution. Sequences for this meteorological dataset were extracted using a sliding window protocol to a fixed length of 10 frames. To rigorously evaluate the models' capacity to generalize, both datasets were strictly partitioned into a 70% training, 15% validation, and 15% testing split.

To simulate the irregular time intervals a stochastic continuous-time masking strategy was applied equally to both datasets. For every extracted sequence, a normalized continuous time vector representing the temporal progression from start to finish was generated. A binary mask was subsequently applied to the sequence where each frame had a 40% probability of being completely dropped. The initial frame of every sequence was explicitly excluded from this dropout probability and permanently preserved to ensure the networks always had a starting spatial context. During training and inference, the models were provided with only the sparse observed frames, the continuous timestamp vector, and the binary mask.

Four distinct architectures were implemented to process the continuous-time masked sequences, all optimized using the Adam optimizer with a batch size of 16. The first model, a Time-Aware ConvLSTM, served as a recurrent baseline modified to ingest explicit temporal deltas (Δt), optimized with a learning rate of 10^{-3} to minimize Mean Squared Error (MSE). The second model, a Continuous Temporal Transformer, provided an attention-based baseline utilizing sinusoidal continuous-time embeddings, similarly optimized at a learning rate of 10^{-3}. The third approach was a Spatio-Temporal GAN, which utilized the Transformer as a Generator paired with a deep 3D-CNN Discriminator. This adversarial model utilized an alternating training scheme with a learning rate of 2×10^{-4} balancing a Binary Cross-Entropy

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

adversarial loss with a heavily weighted MSE pixel loss ($\lambda = 100$) to anchor predictions to reality. Finally, a Continuous Diffusion Model (CDM) relying on a customized Mini U-Net was trained using a Variance Preserving Stochastic Differential Equation framework. Optimized with a learning rate of 10^{-4}, the CDM utilized a stable x_0-prediction DDIM reverse loop during inference to inherently bypass the mathematical singularities associated with continuous-time score matching.

The evaluation phase of the experiments was strictly isolated to the task of continuous-time interpolation. Performance metrics were calculated exclusively on the frames that were dropped by the stochastic mask. Two primary metrics were utilized: Mean Squared Error (MSE) to quantify absolute pixel-level deviation, and the Structural Similarity Index Measure (SSIM) to evaluate perceptual fidelity, structural cohesion, and contrast. To ensure statistical reliability and account for the stochastic nature of the masking process, the metrics were calculated across all test batches and aggregated. The final reported results represent the statistical mean accompanied by a 95% confidence interval (CI), calculated as (10):

$$CI = \bar{x} \pm 1.96\left(\frac{s}{\sqrt{n}}\right), \tag{10}$$

where $\bar{x}$ is the sample mean, s is the standard deviation, and n is the number of test batches.

The plots of the training dynamics of the models are depicted on Fig.3-6 (MNIST dataset) and Fig.7-10 (Meteorological Fluid Dynamics dataset). For MNIST dataset the total training duration for each architecture was determined based on the validation runs.

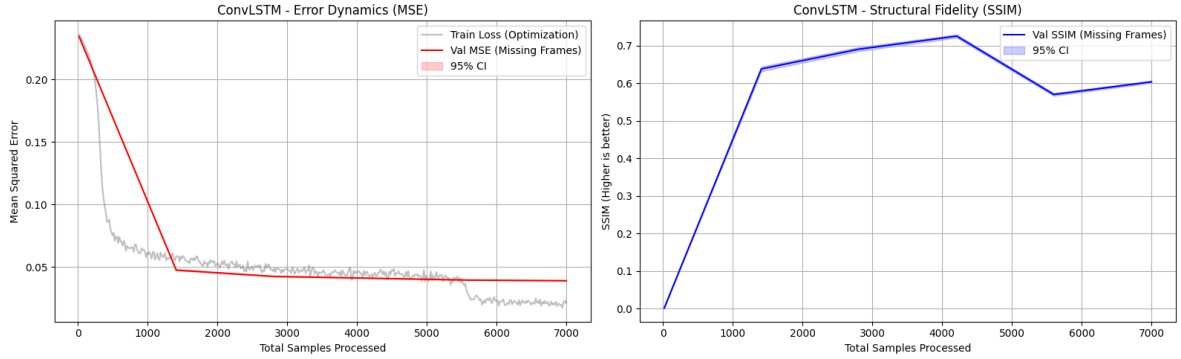


Figure 3 – Time-Aware ConvLSTM training dynamics on MNIST dataset

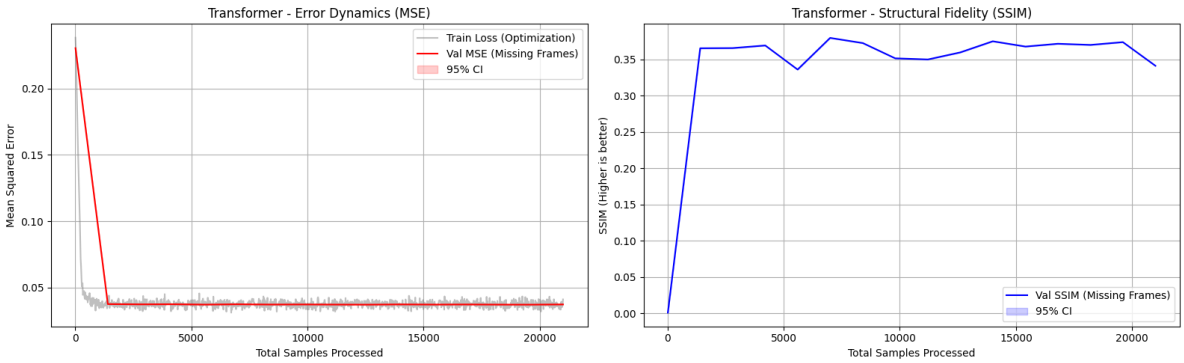


Figure 4 – Continuous Temporal Transformer training dynamics on MNIST dataset

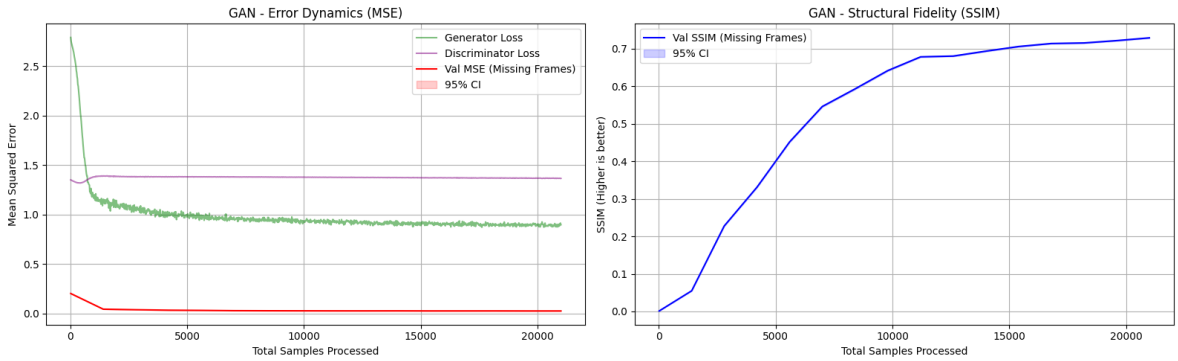


Figure 5 – GAN training dynamics on MNIST dataset

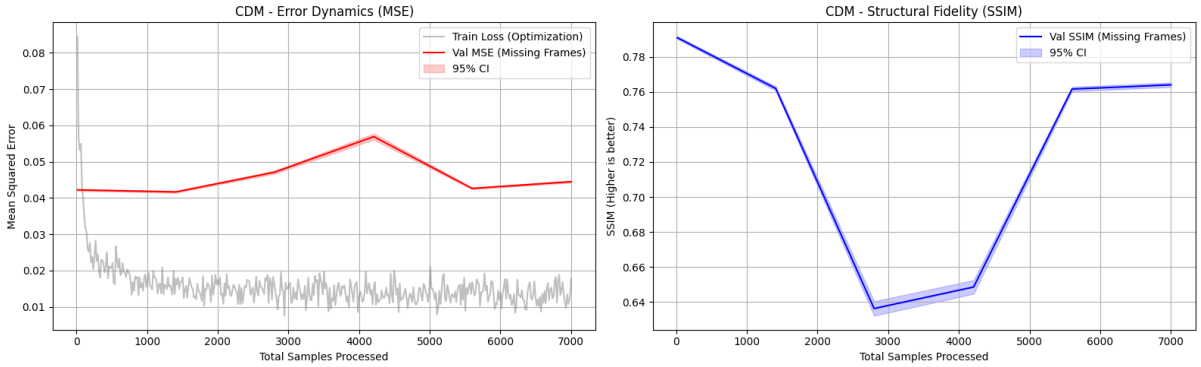


Figure 6 – CDM training dynamics on MNIST dataset

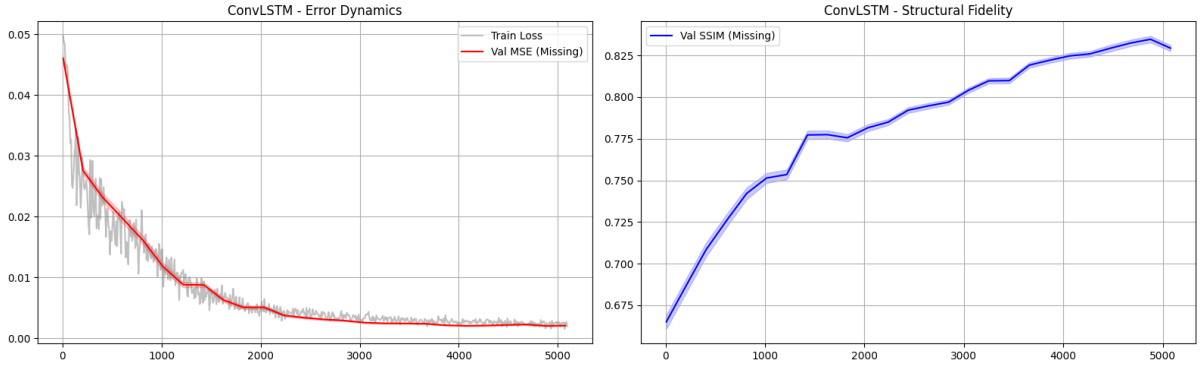


Figure 7 – Time-Aware ConvLSTM training dynamics on air temperature dataset

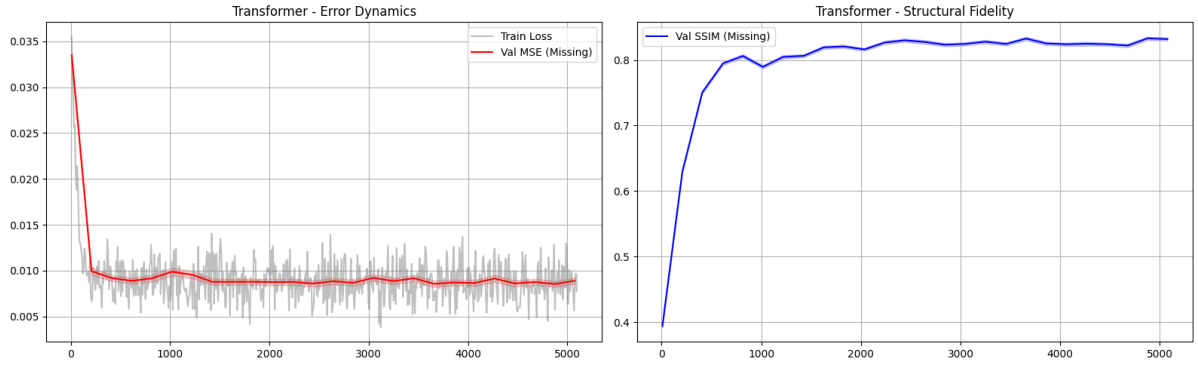


Figure 8 – Continuous Temporal Transformer training dynamics on air temperature dataset

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

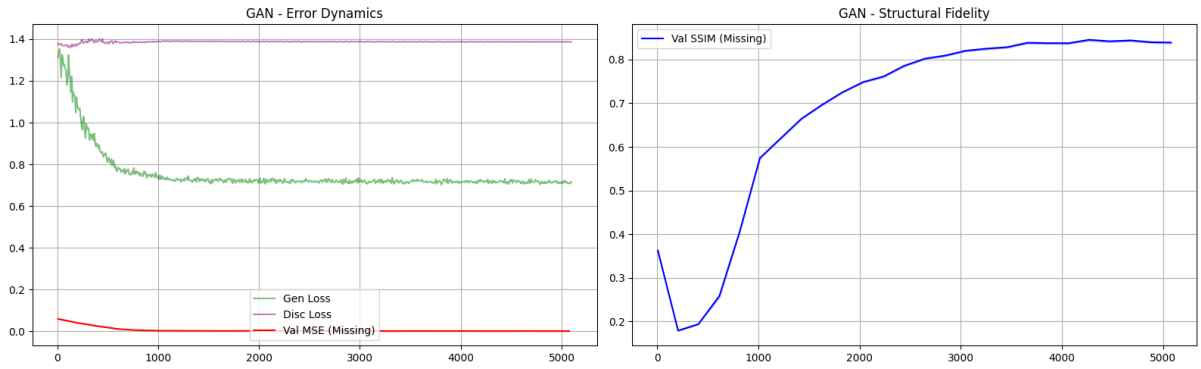


Figure 9 – GAN training dynamics on air temperature dataset

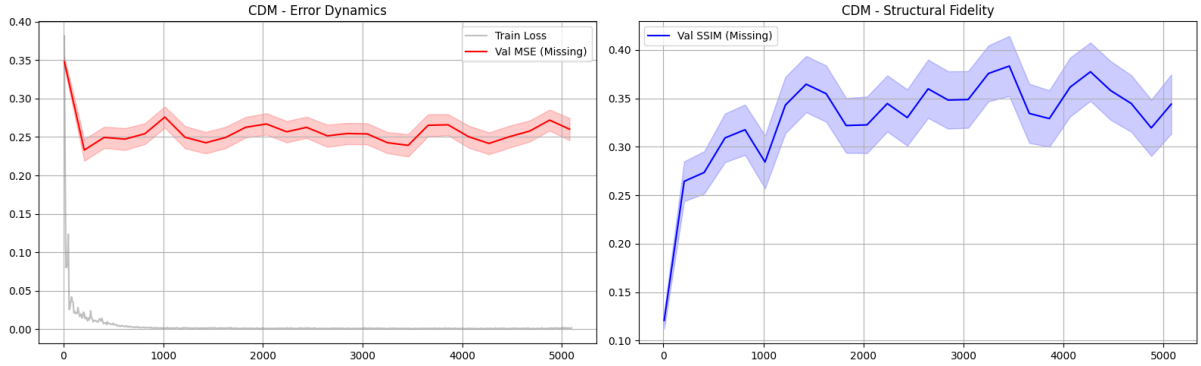


Figure 10 – CDM training dynamics on air temperature dataset

Table 1 and Table 2 summarize the final quantitative evaluation of all architectures across the two datasets.

Table 1

Forecast performance on irregular Moving MNIST

	Mean Squared Error (MSE) ± 95% CI	Structural Similarity (SSIM) ± 95% CI
Time-Aware ConvLSTM	0.0390 ± 0.0002	0.6037 ± 0.0030
Continuous Temporal Transformer	0.0374 ± 0.0002	0.3413 ± 0.0002
Spatio-Temporal GAN	0.0257 ± 0.0002	0.7289 ± 0.0011
Continuous Diffusion Model (CDM)	0.0445 ± 0.0003	0.7640 ± 0.0013

Table 2

Forecast performance on irregular Meteorological Fluid Dynamics dataset

	Mean Squared Error (MSE) ± 95% CI	Structural Similarity (SSIM) ± 95% CI
Time-Aware ConvLSTM	0.0021 ± 0.0001	0.8294 ± 0.0018
Continuous Temporal Transformer	0.0089 ± 0.0003	0.8320 ± 0.0025
Spatio-Temporal GAN	0.0018 ± 0.0001	0.8391 ± 0.0021
Continuous Diffusion Model (CDM)	0.2601 ± 0.0144	0.3441 ± 0.0305

On the meteorological dataset comprising atmospheric climate records, the ConvLSTM demonstrated highly stable, monotonic convergence. Across the 5,076 processed samples, the Validation MSE underwent a rapid exponential decay, ultimately settling at an 0.0021 ± 0.0001. Correspondingly, the structural fidelity climbed in a smooth, near-linear trajectory, achieving a final SSIM of 0.8294 ± 0.0018 on the missing frames. Because ConvLSTMs are optimized via MSE, they inherently produce spatial averages when uncertain about exact future states. When forecasting the shifting gradients of near-surface air temperatures, a slightly blurred spatial average still maintains good structural cohesion and gradient continuity, thereby yielding low MSE and high SSIM scores. Conversely, the limitations of deterministic MSE-optimized sequence modeling became apparent on the Moving MNIST dataset. ConvLSTM plateaued quite early, achieving a final MSE of 0.0390 ± 0.0002 and SSIM of 0.6037 ± 0.0030. When the ConvLSTM is uncertain about the exact linear trajectory or the precise angle of an elastic collision during a temporal gap, it hedges its prediction by outputting an overlapping average of all possible future positions. While this mathematically minimizes the worst-case MSE penalty, it results visually in severe blurring or "ghosting" of the final prediction.

The Continuous Temporal Transformer served as the non-recurrent, attention-based baseline. The training dynamics of the Transformer on the Moving MNIST dataset exhibited a rapid initial convergence followed by a plateau. As observed in the error dynamics, the Validation MSE plummeted almost instantly within the first 1,500 samples, dropping to its minimum boundary and remaining entirely flat for the remainder of the 21,000-sample training cycle. The final recorded MSE on the missing frames was 0.0374 ± 0.0002, which is mathematically competitive with the recurrent ConvLSTM baseline. However, an analysis of the SSIM reveals a critical vulnerability in the attention-based approach for predicting the trajectories of rigid bodies. While the SSIM initially spiked alongside the dropping MSE, it quickly plateaued and stagnated, ultimately finishing at a score of 0.3413 ± 0.0002. In contrast to its performance on rigid body dynamics, the Continuous Temporal Transformer demonstrated high capability when applied to the meteorological dataset of atmospheric climate records.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

The training dynamics reveal a highly efficient convergence profile. Within the first 1,000 samples, the Validation MSE plummeted to its near-minimum, while the SSIM simultaneously spiked from 0.4 to over 0.8. By the conclusion of the 5,076-sample training cycle, the Transformer achieved a final MSE of 0.0089 ± 0.0003 and a SSIM of 0.8320 ± 0.0025 on the stochastically dropped frames.

The training dynamics of the GAN on the Moving MNIST dataset illustrate the mechanics of adversarial equilibrium. Looking at the error dynamics, the Generator loss initially drops sharply as it quickly learns the general spatial distribution of the moving objects in the image sequences, while the Discriminator loss remains relatively stable, indicating a healthy adversarial tension without mode collapse. Correspondingly, the Validation MSE plummets in the first few thousand samples, flatlining at 0.0257 ± 0.0002 across the 21,000-sample training cycle. The structural fidelity climbs monotonically and smoothly throughout the training process, ultimately achieving a final SSIM of 0.7289 ± 0.0011 on the stochastically dropped frames. When applied to the meteorological dataset, the Spatio-Temporal GAN produced the most mathematically precise and structurally cohesive reconstructions of any evaluated architecture. The training dynamics reveal a sequence of adversarial adaptation. In the initial phases of training (the first 1,000 samples), the Validation SSIM experienced a sharp, transient dip, falling below 0.20. The Generator loss steadily decayed to a stable plateau, and the Validation MSE plunged to near absolute zero. Simultaneously, the SSIM reversed its initial decline and executed a smooth, monotonic climb, ultimately reaching a peak Validation SSIM of 0.8391 ± 0.0021 alongside the MSE of 0.0018 ± 0.0001 by the end of the 5,076-sample training cycle.

Unlike the smooth, monotonic curves of the ConvLSTM or GAN, the CDM's structural fidelity exhibits a "V-shape" trajectory on the Moving MNIST dataset. In the earliest stages of training, the Validation SSIM starts high, before plummeting to roughly 0.63 near the 3,000-sample mark. Initially, the untrained U-Net fails to properly estimate the noise gradients, actively destroying the image during the reverse diffusion loop. However, once the network's gradients align with the true data manifold, the model experiences a rapid recovery. The SSIM climbs, ultimately finishing at 0.7640 ± 0.0013 by the end of the 7,000-sample cycle. The Validation MSE for the CDM does not achieve the near-zero values seen in the ConvLSTM or GAN, instead fluctuating and ultimately settling at a higher 0.0445 ± 0.0003. While a deterministic model outputs a blurry average that strictly minimizes MSE, the CDM tries to generate a highly realistic prediction. However, if that generated trajectory is shifted by even one or two pixels from the exact ground-truth trajectory, the rigid MSE metric penalizes it heavily. While the CDM demonstrated a remarkable capacity to reconstruct the rigid boundaries and linear momentum of Moving MNIST, its application to the atmospheric climate records revealed a severe architectural vulnerability. When tasked with interpolating chaotic, low-frequency temperature gradients, the CDM exhibited the poorest mathematical and structural performance of all evaluated models. By the end of the 5,076-sample training cycle, the model achieved a Validation MSE of 0.2601 ± 0.0144 and a Validation SSIM of

0.3441 ± 0.0305 on the missing frames. The training loss decayed rapidly to near-zero, proving that the U-Net successfully learned to minimize the denoising objective on the training set. However, the Validation MSE hit a hard plateau almost immediately, hovering at an exceptionally high error rate. Simultaneously, the structural fidelity (SSIM) trajectory exhibited extreme instability, oscillating with large 95% CI indicating a complete lack of predictive consensus across the validation batches. When the CDM attempts to denoise these atmospheric states across an irregular temporal gap, it struggles to differentiate between the artificial Gaussian noise it must remove and the natural, dispersion of the weather system it must preserve. Furthermore, because unguided diffusion is inherently stochastic, the model generates plausible but mathematically incorrect weather states that diverge largely from the true physical trajectory, leading to the high absolute pixel error. These results indicate that while standard continuous-time diffusion is highly effective for predicting trajectories of rigid objects, it is fundamentally ill-suited for interpolating chaotic, low-frequency atmospheric climate data without explicit physics-informed regularization.

Conclusions. The empirical results confirm that the efficacy of continuous-time forecasting architectures is fundamentally linked to the frequency domain and physical nature of the target data. Adapted discrete-time models, such as Time-Aware ConvLSTMs, struggle significantly with high-frequency rigid body dynamics, frequently suffering from severe structural degradation due to the averaging of uncertain futures. Similarly, pure Continuous Temporal Transformers fail to preserve sharp boundaries when predicting linear trajectories, smearing predictions to minimize absolute error. However, for chaotic, low-frequency atmospheric climate records, the Transformer's global attention mechanism excels. Ultimately, the Spatio-Temporal GAN proved to be the robust architecture, utilizing adversarial regularization to prevent blurring and achieve state-of-the-art structural fidelity and pixel precision in across both rigid-body trajectories and complex atmospheric environments. Conversely, the CDM demonstrated high structural reconstruction for rigid body trajectoris but failed on smooth atmospheric temperature gradients due to the unintended injection of structural noise.

While the experiments conducted on the stochastic Moving MNIST and atmospheric climate record datasets successfully highlight the foundational differences between these algorithmic paradigms across both linear and non-linear physics, they operate within 2D constraints. A critical next step for future research is to validate these findings on high-dimensional, real-world datasets that inherently feature unequal time intervals and volumetric geometry. Specifically, extending these continuous-time architectures into 3D spatial environments is essential for emerging XR systems, where robust spatio-temporal forecasting is strictly required to anchor digital objects seamlessly despite sensor interruptions, asynchronous polling, and network latency. Additionally, testing these models on longitudinal 3D MRI medical scans to track disease progression, or on high-resolution satellite imagery, is necessary to fully assess their computational scalability and practical clinical or environmental utility.

Furthermore, the contrasting strengths and vulnerabilities of the evaluated models suggest a highly promising avenue for future architectural development: a forecasting-generation

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

framework. Because Continuous Temporal Transformers excel at modeling long-range global dependencies and tracking smooth, interconnected gradients, they are well suited to act as the temporal encoder. This Transformer could process the irregularly sampled historical context to output a highly accurate conditional latent vector for a specific target time. To overcome the Transformer's inability to render sharp boundaries, this condition could then be fed into a GAN acting as the spatial decoder. By utilizing the Transformer to resolve the continuous-time temporal dynamics and the GAN to solve the reverse stochastic differential equation, such model could theoretically achieve both the rigorous temporal accuracy of attention mechanisms and the high structural fidelity of score-based generative modeling.

REFERENCES

1. Shi, X., Chen, Z., Wang, H., Yeung, D. Y., Wong, W. K., & Woo, W. C. (2015). Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting. *Advances in Neural Information Processing Systems*, *28*.

2. Neil, D., Pfeiffer, M., & Liu, S. C. (2016). Phased LSTM: Accelerating Recurrent Network Training for Long or Event-based Sequences. *Advances in Neural Information Processing Systems*, *29*.

3. Cui, Z., Henrickson, K., Ke, R., & Wang, Y. (2019). Traffic Graph Convolutional Long Short-Term Memory Neural Network for Urban Traffic Flow Prediction. *Transportation Research Part C: Emerging Technologies*, 105, 361-374.

4. Vaswani, A., Shazeer, N., Parmar, N.,Uszkoreit, J., Jones, L.,Gomez, A.N.,& Polosukhin,I.(2017).Attention is all you need. *Advances in neural information processing systems, 30*.

5. Fonseca, A. H. D. O., Zappala, E., Caro, J. O., & Van Dijk, D. (2023). Continuous spatio-temporal transformers. *arXiv preprint* arXiv:2301.13338.

6. Rußwurm, M., & Körner, M. (2020). Self-Attention for Raw Optical Satellite Time Series Classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 169, 421-435.

7. Tipirneni, S., & Reddy, C. K. (2022). Self-Supervised Transformer for Sparse and Irregularly Sampled Multivariate Clinical Time Series. *ACM Transactions on Knowledge Discovery from Data*, 16(6), 1-17.

8. Rubanova, Y., Chen, R. T. Q., Bettencourt, J., & Duvenaud, D. (2018). Neural Ordinary Differential Equations. *Advances in Neural Information Processing Systems*, 31.

9. De Brouwer, E., Simm, J., Arany, A., & Moreau, Y. (2019). GRU-ODE-Bayes: Continuous Modeling of Sporadically-Observed Time Series. *Advances in Neural Information Processing Systems*, 32.

10. Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2020). Generative Adversarial Networks. *Communications of the ACM*, *63*(11), 139–144. https://doi.org/https://doi.org/10.1145/3422622

11. Xie, Y., Franz, E., Chu, M., & Thuerey, N. (2018). tempoGAN:A Temporally Coherent, Volumetric GAN for Super-resolution Fluid Flow. *ACM Transactions on Graphics*, 37(4), 1-15.

12. Song, Y., Sohl-Dickstein, J., Kingma, D. P., Kumar, A., Ermon, S., & Poole, B. (2020). Score-based generative modeling through stochastic differential equations. *arXiv preprint* arXiv:2011.13456.

13. Price, I., Sanchez-Gonzalez, A., Alet, F., et al. (2023). GenCast: Diffusion-based Ensemble Forecasting for Medium-Range Weather. *arXiv preprint* arXiv:2312.15796.

14. Srivastava, N., Mansimov, E., & Salakhutdinov, R. (2015). Unsupervised learning of video representations using LSTMs. Proceedings of the 32nd International Conference on Machine Learning (ICML). http://www.cs.toronto.edu/~nitish/unsupervised_video/
15. Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., & Joseph, D. (2018). The NCEP/NCAR 40-year reanalysis project. *Renewable energy*, 1, 146-194.

Received 24.06.2026.
Accepted 25.08.2026.
Published 30.10.2026

Часово-просторове прогнозування цифрових зображень з неоднорідними часовими інтервалами: огляд глибоких нейромережевих підходів

У дослідженні оцінено чотири архітектури глибокого навчання (ConvLSTM, Transformer, Spatio-Temporal GAN та дифузійна модель) для прогнозування послідовностей зображень через нерегулярні часові інтервали. Тестування на наборах Moving MNIST таатмосферних кліматичних даних показало, що детерміновані моделі розмивають траєкторії жорстких об'єктів, тоді як GAN забезпечує високу структурну точність на обох наборах даних. Дифузійна модель чудово генерує траєкторії жорстких об'єктів, але не справляється з плавними атмосферними градієнтами. Запропоновано концепцію Transformer-GAN архітектури для забезпечення високоточного просторово-часового прогнозування цифрових зображень з неоднорідними часовими інтервалами.

Емпіричні результати підтверджують, що ефективність архітектур прогнозування в безперервному часі фундаментально пов'язана з частотною областю та фізичною природою цільових даних. Адаптовані дискретні моделі, такі як Time-Aware ConvLSTM, значно стикаються з труднощами при прогнозуванні високочастотної динаміки твердого тіла, часто страждаючи від серйозної структурної деградації через усереднення невизначеного майбутнього.

Федорчук Іван Васильович – аспірант кафедри програмного забезпечення комп’ютерних систем, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Україна.
ORCID: http://orcid.org/0009-0009-4429-1683
Шкурат Оксана Сергіївна – доцент кафедри програмного забезпечення комп’ютерних систем, Національний технічний університет України “Київський політехнічний інститут імені Ігоря Сікорського”, Київ.
ORCID: http://orcid.org/0000-0001-7633-9121

Fedorchuk Ivan – PhD student of the Department of Computer Systems Software, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Ukraine.
ORCID: http://orcid.org/0009-0009-4429-1683
Shkurat Oksana – Associate Professor of Computer Systems Software Department, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Kyiv.
ORCID: http://orcid.org/0000-0001-7633-9121

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

DOI 10.34185/1562-9945-5-166-2026-21
УДК 004.8:37.018.43

К.О. Палагута, О.Є. Савон

ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ПІДТРИМКИ НАВЧАЛЬНИХ ТРАЄКТОРІЙ НА ОСНОВІ РЕКУРЕНТНИХ НЕЙРОННИХ МЕРЕЖ І СИТУАЦІЙНО-СЦЕНАРНОГО ПІДХОДУ

Анотація. Актуальність дослідження зумовлена необхідністю підвищення ефективності підтримки навчальних траєкторій студентів у системах дистанційного навчання на основі інтелектуального аналізу накопичених освітніх даних. Сучасні освітні платформи забезпечують значні обсяги інформації про результати та активність студентів, однак прогнозування успішності саме по собі не визначає подальших педагогічних дій. Проблема дослідження полягає у необхідності переходу від отримання прогнозних показників до їх ситуаційної інтерпретації та формування адаптивних сценаріїв підтримки. Метою роботи є розроблення інформаційної технології підтримки навчальних траєкторій на основі поєднання рекурентних нейронних мереж і ситуаційно-сценарного підходу. Використано методи нейромережевого прогнозування часових послідовностей, ситуаційного аналізу, продукційні правила, порогові механізми прийняття рішень та експериментальне моделювання. Запропоновано багаторівневу архітектуру технології, що охоплює збір і підготовку даних, прогнозування, ситуаційну інтерпретацію, формування сценаріїв, звітність і візуалізацію. RNN-модель забезпечує прогнозування підсумкової оцінки та ймовірності успішного завершення курсу, а ситуаційно-сценарний модуль визначає стан навчальної траєкторії та формує відповідний набір педагогічних дій. Розроблено програмний прототип, інтегрований з LMS Moodle, та проведено його апробацію на модельних даних дисципліни «Основи програмування». Результати підтвердили працездатність повного технологічного циклу та можливість формування диференційованих сценаріїв підтримки з урахуванням прогнозованих результатів і показників навчальної активності студентів. Запропонована технологія створює передумови для своєчасного виявлення ризиків, адаптивного супроводу навчальних траєкторій та підтримки прийняття обґрунтованих педагогічних рішень у дистанційному навчанні.

Ключові слова: дистанційне навчання, освітня платформа, навчальна траєкторія, рекурентна нейронна мережа, RNN, GRU, прогнозування успішності, ситуаційно-сценарний підхід, адаптивна підтримка, LMS Moodle.

Постановка проблеми. Стрімкий розвиток технологій дистанційного навчання та поширення освітніх платформ зумовлюють накопичення значних обсягів даних про навчальну діяльність студентів, зокрема результати оцінювання, інтенсивність взаємодії з навчальними матеріалами, виконання тестових і практичних завдань, дотримання тер-

© Палагута К.О., Савон О.Є., 2026
Copyright for this paper by its authors. Use permitted under License CC BY 4.0.
ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

мінів та інші поведінкові характеристики. Використання цих даних створює передумови для переходу від традиційного ретроспективного оцінювання результатів навчання до прогнозування подальшої успішності та своєчасного виявлення ризиків відставання. Особливий потенціал у цьому контексті мають рекурентні нейронні мережі, здатні враховувати часову послідовність навчальних подій та динаміку поведінки студента.

Водночас отримання прогнозу саме по собі не забезпечує ефективної підтримки навчальної траєкторії. Прогнозована підсумкова оцінка або ймовірність успішного завершення курсу потребують подальшої інтерпретації з урахуванням фактичної активності студента та перетворення на конкретні педагогічні рекомендації й керуючі дії. Тому актуальною є задача інтеграції засобів нейромережевого прогнозування із ситуаційно-сценарним підходом, який дозволяє ідентифікувати поточну навчальну ситуацію та сформувати відповідний сценарій підтримки.

Отже, виникає необхідність розроблення інформаційної технології, яка забезпечувала б цілісний цикл опрацювання освітніх даних: від аналізу послідовності навчальних подій і прогнозування результатів засобами RNN до ситуаційної інтерпретації прогнозу, вибору сценарію та формування адаптивних рекомендацій у середовищі дистанційного навчання. Розв’язання цієї задачі створює підґрунтя для своєчасного виявлення несприятливих змін у навчальній траєкторії та підтримки прийняття обґрунтованих педагогічних рішень.

Аналіз останніх досліджень і публікацій. Цифрова трансформація освіти сприяла активному розвитку інтелектуальних освітніх платформ, здатних не лише накопичувати дані про навчальну діяльність здобувачів освіти, а й адаптувати освітній контент відповідно до індивідуальних потреб користувачів. Одним із найбільш перспективних напрямів розвитку таких платформ є використання рекурентних нейронних мереж (Recurrent Neural Networks, RNN) та сценарно-орієнтованого навчання (Scenario-Based Learning, SBL). Останні дослідження демонструють можливість поєднання цих підходів для створення адаптивних освітніх середовищ нового покоління.

Однією із фундаментальних праць у сфері застосування RNN в освіті є дослідження Deep Knowledge Tracing [1], у якому було запропоновано використовувати рекурентні нейронні мережі для моделювання динаміки навчання студентів. Автори довели, що RNN здатні ефективно відстежувати зміни рівня знань здобувача освіти на основі послідовності виконаних завдань та прогнозувати результати майбутньої діяльності. Запропонований підхід став основою для розвитку сучасних адаптивних систем навчання та технологій learning analytics.

Подальшого розвитку цей напрям набув у роботі [2], де було розроблено систему раннього попередження про академічні ризики студентів у STEM-курсах. Автори використали дані навчальної платформи, зокрема частоту входів до системи, активність на форумах та виконання домашніх завдань. Результати продемонстрували, що модель RNN забезпечує більш високу точність виявлення студентів групи ризику порівняно з традиційними алгоритмами машинного навчання. Особливою перевагою запропонова-

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

ного підходу є можливість використання обмеженого набору характеристик, доступних практично в будь-якій LMS.

Значна увага в сучасних дослідженнях приділяється також аналізу поведінкових характеристик студентів. Так, у роботі Jeong і Cho [3] досліджувалася можливість визначення рівня концентрації студентів під час дистанційного навчання за допомогою моделей RNN, LSTM та GRU. Для аналізу використовувалися дані відеоспостереження, включаючи напрямок погляду та міміку користувачів. Автори встановили, що модель GRU забезпечує найвищу точність прогнозування концентрації, що відкриває перспективи для створення адаптивних систем підтримки навчання в реальному часі.

Важливий напрям сучасних досліджень пов'язаний із персоналізацією освітнього контенту. У роботі [4] запропонована рекомендаційна система, яка аналізує навчальну активність користувачів віртуального освітнього середовища та формує персоналізовані рекомендації щодо подальшого проходження навчального матеріалу. Автори показали, що використання модифікованої архітектури LSTM дозволяє суттєво підвищити точність рекомендацій та забезпечити ефективніше індивідуалізоване навчання.

Окремий напрям формується навколо прогнозування академічної успішності студентів. У роботі Fazil, Rísquez та Halpin [5] представлено модель ASIST, яка поєднує механізм уваги (Attention Mechanism) із багатошаровими двонапрямними мережами BiLSTM для аналізу даних віртуального навчального середовища. Отримані результати показали високі значення AUC (0,86–0,90) та можливість виявлення студентів групи ризику на ранніх етапах навчання.

Схожі результати отримано в дослідженні Kalita та співавторів [6], де для прогнозування середнього балу студентів було використано двонапрямну мережу Bi-LSTM. Особливістю роботи стало застосування SHAP-аналізу для інтерпретації результатів моделі. Автори довели, що пояснюване машинне навчання дозволяє не лише здійснювати точні прогнози, але й виявляти фактори, які найбільше впливають на успішність студентів.

Паралельно з розвитком технологій штучного інтелекту активно розвивається сценарний підхід до організації навчання. Дослідження Morgan [7] демонструє ефективність інтеграції Scenario-Based Learning та Universal Design for Learning в асинхронних онлайн-курсах. Автор показує, що навчальні сценарії, побудовані на реальних професійних ситуаціях, сприяють підвищенню залученості студентів та розвитку практичних компетентностей.

Подальший розвиток сценарного підходу представлений у праці Younas та співавторів [8], де досліджується поєднання цифрового сценарного навчання та генеративного штучного інтелекту. Результати експерименту підтвердили покращення комунікативних навичок студентів, підвищення їх мотивації та адаптивності освітнього процесу. Особливого значення набуває можливість автоматичного генерування індивідуальних навчальних сценаріїв залежно від потреб конкретного студента.

Аналіз сучасних досліджень дозволяє зробити висновок, що розвиток інтелектуальних освітніх платформ відбувається у напрямі інтеграції двох взаємодоповнювальних підходів. RNN та їх похідні (LSTM, GRU, Bi-LSTM) забезпечують аналіз часових пос-

лідовностей навчальної активності, прогнозування результатів і побудову адаптивних рекомендацій. Сценарний підхід, своєю чергою, створює умови для практико-орієнтованого навчання, розвитку критичного мислення та формування професійних компетентностей. Перспективним напрямом подальших досліджень є створення освітніх платформ, у яких нейронні мережі автоматично формуватимуть та адаптуватимуть сценарії навчання відповідно до індивідуальної траєкторії здобувача освіти.

Мета дослідження. Метою дослідження є розроблення інформаційної технології підтримки навчальних траєкторій у системах дистанційного навчання на основі поєднання рекурентних нейронних мереж і ситуаційно-сценарного підходу, що забезпечує прогнозування результатів навчальної діяльності, інтерпретацію прогнозних і поведінкових показників, визначення поточної навчальної ситуації та формування адаптивних сценаріїв педагогічної підтримки.

Виклад основного матеріалу. У межах розробки освітньої платформи дистанційного навчання однією з ключових задач є побудова нейромережної моделі, яка дозволяє здійснювати інтелектуальний аналіз навчального процесу, прогнозування й адаптацію контенту до потреб студента. Нехай для кожного користувача задана послідовність дій у часі:

$$X = \{x_1, x_2, ..., x_T\}, \quad (1)$$

де $x_t \in R^n$ — вектор параметрів користувача на кроці t. Кожен вектор включає: оцінки (g_t), час активності (a_t), перегляд матеріалів ($_{mt)}$, виконання тестів (t_t), тему заняття (s_t).

Побудована модель f реалізує:

- регресійне прогнозування оцінки: $\hat{y}_T = f(X)$;
- класифікацію ймовірності успішного завершення курсу: $\hat{y}_T = P$(успіх | X).

У рамках проєктування інтелектуального модуля для платформи дистанційного навчання обрано архітектуру нейромережі, яка: здатна обробляти послідовності подій у часі; ефективно враховує контекст навчального процесу; забезпечує точне прогнозування навчальних результатів; дозволяє генерувати рекомендації та підтримує адаптивну поведінку системи. Доцільним є використання рекурентної нейронної мережі (RNN), яка забезпечує внутрішню пам'ять системи завдяки механізму зворотного зв’язку. Кожен крок у послідовності впливає на наступний за допомогою прихованого стану, що дозволяє враховувати динаміку поведінки студента в часі. Розглянемо математичний опис RNN.

Умовні позначення:

$x_t \in R^n$ — вектор вхідних даних на кроці часу t;

$h_t \in R^m$ — прихований стан мережі на кроці t;

$y_t \in R^k$ — вихід мережі на кроці t;

$W_{xh} \in R^{mxn}$ — матриця ваг від входу до прихованого шару;

$W_{hh} \in R^{mxm}$ — матриця ваг між прихованими станами;

$W_{hy} \in R^{mxk}$ — матриця ваг від прихованого стану до виходу;

$b_h \in R^m$, $b_y \in R^k$ — вектори зміщення (bias terms);

σ_h — функція активації прихованого шару (наприклад, tanh);

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

σ_y — функція активації виходу (наприклад, softmax).

Оновлення стану здійснюється в такий спосіб - на кожному кроці часу t прихований стан h_t оновлюється за допомогою вхідного вектору x_t та попереднього стану h_{t-1}:

$$h_t = \sigma_h(W_{xh} * x_t + W_{hh} * h_{t-1} + b_h), \quad (2)$$

де σ_h – нелінійна активаційна функція.

Вихід мережі y_t обчислюється як:

$$y_t = \sigma_y(W_{hy} * h_t + b_y), \quad (3)$$

де σ_y – функція активації.

При навчанні на послідовності довжини T загальна функція втрат має вигляд:

$$\mathcal{L} = \sum_{t=1}^{T} \mathcal{L}_t(\hat{y}_t, y_t) \quad (4)$$

де $\hat{y}_t$ — справжнє (еталонне) значення виходу на кроці t.

Навчання RNN проводиться за допомогою алгоритму зворотного поширення помилки у часі (Backpropagation Through Time, BPTT), який враховує залежності між усіма кроками часу:

$$\frac{\partial \mathcal{L}}{\partial W} = \sum_{t=1}^{T} \frac{\partial \mathcal{L}_t}{\partial \hat{y}_t} \cdot \frac{\partial \hat{y}_t}{\partial h_t} \cdot \frac{\partial h_t}{\partial W} \quad (5)$$

У системі дистанційного навчання, що спроєктована, рекурентна нейронна мережа (RNN) використовується для аналізу послідовності дій студента з метою прогнозування його успішності, адаптації контенту, виявлення ризиків відставання тощо.

Кожен студент формує послідовність подій $\{x_1, x_2, ..., x_T\}$, де кожен x_t — це вектор ознак, наприклад:

- g_t — оцінка за практичну роботу
- a_t — час активності в системі
- v_t — час роботи з теоретичним матеріалом
- t_t — виконання тесту (0 або 1)
- s_t — тема заняття (у вигляді one-hot encoding)

Мережа обробляє послідовність $\{x_1, x_2, ..., x_T\}$, оновлюючи прихований стан h_t і формуючи прогноз y_t — наприклад, ймовірність успішного завершення модуля або очікувана підсумкова оцінка.

Стан прихованого шару оновлюється на кожному кроці часу:

$$h_t = \tanh(W_{xh} * x_t + W_{hh} * h_{t-1} + b_h) \quad (6)$$

Прогнозований вихід:

$$\hat{y}_t = \sigma(W_{hy} * h_t + b_y), \quad (7)$$

де W_{xh} — ваги входу;

W_{hh} — рекурентні ваги;

W_{hy} — ваги виходу;

σ — сигмоїда.

Задачі прогнозування:

- Прогноз оцінки:

$$\hat{y}_T = \text{expected grade} \in [0,100] \quad (8)$$

- Ймовірність завершити курс:

$$\hat{y}_T = P(\text{успішне завершення курсу} \mid x_1, x_2, \ldots, x_T) \quad (9)$$

Функція втрат на всій послідовності має вигляд:

$$\mathcal{L} = \sum_{t=1}^{T} \mathcal{L}_t(\hat{y}_t, y_t) \quad (10)$$

Зворотне поширення помилки у часі (BPTT) дозволяє оновлювати ваги W_{xh}, W_{hh}, W_{hy}.

Модель оцінюється за метриками: MAE, RMSE — для оцінок, Accuracy, F1-score — для класифікації успішності, AUC-ROC — для ймовірнісних прогнозів.

Сценарний модуль освітньої платформи, що пропонується, не виконує прогнозування, а інтерпретує прогноз разом із фактичними показниками активності та визначає ситуацію і набір керуючих дій.

Тобто вхідний вектор ситуаційно-сценарного модуля доцільно визначити як

$$\mathbf{x}_i(t) = \left[\hat{G}_i, \hat{P}_i, R_i, A_i(t), T_i(t), D_i(t), L_i(t), \Delta A_i(t), \ldots\right], \quad (11)$$

де $\hat{G}_i$— прогнозована підсумкова оцінка; $\hat{P}_i$ — прогнозована ймовірність успішного завершення курсу; R_i — рівень ризику; $A_i(t)$ — показник навчальної активності; $T_i(t)$ — результати тестування; $D_i(t)$ — показник дотримання термінів; $L_i(t)$ — показник роботи з навчальними матеріалами; $\Delta A_i(t)$ — зміна активності порівняно з попереднім періодом.

Далі працює rule engine, тобто ситуація визначається не нейронною мережею, а системою впорядкованих правил:

$$r_j : C_j(\mathbf{x}_i(t), \theta) \rightarrow S_j, \quad (12)$$

де C_j — логічна умова правила; θ — множина порогових значень із thresholds; S_j — відповідна навчальна ситуація.

Оскільки правила мають пріоритети, механізм вибору ситуації можна записати як

$$S_i(t) = S_{j^*}, \qquad j^* = arg \max_{j : C_j(\mathbf{x}_i(t))=1} p_j, \quad (13)$$

де p_j — пріоритет правила. Таким чином, якщо одночасно виконуються умови кількох правил, обирається ситуація, пов'язана з правилом найвищого пріоритету.

Наприклад, для критичного ризику правило може мати вигляд:

$$(\hat{P}_i < \theta_P) \wedge (\hat{G}_i < \theta_G) \wedge (A_i < \theta_A) \rightarrow S3_{\text{RISK_CRITICAL}}$$

Для зниження активності:

$$\Delta A_i < -\theta_{\Delta A} \wedge \hat{P}_i \geq \theta_P \rightarrow S2_{\text{ACTIVITY_DECLINE}}$$

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

А для ситуації відновлення особливо важливо використовувати динаміку, а не лише поточне значення:

$$A_i(t) > A_i(t-1) \wedge T_i(t) > T_i(t-1) \wedge R_i(t) > R_i(t-1) \rightarrow S4_{\text{RECOVERY}}$$

Після визначення ситуації застосовується відображення

$$\Phi: S_j \rightarrow \mathcal{A}_j,$$

де $\mathcal{A}_i = \{a_1, a_2, \dots, a_m\}$ — набір сценарних дій. Наприклад,

$$S3_{\text{RISK_CRITICAL}} \rightarrow \{a_{\text{student-message}}, a_{\text{teacher-notify}}, a_{\text{resource}}, a_{\text{consultation}}\}$$

На цьому етапі scenarios.json фактично виконує роль формалізованої бази знань сценарного модуля: thresholds → параметри прийняття рішень; rules → продукційні правила виду «умова → ситуація»; situations → семантичний опис станів навчальної траєкторії; scenarios → відображення «ситуація → набір дій»; moodle_action_mapping → відображення «сценарна дія → механізм реалізації в Moodle».

Всю модель сценарного підходу можна компактно формалізував як:

$$\mathbf{x}_i(t) \xrightarrow[\theta]{\mathcal{R}} S_i(t) \xrightarrow{\Phi} \mathcal{A}_j(t) \xrightarrow{\psi} \mathcal{M}_j(t), \qquad (14)$$

де $\mathcal{R}$ — множина правил; $S_i(t)$ — визначена ситуація; Φ — функція вибору сценарних дій; ψ — функція відображення сценарних дій; $\mathcal{M}_j(t)$ — множина фактично виконуваних дій у середовищі дистанційного навчання.

Отже, у роботі запропоновано механізм формалізованого перетворення прогнозно-поведінкового стану студента на педагогічну ситуацію, а ситуації — на адаптивний сценарій підтримки.

Архітектурно запропонована технологія реалізується як багатомодульний програмний комплекс, інтегрований із LMS (наприклад, Moodle) через механізми обміну даними (REST API, журнали подій, результати оцінювання).

Структурно виділено такі функціональні рівні:

1. Рівень збору та підготовки даних — забезпечує отримання журналів активності та результатів навчання, їх очищення, агрегування та перетворення у формат, придатний для аналітичної обробки.
2. Аналітичний рівень (прогнозування) — реалізує обчислення прогнозних показників успішності та ризику на основі рекурентної нейромережевої моделі.
3. Логічний рівень (ситуаційна інтерпретація) — здійснює класифікацію поточного стану навчальної траєкторії відповідно до визначених правил та порогових значень.
4. Сценарний рівень — формує набір педагогічних рекомендацій та управлінських дій залежно від встановленої ситуації.
5. Рівень звітності та візуалізації — забезпечує генерацію аналітичних звітів та графічних матеріалів для підтримки прийняття рішень викладачем.

Така багаторівнева структура дозволяє чітко розмежувати функціональні ролі компонентів і забезпечити масштабованість системи.

Функціонування технології організовано у вигляді замкненого адаптивного циклу, що включає: накопичення даних про навчальну діяльність; аналітичну обробку та прогнозування; інтерпретацію результатів; формування рекомендацій; реалізацію педагогічного впливу через LMS; подальше накопичення нових даних. Таким чином, створюється контур зворотного зв'язку, в якому результати прогнозування впливають на освітню поведінку, що, у свою чергу, змінює подальші аналітичні оцінки (рис.1).

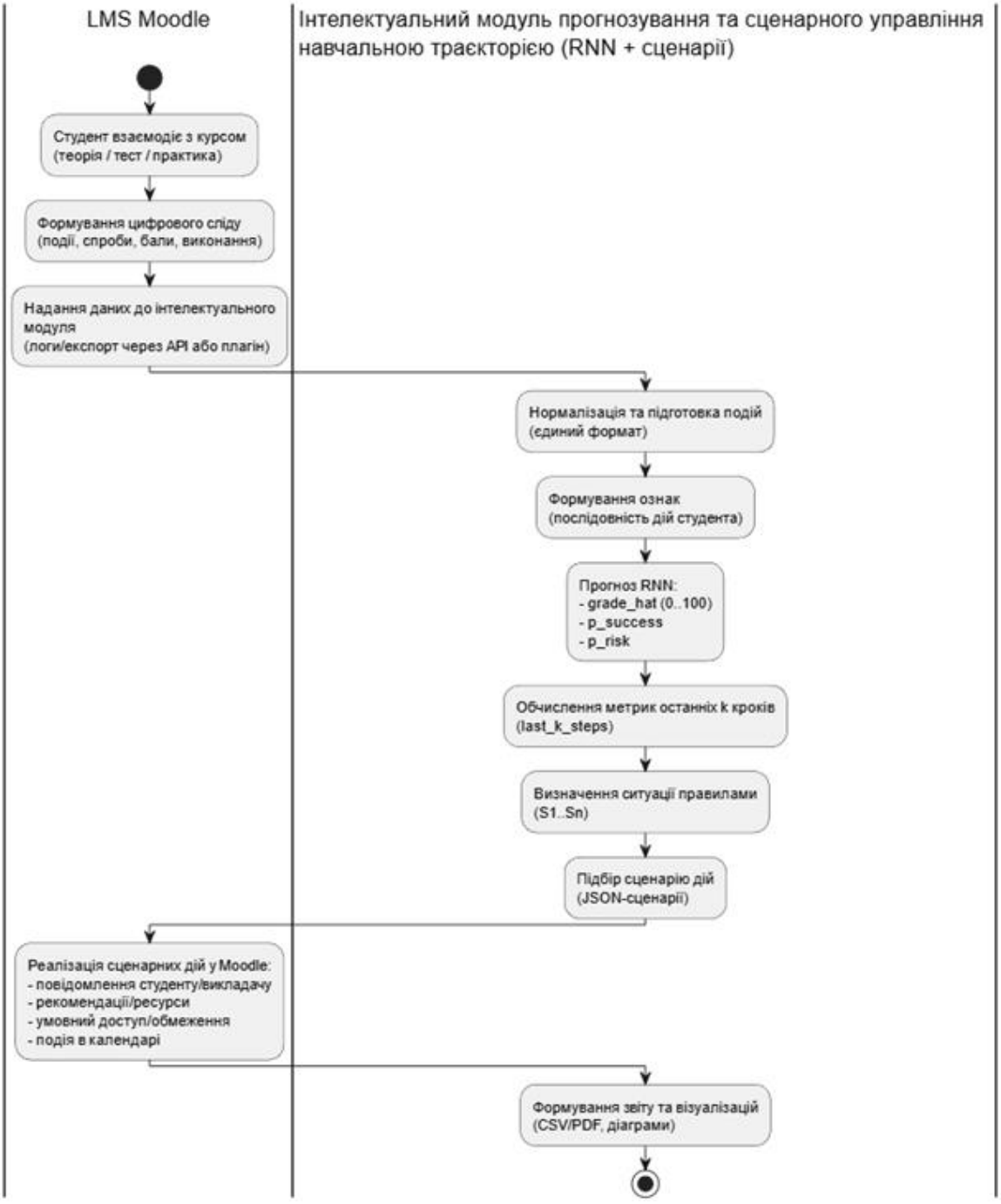


Рисунок 1 – Узагальнений цикл функціонування технології ситуаційно-сценарного управління навчальною траєкторією з використанням RNN

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Реалізація інформаційної технології освітньої платформи дистанційного навчання здійснена у вигляді програмного прототипу, який інтегрується з LMS Moodle та забезпечує повний цикл аналітичної обробки навчальної активності студентів: від збору подій до формування управлінських сценаріїв і звітів. Архітектурно система реалізована як багатомодульний програмний комплекс, що включає модуль підготовки даних, модуль прогнозування на основі RNN, модуль ситуаційно-сценарного аналізу та модуль формування аналітичних звітів (CSV/PDF та графічні діаграми). Розроблено програмне забезпечення, яке виконує побудову датасету та навчання RNN-моделі у прототипі (рис.2).

```python
from __future__ import annotations
import argparse, csv
from collections import defaultdict
from dataclasses import dataclass
from typing import Dict, List, Tuple, Optional

import torch
import torch.nn as nn
from torch.utils.data import Dataset, DataLoader
from torch.nn.utils.rnn import pack_padded_sequence


# ----------------------------
# Конфіг
# ----------------------------
@dataclass
class ModelConfig:
    num_features: int
    cat_features: Dict[str, Tuple[int, int]]
    rnn_type: str = "gru"
    hidden_size: int = 128
    num_layers: int = 1
    dropout: float = 0.2
    bidirectional: bool = False
    predict_risk: bool = True
    grade_min: float = 0.0
    grade_max: float = 100.0


# ----------------------------
# Dataset + collate
# ----------------------------
class MoodleseqDataset(Dataset):
    def __init__(self, samples: List[Dict], cat_feature_names: List[str], include_risk: bool = True):
        self.samples = samples
        self.cat_feature_names = cat_feature_names
        self.include_risk = include_risk

    def __len__(self) -> int:
        return len(self.samples)
```

Рисунок 2 – Фрагмент коду модуля підготовки датасету і навчання RNN-моделі

Після завершення навчання здійснюється оцінювання якості моделі на валідаційній вибірці. За результатами оцінювання обирається найкраща версія моделі, яка зберігається у вигляді файлу ваг. Ця версія моделі передається в режим експлуатації та використовується для прогнозування навчальних результатів у реальному часі. Надалі, у міру накопичення нових даних, може здійснюватися періодичне донавчання або повне

перенавчання моделі, що дозволяє адаптувати її до змін у навчальних траєкторіях студентів, структурі курсу або педагогічних підходах.

Режим експлуатації є основним режимом функціонування інтелектуального модуля у складі освітньої платформи. У цьому режимі модель не навчається, а використовується для прогнозування та підтримки прийняття педагогічних рішень. Джерелом даних виступають поточні події навчальної діяльності студентів у LMS Moodle. Ці події передаються до модуля підготовки даних, де здійснюється їх нормалізація, агрегування та формування послідовностей ознак, що відповідають поточній навчальній траєкторії студента. Підготовлені дані подаються на вхід навченої рекурентної моделі, яка формує прогнозні показники: очікувану підсумкову оцінку, ймовірність успішного завершення курсу та ймовірність ризику неуспішності. Отримані значення передаються до модуля ситуаційно-сценарного аналізу, де на основі заданих правил і порогових значень визначається поточна навчальна ситуація студента. Відповідно до визначеної ситуації формується сценарій дій, який може включати рекомендації щодо повторення матеріалу, нагадування про необхідність виконання завдань, пропозицію консультації або інші підтримувальні заходи.

Для експерименту використано дисципліну **«**Основи програмування», що структурована на 10 тем, кожна з яких містить три компоненти: теоретичний матеріал (умовно — читання/перегляд); тестування (quiz); практичне завдання (assignment/practice). Для відтворення логіки Moodle сформовано події типу theory / quiz / practice. Кожна подія характеризується часовими, поведінковими та результативними ознаками. Важливим елементом експерименту стала формалізація системи оцінювання, яка прив’язує дії студента до балів, максимально наближених до реальної практики викладання: опрацювання теоретичного матеріалу не менше 30 хв → 2 бали; тест → максимум 4 бали, при цьому кожна додаткова спроба зменшує результат на 10%; практичне завдання → максимум 4 бали, при запізненні мінус 10%; сумарна оцінка за дисципліну — 100 балів. Параметри вибірки: кількість студентів: 50; кількість тем дисципліни: 10; кількість активностей на тему: 3 (теорія, тест, практика); максимальна підсумкова оцінка: 100 балів; частка студентів, що завершують курс: ≈70 %.

Для кожного студента формувалася часово впорядкована послідовність подій довжиною до 30 кроків. Було проведено навчання RNN, побудовані графіки навчання: **к**рива навчальної втрати (рис.3), крива валідаційної втрати (рис.4). Також досліджено інференс системи на даних про вивчення перших 6-ти тем дисципліни, отримані результати ситуаційно-сценарного аналізу (рис.5), сформовані звіти (рис.6).

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

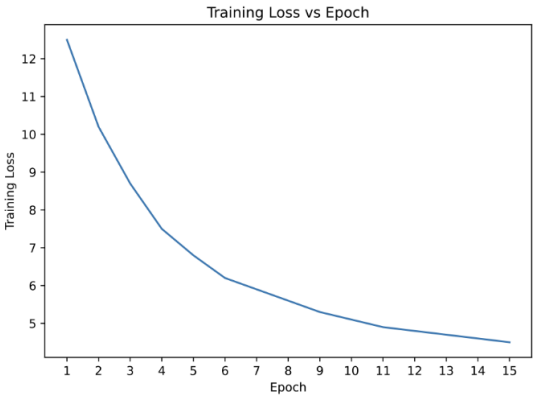


Рисунок 3 – Крива навчальної втрати

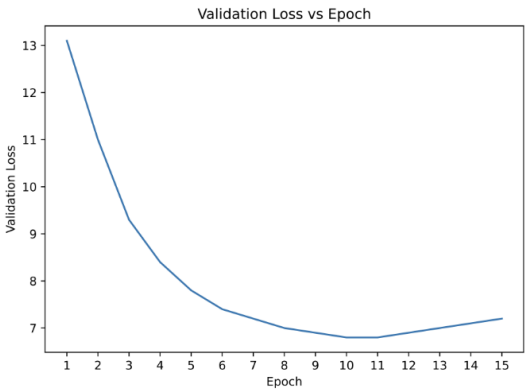


Рисунок 4 – Крива валідаційної втрати

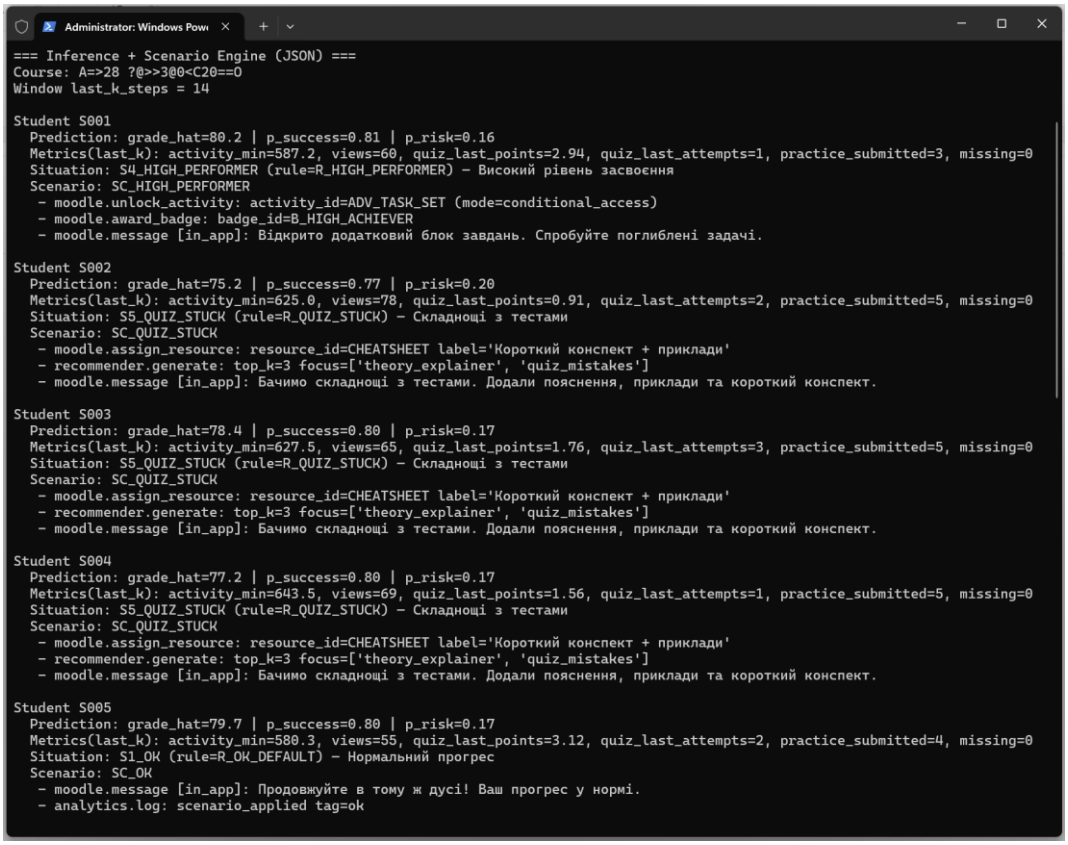
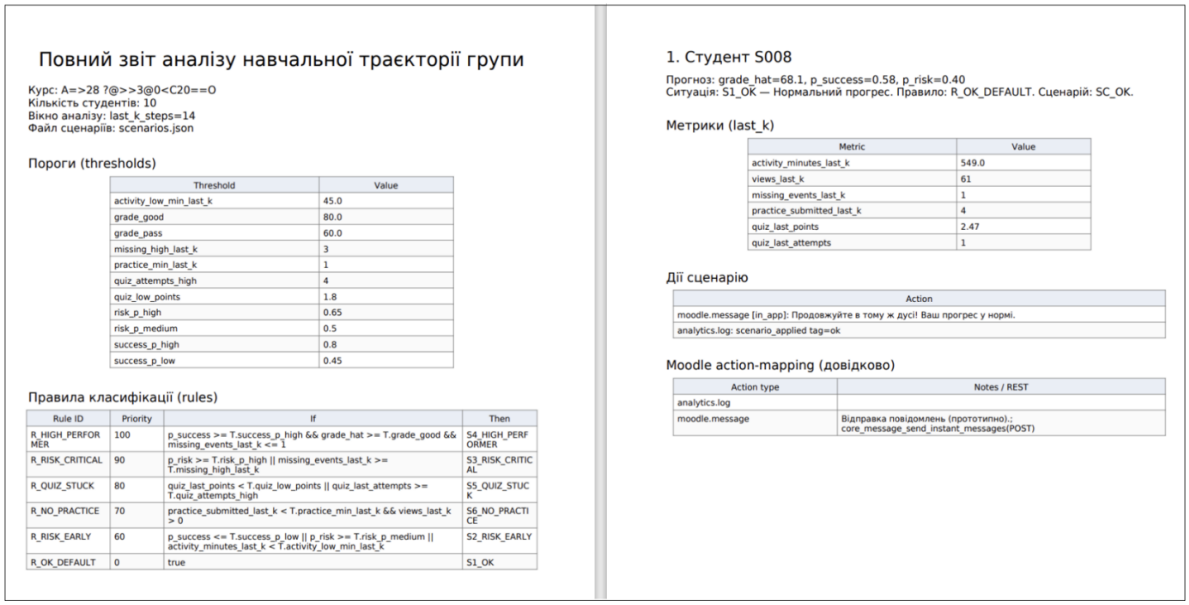


Рисунок 5 – Результати ситуаційно-сценарного аналізу



Повний звіт аналізу навчальної траєкторії групи

Курс: A=>28 ?@>>3@0<C20==O
Кількість студентів: 10
Вікно аналізу: last_k_steps=14
Файл сценаріїв: scenarios.json

Пороги (thresholds)

Threshold	Value
activity_low_min_last_k	45.0
grade_good	80.0
grade_pass	60.0
missing_high_last_k	3
practice_min_last_k	1
quiz_attempts_high	4
quiz_low_points	1.8
risk_p_high	0.65
risk_p_medium	0.5
success_p_high	0.8
success_p_low	0.45

Правила класифікації (rules)

Rule ID	Priority	If	Then
R_HIGH_PERFORMER	100	p_success >= T.success_p_high && grade_hat >= T.grade_good && missing_events_last_k <= 1	S4_HIGH_PERFORMER
R_RISK_CRITICAL	90	p_risk >= T.risk_p_high || missing_events_last_k >= T.missing_high_last_k	S3_RISK_CRITICAL
R_QUIZ_STUCK	80	quiz_last_points < T.quiz_low_points || quiz_last_attempts >= T.quiz_attempts_high	S5_QUIZ_STUCK
R_NO_PRACTICE	70	practice_submitted_last_k < T.practice_min_last_k && views_last_k > 0	S6_NO_PRACTICE
R_RISK_EARLY	60	p_success <= T.success_p_low || p_risk >= T.risk_p_medium || activity_minutes_last_k < T.activity_low_min_last_k	S2_RISK_EARLY
R_OK_DEFAULT	0	true	S1_OK

1. Студент S008

Прогноз: grade_hat=68.1, p_success=0.58, p_risk=0.40
Ситуація: S1_OK — Нормальний прогрес. Правило: R_OK_DEFAULT. Сценарій: SC_OK.

Метрики (last_k)

Metric	Value
activity_minutes_last_k	549.0
views_last_k	61
missing_events_last_k	1
practice_submitted_last_k	4
quiz_last_points	2.47
quiz_last_attempts	1

Дії сценарію

Action
moodle.message [in_app]: Продовжуйте в тому ж дусі! Ваш прогрес у нормі.
analytics.log: scenario_applied tag=ok

Moodle action-mapping (довідково)

Action type	Notes / REST
analytics.log	
moodle.message	Відправка повідомлень (прототипно).: core_message_send_instant_messages(POST)

Рисунок 6 – Фрагмент звіту з результатами ситуаційно-сценарного аналізу

Висновки. У роботі запропоновано інформаційну технологію підтримки навчального процесу в освітніх платформах дистанційного навчання, що поєднує прогнозування на основі рекурентних нейронних мереж із ситуаційно-сценарним механізмом інтерпретації результатів та формування педагогічних рекомендацій. RNN-модель забезпечує прогнозування підсумкової оцінки та ймовірності успішного завершення курсу, тоді як ситуаційно-сценарний модуль на основі прогнозних і фактичних показників, системи правил, порогових значень та пріоритетів визначає поточну навчальну ситуацію і формує відповідний сценарій підтримки.

Запропоновано багаторівневу архітектуру технології, що охоплює збір і підготовку даних, прогнозування, ситуаційну інтерпретацію, формування сценаріїв, звітність та візуалізацію. Її практичну реалізацію здійснено у вигляді програмного прототипу, інтегрованого з LMS Moodle та організованого у двох взаємопов’язаних режимах — навчання моделі та експлуатації.

Експериментальна апробація на модельних даних дисципліни «Основи програмування» підтвердила працездатність повного технологічного циклу — від обробки навчальних подій і прогнозування результатів до визначення ситуацій, формування диференційованих рекомендацій та аналітичної звітності. Отримані результати підтверджують можливість використання запропонованої технології для підтримки прийняття педагогічних рішень та індивідуалізації навчальних траєкторій. Подальші дослідження доцільно спрямувати на апробацію технології на реальних даних LMS, удосконалення механізмів налаштування сценарних правил та оцінювання впливу сформованих рекомендацій на результати навчання.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

ЛІТЕРАТУРА / REFERENCES

1. Piech C., Spencer J., Huang J., Ganguli S., Sahami M., Guibas L., Sohl-Dickstein J. Deep Knowledge Tracing. Advances in Neural Information Processing Systems. 2015. Vol. 28. DOI: https://doi.org/10.48550/arXiv.1506.05908.
2. Yu C.-C., Wu Y. Early Warning System for Online STEM Learning—A Slimmer Approach Using Recurrent Neural Networks. Sustainability. 2021. Vol. 13, No. 22. Article 12461. DOI: https://doi.org/10.3390/su132212461.
3. Jeong Y.-S., Cho N.-W. Evaluation of e-learners' concentration using recurrent neural networks. Journal of Supercomputing. 2023. Vol. 79. P. 4146–4163.
DOI: https://doi.org/10.1007/s11227-022-04804-w.
4. Dynamic Educational Recommender System Based on Improved LSTM Neural Network. Scientific Reports. 2024. Vol. 14. DOI: https://doi.org/10.1038/s41598-024-54729-y.
5. Fazil M., Rísquez A., Halpin C. A Novel Deep Learning Model for Student Performance Prediction Using Engagement Data. Journal of Learning Analytics. 2024. Vol. 11, No. 2. P. 23–41. DOI: https://doi.org/10.18608/jla.2024.7985.
6. Kalita E., Alfarwan A. M., El Aouifi H., Kukkar A., Hussain S., Ali T., Gaftandzhieva S. Predicting Student Academic Performance Using Bi-LSTM: A Deep Learning Framework with SHAP-Based Interpretability and Statistical Validation. Frontiers in Education. 2025. Vol. 10. DOI: https://doi.org/10.3389/feduc.2025.1581247.
7. Morgan A. Enhancing Access in an Online Course Using Universal Design for Learning (UDL) and Scenario-Based Learning (SBL). TechTrends. 2024.
DOI: https://doi.org/10.1007/s11528-024-00981-y.
8. Younas M., Ismayil I., El-Dakhs D. A. S., Anwar B., Noor U. Evaluating the Effectiveness of Digital Scenario-Based English Teaching at the University Level Using Artificial Intelligence Generated Content. Frontiers in Education. 2025. Vol. 10.
DOI: https://doi.org/10.3389/feduc.2025.1670892.
9. Brown M. Scenario-based Learning. In: How to Use Digital Learning with Confidence and Creativity. Cheltenham : Edward Elgar Publishing, 2024. P. 260–268.
DOI: https://doi.org/10.4337/9781035311293.00047.
10. Ören F. Ş., Karapınar A., Sarı K., Demirer T. The Effect of Scenario-Based Learning on Eighth-Grade Students' Perceptions of Scientists. Journal of Educational Research and Practice. 2023. Vol. 13, No. 1. P. 99–122. DOI: https://doi.org/10.5590/JERAP.2023.13.1.09.

Received 24.06.2026.
Accepted 25.08.2026.
Published 30.10.2026

Information technology for supporting learning trajectories based on recurrent neural networks and a situation-scenario approach

Abstract. The relevance of this study is determined by the need to improve the effectiveness of supporting students' learning trajectories in distance learning systems through intelligent analysis of accumulated educational data. Modern educational platforms provide substantial amounts of information on student performance and activity; however, academic performance prediction alone does not determine subsequent pedagogical actions. The research problem lies in the need to move from obtaining predictive indicators to their situational in-

terpretation and the formation of adaptive support scenarios. The aim of the study is to develop an information technology for supporting learning trajectories based on the integration of recurrent neural networks and a situation-scenario approach. The research employs methods of neural network forecasting of time series, situational analysis, production rules, threshold-based decision-making mechanisms, and experimental modeling. A multi-level architecture of the proposed technology is developed, encompassing data collection and preparation, forecasting, situational interpretation, scenario generation, reporting, and visualization. The RNN model provides predictions of the final grade and the probability of successful course completion, while the situation-scenario module determines the state of a learning trajectory and generates an appropriate set of pedagogical actions. A software prototype integrated with LMS Moodle was developed and tested using model data from the course Fundamentals of Programming. The results confirmed the operability of the complete technological cycle and demonstrated the possibility of generating differentiated support scenarios based on predicted outcomes and indicators of students' learning activity. The proposed technology creates prerequisites for the timely detection of risks, adaptive support of learning trajectories, and informed pedagogical decision-making in distance education.

Keywords: distance learning, educational platform, learning trajectory, recurrent neural network, RNN, GRU, academic performance prediction, situation-scenario approach, adaptive support, LMS Moodle.

Палагута Катерина Олексіївна - кандидат економічних наук, доцент, доцент кафедри інженерії програмного забезпечення та кібербезпеки, Державний торговельно-економічний університет, м. Київ, Україна.
ORCID: https://orcid.org/0000-0003-1167-9509
Савон Олексій Євгенійович – аспірант кафедри інженерії програмного забезпечення та кібербезпеки, Державний торговельно-економічний університет, м. Київ, Україна.
ORCID: https://orcid.org/0009-0006-4905-8969

Palaguta Kateryna Oleksiivna – PhD in Economics, Associate Professor, Associate Professor of the Department of Software Engineering and Cybersecurity, State University of Trade and Economics, Kyiv, Ukraine.
ORCID: https://orcid.org/0000-0003-1167-9509
Savon Oleksii Yevheniiovych – Postgraduate Student (PhD Student) of the Department of Software Engineering and Cybersecurity, State University of Trade and Economics, Kyiv, Ukraine.
ORCID: https://orcid.org/0009-0006-4905-8969

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

DOI 10.34185/1562-9945-5-166-2026-22
УДК 004.65:502.51

О.Г. Рибальченко, А.М. Гриценко, С.В. Білашенко, В.С. Сорокін

ПРОЄКТУВАННЯ АРХІТЕКТУРИ ТА МОДЕЛІ ДАНИХ ІНФОРМАЦІЙНО-АНАЛІТИЧНОЇ СИСТЕМИ МОНІТОРИНГУ ВОДНИХ РЕСУРСІВ

Анотація. Управління екологічними даними на підприємствах здебільшого здійснюється вручну та децентралізовано, що суттєво ускладнює оперативне виявлення наднормативних скидів та підвищує ймовірність значних фінансових ризиків. Проблема полягає у виникненні архітектурного конфлікту під час створення системи моніторингу: жорсткі вимоги екологічного комплаєнсу щодо збереження безперервного аудиторського сліду суперечать потребі швидкого генерування аналітичних звітів, що вимагає денормалізації даних. Метою дослідження є розробка архітектури та моделі даних системи, здатної розв'язати суперечність між юридичною незмінністю ретроспективних масивів та високою швидкодією аналітики. У роботі застосовано методи системного аналізу, декомпозиції предметної області, бітемпорального моделювання, а також архітектурний патерн CQRS. Запропоновано концептуальну модель на основі стандарту бітемпоральних таблиць, яка запобігає незворотній втраті історичних станів та підтримує транзакційну цілісність під час зміни нормативів. Для прискорення просторово-часових об'єднань використано кластерні стовпчикові індекси з алгоритмами відсікання сегментів. Імплементація CQRS через матеріалізовані представлення фізично ізолювала аналітичне навантаження від операційного. Розроблена архітектура мінімізує ризик транзакційних блокувань та деградації продуктивності, формуючи технологічне підґрунтя для функціонування системи як інструменту оперативного контролю водокористування.
Ключові слова: архітектура бази даних, бітемпоральне моделювання, екологічний моніторинг, CQRS, матеріалізовані представлення, аудиторський слід, кластерні стовпчикові індекси, інтервальна алгебра.

Постановка проблеми. Станом на 2026 рік процеси екологічного моніторингу на підприємствах характеризуються глибокою децентралізацією. Управління інформацією здебільшого спирається на локальні електронні таблиці та паперові носії. Відсутність інтеграції з державними реєстрами зумовлює необхідність ручного пошуку актуальних екологічних норм безпосередньо у законодавчих актах. Підготовка звітності та перевірка дотримання квот вимагають ручного агрегування показників із численних розрізнених лабораторних журналів. Як наслідок, існуюча децентралізована обробка екологічних даних призводить до критичних затримок в ідентифікації наднормативних скидів забруднювачів, що суттєво ускладнює оперативне реагування на інциденти та створює значні фінансові ризики для підприємств.

© Рибальченко О.Г., Гриценко А.М., Білашенко С.В., Сорокін В.С., 2026
Copyright for this paper by its authors. Use permitted under License CC BY 4.0.
ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Під час проєктування централізованої інформаційно-аналітичної системи для розв'язання цієї проблеми виникає фундаментальний архітектурний конфлікт. Вимоги екологічного комплаєнсу диктують необхідність збереження безперервного аудиторського сліду, що передбачає зберігання кожної версії дозвільного документа та кожного внесеного результату вимірювання без права на фізичне видалення чи модифікацію, гарантуючи юридичну незмінність історичного контексту водокористування. Це зумовлює необхідність побудови суворо нормалізованих структур збереження інформації.

Водночас, управління екологічними ризиками потребує швидкої перевірки дотримання лімітів гранично допустимих скидів, що вимагає безперервної агрегації розрізнених ретроспективних масивів. Збереження великих обсягів історичних даних у нормалізованому вигляді спричиняє значне зниження швидкодії під час генерації складних багатовимірних аналітичних звітів.

Аналіз останніх досліджень і публікацій. Сучасні дослідження в галузі інженерії програмного забезпечення фіксують необхідність переходу від децентралізованого ручного обліку екологічних даних до централізованих реляційних архітектур. Традиційне використання електронних таблиць та ізольованих лабораторних журналів не гарантує дотримання властивостей ACID, що забезпечують транзакційну цілісність баз даних, і генерує критичні аномалії модифікації даних. Праці з управління водними ресурсами обґрунтовують міграцію інформаційних потоків водокористувачів до систем планування ресурсів підприємства (ERP) та систем екологічного менеджменту (EMS) [1, 2]. Проєктування нормалізованих реляційних схем із застосуванням жорстких декларативних обмежень на рівні ядра системи управління базами даних (СУБД) мінімізує структурну гетерогенність та забезпечує надійну валідацію [3, 4].

Основоположною вимогою екологічного комплаєнсу є забезпечення юридичної значущості інформації через незмінний аудиторський слід. Використання класичної реляційної алгебри з деструктивними операторами оновлення та видалення призводить до втрати ретроспективного контексту водокористування. Сучасні інженерні підходи розв'язують цю проблему шляхом впровадження бітемпорального моделювання, яке оперує двома ортогональними вимірами: часом дійсності та транзакційним часом [5]. Заміна операцій фізичного оновлення на логіку генерації нових версій записів забезпечує збереження посилальної цілісності між усіма історичними ітераціями дозволів та результатами лабораторного контролю [6].

Накопичення незмінних історичних масивів інформації спричиняє деградацію продуктивності під час виконання просторово-часових об'єднань. Аналітичні запити для перевірки дотримання квот на спеціальне водовикористання виконуються з низькою ефективністю за допомогою класичних B-Tree індексів через специфіку інтервальної алгебри [7]. Для подолання цього структурного конфлікту обґрунтовано застосування архітектурного патерна розділення відповідальності за команди та запити CQRS (Command Query Responsibility Segregation). Фізичне розділення нормалізованої бітемпоральної моделі запису (для збереження аудиту) та денормалізованої моделі читання

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

на базі матеріалізованих або віртуальних представлень забезпечує високопродуктивну агрегацію без блокування транзакцій [8, 9].

Мета дослідження. Метою дослідження є розробка архітектури та моделі даних інформаційно-аналітичної системи підприємства, яка забезпечить перехід від децентралізованого ручного обліку екологічних показників до централізованої реляційної структури з розширеними можливостями ретроспективного аналізу. Головне завдання полягає у збалансуванні архітектурного конфлікту між вимогою збереження юридичної незмінності історичного контексту водокористування та необхідністю високої швидкодії під час генерації багатовимірних аналітичних звітів.

Викладення основного матеріалу дослідження. Проєктування архітектури сучасної інформаційно-аналітичної системи екологічного моніторингу вимагає подолання високої структурної складності предметної області. Управління водними ресурсами характеризується високою складністю бізнес-правил та паралельним функціонуванням різних нормативно-правових вимог (національних гранично допустимих концентрацій забруднювальних речовин ГДК та європейської нормативної бази екологічних стандартів якості середовища EQS). Для забезпечення ізоляції бізнес-правил та управління складністю концептуальну модель системи декомпозитовано на чотири семантично ізольовані простори: довідковий, топологічний, нормативний та транзакційний.

Довідковий блок (Рис. 1) містить фундаментальні структури даних, що характеризуються низькою частотою змін, високою частотою зчитування та слугують централізованим сховищем еталонних даних для системи. Ключовим елементом є реєстр забруднювальних речовин (Pollutant), ідентифікація в якому базується на складеній системі міжнародних та національних кодів: CasNumber, EuNumber та StateCode. Реалізовано механізм часткової унікальності, який унеможливлює дублювання записів навіть за відсутності одного з кодів. До складу блоку також інтегровано довідник галузей економіки (IndustryBranch), де для підтримки ієрархічної класифікації застосовано реляційний патерн списку суміжності. Просторова гідрологічна ієрархія забезпечується довідником річкових басейнів (RiverBasin). Нормативні коефіцієнти втрат води фіксуються у класифікаторі видів робіт водокористування (WaterManagementWork).

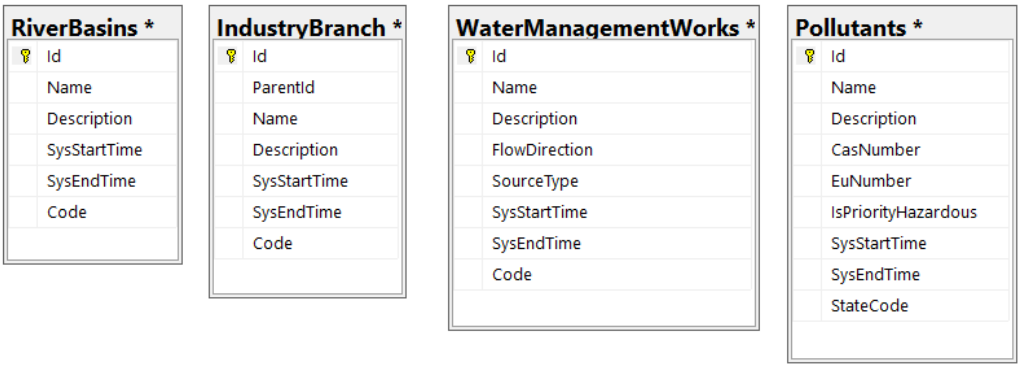


Рисунок 1 - ER-діаграма концептуальної моделі довідкового блоку

Топологічний блок (Рис. 2) моделює фізичну та адміністративну мережу екологічного контролю, розмежовуючи юридичних осіб та фізичні локації впливу на довкілля.

Підприємства-водокористувачі (MonitoringSubject) виступають базовою сутністю для всієї подальшої звітності. Цілісність цих даних забезпечується алгоритмічною перевіркою контрольної суми ідентифікатора EdrpouCode та географічною нормалізацією через код адміністративно-територіальної одиниці KatottgCode. Фізичні водні об'єкти (MonitoringObject), такі як водозабори чи випуски, зберігають географічні координати, що дозволяє системі управління базами даних виконувати просторові запити для розрахунку зон впливу. Кожен об'єкт має обов’язкову структурну прив'язку до гідрографічної мережі, забезпечуючи точну топологічну маршрутизацію забруднень.

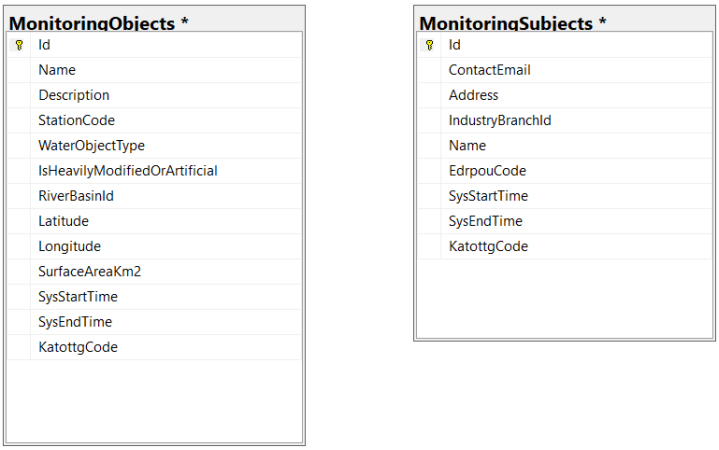


Рисунок 2 - ER-діаграма концептуальної моделі топологічного блоку

Нормативний блок (Рис. 3) визначає гранично допустимі параметри впливу на довкілля. Архітектура сутності державних екологічних стандартів якості (EnvironmentalQualityStandard) спирається на темпоральну модель даних. Підтримка часової дійсності нормативів реалізована через атрибути ValidFrom та ValidTo, що забезпечує збереження історії змін і дозволяє системі відтворювати розрахунки для минулих періодів. Сутність екологічного дозволу (EcologicalPermit) містить набір індивідуальних нормативів (лімітів забору чи скиду), які підпорядковані документу.

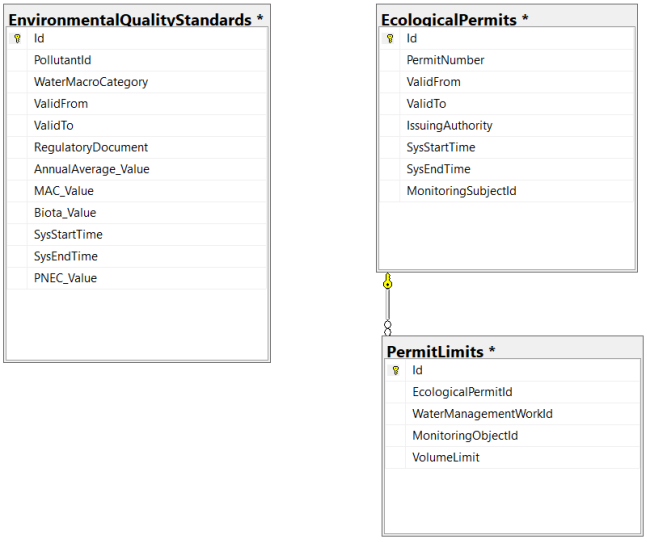


Рисунок 3 - ER-діаграма концептуальної моделі нормативного блоку

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Транзакційний блок (Рис. 4) консолідує масиви операційних вимірювань і розрахунків, блокуючи редагування записів після їх фіксації на рівні застосунку для збереження аудиторського сліду. Лабораторні протоколи (WaterQualityProtocol) агрегують метадані процедури та колекцію результатів вимірювань якості води. Звіти водоспоживання (WaterUsageReport) об'єднують дані державного статистичного спостереження, гарантуючи недоторканність поданих історичних показників.

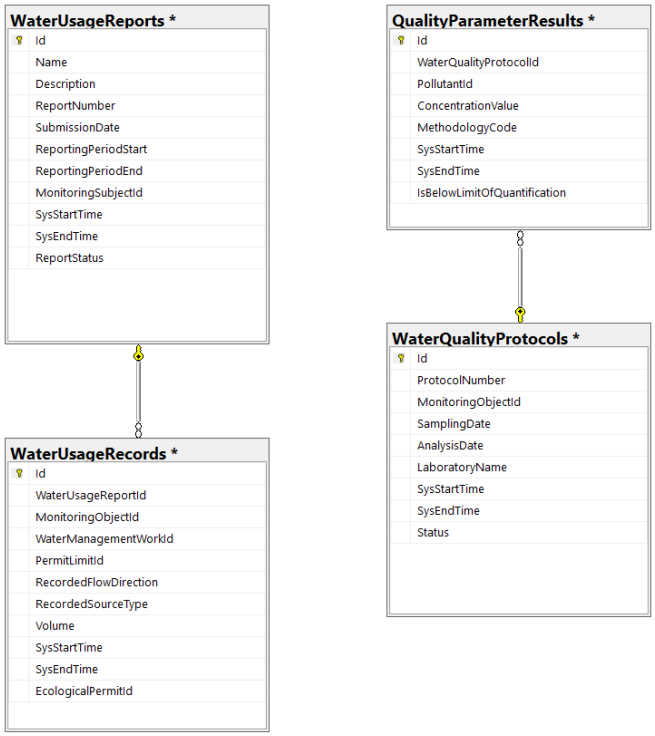


Рисунок 4 - ER-діаграма концептуальної моделі транзакційного блоку

Екологічний моніторинг оперує даними, що мають юридичну та фіскальну значущість, тому фізичне видалення записів із бази даних є неприпустимим. Використання стандартної операції DELETE призводить до безповоротної втрати інформації, що унеможливлює проведення довгострокового аудиту та руйнує історію змін, необхідну для перевірок.

Для розв'язання цієї проблеми технічна реалізація моделі даних спирається на стандарт бітемпоральних реляційних таблиць [10]. У межах спроєктованої архітектури даних кожне відношення фізично розділяється на дві структури: актуальну таблицю (поточний стан) та історичну таблицю. Будь-яка ініційована операція модифікації стану бази даних автоматично перехоплюється СУБД, яка фіксує попередню версію запису в історичній таблиці з точними мітками початку та кінця транзакції. Це дозволяє системі гарантувати повне алгоритмічне відтворення стану бази даних на будь-який момент часу в минулому.

Для довідкового простору застосовується унітемпоральна модель (лише System Time). У разі виконання команд оновлення або видалення (UPDATE або DELETE) СУБД автоматично здійснює зсув рядків таблиці. Наприклад, при видаленні тестового запису (ID 3) та оновленні статусу небезпеки для запису (ID 1), актуальна таблиця

(Таблиця 1) міститиме лише дійсні сутності, тоді як історична (Таблиця 2) збереже повний аудиторський слід знищених та змінених даних із точними транзакційними мітками закриття періоду (SysEndTime).

Таблиця 1

Актуальна таблиця

Id	CasNumber	IsPriorityHazardous	SysStartTime	SysEndTime
1	7440-43-9	TRUE	10.05.2026 14:00	31.12.9999 23:59
2	7439-97-6	TRUE	12.05.2026 9:15	31.12.9999 23:59

Таблиця 2

Історична таблиця

Id	CasNumber	IsPriorityHazardous	SysStartTime	SysEndTime
1	7440-43-9	FALSE	01.01.2026 10:00	10.05.2026 14:00
3	Test	FALSE	15.02.2026 11:30	10.05.2026 14:05

Окрім збереження транзакційної історії через механізми логічного видалення, архітектура бази даних має адекватно обробляти динаміку зміни екологічного законодавства, що реалізується шляхом застосування концепції темпоральності для нормативної бази. Екологічні нормативи, зокрема стандарти якості середовища та гранично допустимі концентрації для конкретних забруднювальних речовин, підлягають регулярному перегляду та оновленню в контексті імплементації європейських водоохоронних директив та національних нормативно-правових актів. Пряма модифікація існуючих значень екологічних лімітів у таблицях нормативів призвела б до непоправної фальсифікації результатів верифікації для історичних лабораторних протоколів, порушуючи семантичну цілісність минулих розрахунків та нівелюючи доказову базу системи.

Для гарантування консистентності ретроспективної аналітики в умовах зміни законодавства використовується механізм системного версіонування, реалізований через темпоральну модель даних. Архітектура сутності екологічних стандартів якості розширюється атрибутами визначення періоду валідності, що формують закритий часовий інтервал, протягом якого конкретне числове значення нормативу вважається юридично чинним. Коли державні органи затверджують нові ліміти, адміністративний модуль системи не здійснює деструктивного перезапису; натомість попередньому нормативу встановлюється точна дата завершення дії шляхом оновлення атрибута ValidTo, і паралельно формується новий запис із актуальними числовими показниками нормативів та відповідною датою набуття чинності в атрибуті ValidFrom.

При зміні нормативу з 1.5 мг/л на 1.0 мг/л, СУБД виконує транзакцію, що обмежує час дійсності (ValidTo) старого запису датою набрання чинності новим законом та створює новий запис (Таблиця 3). Одночасно автоматично генерується аудиторський слід в історичній таблиці (Таблиця 4).

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Таблиця 3

Бітемпоральна актуальна таблиця лімітів

ID	Pollutant_ID	Limit_Value	ValidFrom	ValidTo	SysStartTime	SysEndTime
1	42	1.5	01.01.2024	01.08.2026	28.07.2026 10:30	31.12.9999 23:59
2	42	1.0	01.08.2026	31.12.9999	28.07.2026 10:30	31.12.9999 23:59

Таблиця 4

Бітемпоральна історична таблиця лімітів

ID	Pollutant_ID	Limit_Value	ValidFrom	ValidTo	SysStartTime	SysEndTime
1	42	1.5	01.01.2024	31.12.9999	15.01.2024 9:00	28.07.2026 10:30

Ця хронологічна безперервність дозволяє системі виконувати просторово-часові запити, застосовуючи релевантне значення нормативу для оцінки факту перевищення, навіть якщо дата відбору проби припадає на період дії старих законодавчих норм. Таким чином забезпечується точність автоматизованих інспекційних висновків.

Транзакційні таблиці, які акумулюють первинні результати вимірювань, зберігають дані без можливості фізичного видалення на рівні прикладного застосунку, виконуючи роль аудиторського сліду. Якщо виникає необхідність виправити технічну помилку у затвердженому протоколі якості води, система відхиляє пряму модифікацію значень. Замість перезапису генерується новий транзакційний запис з інкрементованою версією, який за допомогою вказівника логічно пов'язується з попередньою версією протоколу. Це забезпечує збереження історії змін та унеможливлює приховану фальсифікацію ретроспективних масивів через інтерфейс системи.

Інформаційно-аналітичні системи моніторингу водних ресурсів характеризуються значною асиметрією навантажень на читання та запис. Традиційний підхід на основі базових універсальних операцій створення, читання, оновлення та видалення даних використовує єдину доменну модель як для модифікації стану системи, так і для формування багатовимірних вибірок. Використання такої універсальної моделі генерує структурний конфлікт: масиви даних, оптимізовані для безпечного транзакційного запису та збереження цілісності, виявляються неефективними під час виконання складних аналітичних агрегацій. Для розв'язання цього конфлікту у спроєктованій БД застосовано архітектурний патерн CQRS, який забезпечує незалежну оптимізацію операцій запису та читання, підвищуючи стійкість системи під навантаженням.

Обґрунтування розділення системи на два незалежних компоненти базується на необхідності розмежування життєвих циклів даних, де модель команд проєктується з орієнтацією на предметну область і відповідає за виконання модифікації стану системи. Її головним завданням є прийом вхідних даних, валідація формальних вимог та перевірка ключових бізнес-правил та обмежень доменної моделі. Модель запитів функціонує паралельно та незалежно, оперуючи структурами даних, оптимізованими для швидкої агрегації звітів. Розділення моделей запису та читання вирішує проблему асиметричного навантаження та усуває транзакційні блокування під час формування звітності. У

традиційній архітектурі виконання складних аналітичних запитів призводило б до блокування таблиць, перешкоджаючи паралельному збереженню нових даних.

Логічне розділення операцій вимагає механізму маршрутизації для взаємодії між зовнішніми інтерфейсами та ядром системи. Для цього в архітектуру інтегровано диспетчер повідомлень на базі патерна Mediator. Таке рішення забезпечує слабку зв'язність між точками входу та прикладною бізнес-логікою, централізуючи взаємодію компонентів. Використання патерна Mediator забезпечує абстрагування презентаційного шару. Робота точок входу зводиться до отримання запиту, його перетворення на повідомлення та передачі до шини маршрутизації. Це звільняє зовнішній рівень від логіки валідації та умовних розгалужень. Уся бізнес-логіка зосереджена в ізольованих обробниках, кожен з яких виконує окремий сценарій використання. Інфраструктурні аспекти (авторизація, управління транзакціями) винесено у перехоплювачі, що дозволяє обробникам працювати виключно з предметною логікою.

Хоча застосований патерн CQRS вирішив проблему асиметрії навантаження для регламентованої звітності, він не усунув фундаментальної обчислювальної складності під час первинної генерації матеріалізованих представлень або виконання нерегламентованих ретроспективних аудитів. Для виконання цих завдань необхідно здійснювати просторово-часові об'єднання за правилами інтервальної алгебри. Пошук історично достовірного нормативу в бітемпоральній моделі вимагає перевірки часових перетинів у двох вимірах: часі дійсності та транзакційному часі (за умови виконання нерівностей $ValidFrom \leq t \leq ValidTo$ та $SysStartTime \leq t \leq SysEndTime$). Оскільки класичні індекси B-Tree призначені для одновимірних діапазонів, їх застосування для складних багатовимірних перетинів інтервалів у бітемпоральній моделі є неефективним і створює значне навантаження при обробці багаторічних архівів.

Для ефективної обробки інтервальних запитів в архітектуру інтегровано механізми індексування. При пошуку перетинів у двох часових вимірах для історичних таблиць застосовано кластерні стовпчикові індекси (Clustered Columnstore Index). Під час виконання аналітичних агрегацій оптимізатор запитів використовує алгоритм відсікання сегментів на основі метаданих про мінімальні та максимальні часові межі кожного блоку. Це дозволяє ядру СУБД ігнорувати нерелевантні історичні періоди та обробляти відфільтровані масиви у векторному режимі. Такий підхід запобігає повному скануванню таблиць і забезпечує стабільну швидкодію реляційних об'єднань незалежно від накопиченого обсягу аудиторського сліду.

Використання кластерних стовпчикових індексів для матеріалізованих представлень та історичних масивів не лише забезпечує векторну обробку просторово-часових запитів, але й реалізує стиснення інформації. Стовпчиковий формат зберігання забезпечує коефіцієнт стиснення до 10 разів порівняно з традиційними структурами B-Tree, що радикально зменшує вимоги до дискового простору. Додаткова оптимізація вартості апаратних ресурсів досягається шляхом секціонування історичних таблиць. Цей механізм дозволяє ізолювати масиви даних за хронологічним критерієм і фізично переміщу-

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

вати архіви минулих років на дешевші носії, зберігаючи швидкодію системи під час обробки поточних операцій

Формування багатовимірних аналітичних звітів у системі вимагає перехресного з'єднання транзакційних масивів із бітемпоральною нормативною базою, що також створює ресурсоємне навантаження СУБД. Щоб мінімізувати це навантаження, модель читання патерна CQRS фізично реалізована через систему матеріалізованих представлень. Вони зберігають агреговані екологічні показники у вигляді денормалізованих вибірок, мінімізуючи потребу в динамічних реляційних з'єднаннях під час обробки запитів фахівців-екологів.

Оскільки затверджені документи формують незмінний аудиторський слід, алгоритм синхронізації матеріалізованих представлень у системі розділено на два незалежні процеси. Для завершених розрахункових періодів, масиви яких є незмінними, представлення генеруються одноразово із відключенням механізмів подальшої синхронізації, що звільняє ресурси СУБД. Для активних розрахункових періодів застосування синхронного оновлення представлень відхилено через ризик транзакційних блокувань на рівні моделі команд. Натомість інкрементальний перерахунок активного представлення ініціюється системою виключно асинхронно – у момент переведення нового транзакційного документа у статус затвердження. Цей алгоритм архітектурно ізолює навантаження на читання від навантаження на запис, гарантуючи сувору консистентність даних для минулих періодів та кінцеву узгодженість для поточних операцій.

Додаткове прискорення агрегації метрик у масштабах усього річкового басейну досягнуто за рахунок побудови матеріалізованих представлень на базі кластерних стовпчикових індексів. Замість порядкового читання, вони стискають дані по стовпцям і дозволяють рушію СУБД виконувати векторну обробку запитів (Таблиця 5).

Таблиця 5

Приклад структури матеріалізованого представлення

Reporting Year	**Reporting Month**	**WaterObject Name**	**Pollutant Name**	**TotalDischarged Mass_kg**	**Allowed Limit_kg**	**Exceedance_kg**	**IsViolation**
2026	7	Випуск №1 (р. Інгулець)	Залізо загальне	1250.5	1000.0	250.5	TRUE
2026	7	Випуск №1 (р. Інгулець)	Сульфати	4500.0	5000.0	0.0	FALSE

Структура повністю ізольована від нормалізованих транзакційних таблиць. Матеріалізоване представлення зберігає вже розраховані суми (TotalDischargedMass_kg), математичні дельти (Exceedance_kg) та готові булеві індикатори порушень (IsViolation). Відсутність атрибутів транзакційного часу чи часу дійсності свідчить про те, що правила інтервальної алгебри вже були застосовані на етапі генерації зліпка.

Проте асинхронний перерахунок активних представлень створює часову затримку, що спричиняє ризик розбіжностей під час одночасної генерації звітів і затвердження нових протоколів. Оскільки екологічний моніторинг оперує юридично значущими даними, для усунення ризику отримання неактуальної інформації реалізовано механізм відкладеного читання. Якщо користувач ініціює формування звіту в момент наявності необроблених транзакцій, система призупиняє видачу результату до завершення фонового оновлення матеріалізованого представлення. Для запобігання блокуванню процесу через паралельне введення даних іншими користувачами алгоритм фіксує мітку часу на момент створення запиту. Система очікує на агрегацію виключно тих протоколів, які передують цій мітці. Усі наступні транзакції ізолюються і не збільшують час поточного очікування.

Практична імплементація запропонованої архітектури спирається на реляційну СУБД з Microsoft SQL Server та інфраструктуру платформи .NET. Вибір цієї СУБД зумовлений комплексною підтримкою інструментів, необхідних для реалізації патерна CQRS. Нативна підтримка темпоральних таблиць дозволяє делегувати контроль неперервності аудиторського сліду базі даних, виключаючи необхідність реалізації механізмів версіонування на рівні прикладної бізнес-логіки мовою C#. Водночас наявність технології кластерних стовпчикових індексів із можливістю векторної обробки та підтримкою матеріалізованих структур на базі індексованих представлень формує технологічний фундамент для фізичної ізоляції моделі читання та усунення транзакційних блокувань під час багатовимірної просторово-часової агрегації.

Висновки. Запропонована концептуальна та фізична модель бази даних інформаційно-аналітичної системи спрямована на зменшення затримок в ідентифікації екологічних інцидентів, спричинену децентралізованим ручним обліком даних. Реляційна архітектура демонструє підхід до балансування конфлікту між суворими нормативними вимогами екологічного комплаєнсу та потребою швидкої генерації багатовимірної звітності.

Відмова від деструктивних операторів модифікації на користь бітемпорального моделювання гарантує формування безперервного аудиторського сліду, що забезпечує юридичну достовірність історії водокористування навіть в умовах динамічної зміни державних екологічних нормативів. Водночас ізоляція аналітичного навантаження шляхом формування денормалізованої моделі читання знижує ризик деградації продуктивності та ймовірність транзакційних блокувань під час ретроспективної перевірки квот на спеціальне водовикористання.

Спроєктована архітектура перетворює систему з пасивного сховища розрізнених документів на інструмент оперативного моніторингу, що створює передумови для мінімізації фінансових ризиків підприємства за наднормативні скиди забруднювальних речовин.

ЛІТЕРАТУРА

1. Кірейцева Г., Циганенко-Дзюбенко І., Замула І., Демчук Л. Аналіз стану та моніторинг поверхневих водних об’єктів чернігівської області. Вісник КрНУ імені Михайла Остроградського. 2024. Вип. 1 (144). С. 84–91.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

DOI: https://doi.org/10.32782/1995-0519.2024.1.11.
2. Joy-Camacho W., Thornhill I. Opportunities and limitations to environmental management system (EMS) implementation in UK SMEs. Journal of Environmental Management. 2024. Vol. 367. DOI: https://doi.org/10.1016/j.jenvman.2024.121749.
3. Васенко О., Брук В., Свиридов Ю. Екологічний моніторинг водних ресурсів. Eastern-European Journal of Enterprise Technologies. 2020. Vol. 6, no. 10 (108). P. 6–13.
DOI: https://doi.org/10.15587/1729-4061.2020.219556.
4. Власова О., Шевченко І., Козицький О. Картографічне відображення екологічного стану поверхневих водних об'єктів Правобережжя України. Водна безпека: наукові рішення в умовах криз та трансформацій: збірник тез XIV Міжнародної наук.-практ. конф. (м. Київ, 19–20 берез. 2026 р.). Київ: ІВПІМ НААН, 2026. С. 58–60.
DOI: https://doi.org/10.31073/mivg2026.
5. Chau V. T. N., Chittayasothorn S. A Bitemporal SQL Database Design Method from the Enhanced Entity-Relationship Model. 2021 7th International Conference on Engineering, Applied Sciences and Technology (ICEAST). Pattaya, Thailand, 2021. P. 85–90.
DOI: https://doi.org/10.1109/ICEAST52143.2021.9426270.
6. Гнедіна К. В. Сучасні підходи до управління водними ресурсами та механізми їх екологічного моніторингу. Екологічні науки. 2021. № 4 (37). С. 33–39.
DOI: https://doi.org/10.32846/2306-9716/2021.eco.4-37.4.
7. Henneberg J., Schuhknecht F. M., Reutter P., Brast N., Spichtinger P. Northlight: Declarative and Optimized Analysis of Atmospheric Datasets in SparkSQL. Proceedings of the 34th International Conference on Scientific and Statistical Database Management (SSDBM '22). 2022. Art. 13. DOI: https://doi.org/10.1145/3538712.3538715.
8. Iraqi, O., El Kadiri El Hassani, M., & Zouine, A. (2025). Microhooks: A novel framework to streamline the development of microservices. Computers, 14(4), 139.
DOI: https://doi.org/10.3390/computers14040139.
9. Schuhknecht F. M., Henneberg J., May N., Tatbul N. Accelerating Main-Memory Table Scans with Partial Virtual Views. Proceedings of the 19th International Workshop on Data Management on New Hardware (DaMoN '23). 2023. P. 89–93.
DOI: https://doi.org/10.1145/3592980.3595315.
10. Kulkarni, K., & Michels, J.-E. (2012). Temporal features in SQL:2011. SIGMOD Record, 41(3), 34–43. DOI: https://doi.org/10.1145/2380776.2380786.

REFERENCES

1. Kireitseva, H., Tsyhanenko-Dziubenko, I., Zamula, I., & Demchuk, L. (2024). Analiz stanu ta monitorynh poverkhnevykh vodnykh obiektiv Chernihivskoi oblasti [Analysis of the state and monitoring of surface water bodies of Chernihiv region]. Visnyk KrNU imeni Mykhaila Ostrohradskoho, 1(144), 84–91.
DOI: https://doi.org/10.32782/1995-0519.2024.1.11.
2. Joy-Camacho, W., & Thornhill, I. (2024). Opportunities and limitations to environmental management system (EMS) implementation in UK SMEs. Journal of Environmental Management, 367, 121749. DOI: https://doi.org/10.1016/j.jenvman.2024.121749.
3. Vasenko, O. G., Bruk, V. V., & Svyrydov, Y. (2020). Ekolohichnyi monitoryng vodnykh resursiv [Ecological monitoring of water resources]. Eastern-European Journal of Enterprise Technologies, 6(10(108)), 6–13. DOI: https://doi.org/10.15587/1729-4061.2020.219556.

4. Vlasova, O., Shevchenko, I., & Kozytskyi, O. (2026). Kartohrafichne vidobrazhennia ekolohichnoho stanu poverkhnevykh vodnykh obiektiv Pravoberezhzhia Ukrainy [Cartographic representation of the ecological state of surface water bodies of the Right Bank of Ukraine]. In Vodna bezpeka: naukovi rishennia v umovakh kryz ta transformatsii: zbirnyk tez XIV Mizhnarodnoi naukovo-praktychnoi konferentsii (pp. 58–60). IVPIM NAAN.
DOI: https://doi.org/10.31073/mivg2026.
5. Chau, V. T. N., & Chittayasothorn, S. (2021). A bitemporal SQL database design method from the enhanced entity-relationship model. In 2021 7th International Conference on Engineering, Applied Sciences and Technology (ICEAST) (pp. 85–90). IEEE.
DOI: https://doi.org/10.1109/ICEAST52143.2021.9426270.
6. Gnedina, K. V. (2021). Suchasni pidkhody do upravlinnia vodnymi resursami ta mekhanizmy yikh ekolohichnoho monitorynhu [Modern approaches to water resources management and mechanisms of their ecological monitoring]. Ekolohichni nauky, (4(37)), 33–39. DOI: https://doi.org/10.32846/2306-9716/2021.eco.4-37.4.
7. Henneberg, J., Schuhknecht, F. M., Reutter, P., Brast, N., & Spichtinger, P. (2022). Northlight: Declarative and optimized analysis of atmospheric datasets in SparkSQL. In SSDBM '22: Proceedings of the 34th International Conference on Scientific and Statistical Database Management (Article 13). Association for Computing Machinery.
DOI: https://doi.org/10.1145/3538712.3538715.
8. Iraqi, O., El Kadiri El Hassani, M., & Zouine, A. (2025). Microhooks: A novel framework to streamline the development of microservices. Computers, 14(4), 139.
DOI: https://doi.org/10.3390/computers14040139.
9. Schuhknecht, F. M., Henneberg, J., May, N., & Tatbul, N. (2023). Accelerating main-memory table scans with partial virtual views. In DaMoN '23: Proceedings of the 19th International Workshop on Data Management on New Hardware (pp. 89–93). Association for Computing Machinery. DOI: https://doi.org/10.1145/3592980.3595315.
10. Kulkarni, K., & Michels, J.-E. (2012). Temporal features in SQL:2011. SIGMOD Record, 41(3), 34–43. DOI: https://doi.org/10.1145/2380776.2380786.

Received 29.06.2026.
Accepted 28.08.2026.
Published 30.10.2026

Design of the architecture and data model for an information and analytical water resources monitoring system

Abstract. Environmental data management at enterprises is characterized by a high degree of decentralization, dominated by disjointed spreadsheets and manual aggregation of indicators. This approach significantly complicates the prompt detection of excessive discharges, increasing the risk of significant financial penalties for companies. When attempting automation, an architectural dilemma arises: ensuring environmental compliance requires maintaining a continuous audit trail in rigidly normalized structures, while risk management dictates the need for rapid generation of multidimensional analytical reports, the performance of which degrades under a high level of normalization.

An analysis of engineering approaches highlights the limitations of classical relational models with destructive operations for preserving historical context. Studies suggest using bitemporal modeling as a means of recording chronology. However, this encounters the problem of computational degradation during spatiotemporal joins, as classical indexes are not adapted to interval algebra. The application of the CQRS pattern is considered a way to sep-

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

arate write and read operations. The aim of the work is to develop the architecture and data model of a centralized system that balances the trade-off between the requirement for the integrity of retrospective arrays and the need for high performance of analytical queries.

To achieve this goal, the problem domain was segmented into reference, topological, regulatory, and transactional spaces. The core of the architecture is the implementation of the bitemporal table standard. The variability of environmental legislation is handled through validity period attributes, ensuring the correctness of retrospective calculations. Asymmetric loading is distributed using the CQRS pattern and the Mediator dispatcher, which physically isolated transactions from analytics. To accelerate interval computations, clustered columnstore indexes were integrated. Segment elimination algorithms and batch processing minimized array scanning and ensured data compression. Reading is performed through materialized views with an asynchronous update mechanism. To reduce the risk of generating outdated reports, a deferred reading algorithm with a recorded query timestamp was developed.

The developed model establishes the technological foundation for utilizing the environmental monitoring system as an operational control tool. Abandoning destructive changes in favor of bitemporality ensures data traceability, and workload isolation minimizes the probability of transactional locks. This establishes an infrastructural foundation for timely response to incidents and the reduction of enterprise financial risks.

Keywords: database architecture, bitemporal modeling, ecological monitoring, CQRS, materialized views, audit trail, clustered columnstore indexes, interval algebra.

Рибальченко Олена Геннадіївна – ст. викладач кафедри моделювання та програмного забезпечення, Криворізький національний університет, Кривий Ріг, Україна.
ORCID: http://orcid.org/0000-0001-8691-5401
Гриценко Андрій Миколайович – к.т.н., доцент кафедри моделювання та програмного забезпечення, Криворізький національний університет, Кривий Ріг, Україна.
ORCID: https://orcid.org/0000-0001-8665-853X
Білашенко Світлана Вікторівна – викладач кафедри моделювання та програмного забезпечення, Криворізький національний університет, Кривий Ріг, Україна.
ORCID: https://orcid.org/0000-0002-4331-7425
Сорокін Віталій Сергійович – магістрант, кафедра моделювання та програмного забезпечення, Криворізький національний університет, Кривий Ріг, Україна.
ORCID: https://orcid.org/ 0009-0004-4899-5753

Rybalchenko Olena – Senior Lecturer at the Department of Modeling and Software Engineering, Kryvyi Rih National University, Kryvyi Rih, Ukraine.
ORCID: http://orcid.org/0000-0001-8691-5401
Hrytsenko Andrii – PhD in Technical Sciences, Associate Professor at the Department of Modeling and Software Engineering, Kryvyi Rih National University, Kryvyi Rih, Ukraine.
ORCID: https://orcid.org/0000-0001-8665-853X
Bilashenko Svitlana – Lecturer, Department of Modeling and Software at the Department of Modeling and Software Engineering, Kryvyi Rih National University, Kryvyi Rih, Ukraine.
ORCID: https://orcid.org/0000-0002-4331-7425
Sorokin Vitalii, Master's student, Department of Modeling and Software at the Department of Modeling and Software Engineering, Kryvyi Rih National University, Kryvyi Rih, Ukraine.
ORCID: https://orcid.org/ 0009-0004-4899-5753

DOI 10.34185/1562-9945-5-166-2026-23
UDC 004.89:528.8

I.M. Udovyk

METHODOLOGY FOR CREATING INTELLIGENT INFORMATION TECHNOLOGIES FOR RECOGNITION AND PROCESSING OF MULTISPECTRAL AEROSPACE IMAGES

Abstract. Heterogeneous Earth-observation data and rapidly expanding machine-learning, multiagent, and multimodal models require a methodology that creates coherent intelligent information technologies rather than isolated processing pipelines. The purpose of the study is to develop and formalize a methodology for creating intelligent information technologies for recognition and processing of multispectral aerospace images through synergistic processing of heterogeneous data. The methodology combines systems analysis, set-theoretic mappings, graph models, multilevel integration, multi-criteria configuration selection, and structural-functional verification. A scene is represented as a multidimensional information object with spatial, spectral, textural, contextual, temporal, and semantic characteristics. The principal scientific result is a methodology that integrates a system of creation principles, the Ψ–Ф–Л cycle of data harmonization, representation formation, and evidence-aware decision integration, adaptive task-dependent configuration, multiagent cooperation, multimodal interpretation, and end-to-end control of quality, uncertainty, and provenance. The creation procedure includes problem formalization, admissible source and component selection, data harmonization, integrated representation, intelligent recognition, evidence reconciliation, explained decision formation, and feedback-driven reconfiguration. Structural-functional verification establishes mechanisms for completeness, consistency, modularity, extensibility, adaptability, traceability, and interpretability. The methodology provides a design basis for monitoring and decision-support systems for natural and anthropogenic objects.

Keywords: *multispectral aerospace images, intelligent information technology, methodology for creation, heterogeneous data, synergistic processing, multilevel integration, spatial-spectral-semantic representation, multiagent analysis, multimodal interpretation, decision support.*

Statement of the problem. Earth observation systems generate multispectral, hyperspectral, radar, thermal, airborne, unmanned-aerial, and very-high-resolution imagery. Their coordinated use is constrained by geometric misalignment, unequal spatial resolution, sensor-specific spectral response, variable acquisition conditions, missing channels, and incompatible taxonomies [1–4]. These factors propagate from source-data quality to feature extraction, recognition, interpretation, and decision support. Therefore, a collection of high-performance

© Udovyk I.M., 2026
Copyright for this paper by its authors. Use permitted under License CC BY 4.0.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

algorithms does not constitute an intelligent information technology unless the route from observation to an explained and traceable result is methodologically defined.

Multisensor fusion addresses integration at the data, feature, or decision level [5–7], while machine-learning and deep-learning models improve classification, segmentation, and object recognition [8–13]. Foundation models extend these capabilities through multispectral, multitemporal, and multimodal representations [14–18]. However, these advances usually optimize individual stages. Creating a complete technology additionally requires source admissibility, integrated representation, component selection, evidence reconciliation, quality and uncertainty control, provenance preservation, and feedback.

The methodological contradiction is between the expanding diversity of data, models, and application requirements and the predominantly fixed organization of existing processing pipelines. Local improvement of preprocessing, fusion, or recognition cannot guarantee validity of the complete process. The methodological problem is therefore to define invariant principles, information models, interlevel operators, admissibility rules, and feedback mechanisms that jointly determine the creation and operation of an intelligent information technology.

In this study, creation denotes the full methodological lifecycle: formulation of the problem and requirements, construction of the information space, selection and coordination of components, execution of the processing route, verification of the result, and subsequent adaptation. Controlled synthesis is its operational mechanism, which assembles a task-dependent configuration from admissible data sources, transformations, representations, recognition models, agents, multimodal components, and decision rules. The methodology must preserve physical meaning at the data level, information content at the representation level, and evidence consistency at the decision level.

Analysis of recent research and publications. Classical remote-sensing methodology organizes radiometric and geometric correction, spectral analysis, classification, and thematic mapping as consecutive stages [1, 2]. Fusion research distinguishes data-, feature-, and decision-level integration, but these levels are often treated separately and their effect on downstream reliability is not made explicit [5–7].

Random forests and support vector machines remain effective for well-defined feature spaces and limited training samples [8, 9], whereas encoder-decoder networks, atrous models, transformers, and promptable models extend automated interpretation [10–13]. Their performance depends on compatibility of the representation, physical scale, taxonomy, sensor, season, and application domain. Self-supervised and foundation models add multitemporal pretraining, vision-language alignment, and multimodal representations [14–18], but also require control of physical meaning, calibration, evidential support, and reproducibility.

Multiagent systems decompose complex analysis into specialized functions and coordinate local hypotheses [19], while multimodal models combine imagery, text, metadata, and domain knowledge [20]. Previous studies involving the author addressed multichannel preprocessing [21], data integration [22], multiagent object detection [23], multimodal land-cover interpretation [24], and post-training adaptation [25]. These studies demonstrate component feasibility but do not provide a unified methodology that simultaneously defines the scene

model, synergistic Ψ–Φ–Λ integration, adaptive configuration, and end-to-end quality, uncertainty, evidence, and provenance control.

Purpose of the study. The purpose of the study is to develop and formalize a methodology for creating intelligent information technologies for recognition and processing of multispectral aerospace images based on synergistic processing and analysis of heterogeneous data. The object is the process of transforming heterogeneous Earth observation data into recognized, interpreted, and traceable information products. The subject is the system of principles, information models, integration operators, admissibility conditions, and structural-functional relations that determine the creation of a task-dependent intelligent information technology.

Systems analysis is used to decompose the technology and establish interlevel relations; set theory and mappings formalize information spaces and integration operators; graph models represent spatial, temporal, functional, and semantic relations; multi-criteria optimization selects a technological configuration; and comparative structural-functional analysis verifies methodological properties.

The study addresses four tasks: to establish the principles for creating intelligent information technologies; to formalize the multispectral scene and its integrated representation; to develop synergistic multilevel integration of data, representations, and decisions; and to define an adaptive creation procedure together with conditions for analytical verification. The scientific novelty consists in a methodology that unites synergistic processing, spatial-spectral-semantic representation, intelligent recognition, multiagent cooperation, multimodal interpretation, and evidence-aware decision support within a single quality-, uncertainty-, and provenance-controlled model.

Presentation of the main research material. The proposed methodology is a system of interconnected conceptual, informational, operational, analytical, and verification provisions that determine how an intelligent information technology is created for a specific recognition or processing task. Its information basis is a multispectral aerospace scene considered as a multidimensional information object rather than as a set of independent channels. The corresponding information space is defined as

$$I = X \times S \times T \times C \times M, \tag{1}$$

where X is the space of spatial and geometric characteristics; S is the spectral feature space; T is the textural feature space; C is the space of contextual, relational, and temporal characteristics; and M is the semantic space containing class taxonomies, textual attributes, and domain knowledge.

For the i-th object, segment, or spatial unit, the integrated representation is formed as

$$F_i = \Phi(X_i, S_i, T_i, C_i, M_i), \tag{2}$$

where Φ denotes the operator for forming, normalizing, selecting, and integrating heterogeneous characteristics after the source data have been harmonized by Ψ. The representation F_i may include a dense feature tensor, object descriptors, class distributions, textual attributes, uncertainty estimates, and links to source metadata. A scene-level representation is defined by

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

the set of object representations, the graph G of spatial and semantic relations, and cross-cutting quality, uncertainty, and provenance information.

Synergistic processing is not interpreted as mechanical concatenation of all available channels or model outputs. It denotes coordinated interaction of complementary data and analytical mechanisms: geometric and spectral harmonization constrains feature formation; feature quality and availability constrain activation of recognition models; model and agent outputs are reconciled together with their uncertainty and provenance; and the final decision provides feedback for reconfiguration. The effect of such coordination is evaluated at the task level by informational completeness, reliability, interpretability, and resource feasibility rather than by the amount of combined data alone.

Table 1

Notation and principal components of the proposed methodology

Notation	Meaning
$D_1, \ldots, D_k$	Heterogeneous aerospace images, auxiliary data, metadata, and reference sources.
X, S, T, C, M	Spatial, spectral, textural, contextual/temporal, and semantic feature spaces.
F_i	Integrated representation of the i-th object, segment, or spatial unit.
G	Graph of spatial, temporal, functional, or semantic relations between scene objects.
Ψ	Operator for compatibility control, harmonization, and integration of primary data.
Φ	Operator for feature formation, normalization, selection, and integrated spatial-spectral-semantic representation.
Λ	Operator for reconciliation of model, agent, multimodal, and domain-rule outputs.
Q, U, μ	Quality, uncertainty, and metadata/provenance propagated throughout the processing route.

The methodology is based on principles that define invariant requirements for creating and operating the technology rather than prescribing a fixed sequence of software procedures. The principles form three interconnected groups: the information basis, intelligent analysis, and decision formation. Quality Q, uncertainty U, and provenance μ are cross-cutting dimensions that accompany all groups and all intermediate products.

The principle of multisource operation requires the technology to accept data of different physical nature and origin, including optical, radar, thermal, multispectral, hyperspectral, airborne, unmanned-aerial, and ground observations. It does not require simultaneous use of every source; an admissible complementary subset is selected according to the task, availability, compatibility, quality, information value, and access constraints.

The principle of multilevel integration establishes explicit links between the data, feature, and decision levels. A lower-level transformation is justified only when it preserves physical meaning and increases or maintains the information value of subsequent recognition and interpretation. For example, visual sharpening is not considered beneficial when it distorts spectral characteristics required for thematic classification.

The principle of spatial-spectral-semantic representation requires object and scene descriptions to combine geometry, spectra, texture, context, temporal relations, and semantic attributes. The principles of adaptability, multiagent cooperation, and multimodality determine how sources, scales, feature composition, recognition models, agent coalitions, and interpretation mechanisms are changed under variable conditions.

The principle of interpretability requires the final result to be linked to verifiable evidence, alternatives, and explicitly stated limitations. Decision support extends the output from a class, mask, or detected object to a recommended action. Trust and traceability require the recording of data sources, model versions, parameters, intermediate products, evidence, and the route by which the decision was obtained.

Synergistic processing and multilevel integration of heterogeneous data. The operational core of the methodology is represented by a coordinated composition of three operators. The Ψ operator performs compatibility control, harmonization, and integration of primary sources. The Φ operator constructs the integrated spatial-spectral-textural-contextual-semantic representation. The Λ operator reconciles the outputs of statistical and machine-learning models, specialized agents, multimodal components, and domain rules. Their coordinated composition realizes synergy across the data, representation, and decision levels. The generalized representation is

$$SI = \Lambda \circ \Phi \circ \Psi(D1, \ldots, Dk), \quad (3)$$

where $D_1, \ldots, D_k$ are the available sources and SI is the integrated scene representation associated with the resulting recognition or decision. The composition does not imply an irreversible linear sequence. Quality, uncertainty, or evidence control may initiate a return to source selection, harmonization parameters, feature composition, spatial scale, or the set of active models and agents.

At the Ψ level, the technology verifies spatial, temporal, radiometric, and spectral compatibility; transforms data to a common spatial reference; constructs validity masks; and preserves physically interpretable channels. Its output can be a harmonized multichannel data cube or a set of aligned modalities with documented relationships rather than a single fused image.

At the Φ level, spatial, spectral, textural, contextual, temporal, and semantic characteristics are formed, normalized, selected, and linked to object or graph structures. Every feature retains information about its source, spatial support, transformation method, validity, and quality. This prevents an integrated representation from becoming detached from the physical and procedural origin of its elements.

At the Λ level, class distributions, object hypotheses, agent evidence, multimodal descriptions, and domain rules are reconciled. The output contains the preferred class or state,

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

alternatives, reliability, uncertainty, supporting and contradicting evidence, an explanation, and a recommended action. A multimodal output is treated as one evidence source and cannot replace evidence-aware reconciliation.

The complete methodological model for creating an intelligent information technology is defined by the tuple

$$M = (P, I, \Psi, \Phi, A, V, \Lambda, Q, U, \mu), \tag{4}$$

where P is the system of methodological principles; I is the information space; Ψ, Φ, and Λ are the integration operators; A is the set of specialized recognition models and agents; V is the multimodal component; Q is the quality-control system; U is uncertainty; and μ contains metadata, evidence links, and provenance. The tuple specifies the minimum set of components and relations that must be instantiated in a concrete intelligent information technology.

The structural and functional relations are summarized in Fig. 1. The methodology is not a fixed linear pipeline but a controlled process for creating a task-dependent technology. The principles impose invariant requirements; the Ψ–Φ–Λ cycle implements synergistic integration at the data, representation, and decision levels; and Q, U, and μ accompany every transformation and constrain acceptance of the result.

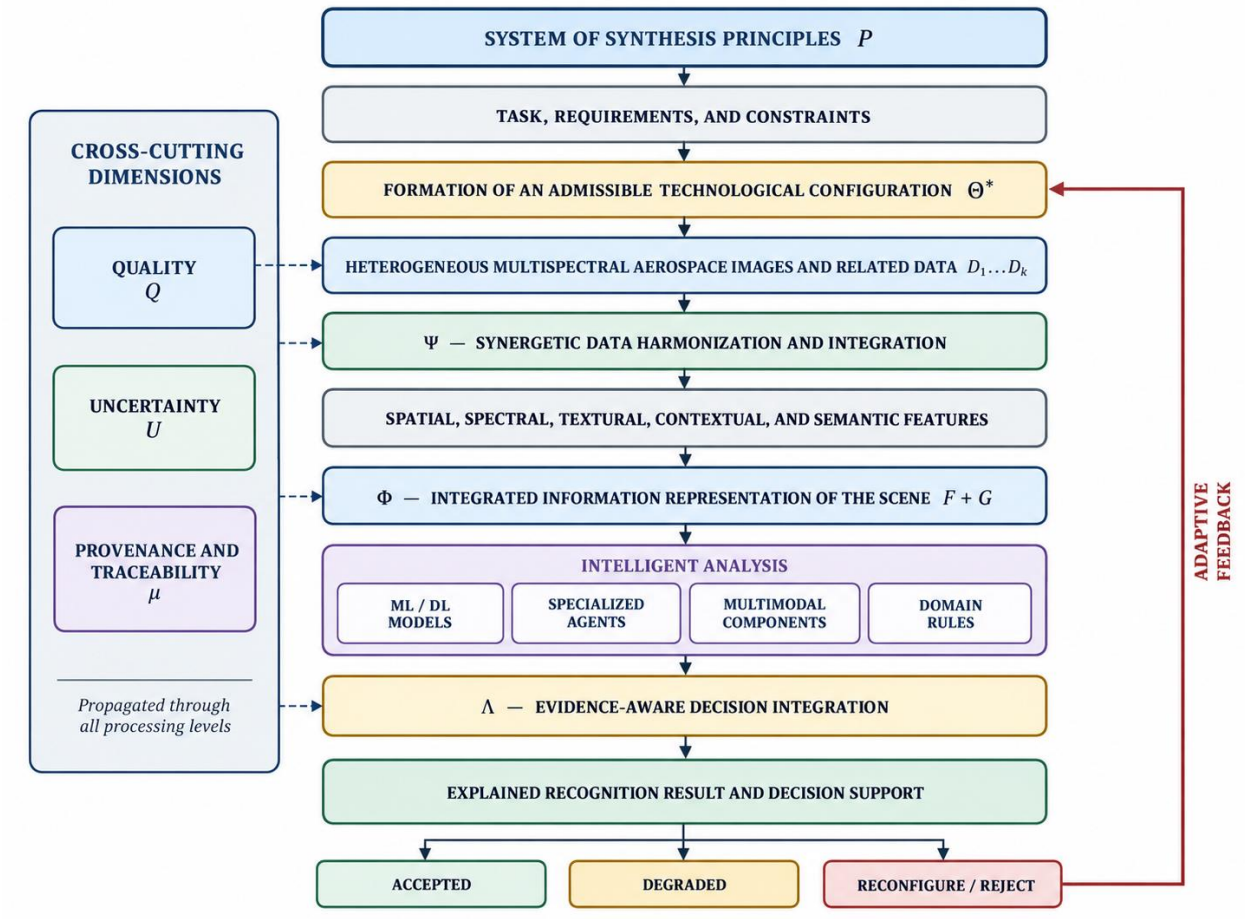


Figure 1 – Methodology for creating intelligent information technologies for recognition and processing of multispectral aerospace images

Figure 1 separates the design and operational levels. The design level establishes principles, task, requirements, constraints, and the admissible configuration Θ*. The operational level implements Ψ–Φ–Λ through ML/DL models, specialized agents, multimodal components, and domain rules. Its output is an explained recognition result with alternatives, uncertainty, supporting and contradicting evidence, limitations, and a recommendation. Accepted results satisfy critical quality, evidence, and traceability conditions; degraded results require explicit limitations or expert verification; and reconfigure/reject activates feedback to data, scale, parameters, features, models, agents, rules, or thresholds. Thus, adaptability is embedded in the creation methodology.

Procedure for creating a task-oriented intelligent information technology. For a specific task of recognition or processing of multispectral aerospace images, creation is operationalized as controlled synthesis and selection of a configuration Θ from an admissible set Ω. A configuration specifies selected data sources, preprocessing operators, feature composition, recognition models, agents, multimodal components, domain rules, resource profile, acceptance thresholds, and trust constraints. The selection criterion is

$$\Theta opt = \arg\max J(\Theta), \Theta \in \Omega, \quad (5)$$

where J(Θ) is a normalized multi-criteria objective combining expected decision reliability *Qdec*, information completeness *I*, explainability *E*, adaptability and transferability *A*, uncertainty *U*, and resource cost *Cres* with application-dependent weights. The expression is a methodological objective rather than a universal numerical metric. Each implementation must define normalized components, critical constraints, and admissible trade-offs for its subject domain.

The procedure for creating a task-oriented intelligent information technology consists of the following steps:

1. define the recognition or processing task, required information products, admissible risks, and acceptance criteria;

2. build registries of available data, transformations, features, models, agents, multimodal components, rules, and computational resources;

3. exclude incompatible components using spatial, spectral, taxonomic, quality, resource, provenance, and access constraints;

4. generate the admissible configuration set Ω and select Θopt according to the multi-criteria objective;

5. construct and execute the Ψ → Φ → analytical components → Λ route while propagating Q, U, and μ;

6. verify recognition quality, evidence completeness, and traceability and assign an accepted, degraded, or reconfigure/reject status;

7. update data requirements, parameters, components, or the processing route using feedback from quality control and expert validation.

This status logic prevents a high confidence score from compensating for low source validity, uncalibrated uncertainty, or incomplete evidence.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

The methodology is verified structurally and functionally because its principal result is a design framework rather than performance in one application. Verification checks whether the model contains mechanisms required for a coherent, reproducible, adaptive, and explainable information technology. Table 2 compares it with fixed or model-centred approaches.

Table 2

Comparison of fixed approaches and the proposed methodology

Criterion	**Typical fixed or model-centred approach**	**Proposed methodology**
Input organization	One sensor or a predefined data combination.	Task-dependent multisource configuration assembled from registered components.
Integration	One level is usually dominant: pixels, features, or decisions.	Synergistic Ψ–Φ–Λ integration at data, representation, and decision levels.
Scene representation	Pixels, a single feature tensor, or isolated objects.	Object and scene representations linked by graph G and provenance μ.
Analytical components	One model or a fixed ensemble.	Recognition models, specialized agents, multimodal components, and domain rules.
Adaptation	Manual reconfiguration or model retraining.	Configuration selection, status-driven reprocessing, and feedback.
Quality control	Primarily final task accuracy.	End-to-end validity, quality Q, uncertainty U, evidence, and provenance.
Interpretation	Class label, score, or segmentation mask.	Decision, alternatives, evidence, limitations, explanation, and recommendation.
Reproducibility	Parameters may be stored separately from outputs.	Sources, versions, parameters, route, and intermediate provenance are retained.

Structural completeness follows from the explicit presence of all levels from primary data and task requirements to recognition, interpretation, and decision support. Functional consistency is achieved through the defined inputs and outputs of Ψ, Φ, and Λ. Modularity follows from the possibility of replacing an algorithmic implementation without changing the methodological role of the component. Extensibility is provided by registries and admissibility rules that allow new sensors, features, recognition models, agents, multimodal services, and domain rules to be added.

Adaptability is supported by task-dependent configuration selection, status-driven reprocessing, and feedback. Traceability is ensured because quality, uncertainty, evidence, metadata, and provenance remain associated with each intermediate and final product. Interpretability is not reduced to a generated textual description; it is defined as the correspondence between a conclusion, supporting and contradicting evidence, alternatives, limitations, and the processing route.

Table 3

Analytical verification of the methodology

Property	**Verification condition**	**Methodological mechanism**
Structural completeness	All required stages from data acquisition to decision support are represented.	The methodological model includes principles, information space, Ψ–Φ–Λ operators, analytical components, Q, U, and μ.
Functional consistency	Outputs of one level are valid inputs to the next and retain their meaning.	Formal roles of Ψ, Φ, and Λ and documented interlevel contracts.
Modularity	An implementation can be replaced without changing the role of the component.	Registries, admissibility rules, and operator-level interfaces.
Extensibility	New data sources, features, models, or agents can be added.	Open component sets A and V and task-dependent configuration Θ created from registered components.
Adaptability	The route changes when data quality, resources, or task constraints change.	Multi-criteria selection, feedback, and accepted/degraded/reconfigure statuses.
Traceability	Every decision can be linked to sources, transformations, and versions.	Provenance μ and end-to-end recording of the processing route.
Interpretability	The decision is supported by evidence, alternatives, and limitations.	Evidence-aware Λ reconciliation and structured decision output.

Component-level feasibility is supported by the author's previous studies: multichannel preprocessing [21], data harmonization and fusion [22], multiagent object detection [23], multimodal land-cover interpretation [24], and post-training neural adaptation [25]. These works confirm the implementability of the principal methodological levels, but they do not substitute for unified evaluation of a complete task-dependent configuration.

The methodology has limitations. The configuration criterion and its weights must be instantiated for a particular domain; reliability depends on complete model registries, quality

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

metadata, provenance records, and domain-specific uncertainty calibration. Synergistic interaction does not guarantee automatic accuracy improvement. The contribution is to define the conditions, information relations, and feedback mechanisms under which heterogeneous data, models, agents, and evidence are combined in a controlled, reproducible, and explainable manner.

Conclusions. A methodology for creating intelligent information technologies for recognition and processing of multispectral aerospace images has been developed. Its scientific novelty lies in replacing a fixed pipeline with a controlled, task-dependent creation process based on synergistic processing of heterogeneous data. A system of principles and a multidimensional scene model have been formulated. They combine multisource operation, multilevel integration, spatial-spectral-semantic representation, adaptability, multiagent cooperation, multimodality, interpretability, decision support, quality, uncertainty, trust, and traceability. The Ψ–Φ–Λ model coordinates harmonization of primary data, formation of integrated spatial-spectral-textural-contextual-semantic representations, and evidence-aware reconciliation of recognition models, agents, multimodal components, and domain rules. Creation of a task-oriented technology has been formalized as multi-criteria selection from an admissible configuration set, followed by quality-aware execution, evidence verification, status assignment, and feedback-driven reconfiguration. Structural-functional verification established mechanisms for completeness, consistency, modularity, extensibility, adaptability, traceability, and interpretability, and clarified the distinction between methodological validity and performance in a particular application. Previous component-level studies confirm implementability of the principal levels. Further research should provide unified experimental evaluation of complete configurations, domain-dependent uncertainty calibration, and quantitative assessment of recognition reliability, explanation quality, and decision-support effectiveness.

Acknowledgments. This research was funded by the Ministry of Education and Science of Ukraine from the state budget under grant No. 0126U000995, ‘Intelligent technologies for the analysis of spatiotemporal changes in aerospace imagery to support decision-making under conditions of armed aggression’.

Declaration on the use of generative artificial intelligence tools. OpenAI tools were used for language editing, and final preparation of the figure. The author critically reviewed and revised all content and assumes full responsibility for the scientific statements, references, interpretations, and conclusions.

LITERATURE

1. Richards J. A. Remote Sensing Digital Image Analysis: An Introduction. 5th ed. Berlin; Heidelberg: Springer, 2013. 494 p. DOI: https://doi.org/10.1007/978-3-642-30062-2.

2. Schowengerdt R. A. Remote Sensing: Models and Methods for Image Processing. 3rd ed. Burlington: Academic Press, 2007. 560 p.

3. Zhu X. X., Tuia D., Mou L., Xia G.-S., Zhang L., Xu F., Fraundorfer F. Deep learning in remote sensing: a comprehensive review and list of resources. IEEE Geoscience and Remote Sensing Magazine. 2017. Vol. 5, no. 4. P. 8-36.
DOI: https://doi.org/10.1109/MGRS.2017.2762307.

4. Ma L., Liu Y., Zhang X., Ye Y., Yin G., Johnson B. A. Deep learning in remote sensing applications: a meta-analysis and review. ISPRS Journal of Photogrammetry and Remote Sensing. 2019. Vol. 152. P. 166-177. DOI: https://doi.org/10.1016/j.isprsjprs.2019.04.015.
5. Pohl C., Van Genderen J. L. Multisensor image fusion in remote sensing: concepts, methods and applications. International Journal of Remote Sensing. 1998. Vol. 19, no. 5. P. 823-854. DOI: https://doi.org/10.1080/014311698215748.
6. Li S., Kang X., Fang L., Hu J., Yin H. Pixel-level image fusion: a survey of the state of the art. Information Fusion. 2017. Vol. 33. P. 100-112.
DOI: https://doi.org/10.1016/j.inffus.2016.05.004.
7. Luo X., Tong X., Pan H., Shen H., Du P. A critical review on multi-sensor and multi-platform remote sensing data fusion approaches: current status and prospects. International Journal of Remote Sensing. 2025. Vol. 46, no. 3. P. 1327-1402.
DOI: https://doi.org/10.1080/01431161.2024.2429784.
8. Belgiu M., Drăguţ L. Random forest in remote sensing: a review of applications and future directions. ISPRS Journal of Photogrammetry and Remote Sensing. 2016. Vol. 114. P. 24-31. DOI: https://doi.org/10.1016/j.isprsjprs.2016.01.011.
9. Mountrakis G., Im J., Ogole C. Support vector machines in remote sensing: a review. ISPRS Journal of Photogrammetry and Remote Sensing. 2011. Vol. 66, no. 3. P. 247-259. DOI: https://doi.org/10.1016/j.isprsjprs.2010.11.001.
10. Ronneberger O., Fischer P., Brox T. U-Net: convolutional networks for biomedical image segmentation. Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015. Cham: Springer, 2015. P. 234-241.
DOI: https://doi.org/10.1007/978-3-319-24574-4_28.
11. Chen L.-C., Zhu Y., Papandreou G., Schroff F., Adam H. Encoder-decoder with atrous separable convolution for semantic image segmentation. Computer Vision - ECCV 2018. Cham: Springer, 2018. P. 833-851. DOI: https://doi.org/10.1007/978-3-030-01234-2_49.
12. Xie E., Wang W., Yu Z., Anandkumar A., Alvarez J. M., Luo P. SegFormer: simple and efficient design for semantic segmentation with transformers. Advances in Neural Information Processing Systems. 2021. Vol. 34. P. 12077-12090.
13. Kirillov A., Mintun E., Ravi N. et al. Segment Anything. Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023. P. 4015-4026.
DOI: https://doi.org/10.1109/ICCV51070.2023.00371.
14. Cong Y., Khanna S., Meng C., Liu P., Rozi E., He Y., Burke M., Lobell D. B., Ermon S. SatMAE: pre-training transformers for temporal and multi-spectral satellite imagery. Advances in Neural Information Processing Systems. 2022. Vol. 35. P. 197-211.
15. Liu F., Chen D., Guan Z., Zhou X., Zhu J., Ye Q., Fu L., Zhou J. RemoteCLIP: a vision-language foundation model for remote sensing. IEEE Transactions on Geoscience and Remote Sensing. 2024. Vol. 62. P. 1-16. DOI: https://doi.org/10.1109/TGRS.2024.3390838.
16. Guo X., Lao J., Dang B. et al. SkySense: a multi-modal remote sensing foundation model towards universal interpretation for Earth observation imagery. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024. P. 27672-27683.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

17. Huo C., Chen K., Zhang S. et al. When remote sensing meets foundation model: a survey and beyond. Remote Sensing. 2025. Vol. 17, no. 2. Article 179.
DOI: https://doi.org/10.3390/rs17020179.
18. Xiao A., Xuan W., Wang J., Huang J., Tao D., Lu S., Yokoya N. Foundation models for remote sensing and Earth observation: a survey. IEEE Geoscience and Remote Sensing Magazine. 2025. Vol. 13, no. 4. P. 297-324. DOI: https://doi.org/10.1109/MGRS.2025.3576766.
19. Wooldridge M. An Introduction to MultiAgent Systems. 2nd ed. Chichester: John Wiley & Sons, 2009. 461 p.
20. Li C., Gan Z., Yang Z., Yang J., Li L., Wang L., Gao J. Multimodal foundation models: from specialists to general-purpose assistants. Foundations and Trends in Computer Graphics and Vision. 2024. Vol. 16, no. 1-2. P. 1-214. DOI: https://doi.org/10.1561/0600000110.
21. Олевський В. І., Удовик І. М., Грищак Д. Д., Олевська Ю. Б., Журавльова Ю. С. Методологічні основи інформаційних технологій аналізу багатоканальних аерокосмічних зображень. Information Technology: Computer Science, Software Engineering and Cyber Security. 2024. № 4. С. 159-168. DOI: https://doi.org/10.32782/IT/2024-4-19.
22. Гнатушенко В. В., Грищак Д. Д., Олевська Ю. Б., Олевський В. І., Удовик І. М. Методологія об’єднання даних багатоканальних зображень. Системні технології. 2025. Вип. 6 (161). С. 158-171. DOI: https://doi.org/10.34185/1562-9945-5-161-2025-16.
23. Литвинов Є. В., Гнатушенко В. В., Удовик І. М. Виявлення нелегальних сміттєзвалищ на супутникових зображеннях із використанням мультиагентного підходу. Herald of Khmelnytskyi National University. Technical Sciences. 2026. № 361 (1). С. 523-530.
DOI: https://doi.org/10.31891/2307-5732-2026-361-73.
24. Гончаров О. Г., Удовик І. М., Гнатушенко В. В. Конструювання запитів для класифікації земного покриву без навчальних прикладів за допомогою мультимодальних мовних моделей на знімках Sentinel-2. Системні технології. 2026. Вип. 4 (165). С. 130-147. DOI: https://doi.org/10.34185/1562-9945-4-165-2026-13.
25. Udovyk I., Khabarlak K. Improved small agricultural plant segmentation using post-training adaptive neural network. Computer Systems and Information Technologies. 2026. No. 1. P. 118-124. DOI: https://doi.org/10.31891/csit-2026-1-11.

REFERENCES

1. Richards, J. A. (2013). Remote sensing digital image analysis: An introduction (5th ed.). Springer. https://doi.org/10.1007/978-3-642-30062-2
2. Schowengerdt, R. A. (2007). Remote sensing: Models and methods for image processing (3rd ed.). Academic Press.
3. Zhu, X. X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., & Fraundorfer, F. (2017). Deep learning in remote sensing: A comprehensive review and list of resources. IEEE Geoscience and Remote Sensing Magazine, 5(4), 8-36.
https://doi.org/10.1109/MGRS.2017.2762307
4. Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., & Johnson, B. A. (2019). Deep learning in remote sensing applications: A meta-analysis and review. ISPRS Journal of Photogrammetry and Remote Sensing, 152, 166-177. https://doi.org/10.1016/j.isprsjprs.2019.04.015

5. Pohl, C., & Van Genderen, J. L. (1998). Multisensor image fusion in remote sensing: Concepts, methods and applications. International Journal of Remote Sensing, 19(5), 823-854. https://doi.org/10.1080/014311698215748
6. Li, S., Kang, X., Fang, L., Hu, J., & Yin, H. (2017). Pixel-level image fusion: A survey of the state of the art. Information Fusion, 33, 100-112.
https://doi.org/10.1016/j.inffus.2016.05.004
7. Luo, X., Tong, X., Pan, H., Shen, H., & Du, P. (2025). A critical review on multi-sensor and multi-platform remote sensing data fusion approaches: Current status and prospects. International Journal of Remote Sensing, 46(3), 1327-1402.
https://doi.org/10.1080/01431161.2024.2429784
8. Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. ISPRS Journal of Photogrammetry and Remote Sensing, 114, 24-31. https://doi.org/10.1016/j.isprsjprs.2016.01.011
9. Mountrakis, G., Im, J., & Ogole, C. (2011). Support vector machines in remote sensing: A review. ISPRS Journal of Photogrammetry and Remote Sensing, 66(3), 247-259.
https://doi.org/10.1016/j.isprsjprs.2010.11.001
10. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015 (pp. 234-241). Springer. https://doi.org/10.1007/978-3-319-24574-4_28
11. Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). Encoder-decoder with atrous separable convolution for semantic image segmentation. In Computer Vision - ECCV 2018 (pp. 833-851). Springer. https://doi.org/10.1007/978-3-030-01234-2_49
12. Xie, E., Wang, W., Yu, Z., Anandkumar, A., Alvarez, J. M., & Luo, P. (2021). SegFormer: Simple and efficient design for semantic segmentation with transformers. Advances in Neural Information Processing Systems, 34, 12077-12090.
13. Kirillov, A., Mintun, E., Ravi, N., et al. (2023). Segment Anything. Proceedings of the IEEE/CVF International Conference on Computer Vision, 4015-4026.
https://doi.org/10.1109/ICCV51070.2023.00371
14. Cong, Y., Khanna, S., Meng, C., Liu, P., Rozi, E., He, Y., Burke, M., Lobell, D. B., & Ermon, S. (2022). SatMAE: Pre-training transformers for temporal and multi-spectral satellite imagery. Advances in Neural Information Processing Systems, 35, 197-211.
15. Liu, F., Chen, D., Guan, Z., Zhou, X., Zhu, J., Ye, Q., Fu, L., & Zhou, J. (2024). RemoteCLIP: A vision-language foundation model for remote sensing. IEEE Transactions on Geoscience and Remote Sensing, 62, 1-16. https://doi.org/10.1109/TGRS.2024.3390838
16. Guo, X., Lao, J., Dang, B., et al. (2024). SkySense: A multi-modal remote sensing foundation model towards universal interpretation for Earth observation imagery. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 27672-27683.
17. Huo, C., Chen, K., Zhang, S., et al. (2025). When remote sensing meets foundation model: A survey and beyond. Remote Sensing, 17(2), Article 179.
https://doi.org/10.3390/rs17020179

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

18. Xiao, A., Xuan, W., Wang, J., Huang, J., Tao, D., Lu, S., & Yokoya, N. (2025). Foundation models for remote sensing and Earth observation: A survey. IEEE Geoscience and Remote Sensing Magazine, 13(4), 297-324. https://doi.org/10.1109/MGRS.2025.3576766
19. Wooldridge, M. (2009). An introduction to multiagent systems (2nd ed.). John Wiley & Sons.
20. Li, C., Gan, Z., Yang, Z., Yang, J., Li, L., Wang, L., & Gao, J. (2024). Multimodal foundation models: From specialists to general-purpose assistants. Foundations and Trends in Computer Graphics and Vision, 16(1-2), 1-214. https://doi.org/10.1561/0600000110
21. Olevskyi, V., Udovyk, I., Hryshchak, D., Olevska, Y., & Zhuravlova, Y. (2024). Metodolohichni osnovy informatsiinykh tekhnolohii analizu bahatokanalnykh aerokosmichnykh zobrazhen [Methodological foundations of information technologies for the analysis of multichannel aerospace images]. Information Technology: Computer Science, Software Engineering and Cyber Security, (4), 159-168. https://doi.org/10.32782/IT/2024-4-19 [in Ukrainian].
22. Hnatushenko, V., Hryshchak, D., Olevska, Y., Olevskyi, V., & Udovyk, I. (2025). Metodolohiia obiednannia danykh bahatokanalnykh zobrazhen [Methodology for multichannel image-data fusion]. System Technologies, 6(161), 158-171. https://doi.org/10.34185/1562-9945-5-161-2025-16 [in Ukrainian].
23. Lytvynov, Y., Hnatushenko, V., & Udovyk, I. (2026). Vyiavlennia nelehalnykh smittiezvalyshch na suputnykovykh zobrazhenniakh iz vykorystanniam multyahentnoho pidkhodu [Detection of illegal dumpsites in satellite images using a multiagent approach]. Herald of Khmelnytskyi National University. Technical Sciences, 361(1), 523-530. https://doi.org/10.31891/2307-5732-2026-361-73 [in Ukrainian].
24. Honcharov, O. H., Udovyk, I. M., & Hnatushenko, V. V. (2026). Konstruiuvannia zapytiv dlia klasyfikatsii zemnoho pokryvu bez navchalnykh prykladiv za dopomohoiu multymodalnykh movnykh modelei na znimkakh Sentinel-2 [Prompt engineering for zero-shot land-cover classification using multimodal language models on Sentinel-2 imagery]. System Technologies, 4(165), 130-147. https://doi.org/10.34185/1562-9945-4-165-2026-13 [in Ukrainian].
25. Udovyk, I., & Khabarlak, K. (2026). Improved small agricultural plant segmentation using post-training adaptive neural network. Computer Systems and Information Technologies, (1), 118-124. https://doi.org/10.31891/csit-2026-1-11

Received 29.06.2026.
Accepted 31.08.2026.
Published 30.10.2026

Методологія створення інтелектуальних інформаційних технологій розпізнавання та обробки багатоспектральних аерокосмічних зображень

Аналіз останніх досліджень і публікацій. Сучасні системи дистанційного зондування Землі формують різнорідні багатоспектральні, гіперспектральні, радіолокаційні, теплові й безпілотні дані. Методи їх корекції, інтеграції, розпізнавання, мультиагентного та мультимодального аналізу переважно розвиваються окремо. Невирішеною є проблема єдиної методології, що визначає вибір джерел і компонентів, формування представлень, узгодження рішень, контроль якості й невизначеності, збереження походження та адаптацію маршруту обробки.

Мета дослідження. Розробити й формалізувати методологію створення інтелектуальних інформаційних технологій розпізнавання та обробки багатоспектральних аерокосмічних зображень на основі синергетичної обробки різнорідних даних, багаторівневої інтеграції й адаптивного формування предметно залежної конфігурації.

Виклад основного матеріалу дослідження. Багатоспектральну сцену подано як багатовимірний інформаційний об'єкт, що поєднує просторові, спектральні, текстурні, контекстні, часові й семантичні характеристики. Запропонована методологія інтегрує систему принципів створення, формальну інформаційну модель, цикл Ψ–Ф–Л узгодження даних, формування представлення та доказового узгодження рішень, адаптивний вибір конфігурації, інтелектуальне розпізнавання, мультиагентну кооперацію, мультимодальну інтерпретацію й наскрізний контроль якості Q, невизначеності U та походження μ. Процедура створення охоплює постановку задачі, формування реєстрів компонентів, перевірку сумісності, вибір допустимої конфігурації, виконання маршруту, верифікацію результату й реконфігурацію за зворотним зв'язком.

Висновки. Розроблена методологія, на відміну від фіксованих конвеєрних підходів, визначає керований процес створення предметно залежної технології на основі синергетичної обробки, інтегрованого просторово-спектрально-семантичного представлення, кооперації моделей і агентів, доказового узгодження рішень та адаптивного зворотного зв'язку. Структурно-функціональна верифікація встановила механізми повноти, узгодженості, модульності, розширюваності, адаптивності, простежуваності та інтерпретованості.

Ключові слова: багатоспектральні аерокосмічні зображення, інтелектуальна інформаційна технологія, методологія створення, різнорідні дані, синергетична обробка, багаторівнева інтеграція, просторово-спектрально-семантичне представлення, мультиагентний аналіз, мультимодальна інтерпретація, підтримка прийняття рішень.

Udovyk Iryna Mykhailivna - Candidate of Technical Sciences, Associate Professor, Dean of the Faculty of Information Technologies, Dnipro University of Technology, Dnipro, Ukraine.
ORCID: https://orcid.org/0000-0002-5190-841X

Удовик Ірина Михайлівна - кандидат технічних наук, доцент, деканка факультету інформаційних технологій Національного технічного університету «Дніпровська політехніка», м. Дніпро, Україна.
ORCID: https://orcid.org/0000-0002-5190-841X

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

DOI 10.34185/1562-9945-5-166-2026-24
УДК 004.056.5:004.738

А.А. Кравчук, М.В. Онай

ПРОЄКТУВАННЯ АРХІТЕКТУРИ ПРОГРАМНОЇ СИСТЕМИ ВИЯВЛЕННЯ ТА БЛОКУВАННЯ HTTP LR-DDOS АТАК НА ПЛАТФОРМІ KUBERNETES

Анотація. У статті запропоновано та реалізовано розподілену архітектуру захисту мікросервісів від низькошвидкісних DDoS-атак (LR-DDoS) прикладного рівня у середовищі Kubernetes, без модифікації коду цільових застосунків та без відчутних затримок легітимного трафіку. Архітектурне рішення базується на розділенні площини маршрутизації (Data Plane) та аналізу (Control Plane): атрибути прикладного рівня збираються розширенням EnvoyFilter після дешифрування TLS-трафіку на вхідному шлюзі Istio, а мережеві атрибути – засобами eBPF/XDP, після чого асинхронно передаються до брокера Apache Kafka. Розроблені мовою Go модуль класифікації з горизонтальним масштабуванням та контролер на базі шаблону Kubernetes Operator зберігають джерела DDoS-атак у Redis із механізмом TTL для видалення застарілих даних та оновлюють правила блокування в ресурсі AuthorizationPolicy. Це забезпечує три сценарії реагування на атаку: повну класифікацію невідомих сесій, прискорену перевірку за збереженими сигнатурами та блокування запитів на шлюзі зі складністю O(1). Експериментальні дослідження підтвердили високу продуктивність системи: затримка від EnvoyFilter становить лише 2,933 мс, а час реакції – 3,15 с для нових джерел, 0,8 с для відомих сигнатур та 5,674 мс для блокування на рівні шлюзу. Модуль класифікації обробляє 4000 сесій за 32,5 мс із використанням 3,09 МБ оперативної пам'яті. Порівняно з брандмауером вебзастосунків Coraza WAF система демонструє у 12,3 раза меншу затримку (3,667 мс проти 45 мс) та значно вищу пропускну здатність (1973 проти 473 запитів/с).

Ключові слова: DDoS, низькошвидкісні DDoS-атаки, LR-DDoS, програмна система, інженерія програмного забезпечення, мікросервісна архітектура, Kubernetes, Istio, захист інформації, криптографія.

Постановка проблеми. Розподілені атаки на відмову в обслуговуванні (distributed denial-of-service, DDoS) залишаються однією з найпоширеніших причин порушення безперебійної роботи вебсервісів, причому кількість таких інцидентів продовжує стрімко зростати [1]. Особливу небезпеку в цьому контексті становлять низькошвидкісні атаки прикладного рівня (low-rate DDoS, LR-DDoS), спрямовані на протоколи HTTP/HTTPS. На відміну від об'ємних (volumetric) DDoS-атак, які перевантажують канали зв'язку безперервним потоком мережевих пакетів даних із високою частотою [2], LR-DDoS атаки (наприклад, Slowloris [3], RUDY [4], Slow Read [5]) застосовують принципово інший підхід: вони штучно утримують мережеві сесії відкритими за допо-

© Кравчук А.А., Онай М.В., 2026
Copyright for this paper by its authors. Use permitted under License CC BY 4.0.
ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

могою навмисного затягування процесу передачі або зчитування фрагментів HTTP-запитів, що призводить до вичерпання пулу TCP-з'єднань, переповнення черг очікування, блокування потоків операційної системи та надмірного споживання оперативної пам'яті цільового сервера [6; 7]. Через малу інтенсивність генерації пакетів трафік LR-DDoS атак за своїми параметрами є подібним до активності легітимних користувачів із нестабільним мережевим з'єднанням, що значно ускладнює їх ідентифікацію за допомогою статистичних методів та звичайних механізмів контролю трафіку (зокрема, обмеження інтенсивності запитів), оскільки швидкість надходження зловмисних запитів зазвичай не перевищує встановлених порогових значень [8].

Сучасні програмні системи дедалі частіше проєктуються на основі мікросервісної архітектури [9] та працюють у розподілених хмарних середовищах із застосуванням технологій контейнеризації, де загальноприйнятим стандартом автоматизації та оркестрації є платформа Kubernetes [10]. Розгортання програмних компонентів у межах кластера Kubernetes передбачає об'єднання обчислювальних вузлів у спільний пул ресурсів, де керування контейнерами, балансування навантаження та маршрутизація трафіку між ними здійснюються динамічно [10]. Подібний підхід суттєво підвищує гнучкість та горизонтальну масштабованість систем, проте водночас має ризики для їхньої загальної відмовостійкості. Зокрема, через взаємозв'язки між елементами системи, тимчасове обмеження доступності або вичерпання обчислювальних ресурсів одного критичного компонента (наприклад, централізованого сервісу автентифікації) внаслідок DDoS-атаки здатне спровокувати лавиноподібну деградацію продуктивності всього ланцюга обробки запитів. Окрім того, процес побудови систем захисту від атак ускладнюється зазвичай притаманною мікросервісним рішенням гетерогенністю (що полягає у використанні різних технологій та інструментарію програмування), де окремі компоненти створюються різними командами розробників із застосуванням різних мов програмування, архітектурних шаблонів та вебфреймворків. За таких умов реалізація механізмів протидії LR-DDoS атакам всередині прикладного коду кожного сервісу є неефективною, оскільки вона призводить до розосередження алгоритмів захисту, надмірного дублювання вихідного коду та суттєвого ускладнення централізованого керування політиками фільтрації мережевого трафіку.

З огляду на вищезазначене, актуальною проблемою є проєктування уніфікованої архітектури програмної системи, яка забезпечувала б виявлення та блокування HTTP LR-DDoS атак у кластері Kubernetes без внесення змін до вихідного коду окремих мікросервісів, гарантуючи при цьому збереження їхньої автономності, масштабованості та відмовостійкості. Складність вирішення цієї проблеми полягає в необхідності одночасного здійснення комплексного аналізу вхідного інтернет-трафіку та забезпечення високої швидкості обробки HTTP-запитів. З одного боку, забезпечення високої точності виявлення LR-DDoS атак прикладного рівня зазвичай вимагає використання даних як мережевого, так і прикладного рівнів мережевої моделі OSI [11]. Враховуючи широке використання протоколу HTTPS у сучасних системах, необхідною передумовою є отримання доступу до вмісту HTTP-запитів у дешифрованому вигляді до їхньої передачі на

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

оброблення в цільовий мікросервіс або застосунок. З іншого боку, інтеграція методів виявлення DDoS-атак безпосередньо у вихідний код мікросервісів є неефективною, оскільки це призводить до суттєвого зростання затримок обробки запитів та спричиняє додаткове навантаження на обчислювальні ресурси, що загрожує зниженням загальної продуктивності всієї системи. Відтак, постає задача виокремлення механізмів захоплення трафіку, його аналізу для виявлення DDoS-атак та подальшого блокування, що передбачає їхню реалізацію у вигляді уніфікованого інфраструктурного рішення на рівні кластера. Практична цінність такого підходу полягає в забезпеченні захисту системи без втручання у вихідний код мікросервісів.

Аналіз останніх досліджень і публікацій. Сучасні дослідження, присвячені проблемі виявлення LR-DDoS атак, пропонують широкий спектр методів та алгоритмічних рішень, серед яких найбільшого поширення набули метод статистичного аналізу ентропії [12], метод на основі прогнозування часових рядів [13; 14], а також методи на основі машинного навчання [15], зокрема нейронних мереж [16]. Водночас для підвищення точності ідентифікації таких атак у межах цих підходів можуть залучатися додаткові атрибути прикладного рівня, що своєю чергою передбачає аналіз вмісту HTTP-запитів у дешифрованому вигляді. Зокрема, подібний підхід покладено в основу модифікованого методу на основі часових рядів, який передбачає перехід від аналізу суто мережевих показників до атрибутів прикладного рівня [17], водночас зберігаючи можливість комплексного поєднання даних обох рівнів. Проте більшість розглянутих робіт зосереджена виключно на аналізі ефективності самих методів виявлення DDoS-атак, тоді як питання проєктування програмної архітектури для їхнього впровадження у розподілені мікросервісні системи залишається відкритим.

Аналіз наявних варіантів розгортання засобів захисту від DDoS-атак дозволяє згрупувати їх у три основні категорії. Перша з них базується на використанні традиційних брандмауерів для вебзастосунків (Web Application Firewall, WAF), які розгортаються на зовнішніх точках входу до системи, зокрема на вхідних мережевих шлюзах [18; 19]. Такі програмні рішення, зокрема на основі ModSecurity та OWASP Core Rule Set [20], перехоплюють і послідовно аналізують кожен запит, перед тим як передати його до цільового сервісу. Основний недолік такого підходу полягає в тому, що вхідний шлюз перетворюється на критичний вузол інфраструктури, перевантаження або відмова якого зупиняє роботу всієї системи. Оскільки виявлення LR-DDoS атак потребує ресурсоємного аналізу вмісту та поведінкових характеристик HTTP-запитів, раптове зростання шкідливого трафіку може перевантажити потужності вузла мережевого шлюзу, спричиняючи зростання затримок під час обробки легітимного трафіку та зумовлюючи високий рівень хибнопозитивних спрацьовувань.

Друга категорія передбачає використання вбудованих модулів безпеки або проміжного програмного забезпечення (middleware) [21], коли механізми виявлення DDoS-атак інтегруються безпосередньо у вихідний код мікросервісів. Хоча це нівелює ризик перевантаження вхідного мережевого шлюзу, такий підхід вимагає модифікації вихідного коду вебфреймворків та створює додаткову залежність для бізнес-логіки самого застосунку від впровадженої системи захисту. У системах, де мікросервіси реалізовані з

використанням різних мов програмування та технологічних стеків, супровід таких програмних модулів стає складним завданням. Крім того, цей підхід порушує базовий принцип розділення відповідальності (separation of concerns) [22], оскільки перекладає завдання виявлення і блокування DDoS-атак на рівень прикладного коду застосунків.

Третя категорія охоплює рішення на основі асинхронного аналізу збережених копій захопленого мережевого трафіку або журналів доступу до вебресурсів, сформованих вебсерверами чи мережевими шлюзами для фіксації параметрів HTTP-запитів [23]. За такого сценарію аналіз атрибутів запитів здійснюється окремими програмними компонентами поза основним ланцюгом обробки, що дозволяє уникнути додаткових затримок у роботі мікросервісів. Проте суттєвим недоліком асинхронного підходу є часова затримка між початком атаки та моментом її виявлення. Оскільки аналіз даних виконується у відкладеному режимі, реагування на атаку зазвичай здійснюється із запізненням, що є критичним під час LR-DDoS атак, коли несвоєчасне блокування джерел зловмисних запитів призводить до перевантаження ресурсів системи та недоступності вебзастосунків для легітимних користувачів.

З огляду на проведений аналіз, невирішеною частиною загальної проблеми залишається відсутність цілісної програмної архітектури, яка б, по-перше, отримувала атрибути прикладного рівня безпосередньо після дешифрування HTTPS-запитів без втручання у вихідний код застосунків; по-друге, виконувала ресурсоємний аналіз в асинхронному режимі, мінімізуючи вплив на легітимний трафік, але водночас гарантуючи мінімальний час виявлення DDoS-атак; і по-третє, забезпечувала оперативне блокування джерел зловмисних запитів ще до того, як вони досягнуть цільових вебзастосунків. Проєктуванню та програмній реалізації такої архітектури і присвячено цю статтю.

Мета дослідження. Метою дослідження є проєктування та програмна реалізація розподіленої архітектури системи виявлення і блокування HTTP LR-DDoS атак для Kubernetes-кластерів, яка забезпечує захист від атак без модифікації вихідного коду цільових мікросервісів. Ключовою вимогою до розроблюваної системи є зменшення часу реагування (від моменту початку атаки до блокування її джерел) порівняно з наявними системами виявлення DDoS-атак, причому додаткова затримка в обробленні та маршрутизації легітимного трафіку від вхідного шлюзу до цільового мікросервісу не повинна перевищувати 10 мс.

Для досягнення поставленої мети необхідно вирішити такі завдання: спроєктувати розподілену архітектуру системи захисту від DDoS-атак; здійснити її програмну реалізацію для кластерного середовища Kubernetes; розробити методику експериментального оцінювання; провести тестування розробленої системи та порівняти показники її ефективності з наявними аналогами.

Викладення основного матеріалу дослідження. Зважаючи на ключову вимогу щодо забезпечення захисту від DDoS-атак без втручання у вихідний код мікросервісів, базовим інфраструктурним середовищем для реалізації архітектури обрано технологію сервісної сітки (Service Mesh) [24; 25]. Її застосування передбачає винесення функцій управління мережевим трафіком між мікросервісами в окремий шар інфраструктури,

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

який реалізується за допомогою мережі розподілених проксі-серверів. Одним із найбільш поширених рішень для реалізації цього підходу в кластерах Kubernetes є платформа Istio [26], яка забезпечує централізоване керування трафіком між сервісами, автентифікацію та авторизацію, а також гнучке налаштування політик безпеки та правил маршрутизації. Головною перевагою Istio в контексті цього дослідження є використання високопродуктивних проксі-серверів Envoy [27] для кожного мікросервісу та вхідних шлюзів. Завдяки механізму програмних розширень EnvoyFilter [28], сервісна сітка Istio надає можливість перехоплювати та аналізувати трафік безпосередньо на вхідних мережевих шлюзах кластера вже після дешифрування HTTPS-запитів, що забезпечує доступ до їхніх атрибутів в умовах захищеного з'єднання без внесення змін до вихідного коду застосунків.

Архітектура запропонованої програмної системи. В основу запропонованої архітектури покладено принцип розділення площини захоплення трафіку (Data Plane), яка здійснює його маршрутизацію, та площини аналізу і прийняття рішень (Control Plane). Перехоплення HTTP-запитів виконується безпосередньо в проксі-сервері Envoy на вхідному шлюзі кластера . На цьому етапі із запитів відбираються виключно необхідні атрибути, які без додаткових трансформацій асинхронно надсилаються до наступного компонента системи. Це дозволяє уникнути додаткових затримок під час маршрутизації трафіку до мікросервісів. Оскільки механізм виявлення DDoS-атак повністю винесено в окремий модуль, часові витрати на оброблення легітимних запитів є мінімальними.

На вхідному шлюзі Istio Ingress Gateway, де здійснюється дешифрування HTTPS-трафіку, розгортається розширення EnvoyFilter, реалізоване скриптовою мовою програмування Lua. Воно отримує параметри кожного запиту (HTTP-метод, URI, значення заголовків User-Agent, Accept-Language, Accept-Encoding, перелік імен всіх наявних заголовків у запиті, а також IP-адресу та порт джерела) у дешифрованому вигляді ще до моменту перенаправлення запиту у внутрішню мережу кластера. Важливо зазначити, що реалізований EnvoyFilter виконує виключно функцію експорту: за допомогою асинхронного виклику httpCall дані надсилаються до відповідного сервісу, який збирає атрибути запитів, у фоновому режимі (без очікування на результат обробки). Завдяки цьому критичний шлях оброблення та маршрутизації трафіку до цільових застосунків не залежить від процесу виявлення DDoS-атак.

Оскільки запропоноване розширення EnvoyFilter після дешифрування HTTPS-трафіку не має доступу до атрибутів мережевого та транспортного рівнів (зокрема значення TTL, SYN, розміру вікна TCP), які є важливими для виявлення LR-DDoS атак типу Slow Read [5], до архітектури розроблюваної системи інтегровано додатковий модуль аналізу мережевого стека на рівні протоколів TCP та UDP. Цей модуль реалізовано в середовищі Kubernetes як об'єкт DaemonSet [29], що гарантує його глобальне розгортання на всіх робочих вузлах кластера. Завдяки використанню технологій eBPF/XDP [30] програмний модуль перехоплює мережевий трафік безпосередньо на рівні мережевих інтерфейсів вузлів. Перехоплені пакети групуються за унікальною па-

рою ідентифікаторів (IP-адреса та номер порту відправника), після чого зібрані дані експортуються до іншого сервісу для подальшої обробки.

Архітектура розробленої системи ґрунтується на модульному принципі та використовує механізми просторів імен (namespaces) середовища Kubernetes для логічної ізоляції компонентів [29]. Зокрема, процеси безпосереднього перехоплення трафіку та блокування джерел LR-DDoS атак виконуються у системному просторі імен istio-system на основі вхідного мережевого шлюзу. Усі процеси обробки даних інкапсульовано у виділеному просторі імен ddos-system, який включає захоплення HTTP-запитів і мережевих пакетів, їх оброблення, класифікацію, а також збереження результатів. Для дотримання принципу єдиної відповідальності функції автоматизованого реагування на DDoS-атаки винесено в окремий простір імен ddos-operator-system. Такий підхід гарантує повну незалежність захисних механізмів від просторів імен цільових застосунків, для захисту яких створюється ця система.

Узагальнену UML-діаграму розгортання запропонованої архітектури в кластері Kubernetes та інформаційні потоки між її складовими наведено на рис. 1.

Відповідно до наведеної діаграми (рис. 1), ключові компоненти системи виконують такі функції:

– Istio Ingress Gateway – вхідний мережевий шлюз на основі проксі-сервера Envoy, в якому налаштовані розширення EnvoyFilter для виокремлення та експорту метаданих і заголовків (атрибутів) HTTP-запитів та AuthorizationPolicy для блокування шкідливого трафіку від DDoS-атак.

– HTTP Attributes Collector – сервіс, що отримує експортовані атрибути HTTP-запитів від розширення EnvoyFilter, формує з них структуровані об’єкти клієнтських сесій та надсилає їх брокеру повідомлень для подальшого аналізу модулем DDoS Classifier.

– eBPF Packet Collector – сервіс, що здійснює захоплення мережевих пакетів та надсилає атрибути пакетів на рівні мережевих та транспортних протоколів до брокера повідомлень, які потім будуть оброблені модулем DDoS Classifier.

– Apache Kafka – брокер повідомлень, що забезпечує асинхронну потокову передачу даних і усуває жорстку архітектурну залежність між компонентами системи.

– DDoS Classifier – програмний модуль, який відповідає за реалізацію методу виявлення DDoS-атак: він зчитує з брокера повідомлень атрибути прикладного та мережевого рівнів, здійснює їхню класифікацію для ідентифікації джерел атак і передає виявлені ідентифікатори до сховища Redis.

– Redis – високошвидкісне сховище даних типу “ключ-значення”, призначене для збереження ідентифікованих джерел DDoS-атак (IP-адреса та порт) із механізмом автоматичного видалення застарілих записів (TTL).

– DDoS Protection Operator – контролер, який забезпечує початкове розгортання інфраструктури системи на кластері Kubernetes, а також безперервне зчитування виявлених джерел атак зі сховища та автоматичне оновлення правил блокування AuthorizationPolicy на вхідному шлюзі.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

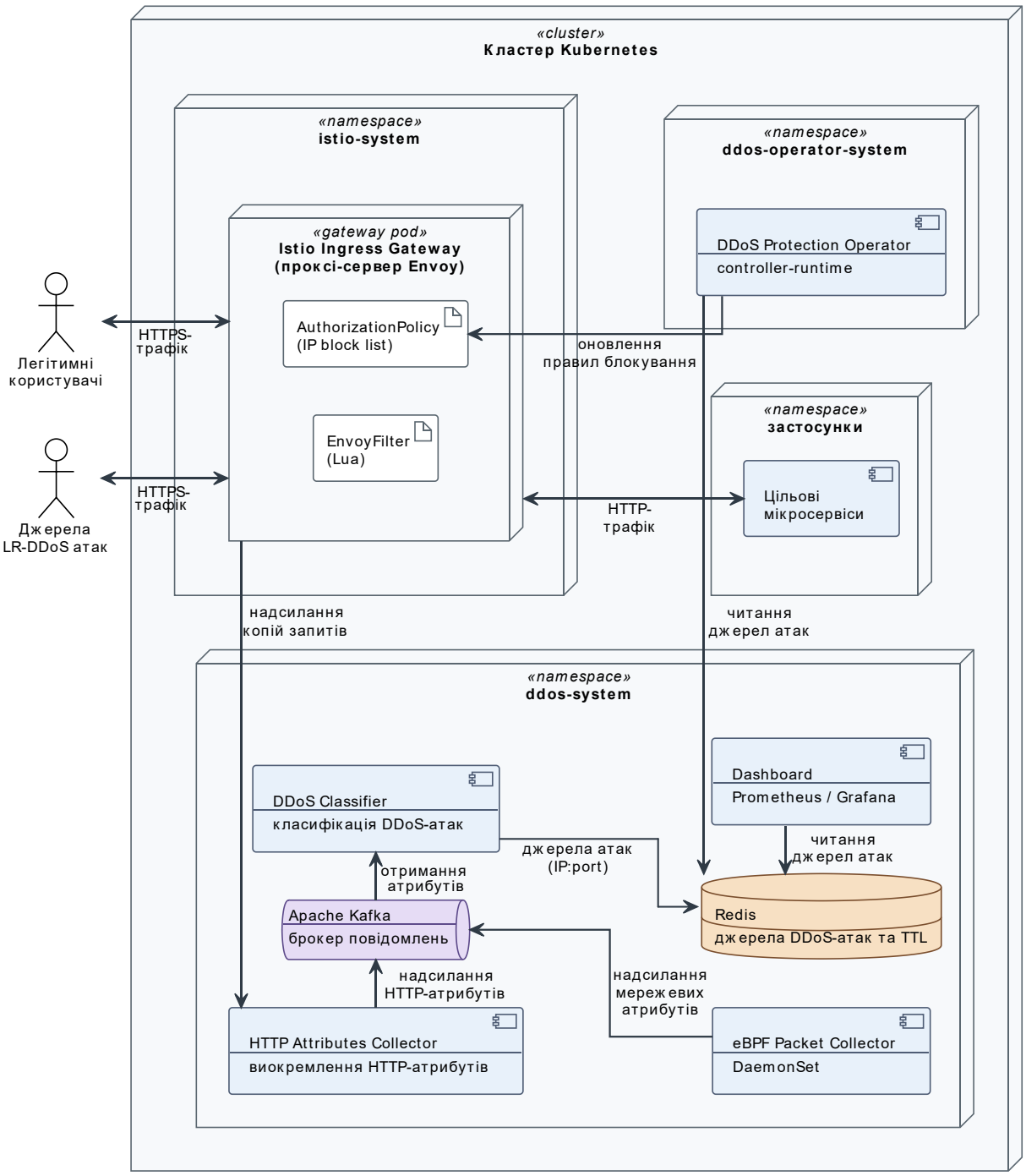


Рисунок 1 – UML-діаграма розгортання програмної системи в кластері Kubernetes

У межах описаної архітектури ресурсоємні сервіси, до яких належать HTTP Attributes Collector та DDoS Classifier, підтримують механізм горизонтального масштабування Horizontal Pod Autoscaler (HPA) для забезпечення стабільної роботи під навантаженням. Застосування брокера повідомлень Apache Kafka запроваджує подійно-орієнтовану взаємодію за шаблоном “публікація-підписка” (publish-subscribe). Це створює слабку зв’язність компонентів і гарантує, що процес швидкого зчитування атрибутів запитів не блокуватиметься тривалим аналізом навіть за умов масованих атак. Також варто зазначити, що використання архітектурного шаблону Kubernetes Operator у

контролері DDoS Protection Operator автоматизує адміністрування розробленої програмної системи та перетворює її на нативне розширення кластера [31].

Особливості програмної реалізації запропонованої системи. Ключові програмні компоненти системи, що відповідають за захоплення мережевих пакетів, оброблення HTTP-запитів, класифікацію DDoS-атак та встановлення правил фільтрації зловмисного трафіку, розроблені мовою Go. Вибір цієї мови обґрунтовується її вбудованими механізмами багатопотоковості, автоматичного керування пам'яттю (garbage collection) та помірним використанням апаратних ресурсів. Для того щоб швидкість аналізу не обмежувала загальну пропускну здатність системи під час DDoS-атак, для оброблення кожної сесії використовуються легкі потоки (goroutines), адже вони не вимагають виділення значних ресурсів, на відміну від потоків на рівні операційної системи. Це дозволяє одночасно аналізувати тисячі клієнтських сесій, а також гарантує стабільну роботу Control Plane без надмірного споживання оперативної пам'яті навіть за екстремальних навантажень.

Розгортання розробленої системи на кластері Kubernetes відбувається за допомогою контролера DDoS Protection Operator, який реалізовано на основі фреймворку controller-runtime [32]. Інтерфейсом конфігурування системи виявлення та блокування DDoS-атак виступає об'єкт для визначення користувацького ресурсу (Custom Resource Definition, CRD) під назвою DDoSProtection. Цей об'єкт розширює API кластера Kubernetes новим класом ресурсів і формалізує структуру даних: як конфігураційні параметри, що можуть бути задані в об'єкті, так і поля для відображення поточного стану системи. На його основі створюється декларативний екземпляр (Custom Resource, CR), який інкапсулює всі налаштування розроблюваної системи: параметри розгортання власних або підключення до зовнішніх екземплярів брокера повідомлень Kafka та сховища даних Redis, назви мережевих інтерфейсів для захоплення трафіку, назву вхідного шлюзу Istio, посилання на ресурс AuthorizationPolicy, інтервали синхронізації для оновлення правил блокування зі сховища Redis тощо. Основним завданням контролера є реалізація циклу узгодження (reconciliation loop) – ідемпотентної функції, яка виключно за декларативним описом ресурсу автоматично генерує та підтримує в актуальному стані всі залежні об'єкти кластера (Deployment, DaemonSet, Service), необхідні для роботи системи. Застосування механізмів локального кешування, наданих фреймворком, дозволяє суттєво знизити навантаження на API-сервер Kubernetes під час моніторингу стану цих об'єктів.

Окрім керування життєвим циклом розробленої системи, контролер DDoS Protection Operator безперервно оновлює правила фільтрації зловмисних джерел у ресурсі AuthorizationPolicy. Цей механізм винесено в окремий паралельний процес, незалежний від головного циклу узгодження компонентів системи. Така ізоляція зумовлена різним характером подій, що ініціюють ці процеси: якщо основний цикл узгодження об'єктів системи реагує на зміну стану розгорнутих ресурсів у кластері, то оновлення правил блокування потребує зчитування нових ідентифікаторів загроз зі сховища Redis, як тільки вони там з'являються. Об'єднання цих завдань у межах єдиного циклу узго-

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

дження унеможливило б миттєве блокування джерел DDoS-атак, оскільки для актуалізації фільтрів трафіку контролеру довелося б очікувати на виникнення подій зміни стану розгорнутих компонентів. Цей паралельний процес передбачає періодичне зчитування поточного переліку заблокованих IP-адрес зі сховища Redis, обчислення бажаного стану ресурсу AuthorizationPolicy та порівняння його з фактичним станом у кластері. Модифіковані правила блокування джерел DDoS-атак застосовуються до об'єкта AuthorizationPolicy лише в разі виявлення розбіжностей, що мінімізує кількість операцій запису до API-сервера Kubernetes. Для взаємодії з ресурсами Istio в кластері використовуються типізовані структури даних із відповідних програмних бібліотек, що забезпечує коректність модифікації об'єктів Istio.

Особливістю реалізації модуля класифікації трафіку DDoS Classifier є використання легких потоків (goroutines), що дозволяє обробляти отримані дані з високою продуктивністю. Кожна клієнтська сесія, що визначається унікальною парою з IP-адреси джерела та номера порту, аналізується в окремому потоці. У межах кожної сесії здійснюється одночасний аналіз атрибутів мережевого та прикладного рівнів, які надходять від сервісів eBPF Packet Collector та HTTP Attributes Collector відповідно. Для аналізу даних мережевого рівня реалізовано метод на основі прогнозування часових рядів [14], а для прикладного рівня – реалізовано модифіковану версію цього методу [17]. Аналіз даних на обох рівнях передбачає бінарну класифікацію клієнтської сесії для її ідентифікації як легітимної або зловмисної. Після завершення роботи паралельних потоків результати класифікації з обох рівнів об'єднуються за допомогою примітива синхронізації sync.WaitGroup. Якщо хоча б один із рівнів класифікував DDoS-атаку, пара з IP-адреси джерела та номера порту цієї клієнтської сесії вноситься до бази сигнатур атак. Формування та перевірка бази сигнатур здійснюються за допомогою структури sync.Map, яка є безпечною для використання в багатопотоковому середовищі, а також оптимізованою для інтенсивного одночасного читання та запису без взаємних блокувань. Це дозволяє виконувати швидку перевірку наявності сигнатури поточної сесії у переліку відомих джерел DDoS-атак.

Також у процесі аналізу в модулі DDoS Classifier кожній клієнтській сесії ставиться у відповідність набір сигнатур, які обчислюються на основі атрибутів мережевого та прикладного рівнів. Використання сигнатур обґрунтовується тим, що під час DDoS-атак вузли одного ботнету часто генерують трафік зі схожими характеристиками, але з різних IP-адрес. Щоб уникнути ресурсоємної класифікації трафіку від нових IP-адрес раніше виявленого ботнету та пришвидшити блокування таких джерел, модуль класифікації спочатку обчислює сигнатури з атрибутів обох рівнів і перевіряє, чи наявні вони в базі сигнатур відомих атак. У разі виявлення збігу результат класифікації повертається миттєво, визначаючи таку сесію як зловмисну. Якщо ж збігів немає, виконується класифікація атрибутів трафіку на обох рівнях. У випадку, коли за результатами класифікації виявлено DDoS-атаку, до бази сигнатур вноситься не лише пара з IP-адреси та номера порту, а й обчислені сигнатури атрибутів обох рівнів, що відповідають цій парі. Такий підхід дозволяє системі значно швидше виявляти та блокувати нові джерела атак за наявності схожих атрибутів трафіку.

Захоплення інтернет-трафіку здійснюється двома спеціалізованими модулями, які працюють паралельно та незалежно один від одного. Сервіс eBPF Packet Collector, реалізований за допомогою бібліотеки cilium/ebpf, перехоплює пакети за технологією XDP (eXpress Data Path) безпосередньо на рівні драйвера мережевого інтерфейсу. На відміну від стандартних засобів захоплення трафіку (як-от tcpdump), це дозволяє зчитувати та аналізувати атрибути мережевого і транспортного рівнів без створення додаткового навантаження на процесор та без повного копіювання вмісту пакетів [30]. На прикладному рівні HTTP Attributes Collector отримує дані на кінцеву точку (endpoint) API /ingest від розширення EnvoyFilter, яке здійснює асинхронний експорт атрибутів HTTP-запитів. Цей сервіс обробляє отримані заголовки та інші атрибути запитів для того, щоб трансформувати їх у структуровані об’єкти клієнтських сесій. Підготовлені дані з обох сервісів публікуються у відповідні теми (topics) брокера повідомлень Kafka за допомогою бібліотеки segmentio/kafka-go для подальшого аналізу.

UML-діаграму компонентів розробленої системи наведено на рис. 2.

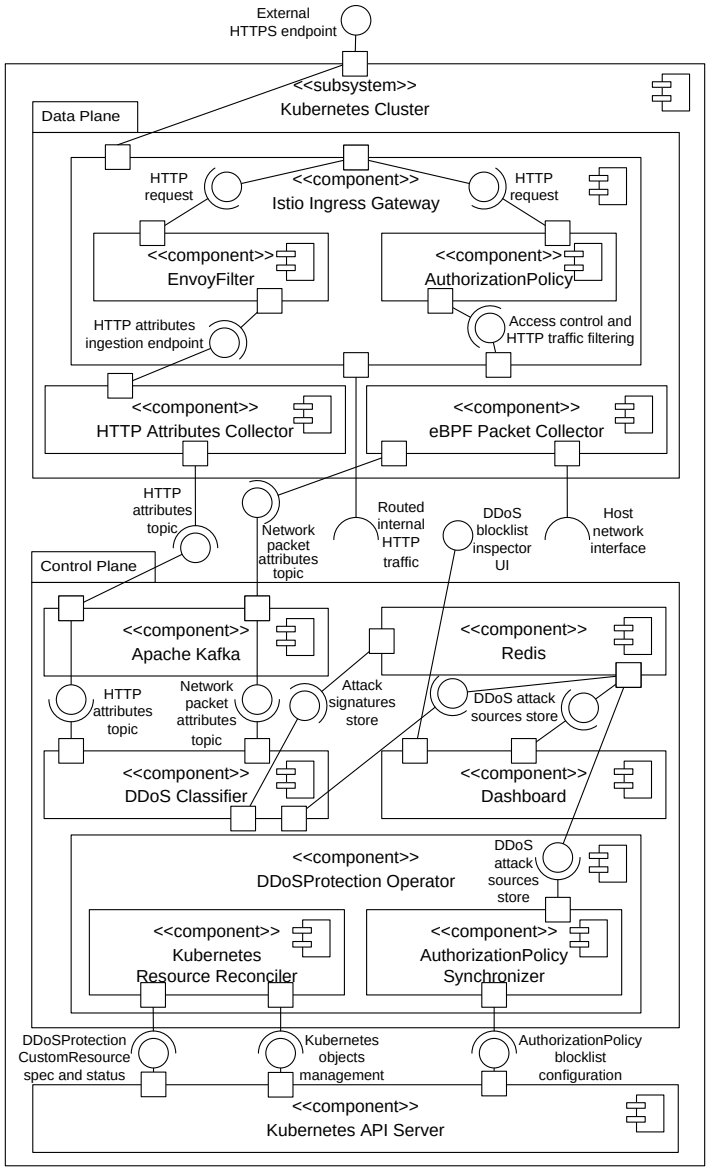


Рисунок 2 – UML-діаграма компонентів системи виявлення і блокування DDoS-атак

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Сценарії оброблення трафіку та оцінка часу реакції на атаки. Оброблення та маршрутизація вхідного трафіку в запропонованій системі здійснюються за одним із трьох сценаріїв, вибір якого залежить від статусу ідентифікації джерела HTTP-запитів та наявності його сигнатури у базі відомих атак. Такий підхід дозволяє оптимально використовувати обчислювальні ресурси та мінімізувати затримки: ресурсоємна класифікація виконується виключно для принципово нових джерел атак, тоді як ідентифікація вузлів відомих ботнетів пришвидшується, а вже заблоковані джерела DDoS-атак відхиляються миттєво.

Для нових, ще не класифікованих клієнтських сесій із невідомими сигнатурами трафіку застосовується сценарій "повного холодного шляху" (full cold path), що передбачає повний первинний аналіз. У межах цього сценарію атрибути асинхронно експортуються до модуля класифікації атак, тоді як сам трафік безперешкодно маршрутизується до цільового застосунку. Модуль класифікації накопичує отримані дані у межах заданого вікна спостереження T_{win}, після чого обчислює сигнатури поточної сесії. Не знайшовши збігів у базі відомих атак, модуль виконує класифікацію запитів клієнтської сесії за допомогою відповідних методів. У разі класифікації джерела як шкідливого його IP-адреса, номер порту та обчислені сигнатури записуються до сховища Redis із визначеним часом життя (TTL), після завершення якого вони автоматично видаляються. Якщо ж аномалій не виявлено, сесія вважається легітимною, і процес її перевірки завершується. Згодом компонент AuthorizationPolicy Synchronizer у контролері DDoS Protection Operator під час чергового циклу синхронізації зчитує ці дані та оновлює правила в AuthorizationPolicy на вхідному шлюзі. Цей процес є повністю асинхронним і не додає затримок під час оброблення легітимних запитів. Однак він потребує скінченного часу, включаючи збирання атрибутів клієнтських сесій у межах вікна спостереження, що необхідно для виявлення LR-DDoS атак прикладного рівня. Загальний час реакції системи T_{mit}^{full} для цього сценарію становить:

$$T_{mit}^{full} = T_{win} + T_{fp} + T_{class} + T_{redis} + T_{sync} + T_{prop}, \quad (1)$$

де T_{mit}^{full} – час реакції розробленої системи на DDoS-атаку для сценарію "повного холодного шляху"; T_{win} – тривалість вікна спостереження; T_{fp} – час обчислення сигнатур клієнтської сесії; T_{class} – час класифікації запитів клієнтської сесії; T_{redis} – час запису адреси та сигнатур джерела DDoS-атаки у сховищі Redis; T_{sync} – інтервал циклу перевірки контролером нових джерел атак у сховищі Redis; T_{prop} – час застосування оновленої конфігурації правил блокування у проксі-сервері.

Для нових IP-адрес, трафік яких генерується вже виявленим ботнетом, застосовується сценарій "прискореного холодного шляху" (fast-track cold path). Оскільки перший запит від нового джерела пропускається вхідним шлюзом, його атрибути так само асинхронно експортуються до модуля класифікації та накопичуються протягом того самого вікна спостереження T_{win}. Після його завершення модуль обчислює сигнатури сесії за час T_{fp} і знаходить збіг у локальній копії базі сигнатур ще до запуску класифікації. У

цьому разі сесія ідентифікується як шкідлива миттєво – без виконання ресурсоємної класифікації даних мережевого та прикладного рівнів, тобто $T_{class} = 0$. Відповідна IP-адреса, номер порту та сигнатури одразу зберігаються у сховищі Redis для подальшого блокування оператором. Для цього сценарію час реакції T_{mit}^{sig} скорочується через вилучення найважчої обчислювальної складової T_{class}:

$$T_{mit}^{sig} = T_{win} + T_{fp} + T_{redis} + T_{sync} + T_{prop}, \quad (2)$$

де T_{mit}^{sig} – час реакції розробленої системи на DDoS-атаку для сценарію “прискореного холодного шляху”.

Для вже виявлених зловмисних джерел запитів застосовується сценарій “гарячого шляху” (hot path), який забезпечує їхнє миттєве блокування. Компонент Authorization-Policy Synchronizer у контролері DDoS Protection Operator періодично зчитує заблоковані адреси зі сховища Redis та оновлює ресурс AuthorizationPolicy, що працює на рівні вхідного шлюзу. Після цього кожен наступний запит від ідентифікованого джерела DDoS-атак відхиляється проксі-сервером із кодом відповіді HTTP 403. Обчислювальна складність такого блокування становить $O(1)$, оскільки воно виконується без залучення модуля класифікації та без звернень до бази даних. Час реакції для цього сценарію є мінімальним, оскільки з нього виключаються абсолютно всі складові аналізу, класифікації та синхронізації:

$$T_{mit}^{hot} = O(1), \quad (3)$$

де T_{mit}^{hot} – час реакції розробленої системи на DDoS-атаку для сценарію “гарячого шляху”.

Отже, система здійснює ресурсоємну класифікацію профілю атаки лише один раз на повному холодному шляху (1), потім використовує обчислені сигнатури для прискореного виявлення інших IP-адрес ботнету (2), після чого фільтрує всі їхні подальші запити на рівні шлюзу за сталий час (3). Послідовність взаємодій компонентів для всіх трьох сценаріїв наведено на рис. 3.

Аналіз отриманих результатів. Експериментальну оцінку розробленої системи проведено в локальному кластері Kind (Kubernetes 1.36, Go 1.26.2) з одним вузлом, розгорнутому на комп’ютері з процесором Intel Core i5-10210U (1,60 ГГц, 4 ядра) та 8 ГБ оперативної пам’яті. У кластері запущено розроблений програмний комплекс системи захисту і блокування DDoS-атак разом із тестовим вебсервісом httpbin [33]. Оцінювання системи здійснювалося за такими метриками: затримка оброблення HTTP-запитів, пропускна здатність, час реакції на загрози та споживання оперативної пам’яті. Легітимний трафік генерувався утилітою Fortio [34], запущеною всередині кластера, що надсилала HTTP-запити до Istio Ingress Gateway. Трафік із характеристиками LR-DDoS атак формувався за допомогою спеціальних утиліт та скриптів.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

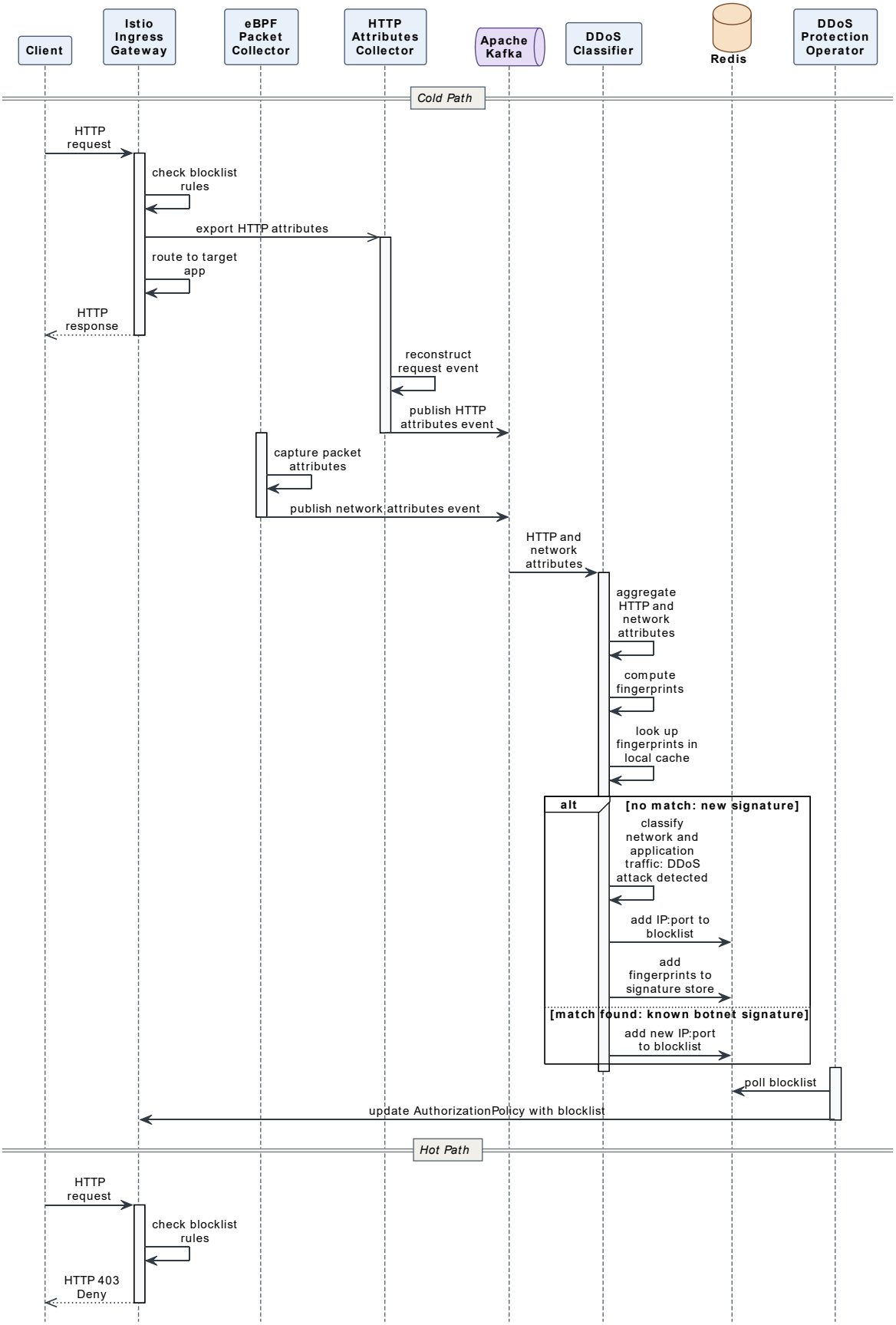


Рисунок 3 – UML-діаграма послідовності взаємодії компонентів для трьох сценаріїв

Зокрема, атака Slow Read відтворювалася за допомогою спеціалізованої утиліти [35], що відкривала TCP-з'єднання з малим приймальним буфером та штучно затримувала зчитування об'ємної відповіді – це дозволило перевірити здатність системи фіксувати аномальне зменшення розміру вікна TCP на транспортному рівні. Окрім цього, розроблено скрипти, які формували серію запитів за допомогою curl із заголовками автоматизованих ботів, а також відтворювали частоту запитів та їхню структуру, типові для LR-DDoS атак. Для оцінки сценарію "повного холодного шляху" скрипт надсилав серію з шести запитів через одне HTTP-з'єднання, при цьому модуль класифікації перед кожним тестовим прогоном запускався із порожньою базою сигнатур. Для сценарію "прискореного холодного шляху" застосовувався автоматизований скрипт, що надсилав запити, параметри яких відповідали попередньо збереженій сигнатурі відомої атаки. Для мінімізації впливу фонових коливань під час вимірювання пропускної здатності та затримок легітимних запитів кожен тестовий запуск тривав 15 секунд і повторювався тричі, а як результуюче значення використовувалася медіана.

Час реакції розробленої системи на DDoS-атаку визначався як часовий інтервал від моменту ініціалізації DDoS-атаки до отримання першої відповіді від вхідного шлюзу з кодом HTTP 403. Для кожного з досліджуваних сценаріїв оброблення трафіку виконано по 10 успішних тестових запусків. Отримані результати вимірювань у вигляді середнього арифметичного значення, а також 50-го (p50, медіана) та 90-го (p90) перцентилів наведено в табл. 1.

Таблиця 1

Час реакції системи на DDoS-атаку для різних сценаріїв оброблення трафіку

Час на DDoS-атаку, с / Сценарій реагування	Full cold path	Fast-track cold path	Hot path
Середнє значення	3,53	1,53	0,005641
p50	3,15	0,8	0,005674
p90	5,22	3,33	0,006576

Аналіз отриманих даних в табл. 1 свідчить, що сценарій "повного холодного шляху" демонструє найбільший час реакції, оскільки перед застосуванням блокування система здійснює ресурсоємну класифікацію раніше невідомої клієнтської сесії. Натомість застосування сценарію "прискореного холодного шляху" дозволяє виключити цей етап завдяки збігу обчисленої сигнатури із базою відомих сигнатур, що забезпечує зниження медіанного часу реакції (p50) майже в 4 рази – до 0,8 с. Мінімальна затримка очікувано спостерігається під час застосування "гарячого шляху", де блокування зловмисного джерела відбувається безпосередньо на рівні вхідного шлюзу без залучення модуля класифікації, забезпечуючи медіанний час блокування запиту на рівні 5,674 мс.

Окремо досліджувалися затримка, пропускна здатність та обчислювальні витрати компонентів Data Plane та модуля класифікації. Для оцінки роботи Istio Ingress Gateway

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

виконано порівняння оброблення ідентичних HTTP-запитів у двох станах: без EnvoyFilter та з увімкненим розширенням EnvoyFilter для експорту атрибутів. Щоб виключити додаткові затримки цільового вебсервісу на результати вимірювань та уникнути пересилання трафіку вглиб кластера, на шлюзі налаштовано генерацію локальної відповіді HTTP 200 для URI /benchmark/local. Завдяки цьому виміряний час відгуку відображає виключно затримку оброблення запиту на рівні шлюзу (зокрема, витрати на виконання EnvoyFilter). Під час експерименту контролер DDoS Protection Operator залишався неактивним, а ресурс AuthorizationPolicy містив адресу-заглушку, що гарантувало проходження всіх запитів без залучення механізмів блокування. Навантаження на шлюз генерувалося утилітою Fortio (сеансами по 15 с): за фіксованої інтенсивності 500 запитів за секунду фіксувалася медіанна затримка, а за максимальної інтенсивності – гранична пропускна здатність. Час виконання та споживання оперативної пам'яті модулем класифікації оцінювалися за допомогою окремого тесту на вибірці з 4000 клієнтських сесій. Результати цих вимірювань наведено в табл. 2.

Таблиця 2

Оцінка накладних витрат шлюзу та модуля класифікації

Метрика	Значення
Медіанна затримка оброблення запиту на шлюзі (без EnvoyFilter), мс	1,71
Медіанна затримка оброблення запиту на шлюзі (з EnvoyFilter), мс	4,643
Максимальна пропускна здатність шлюзу (без EnvoyFilter), запитів/с	6063
Максимальна пропускна здатність шлюзу (з EnvoyFilter), запитів/с	2158
Час класифікації 4000 сесій у модулі класифікації T_{class}, мс	32,5
Обсяг оперативної пам'яті для класифікації 4000 сесій, МБ	3,09

Наведені дані в табл. 2 демонструють обчислювальні накладні витрати на експорт атрибутів на рівні шлюзу та безпосередню класифікацію клієнтських сесій. На розгорнутому тестовому кластері активація EnvoyFilter призвела до збільшення медіанної затримки на 2,933 мс та зниження пропускної здатності шлюзу на 64,4 %. Важливо зазначити, що ці показники відображають виключно витрати на виокремлення та асинхронне відправлення HTTP-атрибутів до модуля збору даних та не включають час на оброблення даних кінцевим вебзастосунком. Значення 32,5 мс та 3,09 МБ характеризують одноразове виконання класифікації для набору даних із 4000 сесій усередині процесу модуля класифікації, а не оброблення однієї сесії чи споживання пам'яті контейнером загалом. Час роботи реалізованих методів виявлення DDoS-атак у модулі класифікації T_{class} для масиву з 4000 сесій становить 32,5 мс. Отримані значення часу та використання оперативної пам'яті (лише 3,09 МБ) є незначними для такого обсягу даних, що підтверджує високу обчислювальну ефективність рішення та мінімізує його навантаження на ресурси середовища Kubernetes.

Для здійснення порівняльного аналізу розробленої системи з наявними аналогами на тому самому вхідному шлюзі розгорнуто рішення Coraza – WAF (Web Application Firewall), що сумісний із ModSecurity та містить повний набір правил OWASP Core Rule

Set (CRS), реалізований як WASM-розширення для проксі-сервера Envoy [36]. Експеримент передбачав дослідження двох альтернативних сценаріїв оброблення запитів: запропонованої системи з асинхронним експортом атрибутів через EnvoyFilter та програми Coraza із синхронною перевіркою правил (без використання EnvoyFilter). Обидва варіанти тестувалися на локальному URI /benchmark/local в шлюзі, що повністю виключало вплив цільового застосунку. Як і в попередніх дослідженнях, кожне вимірювання тривало 15 с і повторювалося тричі, а за підсумковий показник бралося медіанне значення. Важливо зазначити, що медіанну затримку визначали за зниженої інтенсивності навантаження, а саме 10 запитів за секунду. Це зумовлено тим, що при 500 запитів за секунду Coraza вичерпує обчислювальні ресурси, через що до основного часу оброблення додається затримка очікування запитів у черзі, що спотворило б результати порівняння. Водночас гранична пропускна здатність вимірювалася окремо в умовах максимального навантаження. Результати порівняльного оцінювання наведено в табл. 3.

Таблиця 3

Порівняння затримки та пропускної здатності розробленої системи з Coraza WAF

Система захисту	Медіанна затримка оброблення запиту, мс	Максимальна пропускна здатність, запитів/с
Розроблена система	3,667	1973
Coraza WAF	45	473

Наведені в табл. 3 дані демонструють порівняння ключових показників обох рішень за однакових умов експерименту. Відмінність абсолютних значень затримки та пропускної здатності для розробленої системи порівняно з результатами, поданими в табл. 2, пояснюється виключно зниженою інтенсивністю тестового навантаження (10 запитів/с замість 500 запитів/с). Аналіз даних показує, що порівняно із запропонованим рішенням застосування Coraza WAF збільшує медіанну затримку на 41,333 мс (у 12,3 раза) та знижує граничну пропускну здатність на 76 %. Отже, за умов проведеного експерименту синхронна перевірка трафіку з повним набором правил CRS генерує суттєво більші накладні витрати на рівні вхідного шлюзу, ніж розроблений архітектурний підхід на базі асинхронного експорту атрибутів.

Під час аналізу отриманих результатів необхідно враховувати низку обмежень проведених експериментів. По-перше, вимірювання виконувалися на тестовому кластері з одним вузлом за помірної інтенсивності трафіку, тому в розподілених середовищах з декількома вузлами під час масивних атак показники мережевих затримок можуть відрізнятися. По-друге, обчислювальна складова T_{class} безпосередньо залежить від застосованих методів виявлення атак. Оцінка їхньої точності не входила в межі цього дослідження, проте впровадження складніших методів неминуче збільшить час класифікації. Крім того, зі збільшенням кількості виявлених зловмисників зростатиме обсяг списку блокування в ресурсі AuthorizationPolicy, що підвищить споживання оператив-

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

ної пам’яті на шлюзі. Однак цей вплив може ефективно нівелюватися механізмом TTL, який забезпечує автоматичне очищення застарілих записів блокування.

Окремо слід відзначити архітектурну специфіку порівняння розробленого рішення з рішеннями класу WAF (зокрема Coraza). Зазвичай WAF виконує миттєву синхронну перевірку за статичними правилами, що створює значні накладні витрати на кожен запит і потребує ручного налаштування під нові вектори атак. Натомість запропонована система виносить складну логіку класифікації LR-DDoS (яка потребує тривалого вікна спостереження) в окремий асинхронний конвеєр. Через це повний цикл первинного виявлення атаки (cold path) триває довше, ніж перевірка вже наявного правила у WAF. Проте загальна ефективність системи ґрунтується на автоматичній класифікації та оптимізованому способу оброблення запитів: після першої класифікації переважна більшість шкідливого трафіку блокується миттєво безпосередньо на шлюзі (hot path), що робить таку архітектуру оптимальною для захисту застосунків. У подальших дослідженнях планується розширити експериментальну базу для тестування системи в масштабованих середовищах.

Висновки. У статті запропоновано та реалізовано розподілену архітектуру захисту мікросервісів від HTTP LR-DDoS атак у середовищі Kubernetes, яка не потребує модифікації вихідного коду цільових застосунків. Запропоноване архітектурне рішення передбачає виокремлення атрибутів HTTP-запитів після дешифрування TLS-трафіку у вхідному шлюзі Istio Ingress Gateway, їхню асинхронну класифікацію поза основним потоком оброблення та блокування виявлених джерел атак за допомогою Istio-ресурсу AuthorizationPolicy.

Головною особливістю розробленої архітектури є поєднання асинхронного експорту HTTP-атрибутів, збору мережевих атрибутів за допомогою технологій eBPF/XDP та реалізації трьох сценаріїв реагування системи на DDoS-атаки. Завдяки цьому неідентифікований трафік підлягає аналізу за допомогою реалізованих методів класифікації, запити з відомими сигнатурами розпізнаються без повторних обчислень, а HTTP-запити від уже заблокованих джерел атак відхиляються безпосередньо вхідним шлюзом зі сталою обчислювальною складністю $O(1)$.

Результати проведених експериментальних досліджень підтвердили досягнення поставленої мети та виконання всіх встановлених вимог. Застосування розширення EnvoyFilter збільшило затримку оброблення запитів лише на 2,933 мс, що повністю відповідає встановленій вимозі (не більше 10 мс) та не створює відчутних затримок під час маршрутизації легітимного трафіку. Час реагування системи на атаку склав 3,15 с для сценарію “full cold path” (класифікація нових джерел запитів), 0,8 с для “fast-track cold path” (розпізнавання відомих сигнатур) та 5,674 мс для “hot path” (блокування на рівні шлюзу). Модуль класифікації виконав аналіз 4000 сесій за 32,5 мс із використанням 3,09 МБ оперативної пам’яті. У порівнянні з Coraza WAF розроблена система продемонструвала у 12,3 раза меншу медіанну затримку (3,667 мс проти 45 мс) та вищу граничну пропускну здатність (1973 запити за секунду проти 473 запитів за секунду).

Отримані експериментальні дані отримано на тестовому кластері з одним вузлом, тому вони відображають особливості розробленого рішення, а не параметри його роботи у повноцінному промисловому середовищі. Напрямами подальших досліджень є тестування системи у розподіленій конфігурації з багатьма вузлами, налаштування динамічного масштабування класифікатора на основі розміру черги запитів у Kafka та перевірка ефективності функціонування всієї системи на наборах реального мережевого трафіку.

ЛІТЕРАТУРА

1. Yoachimik O., Pacheco J. 2025 Q4 DDoS threat report: A record-setting 31.4 Tbps attack caps a year of massive DDoS assaults. *The Cloudflare Blog*.
URL: https://blog.cloudflare.com/ddos-threat-report-2025-q4/ (date of access: 16.04.2026).

2. Singh Samra R., Barcellos M. DDoS2Vec: flow-level characterisation of volumetric DDoS attacks at scale. *Proceedings of the ACM on networking*. 2023. Vol. 1, CoNEXT3. 13. DOI: 10.1145/3629135.

3. Sabri S., Ismail N., Hazzim A. Slowloris DoS attack based simulation. *IOP conference series: materials science and engineering*. 2021. Vol. 1062, no 1. 012029.
DOI: 10.1088/1757-899x/1062/1/012029.

4. Hands-on denial of service lab exercises using SlowLoris and RUDY / E. Damon et al. *The 2012 information security curriculum development conference*, Kennesaw, Georgia, 12–13 October 2012. P. 21–29. DOI: 10.1145/2390317.2390321.

5. Shafieian S., Zulkernine M., Haque A. CloudZombie: launching and detecting slow-read distributed denial of service attacks from the cloud. *2015 IEEE international conference on computer and information technology; ubiquitous computing and communications; dependable, autonomic and secure computing; pervasive intelligence and computing*, Liverpool, 26–28 October 2015. P. 1733–1740. DOI: 10.1109/cit/iucc/dasc/picom.2015.261.

6. Low-Rate DoS attacks, detection, defense, and challenges: a survey / W. Zhijun et al. *IEEE access*. 2020. Vol. 8. P. 43920–43943. DOI: 10.1109/access.2020.2976609.

7. Tripathi N., Hubballi N. Application layer denial-of-service attacks and defense mechanisms. *ACM computing surveys*. 2021. Vol. 54, no. 4. 86. DOI: 10.1145/3448291.

8. Detection and mitigation of low-rate denial-of-service attacks: a survey / V. D. M. Rios et al. *IEEE access*. 2022. Vol. 10. P. 76648–76668. DOI: 10.1109/access.2022.3191430.

9. Velepucha V., Flores P. A survey on microservices architecture: principles, patterns and migration challenges. *IEEE access*. 2023. Vol. 11. P. 88339–88358.
DOI: 10.1109/access.2023.3305687.

10. Vasireddy I., Ramya G., Kandi P. Kubernetes and docker load balancing: state-of-the-art techniques and challenges. *International journal of innovative research in engineering and management*. 2023. Vol. 10, no. 6. P. 49–54. DOI: 10.55524/ijirem.2023.10.6.7.

11. Deressu T. F., Beyene A. M., Gemechu A. Y. A survey of application layer distributed denial of service detection: approaches, models, and datasets. *Artificial intelligence review*. 2026. Vol. 59, no. 6. 139. DOI: 10.1007/s10462-026-11528-3.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

12. Tian Q., Miyata S. A DDoS attack detection method using conditional entropy based on SDN traffic. *IoT*. 2023. Vol. 4, no. 2. P. 95–111. DOI: 10.3390/iot4020006.
13. Qasim S. S., Nsaif S. M. Advancements in time series-based detection systems for distributed denial-of-service (DDoS) attacks: a comprehensive review. *Babylonian journal of networking*. 2024. Vol. 2024. P. 9–17. DOI: 10.58496/bjn/2024/002.
14. Detection of slow DDoS attacks based on user's behavior forecasting / V. Savchenko et al. *International journal of emerging trends in engineering research*. 2020. Vol. 8, no. 5. P. 2019–2025. DOI: 10.30534/ijeter/2020/90852020.
15. A survey of low rate DDoS detection techniques based on machine learning in software-defined networks / A. A. Alashhab et al. *Symmetry*. 2022. Vol. 14, no. 8. 1563.
DOI: 10.3390/sym14081563.
16. Ilango H. S., Ma M., Su R. A feedforward–convolutional neural network to detect low-rate DoS in IoT. *Engineering applications of artificial intelligence*. 2022. Vol. 114. 105059. DOI: 10.1016/j.engappai.2022.105059.
17. Онай М. В., Кравчук А. А. Вдосконалення та аналіз ефективності методів виявлення DDoS-атак на основі протоколу HTTP. *Прикладна математика та комп'ютинг ПМК' 2024* : зб. тез доп. Сімнадцятої конф. магістрантів та аспірантів, м. Київ, 20–22 листоп. 2024 р. С. 149–153. URL: https://ela.kpi.ua/handle/123456789/71110.
18. Web application firewall for detecting and mitigation of based DDoS attacks using machine learning and blockchain / E. Leka et al. *TEM journal*. 2024. Vol. 13, no. 4. P. 2802–2811. DOI: 10.18421/tem134-17.
19. Dawadi B. R., Adhikari B., Srivastava D. K. Deep learning technique-enabled web application firewall for the detection of web attacks. *Sensors*. 2023. Vol. 23, no. 4. 2073.
DOI: 10.3390/s23042073.
20. OWASP ModSecurity Core Rule Set (CRS): project documentation. *OWASP Foundation*. URL: https://coreruleset.org/docs/ (date of access: 20.04.2026).
21. Alasri A., Sulaiman R. Protection of XML-based denail-of-service and HTTP flooding attacks in web services using the middleware tool. *International journal of engineering and technology*. 2018. Vol. 7, no. 4.7. P. 322–329. DOI: 10.14419/ijet.v7i4.7.20570.
22. Berezhna M. Separation of concerns: a fundamental principle in software development. *DEV Community*. URL: https://dev.to/marie_berezhna/separation-of-concerns-a-fundamental-principle-in-software-development-oaj (date of access: 20.04.2026).
23. Detecting distributed denial-of-service (DDoS) attacks through the log consolidation processing (LCP) framework / M. A. Khairuddin et al. *International journal on informatics visualization*. 2024. Vol. 8, no. 3. P. 1064–1073. DOI: 10.62527/joiv.8.3.2184.
24. Service mesh overview and comparison / A. Kammer et al. *INNOQ*.
URL: https://servicemesh.es/ (date of access: 28.04.2026).
25. Singh P., Ayyasamy S. Exploring, analyzing and tuning service mesh performance: a literature review. 2023. 10 p. (Preprint. TechRxiv). DOI: 10.36227/techrxiv.22776119.v1.
26. Priya Patharlagadda P. Kubernetes traffic management using Istio. *Journal of media & management*. 2022. Vol. 4, no. 1. DOI: 10.47363/jmm/2022(4)e101.
27. Architecture overview and what is Envoy – Envoy documentation. *Envoy proxy*.

URL: https://www.envoyproxy.io/docs/envoy/v1.38.2/intro/what_is_envoy (date of access: 30.04.2026).
28. Envoy Filter: traffic management in Istio – Istio documentation. *Istio*.
URL: https://istio.io/latest/docs/reference/config/networking/envoy-filter/ (date of access: 30.04.2026).
29. Kubernetes documentation: fundamental concepts. *The Kubernetes Authors*.
URL: https://kubernetes.io/docs/concepts/ (date of access: 04.05.2026).
30. Fast Packet Processing with eBPF and XDP / M. A. M. Vieira et al. *ACM computing surveys*. 2020. Vol. 53, no. 1. 16. DOI: 10.1145/3371038.
31. Kubernetes documentation: operator pattern. *The Kubernetes Authors*. URL: https://kubernetes.io/docs/concepts/extend-kubernetes/operator/ (date of access: 04.05.2026).
32. Kubernetes controller-runtime project – Go Packages. *The Go Project*. URL: https://pkg.go.dev/sigs.k8s.io/controller-runtime (date of access: 04.05.2026).
33. Reitz K. PostmanLabs/httpbin: HTTP request & response service, written in python + flask. *GitHub, Inc*. URL: https://github.com/postmanlabs/httpbin (date of access: 04.05.2026).
34. Fortio load testing library, command line tool, advanced echo server in go. *Fortio*.
URL: https://fortio.org/ (date of access: 04.05.2026).
35. SlowHttp – Slow Read DDoS simulator written as a Python script. *GitHub, Inc*.
URL: https://github.com/ojardila/slowhttp (date of access: 04.05.2026).
36. Coraza – open source, enterprise-grade, high performance web application firewall. *Coraza*. URL: https://www.coraza.io/docs/tutorials/introduction/ (date of access: 04.05.2026).

REFERENCES

1. Yoachimik, O., & Pacheco, J. (2026). 2025 Q4 DDoS threat report: A record-setting 31.4 Tbps attack caps a year of massive DDoS assaults. *The Cloudflare Blog*. https://blog.cloudflare.com/ddos-threat-report-2025-q4/
2. Singh Samra, R., & Barcellos, M. (2023). DDoS2Vec: Flow-level characterisation of volumetric DDoS attacks at scale. *Proceedings of the ACM on Networking*, 1(CoNEXT3), Article 13. https://doi.org/10.1145/3629135
3. Sabri, S., Ismail, N., & Hazzim, A. (2021). Slowloris DoS attack based simulation. *IOP Conference Series: Materials Science and Engineering*, 1062(1), Article 012029.
https://doi.org/10.1088/1757-899x/1062/1/012029
4. Damon, E., Dale, J., Laron, E., Mache, J., Land, N., & Weiss, R. (2012). Hands-on denial of service lab exercises using SlowLoris and RUDY. In *The 2012 information security curriculum development conference* (pp. 21–29). ACM Press.
https://doi.org/10.1145/2390317.2390321
5. Shafieian, S., Zulkernine, M., & Haque, A. (2015). CloudZombie: Launching and detecting slow-read distributed denial of service attacks from the cloud. In *2015 IEEE international conference on computer and information technology; ubiquitous computing and communications; dependable, autonomic and secure computing; pervasive intelligence and computing* (pp. 1733–1740). IEEE. https://doi.org/10.1109/cit/iucc/dasc/picom.2015.261

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

6. Zhijun, W., Wenjing, L., Liang, L., & Meng, Y. (2020). Low-Rate DoS attacks, detection, defense, and challenges: A survey. *IEEE Access*, *8*, 43920–43943. https://doi.org/10.1109/access.2020.2976609
7. Tripathi, N., & Hubballi, N. (2021). Application layer denial-of-service attacks and defense mechanisms. *ACM Computing Surveys*, *54*(4), Article 86. https://doi.org/10.1145/3448291
8. Rios, V. D. M., Inacio, P. R. M., Magoni, D., & Freire, M. M. (2022). Detection and mitigation of low-rate denial-of-service attacks: A survey. *IEEE Access*, *10*, 76648–76668. https://doi.org/10.1109/access.2022.3191430
9. Velepucha, V., & Flores, P. (2023). A survey on microservices architecture: Principles, patterns and migration challenges. *IEEE Access*, *11*, 88339–88358. https://doi.org/10.1109/access.2023.3305687
10. Vasireddy, I., Ramya, G., & Kandi, P. (2023). Kubernetes and docker load balancing: State-of-the-art techniques and challenges. *International Journal of Innovative Research in Engineering and Management*, *10*(6), 49–54. https://doi.org/10.55524/ijirem.2023.10.6.7
11. Deressu, T. F., Beyene, A. M., & Gemechu, A. Y. (2026). A survey of application layer distributed denial of service detection: Approaches, models, and datasets. *Artificial Intelligence Review*, *59*(6), Article 139. https://doi.org/10.1007/s10462-026-11528-3
12. Tian, Q., & Miyata, S. (2023). A DDoS attack detection method using conditional entropy based on SDN traffic. *IoT*, *4*(2), 95–111. https://doi.org/10.3390/iot4020006
13. Qasim, S. S., & Nsaif, S. M. (2024). Advancements in time series-based detection systems for distributed denial-of-service (DDoS) attacks: A comprehensive review. *Babylonian Journal of Networking*, *2024*, 9–17. https://doi.org/10.58496/bjn/2024/002
14. Savchenko, V., Ilin, O., Hnidenko, N., Tkachenko, O., Laptiev, O., & Lehominova, S. (2020). Detection of slow DDoS attacks based on user's behavior forecasting. *International Journal of Emerging Trends in Engineering Research*, *8*(5), 2019–2025. https://doi.org/10.30534/ijeter/2020/90852020
15. Alashhab, A. A., Zahid, M. S. M., Azim, M. A., Daha, M. Y., Isyaku, B., & Ali, S. (2022). A survey of low rate DDoS detection techniques based on machine learning in software-defined networks. *Symmetry*, *14*(8), Article 1563. https://doi.org/10.3390/sym14081563
16. Ilango, H. S., Ma, M., & Su, R. (2022). A feedforward–convolutional neural network to detect low-rate DoS in IoT. *Engineering Applications of Artificial Intelligence*, *114*, Article 105059. https://doi.org/10.1016/j.engappai.2022.105059
17. Onai, M. V., & Kravchuk, A. A. (2024). Vdoskonalennia ta analiz efektyvnosti metodiv vyiavlennia DDoS-atak na osnovi protokolu HTTP. In *Prykladna matematyka ta kompiutynh PMK' 2024* (pp. 149–153). https://ela.kpi.ua/handle/123456789/71110
18. Leka, E., Lamani, L., Aliti, A., & Hoxha, E. (2024). Web application firewall for detecting and mitigation of based DDoS attacks using machine learning and blockchain. *TEM Journal*, *13*(4), 2802–2811. https://doi.org/10.18421/tem134-17
19. Dawadi, B. R., Adhikari, B., & Srivastava, D. K. (2023). Deep learning technique-enabled web application firewall for the detection of web attacks. *Sensors*, *23*(4), Article 2073. https://doi.org/10.3390/s23042073

20. OWASP ModSecurity Core Rule Set (CRS): Project Documentation. (2026). *OWASP Foundation*. https://coreruleset.org/docs/
21. Alasri, A., & Sulaiman, R. (2018). Protection of XML-based denail-of-service and HTTP flooding attacks in web services using the middleware tool. *International Journal of Engineering and Technology*, *7*(4.7), 322–329. https://doi.org/10.14419/ijet.v7i4.7.20570
22. Berezhna, M. (2025). Separation of concerns: A fundamental principle in software development. *DEV Community*. https://dev.to/marie_berezhna/separation-of-concerns-a-fundamental-principle-in-software-development-oaj
23. Khairuddin, M. A., Mohd Isa, M. R., Mohd Shukran, M. A., Ismail, M. N., & Maskat, K. (2024). Detecting distributed denial-of-service (DDoS) attacks through the log consolidation processing (LCP) framework. *International Journal on Informatics Visualization*, *8*(3), 1064–1073. https://doi.org/10.62527/joiv.8.3.2184
24. Kammer, A., Koppelt, C., Müller, J., Prinz, H., Schmidt, C., Steinke, M., & Wolff, E. (2024). Service mesh overview and comparison. *INNOQ*. https://servicemesh.es/
25. Singh, P., & Ayyasamy, S. (2023). Exploring, analyzing and tuning service mesh performance: A literature review. *TechRxiv*. https://doi.org/10.36227/techrxiv.22776119.v1
26. Priya Patharlagadda, P. (2022). Kubernetes traffic management using Istio. *Journal of Media & Management, 4*(1). https://doi.org/10.47363/jmm/2022(4)e101
27. Architecture overview and what is Envoy – Envoy documentation. (2026). *Envoy proxy*. https://www.envoyproxy.io/docs/envoy/v1.38.2/intro/what_is_envoy
28. Envoy Filter: Traffic management in Istio – Istio documentation. (2026). *Istio*. https://istio.io/latest/docs/reference/config/networking/envoy-filter/
29. Kubernetes documentation: Fundamental concepts. (2026). *The Kubernetes Authors*. https://kubernetes.io/docs/concepts/
30. Vieira, M. A. M., Castanho, M. S., Pacífico, R. D. G., Santos, E. R. S., Júnior, E. P. M. C., & Vieira, L. F. M. (2020). Fast Packet Processing with eBPF and XDP. *ACM Computing Surveys*, *53*(1), Article 16. https://doi.org/10.1145/3371038
31. Kubernetes documentation: Operator pattern. (2025). *The Kubernetes Authors*. https://kubernetes.io/docs/concepts/extend-kubernetes/operator/
32. Kubernetes controller-runtime project – Go Packages. (2026). *The Go Project*. https://pkg.go.dev/sigs.k8s.io/controller-runtime
33. Reitz, K. (2026). PostmanLabs/httpbin: HTTP request & response service, written in python + flask. *GitHub, Inc*. https://github.com/postmanlabs/httpbin
34. Fortio load testing library, command line tool, advanced echo server in go. (2026). *Fortio*. https://fortio.org/
35. SlowHttp – Slow Read DDoS simulator written as a Python script. (2026). *GitHub, Inc*. https://github.com/ojardila/slowhttp
36. Coraza – open source, enterprise-grade, high performance web application firewall. (2026). *Coraza*. https://www.coraza.io/docs/tutorials/introduction/

Received 29.06.2026.
Accepted 31.08.2026.
Published 30.10.2026

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Designing the architecture of a software system for detecting and blocking HTTP LR-DDoS attacks on the Kubernetes platform

This paper proposes and implements a distributed architecture for protecting microservices against application-layer low-rate DDoS (LR-DDoS) attacks in a Kubernetes environment, without modifying the target applications' code and without introducing noticeable latency to legitimate traffic. The architectural solution is based on separating the routing plane (Data Plane) and the analysis plane (Control Plane): application-layer attributes are collected by the EnvoyFilter extension after TLS decryption at the Istio ingress gateway, and network attributes are gathered using eBPF/XDP technologies, after which they are asynchronously forwarded to the Apache Kafka broker. The horizontally scalable classification module, developed in Go, and the controller based on the Kubernetes Operator pattern store DDoS attack sources in Redis with a TTL mechanism to remove stale data and update blocking rules in the AuthorizationPolicy resource. This enables three attack mitigation scenarios: full classification of unknown sessions, accelerated verification against preserved signatures, and request blocking at the gateway with O(1) complexity. Experimental studies have confirmed the system's high performance: the latency overhead from the EnvoyFilter is only 2.933 ms, and the response time is 3.15 s for new sources, 0.8 s for known signatures, and 5.674 ms for gateway-level blocking. The classification module processes 4,000 sessions in 32.5 ms using 3.09 MB of RAM. Compared to the Coraza Web Application Firewall (WAF), the proposed system demonstrates 12.3 times lower latency (3.667 ms versus 45 ms) and significantly higher throughput (1,973 versus 473 requests/s).

Keywords: DDoS, low-rate DDoS attacks, LR-DDoS, software system, software engineering, microservice architecture, Kubernetes, Istio, information security, cryptography.

Кравчук Аркадій Андрійович – аспірант кафедри програмного забезпечення комп'ютерних систем, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ, Україна.
ORCID: https://orcid.org/0000-0002-6128-206X
Онай Микола Володимирович – к.т.н., доцент, доцент кафедри програмного забезпечення комп'ютерних систем, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ, Україна.
ORCID: https://orcid.org/0000-0002-4938-8355

Kravchuk Arkadii – PhD Student at the Department of Computer System Software, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine.
ORCID: https://orcid.org/0000-0002-6128-206X
Onai Mykola – PhD in Computer Systems and Components, Associate Professor, Associate Professor at the Department of Computer System Software, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine.
ORCID: https://orcid.org/0000-0002-4938-8355

DOI 10.34185/1562-9945-5-166-2026-25
УДК 621.396:004.89

Р.О. Бєляков

ПОРІВНЯЛЬНИЙ АНАЛІЗ ПРОЦЕСІВ ПРОГНОЗНОЇ СИНХРОНІЗАЦІЇ ТА ВИБОРУ ВУЗЛІВ РОЗПОДІЛЕНОЇ АНТЕННОЇ СИСТЕМИ НА ОСНОВІ ВИМІРЯНИХ ХАРАКТЕРИСТИК РАДІОКАНАЛІВ

Анотація. У статті проведено порівняльний аналіз процесів прогнозної часової, частотної та фазової синхронізації і адаптивного вибору вузлів розподіленої антенної системи з одноранговою координацією. Розглядається система з 32 радіовузлів, які формують когерентний інформаційний сигнал для рухомого приймача в умовах втрати актуальності оцінок синхронізації через рух радіовузлів і приймача, нестабільність опорних генераторів радіовузлів і затримки службового обміну. Для проведення дослідження використано виміряні характеристики радіоканалів DICHASUS для 32 антенних елементів на частоті 1,272 ГГц, траєкторії руху БпЛА із набору DeepSense 6G та модель частотно-фазової нестабільності опорних генераторів радіовузлів, параметри якої визначаються на основі девіації Аллана.

Для порівняння методів сформовано математичну модель, яка дозволяє прогнозувати фазовий стан кожного радіовузла, визначати ймовірність виконання фазового допуску та розраховувати очікувану прийняту потужність з урахуванням знака інтерференційного внеску кожної пари радіовузлів. Отримані значення використовуються для вибору складу кластера, застосованого для когерентного формування інформаційного сигналу в точці приймання. За результатами проведеної оцінки ефективності встановлено, що застосування прогнозної синхронізації дозволяє збільшити прийняту потужність інформаційного сигналу на 0,88–9,43 дБ порівняно з реактивною синхронізацією. Вибір складу радіовузлів на основі очікуваної прийнятої потужності дозволяє зберегти рівень потужності, отриманий для повного складу системи, при зменшенні кількості активних радіовузлів на 17,5–30,4 %. На основі проведеного аналізу встановлено, що ентропійний показник не забезпечує окремого приросту прийнятої потужності, а умова надійності на основі межі об'єднання подій у динамічних сценаріях надмірно скорочує склад кластера.

Ключові слова: розподілена антенна система, когерентна передача даних, прогнозна синхронізація, адаптивний вибір вузлів, порівняльний аналіз, фільтр Калмана, консенсусне узгодження, БпЛА, DICHASUS, DeepSense 6G.

Постановка наукового завдання. Мобільні самоорганізовані радіомережі, зокрема мережі з повітряними вузлами, функціонують без постійної базової станції. Однорангова взаємодія визначає спосіб координації радіовузлів, а багатозв'язна топологія

© Бєляков Р.О., 2026
Copyright for this paper by its authors. Use permitted under License CC BY 4.0.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

забезпечує резервування маршрутів. Для когерентної обробки однакового інформаційного сигналу від просторово рознесених радіовузлів необхідно узгодити моменти прибуття сигналів, несучі частоти та фази в точці приймання. В процесі когерентної передачі інформаційного сигналу формується кластер, який складається із радіовузлів, застосованих для формування сигналу в точці приймання. Сукупність узгоджених радіовузлів утворює розподілену антенну систему (РАС), виграш якої залежить від характеристик радіоканалів, складу кластера та обмеження на сумарну потужність [1–3].

Практична реалізація РАС ускладнюється незалежним дрейфом опорних генераторів радіовузлів, похибками визначення положення, часовою мінливістю радіоканалів і затримками службового обміну. Взаємний рух передавачів і приймача спричиняє доплерівський зсув частоти та безперервне накопичення фазової неузгодженості. За таких умов параметри синхронізації, визначені за поточними вимірюваннями, втрачають актуальність до моменту початку когерентної передачі даних. Зазначене обумовлює необхідність дослідження прогнозної синхронізації та вибору вузлів за прогнозованою фазовою сумісністю.

Наукове протиріччя полягає у необхідності підтримувати когерентне формування інформаційного сигналу великої кількості радіовузлів в умовах, коли частота оновлення оцінок синхронізації обмежена ресурсом службового обміну, втрата актуальності зазначених оцінок зростає з динамікою радіовузлів. Збільшення частоти обміну пілотними сигналами зменшує втрату актуальності оцінок, однак збільшує службовий трафік, енергоспоживання та затримки, що суперечить вимогам до автономних радіовузлів [1, 8].

Окремо необхідно визначити внесок прогнозу фазового стану, вибору складу кластера, ентропійного показника та умови надійності за однакових ресурсних обмежень. У відомих підходах результати зазначених процедур переважно розглядаються комплексно, що не дозволяє кількісно визначити вплив кожної з них на прийняту потужність інформаційного сигналу та кількість активних радіовузлів.

З урахуванням наведеного актуальним є наукове завдання, що полягає у порівняльному аналізі процесів прогнозної синхронізації та вибору передавального кластера з кількісною оцінкою внеску кожного процесу на основі виміряних характеристик радіоканалів і траєкторій руху БпЛА.

Аналіз останніх досліджень і публікацій. У роботах [1, 2] систематизовано задачі розподіленого формування променя та наведено вимоги до точності взаємної часової, частотної та фазової синхронізації елементів розподілених фазованих решіток. Результати зазначених досліджень визначають математичну основу когерентної передачі сигналів розподіленими антенними системами. Водночас задача вибору складу кластера за прогнозованою фазовою сумісністю радіовузлів у них окремо не розглядається.

У роботі [3] представлено метод бездротової координації шестиелементної решітки з децентралізованою часовою синхронізацією, застосування якого на частоті 1,05 ГГц забезпечило близько 98 % ідеального когерентного виграшу. В дослідженні [4] наведено метод визначення геометрії решітки за неповними міжвузловими вимірюваннями. У науковій роботі [5] розглянуто формування променя без зво-

ротного зв'язку від кінцевого приймача. Виграш зазначених методів досягається за рахунок підвищення точності окремої підсистеми за заданого складу передавальних радіовузлів, що не дозволяє оцінити вплив вибору радіовузлів на прийняту потужність.

В дослідженні [6] розглянуто інтегроване зондування та зв'язок за асинхронного руху пристроїв, що дозволяє визначати доплерівські параметри радіоканалів. У роботі [7] наведено результати взаємного калібрування та відстеження фазових коефіцієнтів у динамічних сценаріях. У науковій роботі [8] запропоновано метод бездротової фазово-частотної синхронізації розподілених точок доступу BeamSync, застосування якого не потребує кабельного розподілу опорного сигналу. В дослідженні [9] наведено метод частотно-фазової синхронізації розподілених решіток на основі розподіленого фільтра Калмана.

Зазначені методи не обмежуються корекцією за останнім вимірюванням. Однак ресурсні обмеження службового обміну та вибір радіовузлів до процесу синхронізації в них не включено, тому внесок кожного процесу в прийняту потужність окремо не визначається.

У роботі [10] представлено модель інтегрованого мережевого комплексу зв'язку з БпЛА, у якій вибір ретранслятора здійснюється за мережевими показниками, що обмежує її використання для формування когерентного кластера. В дослідженні [11] обґрунтовано переваги безкомірковой архітектури з розподіленими точками доступу над архітектурою малих комірок за рівномірністю обслуговування користувачів. Методи консенсусного узгодження оцінок у мережах з одноранговою взаємодією наведено в роботах [12, 13]. Для опису стабільності опорних генераторів використовується девіація Аллана [14]. У науковій роботі [15] розкрито роль сферичного фронту хвилі у формуванні фокуса скінченної глибини.

Набір даних розподіленої системи вимірювання радіоканалів DICHASUS разом із процедурою бездротової синхронізації її приймачів представлено в роботах [16, 17]. В дослідженні [18] наведено набір DeepSense 6G, який включає синхронізовані мультимодальні вимірювання, зокрема траєкторії БпЛА сценарію 23 [19].

Проведений аналіз показав, що існуючі методи забезпечують когерентне складання сигналів за заданого складу передавальних радіовузлів, а ефект прогнозування, оцінки невизначеності та вибору радіовузлів демонструється сукупно. Водночас питання узгодженого вибору кластера та прогнозних коригувальних впливів із кількісною оцінкою внеску кожного процесу залишається недостатньо дослідженим. Із урахуванням вищенаведеного, у статті проводиться порівняльний аналіз зазначених процесів.

Мета дослідження. Метою статті є порівняльний аналіз процесів прогнозної синхронізації та адаптивного вибору вузлів розподіленої антенної системи на основі виміряних характеристик радіоканалів під час зміни положення БпЛА.

Для досягнення поставленої мети необхідно: узгодити єдину модель сигналу, фазового стану та прийнятої потужності для порівнюваних способів; побудувати відтворюваний експеримент на основі виміряних характеристик радіоканалів і траєкторій ру-

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

ху БпЛА; розділити внески прогнозування, вибору вузлів, ентропійної складової, умови надійності та адаптивного інтервалу синхронізації.

Викладення основного матеріалу дослідження. В статті на першому етапі сформовано математичну модель інформаційного сигналу та фазового стану радіовузлів. На другому етапі визначено очікувану прийняту потужність і критерій вибору складу кластера. На третьому етапі сформовано процедуру прогнозної часової, частотної та фазової синхронізації з подальшою перевіркою готовності радіовузлів до когерентної передачі інформаційного сигналу. На четвертому етапі проведено оцінку ефективності методів із використанням виміряних характеристик радіоканалів DICHASUS, траєкторій руху БпЛА DeepSense 6G та моделі частотно-фазової нестабільності опорних генераторів радіовузлів.

Модель системи та сигналу. Система складається з множини $V = \{v_1, \dots, v_N\}$ автономних мобільних радіовузлів, кожен з яких має власний опорний генератор і засоби оцінки параметрів радіоканалу. В процесі когерентної передачі даних формується кластер $C_k \subseteq V$, який складається із радіовузлів, застосованих для формування інформаційного сигналу. Синхронізація зазначених радіовузлів за часом, частотою та фазою забезпечує когерентне складання сигналів у точці приймання. Топологія однорангової взаємодії радіовузлів у момент дискретного часу k представлена у вигляді динамічного графа (1):

$$G_k = (V, E_k, W_k), \tag{1}$$

де E_k – множина доступних однорангових з'єднань; $W_k = [w_{ij}(k)]$ – матриця ваг, що характеризує придатність радіоканалу між v_i та v_j до службового обміну. Ребро вважається активним за умови $w_{ij}(k) \geq \varepsilon_w$, де ε_w – мінімально допустимий рівень якості з'єднання.

В роботі використовується еквівалентна комплексна низькочастотна модель відносно несучої f_c. Нехай $s(t)$ має одиничну середню потужність, P_i^{tx} – потужність передавача i-го радіовузла, $\beta_i(t)$ – комплексний коефіцієнт радіоканалу без геометричної фази, $\Delta\tau_i^{arr}$ – неузгодженість моментів прибуття огинаючих після часової корекції. Прийнятий інформаційний сигнал представляється у вигляді (2):

$$y(t) = \sum_{i \in C} \sqrt{P_i^{tx}} |\beta_i(t)| s(t - \Delta\tau_i^{arr}) e^{j(\psi_i^0(t) + u_i^{\varphi})} + n(t), \tag{2}$$

Повна фаза сигналу радіовузла до етапу застосування фазової корекції визначається виразом (3):

$$\psi_i^0(t) = \arg\beta_i(t) + \varphi_i^0 + 2\pi \int_{t_0}^{t} \left(\Delta f_i(u) + u_i^f(u)\right) du - 2\pi f_c \tau_i^{prop}(t). \tag{3}$$

де $\arg\beta_i$ – аргумент коефіцієнта радіоканалу; φ_i^0 – початкова фаза опорного генератора; Δf_i – його частотне відхилення; u_i^f – частотна корекція, що враховується від мо-

менту фактичного застосування; τ_i^{prop} – затримка поширення. Після етапу застосування фазової корекції похибка фази радіовузла відносно опорної фази та індивідуальний фазовий допуск визначаються системою (4):

$$\varepsilon_i(t) = \mathrm{wrap}_\pi\left(\psi_i{}^0(t) + u_i{}^\varphi - \psi_{ref}\right),\ |\varepsilon_i(t)| \le \varepsilon_\varphi,\ 0 < \varepsilon_\varphi < \pi/2, \quad (4)$$

де $\mathrm{wrap}_\pi(x) = ((x+\pi)\mathrm{mod}2\pi) - \pi$; ψ_{ref} – єдина опорна фаза. З (4) випливає парний допуск $\left|\mathrm{wrap}_\pi(\varepsilon_i - \varepsilon_j)\right| \le 2\varepsilon_\varphi$. Однаковий для всіх радіовузлів фазовий зсув не змінює когерентного виграшу, тому в експерименті опорною вважається фаза сумарного сигналу. Узгодженість фаз сигналів радіовузлів у точці приймання оцінюється нормованим коефіцієнтом когерентності (5):

$$\eta_{coh} = \frac{\left|\sum_{i\in C}\sqrt{P_i{}^{rx}}e^{j\varepsilon_i}\right|^2}{\left(\sum_{i\in C}\sqrt{P_i{}^{rx}}\right)^2},\ P_i{}^{rx} = P_i{}^{tx}|\beta_i|^2,\ 0 \le \eta_{coh} \le 1. \quad (5)$$

Коефіцієнт (5) характеризує частку ідеального для вибраного кластера фазового складання і не визначає абсолютної прийнятої потужності, тому кластери різного розміру порівнюються за очікуваною потужністю (19).

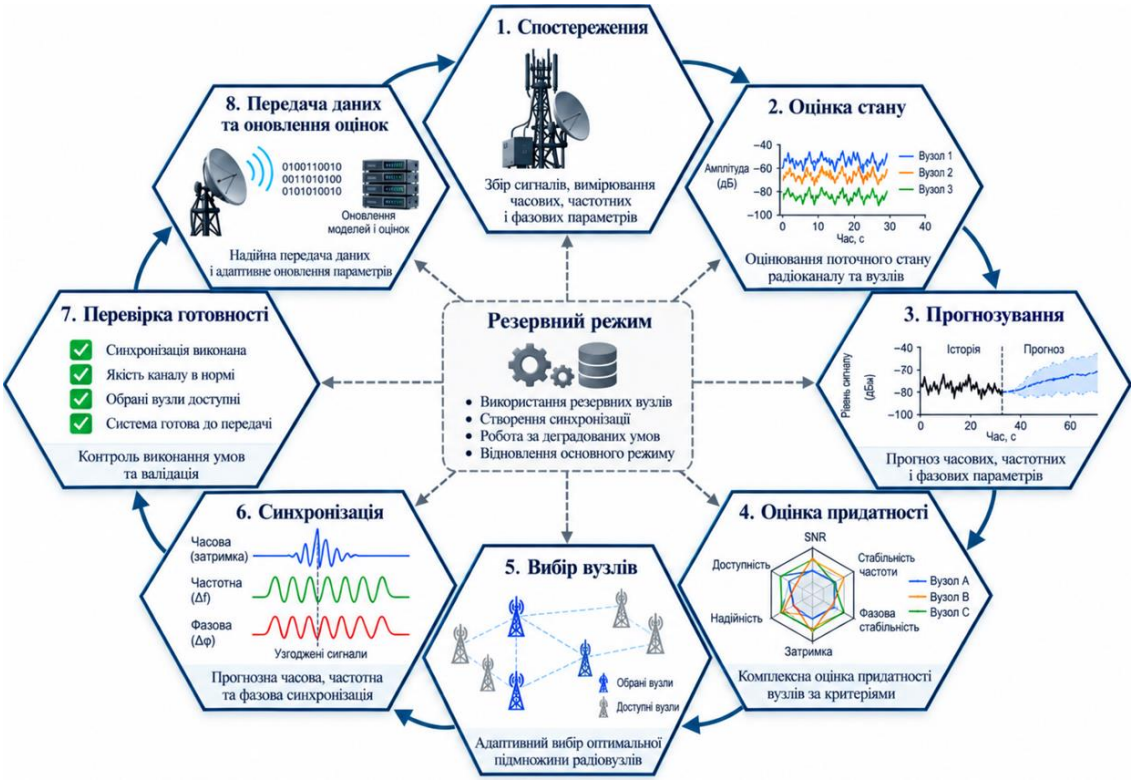


Рисунок 1 – Послідовність етапів циклу координації радіовузлів розподіленої антенної системи

Функціонування системи організовано як замкнений цикл (рис. 1): етап спостереження, оцінка стану, прогнозування, оцінка придатності, вибір кластера, синхронізація, перевірка готовності, передача даних та оновлення оцінок. Перевірка готовності вико-

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

нується після процесу синхронізації та перед початком передачі даних; за невиконання умов готовності система переходить до резервного некогерентного режиму.

Модель ґрунтується на трьох допущеннях. Перше допущення стосується графа та полягає у тому, що в межах одного циклу узгодження граф вимірювань є фіксованим, неорієнтованим і зв'язним, а ваги симетричні та додатні. Друге допущення віднесено каналу передачі даних та полягає у тому, що коефіцієнти радіоканалів між радіовузлами та приймачем доступні з невизначеністю через зворотний зв'язок. Третє допущення стосується фазової моделі та полягає у тому, що використовуються локально розгорнуті фази й вузькосмуговий еквівалент каналу.

Модель прогнозованого стану радіовузла. Для прогнозування параметрів синхронізації стан радіовузла представлено у вигляді вектора (6):

$$\boldsymbol{x}_i(k) = (\boldsymbol{r}_i^{\mathrm{T}}, \boldsymbol{v}_i^{\mathrm{T}}, \delta\tau_i, \delta\varphi_i, \Delta f_i, \nu_i, \gamma_i, e_i, q_i)^{\mathrm{T}}, \tag{6}$$

де $\boldsymbol{r}_i$, $\boldsymbol{v}_i$ – координати та швидкість; $\delta\tau_i$ – зміщення локальної часової шкали; $\delta\varphi_i$ – відносна фаза опорного генератора; ν_i – геометричний доплерівський зсув; γ_i – нормована якість радіоканалу; e_i – енергетичний ресурс; q_i – надійність координаційного зв'язку; $Z_{i,1:k} = \{z_i(1), \dots, z_i(k)\}$ – сукупність вимірювань вузла v_i до моменту k. На горизонті $T_p = lT_s$, де T_s – тривалість циклу координації, прогноз стану визначається виразом (7):

$$\hat{\boldsymbol{x}}_i(k+l|k) = F_i\big(\hat{\boldsymbol{x}}_i(k), Z_{i,1:k}, \boldsymbol{u}_i\big), \tag{7}$$

де F_i – оператор, який визначає модель руху радіовузла, модель опорного генератора та модель вимірювань разом із коваріацією похибки прогнозу. На основі координат і швидкостей руху радіовузла та приймача визначається доплерівський зсув частоти (8), який є похідною геометричної фази з (3):

$$\nu_i = -f_c\,\dot{\tau}_i^{prop} = \frac{f_c}{c}\boldsymbol{u}_{iR}^{\mathrm{T}}(\boldsymbol{v}_i - \boldsymbol{v}_R),\ \boldsymbol{u}_{iR} = \frac{\boldsymbol{r}_R - \boldsymbol{r}_i}{\|\boldsymbol{r}_R - \boldsymbol{r}_i\|}, \tag{8}$$

де c – швидкість поширення електромагнітної хвилі; $\boldsymbol{v}_R$ – швидкість приймача. Після етапу застосування методів компенсації частотного відхилення опорного генератора та доплерівського зсуву визначається частотна похибка радіовузла (9):

$$\delta f_i^{res} = \big(\Delta f_i - \Delta\hat{f}_i\big) + (\nu_i - \hat{\nu}_i). \tag{9}$$

Отримане значення (9) використовується для визначення похибки фази радіовузла на момент прогнозу $t_0 + T_p$, яка описується локальною апроксимацією (10):

$$\varepsilon_i(t_0 + T_p) \approx \varepsilon_i(t_0) + 2\pi T_p \delta f_i^{res} + \pi T_p{}^2 \delta\dot{f}_i^{res} + w_{mod,i}, \tag{10}$$

а її умовна дисперсія визначається виразом (11):

$$\sigma_{\varepsilon i}^2 = \boldsymbol{b}^{\mathrm{T}}\Sigma_i\boldsymbol{b} + \sigma_{mod,i}^2,\ \boldsymbol{b} = \big(1, 2\pi T_p, \pi T_p{}^2\big)^{\mathrm{T}}, \tag{11}$$

де Σ_i – умовна коваріація початкової фазової похибки, похибки компенсації частоти та її похідної, яка містить взаємні коваріації часових, частотних і фазових оцінок; $\sigma^2_{mod,i}$ – дисперсія модельного шуму; $\boldsymbol{b}$ – вектор коефіцієнтів впливу похибок на фазу на горизонті T_p. Оператор F_i реалізується на основі фільтра Калмана з вектором стану $(\varepsilon_i, \delta f_i, \delta \dot{f}_i)$ і точною дискретизацією шумів процесу [20], що забезпечує обчислення (11) без додаткових припущень.

Після етапу застосування прогнозної фазової корекції математичне сподівання похибки фази на момент прогнозу, представленої у вигляді (10), приймається рівним нулю. Із урахуванням зазначеного ймовірність виконання індивідуального фазового допуску за нормальної апроксимації визначається виразом (12):

$$p_{i^{coh}} = \Pr\{|\varepsilon_i(t_0 + T_p)| \le \varepsilon_\varphi | Z_{i,1:k}\} = 2\Phi\left(\frac{\varepsilon_\varphi}{\sigma_{\varepsilon i}}\right) - 1, \quad (12)$$

де Φ – функція розподілу стандартної нормальної величини. Вираз (12) описує процес прогнозування на всьому інтервалі передачі даних, який забезпечується детермінованою оцінкою (39)–(41). Для ранжування радіовузлів використовується евристичний показник невизначеності компенсації доплерівського зсуву (13):

$$D_i = 1 - \exp\left(-\frac{\sigma_{\delta vi} + \alpha_D T_p \sigma_{\delta \dot{v}i}}{v_{ref}}\right), \; D_i \in [0,1), \quad (13)$$

де $\sigma_{\delta vi}$, Гц, і $\sigma_{\delta \dot{v}i}$, Гц/с, – середньоквадратичні відхилення похибки компенсації доплерівського зсуву та її похідної; $\alpha_D \ge 0$; $v_{ref} > 0$ – масштаб у герцах. Значення D_i є оцінкою для ранжування і не є ймовірністю.

Невизначеність прогнозу характеризується розподілом чотирьох взаємно виключних категорій стану $\Omega = \{\Omega_{coh}, \Omega_{corr}, \Omega_{unst}, \Omega_{unav}\}$, ймовірності яких визначаються системою (14):

$$p_{i,coh} = q_i p_{i^{coh}}, \; p_{i,corr} = q_i(p_{i^{(2)}} - p_{i^{coh}}), \; p_{i,unst} = q_i(1 - p_{i^{(2)}}), \; p_{i,unav} = 1 - q_i, \quad (14)$$

де $p_{i^{(2)}}$ – імовірність (12), обчислена для подвоєного допуску $2\varepsilon_\varphi$, у межах якого відхилення усувається додатковим циклом синхронізації; q_i – оцінка доступності радіовузла. Нормована ентропія Шеннона визначається виразом (15):

$$\bar{H}_i = -\frac{1}{\ln M}\sum_{m\in\Omega} p_{i,m} \ln p_{i,m}, \; M = 4, \; 0 \le \bar{H}_i \le 1, \quad (15)$$

Низьке значення $\bar{H}_i$ означає концентрацію прогнозу, однак не гарантує його коректності, тому достовірність імовірностей перевіряється калібруванням (рис. 6).

Критерій придатності та очікувана прийнята потужність. Для ранжування радіовузлів інформаційний стан радіовузла $\Xi_i = (p_{i^{coh}}, \bar{H}_i, D_i, \gamma_i, e_i, q_i)^{\mathsf{T}}$ об'єднується евристичним мультиплікативним показником (16):

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

$$R_i = \left(p_{i^{coh}}\right)^{\omega_p}(1-\bar{H}_i)^{\omega_H}(1-D_i)^{\omega_D}\gamma_{i^{\omega_\gamma}}e_{i^{\omega_e}}q_{i^{\omega_q}},\ \sum\omega=1,\ \omega\geq 0, \quad (16)$$

що є зваженим геометричним середнім і належить до $[0,1]$, також необхідно зазначити, що складова з нульовою вагою в добутку не враховується. Показник (16) не є ймовірністю, а кореляція $p_{i^{coh}}$, $\bar{H}_i$ і D_i може спричиняти повторне врахування одного чинника, що перевіряється експериментально.

Фізична оцінка потужності враховує знаковий парний інтерференційний внесок. На основі багатовимірної нормальної апроксимації розгорнутих фаз радіовузлів коефіцієнт парної сумісності визначається виразом (17):

$$C_{ij} = \mathbb{E}\left[\cos\left(\varepsilon_i-\varepsilon_j\right)|Z\right] = e^{-\sigma_{ij}^2/2}\cos\mu_{ij},\ \sigma_{ij}^2 = \sigma_{\varepsilon i}^2+\sigma_{\varepsilon j}^2-2\mathrm{cov}\left(\varepsilon_i,\varepsilon_j\right), \quad (17)$$

де μ_{ij} – різниця умовних середніх фаз. Від'ємне значення C_{ij} означає деструктивне складання і зберігається у фізичному прогнозі потужності. Коваріація похибок консенсусного узгодження визначається виразом (31). Склад кластера задається бінарним вектором $\boldsymbol{z}$, де $z_i = 1$ для $v_i \in C_k$ і $z_i = 0$ в іншому разі. За фіксованої сумарної потужності P_{tot} використовується розподіл пропорційно до оціненої потужності радіоканалу (18), який максимізує когерентну потужність за узгоджених фаз:

$$P_{i^{tx}}(\boldsymbol{z}) = P_{tot}\frac{z_i|\hat{\beta}_i|^2}{\sum_j z_j|\hat{\beta}_j|^2}. \quad (18)$$

Тоді очікувана прийнята потужність кластера визначається виразом (19):

$$G(\boldsymbol{z}) = \sum_i\sum_j z_i z_j\sqrt{P_{i^{tx}}P_{j^{tx}}}|\beta_i||\beta_j|C_{ij} = P_{tot}\frac{\sum_i\sum_j z_i z_j b_i b_j C_{ij}}{\sum_i z_i b_i},\ b_i = |\beta_i|^2. \quad (19)$$

Величина (19) має одиницю потужності, тому для порівняння кластерів різного розміру використовується нормування на потужність сигналу в точці приймання за ідеальної когерентної передачі даних усіма радіовузлами (20):

$$\tilde{G}(\boldsymbol{z}) = \frac{G(\boldsymbol{z})}{G_{ref}},\ G_{ref} = P_{tot}\sum_{i\in V} b_i. \quad (20)$$

Задача вибору кластера. Цільова функція поєднує фізичну потужність, придатність і витрати координації, що наведено в (21):

$$J(\boldsymbol{z}) = \lambda_G\tilde{G}(\boldsymbol{z}) + \frac{\lambda_R}{N}\sum_i z_i R_i - \frac{\lambda_O}{N}\sum_i z_i,\ \boldsymbol{z}^* \in \mathrm{argmax}_{\boldsymbol{z}\in F}J(\boldsymbol{z}), \quad (21)$$

де ваги λ невід'ємні. Множина допустимих розв'язків F визначається системою обмежень (22):

$$N_{min}\leq\sum_i z_i\leq N_{max},\ z_i(e_i-e_{min})\geq 0,\ \sum_i z_i B_{i^{syn}}\leq B^{syn},\ |u_{if}|\leq u_{max}^f,\ G_C\text{зв'язний}, \quad (22)$$

де $B_{i^{syn}}$ і B^{syn} вимірюються в бітах за цикл. Вимога надійності кластера за потреби задається нерівністю (23):

$$\Pr\{\max_{i:z_i=1}|\varepsilon_i| \le \varepsilon_\varphi\} \ge P_{cl}^{min}, \quad (23)$$

яка за відсутності повної статистики сукупного розподілу похибок фази замінюється консервативною умовою на основі межі об'єднання подій (24):

$$\sum_i z_i\left(1 - p_{i^{coh}}\right) \le 1 - P_{cl}^{min}. \quad (24)$$

Умова (24) не потребує припущення про незалежність станів радіовузлів і використовується як умова вибору, так і умова готовності. Повний перебір має складність 2^N, тому застосовується послідовна евристика з точним перерахунком приросту (25):

$$\Delta J_i = J(\boldsymbol{z} \cup \{i\}) - J(\boldsymbol{z}),\ i: z_i = 0,\ \exists j: z_j = 1, w_{ij} \ge \varepsilon_w. \quad (25)$$

Процедура починається з вузла з максимальним значенням R_i і нарощує зв'язний початковий кластер до N_{min} вузлів. Далі додається суміжний із кластером кандидат із максимальним додатним приростом (25), якщо виконуються (22) та, за потреби, (24), причому вимога суміжності зберігає зв'язність G_C на кожному кроці. Процедура завершується за відсутності допустимого покращення, тобто якщо допустимого кластера не знайдено, передачу даних виконує один вузол з максимальним добутком $b_i p_{i^{coh}}$ у некогерентному резервному режимі. Для запобігання частим перемиканням складу кластера за незначних коливань показників застосовується гістерезис (26):

$$R_i \ge R_{on} - \text{включення},\ R_i \ge R_{off} - \text{збереження},\ R_{off} < R_{on}. \quad (26)$$

Прогнозна синхронізація вибраного кластера. Локальний синхронізаційний стан радіовузла та попарне вимірювання за двобічного обміну представлені у вигляді системи (27):

$$\boldsymbol{s}_i = (\delta\tau_i, \Delta f_i, \delta\varphi_i)^{\mathrm{T}},\ \boldsymbol{y}_{ij} = \boldsymbol{s}_j - \boldsymbol{s}_i + \boldsymbol{n}_{ij}. \quad (27)$$

Попарні вимірювання передбачають попередню оцінку затримки поширення та канальної фази між вузлами; для зворотного напрямку використовується протилежний знак. Вага ребра визначається оберненою дисперсією вимірювання (28):

$$a_{ij} = \frac{1}{\sigma_{ij,meas}^2},\ a_{ij} = a_{ji}. \quad (28)$$

Мультиплікативна вага, що поєднує ймовірнісні, ентропійні та ресурсні ознаки, зміщує оцінку мінімуму зваженої суми квадратів нев'язок, тоді як обернена дисперсія (28) забезпечує мінімальну дисперсію зазначеної оцінки. Задача узгодження з фіксованою опорою $\boldsymbol{s}_r = 0$ записується виразом (29):

$$\hat{S} = \operatorname{argmin}_{S:\boldsymbol{s}_r=0} \sum_{(i,j)\in E_C} a_{ij} \|\boldsymbol{y}_{ij} - (\boldsymbol{s}_j - \boldsymbol{s}_i)\|_{Q^2}, \quad (29)$$

де $\|\boldsymbol{v}\|_{Q^2} = \boldsymbol{v}^{\mathrm{T}} Q \boldsymbol{v}$ з додатно визначеною діагональною матрицею масштабування. Зважена матриця Лапласа визначається виразом (30):

$$L_W = BWB^{\mathrm{T}},\ W = \operatorname{diag}(a_{ij}), \quad (30)$$

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

де B – матриця інцидентності графа кластера. Для зв'язного графа з додатними вагами задача (29) має єдиний розв'язок, коваріація його похибки визначається виразом (31):

$$\mathrm{cov}(\hat{\boldsymbol{s}}) = \left(L_{W^{(r)}}\right)^{-1}, \tag{31}$$

де $L_{W^{(r)}}$ – матриця L_W без рядка та стовпця опорного вузла. Недіагональні елементи (31) задають $\mathrm{cov}(\varepsilon_i, \varepsilon_j)$ у (17). Децентралізована реалізація використовує ітераційне оновлення (32) для всіх вузлів, крім опорного:

$$\boldsymbol{s}_{i^{(n+1)}} = \boldsymbol{s}_{i^{(n)}} + \mu \sum_{j \in N_i} a_{ij}\left(\boldsymbol{s}_{j^{(n)}} - \boldsymbol{s}_{i^{(n)}} - \boldsymbol{y}_{ij}\right), \tag{32}$$

яке є градієнтним спуском для (29). Для фіксованого графа без затримок достатньою є умова (33):

$$0 < \mu < \frac{2}{\lambda_{max}(L_W)}, \tag{33}$$

де λ_{max} – максимальне власне значення матриці Лапласа. Для кругових фаз використовується нелінійна ітерація (34), для якої умова (33) не доводить глобальної збіжності:

$$\varphi_{i^{(n+1)}} = \mathrm{wrap}_\pi\left(\varphi_{i^{(n)}} + \mu_\varphi \sum_{j \in N_i} a_{ij}\mathrm{wrap}_\pi\left(\varphi_{j^{(n)}} - \varphi_{i^{(n)}} - y_{ij}^{\varphi}\right)\right). \tag{34}$$

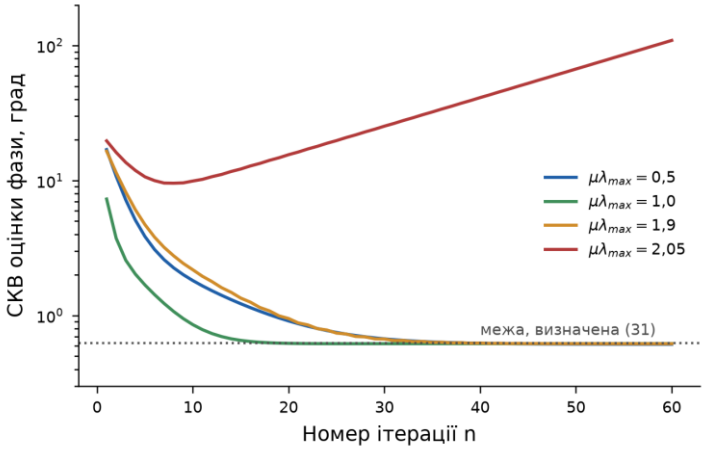


Рисунок 2 – Залежність СКВ оцінки синхронізаційного стану
від кількості ітерацій (32) для різних значень кроку,
граф 32 антенних елементів системи DICHASUS, 200 реалізацій

Перевірку умови (33) проведено для графа, сформованого відповідно до просторового розташування 32 антенних елементів вимірювальної системи DICHASUS. Між двома радіовузлами формувалося ребро за умови, що відстань між ними не перевищувала 12 м. За визначеної конфігурації отримано 304 міжвузлові з'єднання. Для $\mu\lambda_{max} = 1$ середньоквадратичне відхилення оцінки після 20 ітерацій наближалося до 0,63°, що відповідає значенню, визначеному на основі коваріаційної матриці (31), а для $\mu\lambda_{max} = 2{,}05$ отримані значення СКВ суттєво відрізнялися від результатів для попереднього значення кроку (рис. 2), що підтверджує залежність процесу узгодження від вибору параметра ітераційного алгоритму.

Отримані результати підтверджують залежність процесу узгодження від вибору кроку ітераційного алгоритму відповідно до умови (33).1

Коригувальні впливи та перевірка готовності. Нехай t_a – узгоджений момент прибуття сигналів у шкалі опорного вузла, а локальна шкала має модель $t_i = t + \delta\tau_i$. Локальний момент початку передачі сигналу та частотна корекція визначаються системою (35):

$$t_i{}^{tx} = t_a + \delta\hat{\tau}_i - \hat{\tau}_i{}^{prop},\ u_{if} = -\left(\Delta\hat{f}_i + \hat{v}_i\right). \tag{35}$$

Вузли, розташовані далі від приймача, починають процес передачі сигналу раніше. Частоти опорних генераторів оцінюються відносно єдиної опорної частоти з нульовим зміщенням. Нехай t_0 – момент останньої оцінки, $T_p = t_a - t_0$. Прогноз фази та фазовий коригувальний вплив визначаються системою (36):

$$\hat{\psi}_i{}^0(t_a) \approx \psi_i{}^0(t_0) + 2\pi T_p g_i + \pi T_p{}^2 a_{if},\ u_{i\varphi} = \mathrm{wrap}_\pi\left(\psi_{ref} - \hat{\psi}_i{}^0(t_a)\right), \tag{36}$$

де $g_i = \Delta\hat{f}_i + \hat{v}_i$ – оцінка сумарного частотного відхилення, Гц; a_{if} – швидкість його зміни, Гц/с. Геометрична складова вже врахована у виразі (3) і повторно не додається. Готовність фаз перевіряється умовою (24), а готовність огинаючих – нерівністю (37) з індивідуальним ризиком $\alpha_{\tau i}$:

$$|\Delta\hat{\tau}_i{}^{arr}| + \Phi^{-1}(1 - \alpha_{\tau i}/2)\sigma_{\Delta\tau i} \le \varepsilon_\tau{}^{arr},\ i \in C. \tag{37}$$

Інтервал синхронізації встановлюється адаптивно відповідно до системи (38):

$$T_i{}^{syn} = T_{min}^{syn} + \left(T_{max}^{syn} - T_{min}^{syn}\right)p_i{}^{coh}(1 - \bar{H}_i)(1 - D_i),\ T_C{}^{syn} = \min_{i\in C} T_i{}^{syn}. \tag{38}$$

На основі виразу (38) тривалість інтервалу до наступного циклу синхронізації визначається за найменшим допустимим значенням для радіовузлів кластера. Водночас зазначена умова визначає лише момент наступного оновлення параметрів синхронізації та не дозволяє встановити, чи залишатиметься похибка фази в допустимих межах протягом усього інтервалу. Для перевірки виконання фазового допуску на всьому інтервалі використовується детермінована оцінка зміни фазової похибки.

Твердження 1. Нехай на інтервалі $[0, T_p]$ після етапу застосування часової, частотної та фазової корекції похибки радіовузла мають детерміновані межі (39):

$$|\delta f_i{}^{res}| \le \varepsilon_f,\ |\delta\dot{f}_i{}^{res}| \le \varepsilon_{\dot{f}},\ |\delta\tau_i{}^{res}| \le \varepsilon_\tau,\ |\delta\varphi_i{}^{res}| \le \varepsilon_{\varphi 0}. \tag{39}$$

Тоді для всіх $t \in [0, T_p]$ фазова похибка обмежена величиною (40), а індивідуальний допуск виконується на всьому інтервалі, якщо $\varepsilon_{res}^{max} \le \varepsilon_\varphi$:

$$\varepsilon_{res}^{max} = \varepsilon_{\varphi 0} + 2\pi f_c\varepsilon_\tau + 2\pi T_p\varepsilon_f + \pi T_p{}^2\varepsilon_{\dot{f}}. \tag{40}$$

Оцінка (40) виникає в результаті операції інтегрування обмеженої похибки компенсації частоти та нерівності трикутника. Нехай $A = \varepsilon_\varphi - \varepsilon_{\varphi 0} - 2\pi f_c\varepsilon_\tau$. За $A \ge 0$ і

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

$\varepsilon_{\dot{f}} > 0$ допустимий горизонт визначається невід'ємним коренем квадратного рівняння (41):

$$T_{p^{adm}} = \frac{-\varepsilon_f + \sqrt{\varepsilon_{f^2} + \varepsilon_{\dot{f}} A/\pi}}{\varepsilon_{\dot{f}}}, \tag{41}$$

за $\varepsilon_{\dot{f}} = 0$, $\varepsilon_f > 0$ – виразом $T_{p^{adm}} = A/(2\pi\varepsilon_f)$. За $A < 0$ допустима сумарна фазова похибка перевищується вже на початку інтервалу. Для гарантії на пакет інтервал включає проміжок від останнього вимірювання до завершення приймання.

Показники результативності. Відношення сигнал/шум у точці приймання та ймовірність доставки пакета з L бітів за не більш ніж N_{tx} спроб для когерентного приймання BPSK у каналі AWGN визначаються системою (42):

$$\gamma_{out} = \frac{\left|\sum_{i\in C}\sqrt{P_i r_x} e^{j\varepsilon_i}\right|^2}{P_n + P_I}, \; P_b = Q\left(\sqrt{2\gamma_b}\right), \; \gamma_b = \gamma_{out}\frac{B_n}{R_b}, \tag{42}$$

$$P_{del} = 1 - (1 - (1 - P_b)^L)^{N_{tx}}, \tag{43}$$

де B_n – шумова смуга; R_b – бітова швидкість. Рівність (42) стосується AWGN або обґрунтованої гауссової моделі завад; підстановка середнього відношення сигнал/шум у нелінійну формулу (43) загалом некоректна, тому ймовірність обчислюється для кожної реалізації окремо. Середньоквадратична фазова похибка визначається відносно фази сумарного сигналу (44):

$$\mathrm{RMSE}_{\varphi} = \sqrt{\frac{1}{N_C}\sum_{i\in C}\mathrm{wrap}_{\pi^2}\left(\varepsilon_i - \arg\sum_{j\in C}\sqrt{P_j r_x} e^{j\varepsilon_j}\right)}. \tag{44}$$

Визначення (44) не залежить від вибору опорного вузла. Для рівномірно розподілених фаз СКВ (44) не перевищує $\pi/\sqrt{3} \approx 103{,}9°$ і зменшується за малої кількості вузлів через вирівнювання опори за сумарним сигналом.

Експериментальний аналіз. Для перевірки використали три джерела даних. Перше джерело є набір dichasus-cf0x розподіленої системи вимірювання радіоканалів DICHASUS Університету Штутгарта [16, 17]. Вимірювання виконано у виробничому приміщенні дослідницького комплексу ARENA2036: чотири антенні решітки по вісім антен (32 антени) розміщено в кутах Г-подібної ділянки з розсіювачем між ними, що забезпечує наявність каналів як прямої, так і непрямої видимості. Передавач, встановлений на мобільному роботі, рухається ділянкою, приймальні антени реєструють коефіцієнти каналу на несучій 1,272 ГГц у смузі 50 МГц з інтервалом 48,8 мс. Використано файл cf06 тривалістю 674 с (13330 вимірювань), у якому робот рухається зигзагоподібним маршрутом, його положення визначено тахеометром; медіанна швидкість становить 0,26 м/с. За принципом взаємності радіоканалу 32 приймальні антени розглядали як передавальні вузли розподіленої антенної системи, а передавач на роботі – як приймач.

Друге джерело – сценарій 23 набору DeepSense 6G [18, 19]: 11387 відліків 35 польотів БпЛА із тривимірною швидкістю, сумарна тривалість 29 хв, медіанна швидкість

1,71 м/с, максимальна 11,8 м/с. Третє джерело – модель частотно-фазової нестабільності опорних генераторів радіовузлів трьох класів, параметри якої визначаються на основі девіації Аллана [14].

На основі триманих результатів попереднього аналізу встановлено, що виміряні фази системи DICHASUS включають міжприймальну фазову нестабільність: СКВ зміни відносної фази за один інтервал становить 48,8° і не зменшується за нерухомого приймача (65,7° за швидкості до 0,05 м/с). Оскільки зазначена складова не характеризує процес поширення радіосигналу, математичну модель радіоканалу формували таким чином: амплітуди $|\beta_i(t)|$ та положення приймача брали з вимірювань, фазу визначали сферичною хвилею для виміряної геометрії з дифузною складовою просторово узгодженого поля 16 променів.

Коефіцієнт Райса для кожного антенного елемента визначали методом моментів на основі виміряних амплітуд; отримані значення знаходилися в діапазоні 1,4–6,1 дБ. Похибку оцінки фази в діапазоні 1,7–6,2° визначали за різницею парних і непарних піднесучих (рис. 3). Модуль коефіцієнта кореляції відносних фаз за інтервал 48,8 мс становив 0,95 для сформованої моделі радіоканалу та 0,69 для вихідних фазових вимірювань DICHASUS без додаткової корекції.

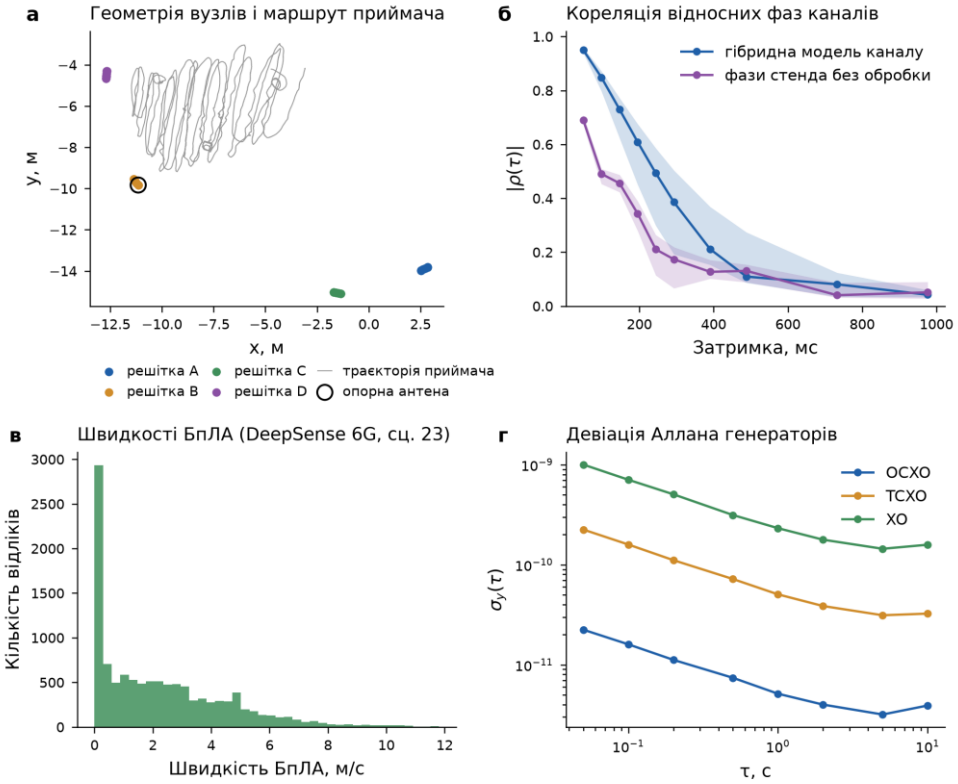


Рисунок 3 – Характеристики вхідних даних експерименту: просторове розташування антенних елементів системи DICHASUS і маршрут приймача (а); залежність модуля коефіцієнта кореляції відносних фаз радіоканалів від часового зсуву, медіана та міжквартильний діапазон за 31 антенним елементом (б); розподіл швидкостей БпЛА DeepSense 6G (в); залежність девіації Аллана від інтервалу усереднення для трьох класів опорних генераторів (г)

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Протокол експерименту показано в таблиці 1. Сценарії відрізнялися кількістю рухомих радіовузлів, складом опорних генераторів, ймовірністю втрати службових повідомлень і горизонтом прогнозу. Для рухомих вузлів використовували профілі швидкості БпЛА з випадковими зсувами початку, а бортову оцінку швидкості моделювали із запізненням 20 мс і похибкою 0,05 м/с. Для кожного сценарію виконували 10 реалізацій по 13 800 циклів; оцінку здійснювали через цикл, перші 10 с кожної реалізації не враховували. Усі способи отримували однакові реалізації каналів, генераторів і втрат повідомлень, тому порівняння є парним.

Таблиця 1

Параметри сценаріїв моделювання

Параметр	**S1**	**S2**	**S3**
Рухомі вузли (траєкторії DeepSense 6G)	0	16	32
Опорні генератори ОСХО / ТСХО / ХО	20 / 10 / 2	12 / 14 / 6	8 / 16 / 8
Імовірність втрати службового повідомлення	0,01–0,10	0,02–0,20	0,05–0,30
Горизонт прогнозу T_p, мс	48,8	48,8	97,7

Для всіх сценаріїв однаковими були: $\sigma_y(1\text{с})$, що дорівнює $5 \cdot 10^{-12}$, $5 \cdot 10^{-11}$ і $2 \cdot 10^{-10}$ для класів ОСХО, ТСХО і ХО; імовірність раптової зміни фази опорного генератора класу ХО $2 \cdot 10^{-3}$ за цикл; $\varepsilon_\varphi = 30°$; $N_{min} = 4$; $L = 1024$ біт; $E_b/N_0 = 10$ дБ для ідеального кластера всіх вузлів; ваги (16) 0,35; 0,15; 0,15; 0,15; 0,10; 0,10; ваги (21) $\lambda_G = 1$, $\lambda_R = 0{,}2$, $\lambda_O = 0{,}02$. Ваги зафіксували до проведення експерименту.

Порівнювали наступні способи: B0 – передача даних без синхронізації; B1 – синхронізація на основі останнього вимірювання; B2 – прогнозна синхронізація на основі фільтра Калмана (7)–(11) з компенсацією доплерівського зсуву за бортовою оцінкою швидкості (35) для всіх вузлів; P – прогнозна синхронізація з вибором вузлів на основі цільової функції (21) з показником придатності (16); P_G – вибір вузлів на основі очікуваної прийнятої потужності (19) без показника придатності ($\lambda_R = 0$); P_R – вибір вузлів з умовою надійності (24) за $P_{cl}^{min} = 0{,}8$ (табл. 2).

Для аналізу впливу окремих складових шляхом їх почергового виключення варіанти P відрізнялися одним елементом: без ентропійної складової ($\omega_H = 0$), з показником придатності на основі ймовірності допуску ($R_i = p_{i^{coh}}$), з показником на основі дисперсії прогнозу ($R_i = \exp(-\sigma_{\varepsilon i}^2)$), з фіксованим після 10 с складом кластера та з адаптивним інтервалом (38). Довірчі інтервали 95 % побудовані на основі методу повторних блокових вибірок [21] за блоками тривалістю 4,9 с і реалізаціями.

Основний виграш забезпечила прогнозна синхронізація B2, яка відносно реактивної синхронізації підвищила прийняту потужність інформаційного сигналу на 0,88 дБ [0,71; 1,10] у S1, на 3,14 дБ [2,93; 3,34] у S2 та на 9,43 дБ [9,18; 9,66] у S3 (табл. 2,

рис. 4). Коефіцієнт когерентності у S3 зріс від 0,134 до 0,730, а ймовірність доставки пакета від 0,014 до 0,861. Отриманий результат обумовлений компенсацією втрати актуальності оцінок синхронізації на основі прогнозу, ефект якої зростає зі збільшенням кількості рухомих радіовузлів.

Таблиця 2

Результати порівняння способів синхронізації та вибору вузлів, 10 реалізацій на сценарій

Сценарій	**Метод**	$\tilde{G}$, дБ [ДІ 95 %]	η_{coh}	RMSE_{φ}, °	P_{out}	P_{del}	N_C
S1	B0	−12,21 [−12,42; −12,00]	0,106	98,1	1,000	0,007	32,0
	B1	−1,10 [−1,29; −0,94]	0,790	31,6	1,000	0,929	32,0
	B2	−0,21 [−0,24; −0,19]	0,959	14,7	0,707	0,989	32,0
	P	−0,20 [−0,23; −0,18]	0,960	14,7	0,707	0,989	31,9
	P_G	−0,21 [−0,23; −0,19]	0,960	13,6	0,602	0,989	26,4
	P_R	−1,17 [−1,32; −1,01]	0,980	9,1	0,162	0,900	17,6
S2	B0	−12,21 [−12,44; −11,97]	0,105	98,1	1,000	0,007	32,0
	B1	−3,65 [−3,86; −3,47]	0,485	63,0	1,000	0,437	32,0
	B2	−0,51 [−0,54; −0,48]	0,901	23,7	0,979	0,972	32,0
	P	−0,50 [−0,53; −0,47]	0,908	23,0	0,976	0,973	31,7
	P_G	−0,51 [−0,54; −0,48]	0,909	20,1	0,878	0,973	24,8
	P_R	−13,89 [−14,47; −13,37]	1,000	0,0	0,000	0,001	1,0
S3	B0	−12,24 [−12,44; −12,00]	0,105	98,2	1,000	0,007	32,0
	B1	−10,93 [−11,17; −10,70]	0,134	94,0	1,000	0,014	32,0
	B2	−1,50 [−1,57; −1,45]	0,730	41,2	1,000	0,861	32,0
	P	−1,49 [−1,56; −1,44]	0,751	39,4	1,000	0,865	30,8
	P_G	−1,51 [−1,58; −1,46]	0,757	34,7	0,997	0,862	22,3
	P_R	−11,26 [−11,80; −10,79]	0,994	2,5	0,018	0,002	1,5

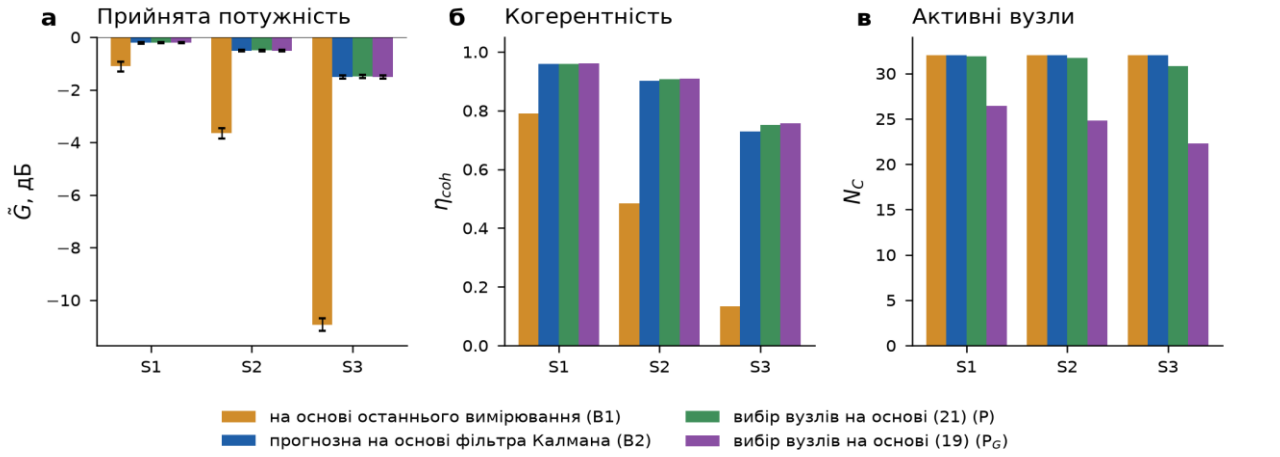


Рисунок 4 – Результати порівняльного аналізу способів синхронізації та вибору радіовузлів за показниками нормованої прийнятої потужності з ДІ 95 % (а), коефіцієнта когерентності (б) та кількості активних радіовузлів (в) для сценаріїв S1–S3

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Вибір вузлів на основі цільової функції (21) з показником придатності (P) змінив потужність відносно B2 на 0,009–0,015 дБ. Зазначена різниця статистично підтверджена, однак становить соті частки децибела і не використовується як самостійний висновок. Варіант P_G зберіг потужність повного кластера: різниця з B2 становить 0,003 дБ [−0,003; 0,011] у S1, −0,002 дБ [−0,008; 0,005] у S2 і −0,008 дБ [−0,016; 0,001] у S3 і статистично не підтверджена.

Водночас $< functionPat0x000001F110188EA0 >_G$ скоротив кількість активних вузлів до 26,4, 24,8 та 22,3 із 32, тобто на 17,5, 22,4 та 30,4 %. До кластера не включалися вузли, внесок яких у (19) не перевищував витрат координації, що зменшує енергоспоживання та службовий обмін за тієї самої прийнятої потужності. Складова λ_R цільової функції (21) призводила до включення таких радіовузлів до кластера, тому у варіанті P кількість активних радіовузлів становила від 30,8 до 31,9.

Таблиця 3

Парні різниці варіантів із почерговим виключенням складових, ДІ 95 %

Порівняння	**S1: $\Delta\vec{G}$, дБ**	**S2: $\Delta\vec{G}$, дБ**	**S3: $\Delta\vec{G}$, дБ**	**S3: ΔN_C**
P – P без ентропійної складової	0,001 [0,000; 0,003]	−0,001 [−0,002; 0,000]	−0,002 [−0,004; 0,000]	−0,37
$R_i = p_{i^{coh}}$ – P	0,000 [−0,003; 0,003]	0,005 [0,000; 0,010]	0,000 [−0,004; 0,004]	0,25
$R_i = e^{-\sigma^2}$ – P	0,000 [−0,004; 0,003]	0,005 [0,000; 0,010]	−0,002 [−0,006; 0,002]	−0,41
фіксований кластер – P	−0,012 [−0,021; −0,005]	−0,043 [−0,103; −0,010]	−0,208 [−0,323; −0,102]	−1,53
адаптивний інтервал – P	−0,034 [−0,041; −0,027]	0,000 [0,000; 0,000]	0,000 [0,000; 0,000]	0,00

Аналіз впливу окремих складових (табл. 3) показав, що виключення ентропійної складової змінило потужність у межах від −0,002 до 0,001 дБ, довірчі інтервали охоплюють нуль або мають ширину тисячних часток децибела. Заміна R_i на $p_{i^{coh}}$ або на $\exp(-\sigma_{\varepsilon i}^2)$ також не змінила результату. Отриманий результат обумовлений надлишковістю ознак: коефіцієнт кореляції $p_{i^{coh}}$ та $\bar{H}_i$ становить від −0,87 до −0,82 (рис. 6, б). Отже, самостійного внеску ентропії не встановлено, тоді як фізичний критерій (19) визначає склад кластера. Фіксований склад кластера знизив потужність на 0,208 дБ [−0,323; −0,102] у S3, що підтверджує потребу в повторному виборі за зміни стану вузлів.

Адаптивний інтервал (38) скоротив кількість циклів синхронізації у S1 на 6,8 % за зниження потужності на 0,034 дБ. У S2 і S3 інтервал не збільшувався, оскільки мінімум (38) за кластером визначається вузлом із мінімальною ймовірністю допуску. Гіпотеза

про зменшення службових витрат адаптивним інтервалом у динамічних сценаріях не підтверджена.

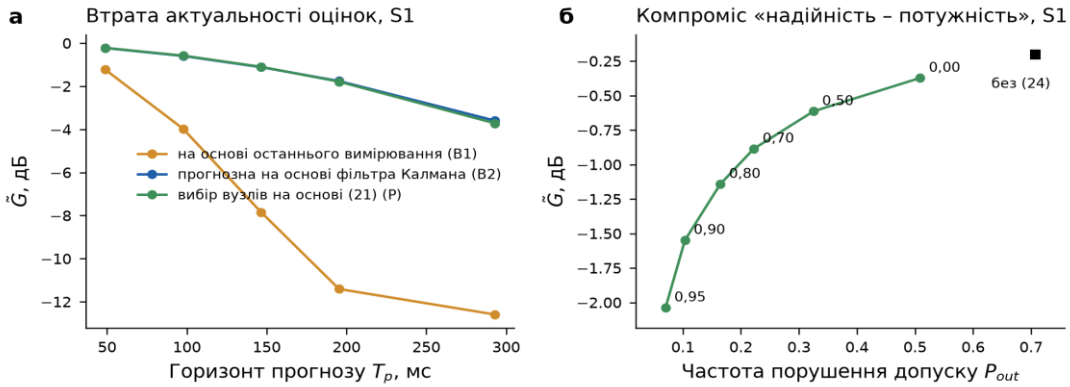


Рисунок 5 – Залежність нормованої прийнятої потужності від горизонту прогнозу у сценарії S1 (а) та залежність нормованої прийнятої потужності від частоти порушення фазового допуску для різних P_{cl}^{min} (б)

Збільшення горизонту прогнозу у S1 від 48,8 до 293 мс знизило потужність B2 від −0,22 до −3,59 дБ, а B1 – від −1,23 до −12,59 дБ (рис. 5, а). Для горизонту 293 мс вибір вузлів знизив потужність відносно B2 на 0,12 дБ, тому за довгих горизонтів вибір за прогнозованою коваріацією потребує попередньої перевірки калібрування.

Умова надійності (24) зменшила частоту порушення допуску, однак знизила прийняту потужність (рис. 5, б): за $P_{cl}^{min} = 0{,}9$ у S1 частота порушення становила 0,104 за потужності −1,55 дБ і 14,3 вузла, тоді як без умови (24) – 0,707 за −0,20 дБ. У S2 і S3 умова (24) за $P_{cl}^{min} = 0{,}8$ залишала допустимими лише поодинокі вузли, і система переходила до резервного режиму. Отже, умова (24) доцільна для квазістаціонарних сценаріїв, а в динамічних сценаріях її слід замінювати вимогою до прийнятої потужності.

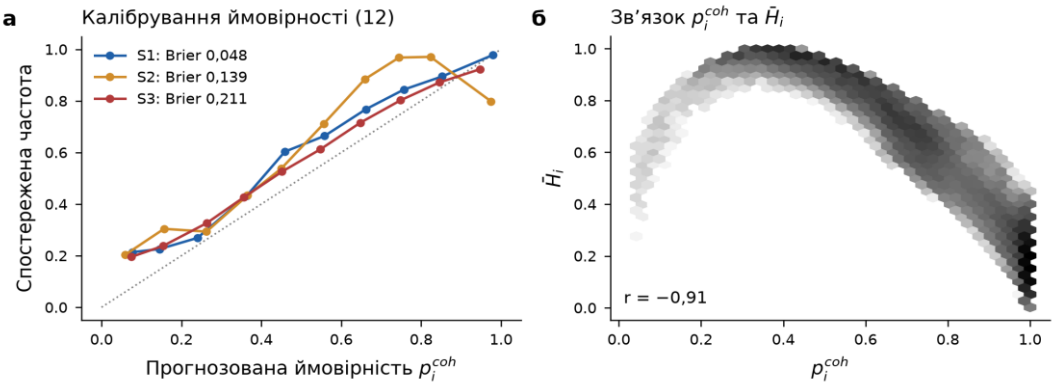


Рисунок 6 – Залежність фактичної частоти виконання фазового допуску від прогнозованої ймовірності (12) для сценаріїв S1–S3 (а) та залежність імовірності допуску від нормованої ентропії, об'єднана вибірка (б)

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Калібрування ймовірності (12) перевіряли на 429,632 прогнозах на сценарій (рис. 6, а). Середня абсолютна похибка калібрування становила 0,018 у S1, 0,194 у S2 і 0,062 у S3, показник Брієра [22] – 0,048, 0,139 та 0,211. Для S2 прогноз недооцінює ймовірність допуску в діапазоні 0,6–0,9; імовірною причиною є неоднорідність вузлів, за якої нерухомі вузли мають меншу фактичну похибку, ніж передбачає єдина для всіх радіовузлів модель шуму процесу. Зазначене пояснення окремо не перевірялося.

Наукова новизна роботи полягає у побудові єдиної узгодженої моделі прогнозованого фазового стану радіовузла, ймовірності виконання фазового допуску та очікуваної прийнятої потужності кластера зі знаковими парними складовими, яка, на відміну від існуючих, дозволяє кількісно розділити внески прогнозної синхронізації, вибору радіовузлів, ентропійної оцінки та умови надійності за однакових ресурсних обмежень.

Висновки. Розроблено математичну модель прогнозної синхронізації та вибору складу кластера розподіленої антенної системи, у якій геометрична фаза радіоканалу враховується один раз, знак парного інтерференційного внеску зберігається під час розрахунку очікуваної прийнятої потужності, а коваріації похибок консенсусного узгодження враховуються під час прогнозу фазового стану радіовузлів. Ваги узгодження визначаються на основі обернених дисперсій вимірювань.

За результатами проведеної оцінки ефективності із використанням виміряних характеристик радіоканалів DICHASUS і траєкторій руху БпЛА DeepSense 6G встановлено, що застосування прогнозної синхронізації дозволяє збільшити прийняту потужність інформаційного сигналу на 0,88, 3,14 та 9,43 дБ порівняно з реактивною синхронізацією для сценаріїв S1–S3. Вибір складу кластера на основі очікуваної прийнятої потужності дозволив зменшити кількість активних радіовузлів на 17,5–30,4 % без суттєвого зменшення потужності, отриманої для повного складу системи. За фіксованого складу кластера у сценарії з найбільшою динамікою прийнята потужність зменшилася на 0,21 дБ, що підтверджує доцільність повторного вибору радіовузлів при зміні їх стану.

Аналіз внеску окремих складових показав, що ентропійний показник не забезпечує окремого приросту прийнятої потужності, оскільки має високий кореляційний зв’язок з імовірністю виконання фазового допуску. Умова надійності на основі межі об’єднання подій дозволяє зменшити частоту порушення фазового допуску у квазістаціонарному сценарії, однак за збільшення динаміки надмірно скорочує кількість радіовузлів, допустимих для формування кластера. Застосування адаптивного інтервалу синхронізації дозволило зменшити кількість службових циклів тільки для сценарію з низькою динамікою.

Подальші дослідження доцільно спрямувати на експериментальну перевірку запропонованої процедури за коротших циклів координації, розробку каліброваної оцінки внеску кожного радіовузла у прийняту потужність інформаційного сигналу та дослідження просторового фокусування для різних конфігурацій розподіленої антенної системи.

ЛІТЕРАТУРА

1. Mudumbai R., Brown D. R., Madhow U., Poor H. V. Distributed transmit beamforming: challenges and recent progress. IEEE Communications Magazine. 2009. Vol. 47, № 2. P. 102–110. DOI: https://doi.org/10.1109/MCOM.2009.4785387

2. Nanzer J. A., Mghabghab S. R., Ellison S. M., Schlegel A. Distributed phased arrays: challenges and recent advances. IEEE Transactions on Microwave Theory and Techniques. 2021. Vol. 69, № 11. P. 4893–4907. DOI: https://doi.org/10.1109/TMTT.2021.3092401

3. Shandi N., Merlo J. M., Nanzer J. A. Distributed beamforming using decentralized time synchronization in a six-element array. arXiv. 2025. arXiv:2505.13248. DOI: https://doi.org/10.48550/arXiv.2505.13248

4. Dula M. J., Shandi N., Nanzer J. A. Decentralized localization of distributed antenna array elements using an evolutionary algorithm. IEEE Transactions on Microwave Theory and Techniques. 2025. Vol. 73, № 10. P. 8013–8023. DOI: https://doi.org/10.1109/TMTT.2025.3564522

5. Nguyen K. D., Lechner G., Ahmad I. Zero-feedback open-loop distributed transmit beamforming. Digital Signal Processing. 2025. Vol. 156. Article 104875. DOI: https://doi.org/10.1016/j.dsp.2024.104875

6. Ventura G., Rossi M., Pegoraro J. AsyMov: integrated sensing and communications with asynchronous moving devices. IEEE Transactions on Wireless Communications. 2026. Vol. 25. P. 11480–11496. DOI: https://doi.org/10.1109/TWC.2026.3653567

7. Jiang Q., Jin J., Wang Q., Tang Y., Cao Y., Kuang B., Dong J., Lv S., Wang D., Huang Y., Wang J., You X. Experimental performance of bidirectional phase coherent transmission and sensing for mmWave cell-free massive MIMO systems with reciprocity calibration. arXiv. 2026. arXiv:2601.14648. DOI: https://doi.org/10.48550/arXiv.2601.14648

8. Ganesan U. K., Sarvendranath R., Larsson E. G. BeamSync: over-the-air synchronization for distributed massive MIMO systems. arXiv. 2023. arXiv:2311.11070. DOI: https://doi.org/10.48550/arXiv.2311.11070

9. Rashid M., Nanzer J. A. High accuracy distributed Kalman filtering for synchronizing frequency and phase in distributed phased arrays. IEEE Signal Processing Letters. 2023. Vol. 30. P. 688–692. DOI: https://doi.org/10.1109/LSP.2023.3283333

10. Сайко В. Г., Зінченко М. О., Комаров В. О., Солодовник В. І., Криволапов Я. В., Криволапов Г. Я. Модель інтегрованого мережевого комплексу зв'язку з використанням БПЛА для підвищення ефективності вирішення цільових завдань спеціального призначення. Вчені записки ТНУ імені В. І. Вернадського. Серія: Технічні науки. 2025. Т. 36 (75), № 2. С. 130–138. DOI: https://doi.org/10.32782/2663-5941/2025.2.1/20

11. Ngo H. Q., Ashikhmin A., Yang H., Larsson E. G., Marzetta T. L. Cell-free massive MIMO versus small cells. IEEE Transactions on Wireless Communications. 2017. Vol. 16, № 3. P. 1834–1850. DOI: https://doi.org/10.1109/TWC.2017.2655515

12. Olfati-Saber R., Fax J. A., Murray R. M. Consensus and cooperation in networked multi-agent systems. Proceedings of the IEEE. 2007. Vol. 95, № 1. P. 215–233. DOI: https://doi.org/10.1109/JPROC.2006.887293

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

13. Xiao L., Boyd S. Fast linear iterations for distributed averaging. Systems & Control Letters. 2004. Vol. 53, № 1. P. 65–78. DOI: https://doi.org/10.1016/j.sysconle.2004.02.022
14. Allan D. W. Statistics of atomic frequency standards. Proceedings of the IEEE. 1966. Vol. 54, № 2. P. 221–230. DOI: https://doi.org/10.1109/PROC.1966.4634
15. Björnson E., Demir Ö. T., Sanguinetti L. A primer on near-field beamforming for arrays and reconfigurable intelligent surfaces. arXiv. 2021. arXiv:2110.06661.
DOI: https://doi.org/10.48550/arXiv.2110.06661
16. Euchner F., Gauger M., Dörner S., ten Brink S. A distributed massive MIMO channel sounder for "Big CSI Data"-driven machine learning. WSA 2021: 25th International ITG Workshop on Smart Antennas. 2021. arXiv:2206.15302.
DOI: https://doi.org/10.48550/arXiv.2206.15302
17. Euchner F., Gauger M. CSI dataset dichasus-cf0x: distributed antenna setup in industrial environment, day 1. DaRUS. 2022. DOI: https://doi.org/10.18419/darus-2854
18. Alkhateeb A., Charan G., Osman T., Hredzak A., Morais J., Demirhan U., Srinivas N. DeepSense 6G: a large-scale real-world multi-modal sensing and communication dataset. IEEE Communications Magazine. 2023. Vol. 61, № 9. P. 122–128.
DOI: https://doi.org/10.1109/MCOM.006.2200730
19. Charan G., Hredzak A., Stoddard C., Berrey B., Seth M., Nunez H., Alkhateeb A. Towards real-world 6G drone communication: position and camera aided beam prediction. GLOBECOM 2022 – IEEE Global Communications Conference. 2022. P. 2951–2956.
DOI: https://doi.org/10.1109/GLOBECOM48099.2022.10000718
20. Bar-Shalom Y., Li X.-R., Kirubarajan T. Estimation with applications to tracking and navigation. New York : Wiley, 2001. 584 p.
21. Künsch H. R. The jackknife and the bootstrap for general stationary observations. The Annals of Statistics. 1989. Vol. 17, № 3. P. 1217–1241.
DOI: https://doi.org/10.1214/aos/1176347265
22. Gneiting T., Raftery A. E. Strictly proper scoring rules, prediction, and estimation. Journal of the American Statistical Association. 2007. Vol. 102, № 477. P. 359–378.
DOI: https://doi.org/10.1198/016214506000001437

REFERENCES

1. Mudumbai, R., Brown, D. R., Madhow, U., & Poor, H. V. (2009). Distributed transmit beamforming: Challenges and recent progress. *IEEE Communications Magazine, 47*(2), 102–110. https://doi.org/10.1109/MCOM.2009.4785387
2. Nanzer, J. A., Mghabghab, S. R., Ellison, S. M., & Schlegel, A. (2021). Distributed phased arrays: Challenges and recent advances. *IEEE Transactions on Microwave Theory and Techniques, 69*(11), 4893–4907. https://doi.org/10.1109/TMTT.2021.3092401
3. Shandi, N., Merlo, J. M., & Nanzer, J. A. (2025). *Distributed beamforming using decentralized time synchronization in a six-element array* [Preprint]. arXiv.
https://doi.org/10.48550/arXiv.2505.13248
4. Dula, M. J., Shandi, N., & Nanzer, J. A. (2025). Decentralized localization of distributed antenna array elements using an evolutionary algorithm. *IEEE Transactions on Microwave Theory and Techniques, 73*(10), 8013–8023. https://doi.org/10.1109/TMTT.2025.3564522

5. Nguyen, K. D., Lechner, G., & Ahmad, I. (2025). Zero-feedback open-loop distributed transmit beamforming. *Digital Signal Processing, 156*, Article 104875. https://doi.org/10.1016/j.dsp.2024.104875
6. Ventura, G., Rossi, M., & Pegoraro, J. (2026). AsyMov: Integrated sensing and communications with asynchronous moving devices. *IEEE Transactions on Wireless Communications, 25*, 11480–11496. https://doi.org/10.1109/TWC.2026.3653567
7. Jiang, Q., Jin, J., Wang, Q., Tang, Y., Cao, Y., Kuang, B., Dong, J., Lv, S., Wang, D., Huang, Y., Wang, J., & You, X. (2026). *Experimental performance of bidirectional phase coherent transmission and sensing for mmWave cell-free massive MIMO systems with reciprocity calibration* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2601.14648
8. Ganesan, U. K., Sarvendranath, R., & Larsson, E. G. (2023). *BeamSync: Over-the-air synchronization for distributed massive MIMO systems* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2311.11070
9. Rashid, M., & Nanzer, J. A. (2023). High accuracy distributed Kalman filtering for synchronizing frequency and phase in distributed phased arrays. *IEEE Signal Processing Letters, 30*, 688–692. https://doi.org/10.1109/LSP.2023.3283333
10. Saiko, V. H., Zinchenko, M. O., Komarov, V. O., Solodovnyk, V. I., Kryvolapov, Ya. V., & Kryvolapov, H. Ya. (2025). Model intehrovanoho merezhevoho kompleksu zviazku z vykorystanniam BPLA dlia pidvyshchennia efektyvnosti vyrishennia tsilovykh zavdan spetsialnoho pryznachennia [Model of an integrated network communication complex using UAVs to increase the efficiency of solving special-purpose tasks]. *Vcheni zapysky TNU imeni V. I. Vernadskoho. Seriia: Tekhnichni nauky, 36*(75)(2), 130–138. https://doi.org/10.32782/2663-5941/2025.2.1/20
11. Ngo, H. Q., Ashikhmin, A., Yang, H., Larsson, E. G., & Marzetta, T. L. (2017). Cell-free massive MIMO versus small cells. *IEEE Transactions on Wireless Communications, 16*(3), 1834–1850. https://doi.org/10.1109/TWC.2017.2655515
12. Olfati-Saber, R., Fax, J. A., & Murray, R. M. (2007). Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE, 95*(1), 215–233. https://doi.org/10.1109/JPROC.2006.887293
13. Xiao, L., & Boyd, S. (2004). Fast linear iterations for distributed averaging. *Systems & Control Letters, 53*(1), 65–78. https://doi.org/10.1016/j.sysconle.2004.02.022
14. Allan, D. W. (1966). Statistics of atomic frequency standards. *Proceedings of the IEEE, 54*(2), 221–230. https://doi.org/10.1109/PROC.1966.4634
15. Björnson, E., Demir, Ö. T., & Sanguinetti, L. (2021). *A primer on near-field beamforming for arrays and reconfigurable intelligent surfaces* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2110.06661
16. Euchner, F., Gauger, M., Dörner, S., & ten Brink, S. (2021). A distributed massive MIMO channel sounder for "Big CSI Data"-driven machine learning. In *WSA 2021: 25th International ITG Workshop on Smart Antennas*. https://doi.org/10.48550/arXiv.2206.15302
17. Euchner, F., & Gauger, M. (2022). *CSI dataset DICHASUS-CF0x: Distributed antenna setup in industrial environment, day 1* [Data set]. DaRUS.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

https://doi.org/10.18419/darus-2854
18. Alkhateeb, A., Charan, G., Osman, T., Hredzak, A., Morais, J., Demirhan, U., & Srinivas, N. (2023). DeepSense 6G: A large-scale real-world multi-modal sensing and communication dataset. *IEEE Communications Magazine, 61*(9), 122–128. https://doi.org/10.1109/MCOM.006.2200730
19. Charan, G., Hredzak, A., Stoddard, C., Berrey, B., Seth, M., Nunez, H., & Alkhateeb, A. (2022). Towards real-world 6G drone communication: Position and camera aided beam prediction. In *2022 IEEE Global Communications Conference (GLOBECOM)* (pp. 2951–2956). IEEE. https://doi.org/10.1109/GLOBECOM48099.2022.10000718
20. Bar-Shalom, Y., Li, X.-R., & Kirubarajan, T. (2001). *Estimation with applications to tracking and navigation*. Wiley.
21. Künsch, H. R. (1989). The jackknife and the bootstrap for general stationary observations. *The Annals of Statistics, 17*(3), 1217–1241. https://doi.org/10.1214/aos/1176347265
22. Gneiting, T., & Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association, 102*(477), 359–378. https://doi.org/10.1198/016214506000001437

Received 30.06.2026.
Accepted 02.09.2026.
Published 30.10.2026

Comparative analysis of predictive synchronization and node selection processes for a distributed antenna system on measured data

Abstract. This paper presents a comparative analysis of predictive time, frequency, and phase synchronization and adaptive node selection in a distributed antenna system with peer-to-peer coordination. The system under consideration comprises 32 radio nodes that coherently form an information-bearing signal for a mobile receiver under conditions where synchronization estimates become outdated due to the motion of both the radio nodes and the receiver, instability of the radio-node reference oscillators, and delays in control information exchange. The study employs measured DICHASUS radio-channel characteristics for 32 antenna elements at a carrier frequency of 1.272 GHz, UAV trajectories from the DeepSense 6G dataset, and a model of frequency and phase instability of the radio-node reference oscillators, whose parameters are determined based on Allan deviation. To compare the considered methods, a mathematical model is developed that enables prediction of the phase state of each radio node, estimation of the probability of satisfying the prescribed phase tolerance, and calculation of the expected received power while accounting for the sign of the pairwise interference contribution between radio nodes. The obtained values are then used to select the cluster of radio nodes involved in coherent formation of the information-bearing signal at the receiver location. The performance evaluation shows that predictive synchronization increases the received information-signal power by 0.88–9.43 dB compared with reactive synchronization. Selecting the radio-node cluster based on the expected received power makes it possible to preserve the power level achieved by the full system while reducing the number of active radio nodes by 17.5–30.4%. The analysis further demonstrates that the entropy-based metric does not provide an additional increase in received power, whereas the reliability con-

straint based on the union bound leads to an excessive reduction in cluster size under dynamic scenarios.

Keywords: distributed antenna system, coherent data transmission, predictive synchronization, adaptive node selection, comparative analysis, Kalman filter, consensus, UAV, DICHASUS, DeepSense 6G.

Бєляков Роберт Олегович – д.т.н., доцент, начальник кафедри радіотехнологій факультету електронних комунікаційних систем, Військовий інститут телекомунікацій та інформатизації імені Героїв Крут, Київ, Україна.
ORCID: https://orcid.org/0000-0001-9882-3088

Bieliakov Robert – doctor of technical sciences, associate professor, head of the department of radio technologies at the faculty of electronic communication systems, Military Institute of Telecommunications and Informatization named after the Heroes of Kruty, Kyiv, Ukraine.
ORCID: https://orcid.org/0000-0001-9882-3088

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

DOI 10.34185/1562-9945-5-166-2026-26
УДК 004.451:004.7

К.Ю. Островська, О.Д. Самуськов

АНАЛІЗ СУЧАСНИХ ПІДХОДІВ ДО МОНІТОРИНГУ ТА УПРАВЛІННЯ НАВАНТАЖЕННЯМ В IT-ІНФРАСТРУКТУРАХ

Анотація. У статті проведено аналіз сучасних підходів до моніторингу та управління навантаженням в IT-інфраструктурах. Досліджено традиційні методи моніторингу ресурсів, концепцію Observability, балансування навантаження, автоматичне масштабування, контейнеризацію та застосування методів машинного навчання для прогнозування навантаження і виявлення аномалій. Проаналізовано особливості реактивного, проактивного та інтелектуального управління ресурсами. Запропоновано узагальнену концептуальну модель інтелектуального управління навантаженням, яка передбачає інтеграцію збору телеметричних даних, їх попередньої обробки, прогнозування навантаження, виявлення аномалій та автоматичного прийняття рішень щодо масштабування ресурсів. Проведено порівняльний аналіз основних підходів за критеріями автоматизації, адаптивності, прогнозних можливостей і вимог до даних. Обґрунтовано доцільність переходу від реактивного до проактивного управління IT-інфраструктурами та визначено перспективи використання технологій машинного навчання і AIOps.
Ключові слова: IT-інфраструктура, моніторинг, управління навантаженням, Observability, хмарні технології, автоматичне масштабування, машинне навчання, прогнозування, виявлення аномалій, AIOps.

Вступ. Сучасний етап цифрової трансформації характеризується активним використанням розподілених інформаційних систем, хмарних обчислень, мікросервісної архітектури та контейнерних технологій. IT-інфраструктури стають дедалі складнішими, а кількість взаємопов’язаних компонентів постійно зростає. До складу сучасної інфраструктури можуть входити фізичні та віртуальні сервери, мережеве обладнання, системи зберігання даних, контейнери, мікросервіси, бази даних, системи балансування та хмарні ресурси.

Зі збільшенням кількості користувачів, інформаційних потоків і обсягів даних зростають вимоги до продуктивності, масштабованості та надійності таких систем. Нерівномірний розподіл навантаження може спричиняти перевантаження окремих компонентів, збільшення часу відповіді, зниження доступності сервісів і виникнення відмов.

Водночас постійне резервування надлишкових ресурсів також є неефективним, оскільки призводить до збільшення витрат на експлуатацію IT-інфраструктури. Особливо актуальною ця проблема є для хмарних середовищ, у яких вартість обчислювальних ресурсів безпосередньо залежить від обсягів їх використання.

© Островська К.Ю., Самуськов О.Д., 2026
Copyright for this paper by its authors. Use permitted under License CC BY 4.0.
ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Традиційні системи моніторингу переважно орієнтовані на контроль поточного стану ресурсів та реагування на вже виявлені проблеми. Такий підхід є ефективним для базового контролю, але має обмежені можливості щодо прогнозування майбутнього навантаження та попередження потенційних відмов.

У зв'язку з цим актуальним є розвиток інтелектуальних підходів, які поєднують моніторинг, аналіз телеметричних даних, прогнозування, виявлення аномалій та автоматичне управління ресурсами.

Постановка проблеми. Основною проблемою управління сучасними IT-інфраструктурами є забезпечення необхідного рівня продуктивності та доступності за одночасного раціонального використання обчислювальних ресурсів.

Навантаження на інфраструктуру не є постійним і може змінюватися залежно від часу доби, дня тижня, кількості користувачів, характеру інформаційних процесів та зовнішніх факторів. У результаті використання статичних правил управління ресурсами може бути недостатньо для забезпечення оптимальної роботи системи.

Проблема ускладнюється переходом до мікросервісної та хмарної архітектури, у якій одна операція користувача може проходити через декілька взаємопов'язаних компонентів. Виявлення джерела деградації продуктивності в такому середовищі потребує аналізу великого обсягу різнорідної телеметричної інформації.

Таким чином, виникає необхідність розроблення підходів, які дозволяють:

- оперативно контролювати поточний стан IT-інфраструктури;
- визначати взаємозв'язки між її компонентами;
- прогнозувати майбутнє навантаження;
- виявляти аномальні режими роботи;
- автоматично змінювати обсяг доступних ресурсів;
- мінімізувати витрати без погіршення якості обслуговування.

3. Аналіз останніх досліджень і публікацій. Проблеми надійності та ефективного управління IT-системами широко досліджуються в межах концепції Site Reliability Engineering (SRE). У роботах B. Beyer та співавторів значна увага приділяється моніторингу, управлінню продуктивністю, автоматизації та забезпеченню надійності програмних систем [1, 2].

Особливого значення в сучасних розподілених системах набуває концепція Observability, яка передбачає використання метрик, логів і трасувань для отримання комплексного уявлення про внутрішній стан системи. На відміну від традиційного моніторингу, Observability орієнтована не лише на виявлення факту проблеми, а й на визначення її можливих причин.

Розвиток контейнеризації та Kubernetes сприяв поширенню автоматичного масштабування. У Kubernetes реалізовані механізми горизонтального та вертикального масштабування, які дозволяють адаптувати кількість або ресурсні характеристики контейнерів відповідно до поточного навантаження [3].

Питання обробки великих обсягів даних та побудови масштабованих розподілених систем розглядаються у працях M. Kleppmann [4] та K. Hwang, G. Fox, J.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Dongarra [7]. Автори наголошують на важливості масштабованості, відмовостійкості та ефективного управління розподіленими ресурсами.

Використання методів машинного навчання для прогнозування та аналізу даних розглядається, зокрема, у працях J. Han, M. Kamber, J. Pei [5] та A. Géron [6]. Зазначені підходи можуть бути адаптовані до задач прогнозування навантаження, класифікації станів інфраструктури та виявлення аномальної поведінки.

Попри значну кількість існуючих досліджень, актуальним залишається питання інтеграції різних методів у єдину систему інтелектуального управління, яка здатна не лише спостерігати за станом інфраструктури, а й прогнозувати його зміни та автоматично формувати відповідні керуючі впливи.

Метою роботи є аналіз сучасних підходів до моніторингу та управління навантаженням в IT-інфраструктурах, визначення їх переваг і обмежень, а також обґрунтування перспектив інтеграції методів машинного навчання, Observability та автоматичного масштабування для побудови проактивних систем управління ресурсами.

Для досягнення поставленої мети необхідно розв’язати такі завдання:

1. проаналізувати традиційні методи моніторингу IT-інфраструктури;
2. дослідити сучасні підходи Observability;
3. розглянути механізми балансування та автоматичного масштабування;
4. проаналізувати можливості машинного навчання для прогнозування навантаження;
5. дослідити методи виявлення аномалій;
6. визначити перспективи формування інтелектуальних систем управління.

Методи дослідження. У роботі використано методи системного та порівняльного аналізу для дослідження сучасних підходів до моніторингу й управління IT-інфраструктурами. Аналіз літературних джерел застосовано для визначення основних тенденцій розвитку SRE, Observability, хмарних і контейнерних технологій.

Для дослідження можливостей інтелектуального управління використано методи машинного навчання, прогнозування часових рядів та виявлення аномалій. Методи узагальнення та систематизації застосовано для формування концептуальної моделі взаємодії моніторингу, аналітики та механізмів автоматичного управління.

Порівняння підходів здійснювалося за такими критеріями: рівень автоматизації, здатність до прогнозування, адаптивність, складність реалізації, вимоги до даних та можливість інтеграції з хмарними середовищами.

Сучасний моніторинг IT-інфраструктури. Моніторинг є базовим компонентом управління IT-інфраструктурою та передбачає безперервний або періодичний збір інформації про стан її апаратних і програмних компонентів.

Основними показниками є:

- завантаження CPU;
- використання оперативної пам’яті;
- дискове навантаження;
- обсяг мережевого трафіку;
- кількість запитів;

– час відповіді;
– кількість помилок;
– кількість активних користувачів;
– доступність сервісів.

Найпростішим підходом є пороговий моніторинг. У цьому випадку визначається допустимий діапазон значень показника, а при його перевищенні генерується попередження.

Наприклад, якщо завантаження CPU перевищує встановлений рівень, система формує повідомлення адміністратору. Перевагами такого підходу є простота, низькі вимоги до обчислювальних ресурсів і зрозумілість правил.

Водночас статичні пороги не враховують часову структуру навантаження. Значення, яке є нормальним у період пікової активності, може бути аномальним у нічний час. Крім того, короткочасні коливання можуть призводити до великої кількості хибних спрацювань.

Тому в сучасних розподілених системах використовується комплексний підхід Observability.

Observability як розвиток традиційного моніторингу. Observability передбачає аналіз трьох основних категорій телеметричних даних: метрик, логів і трасувань.

Метрики відображають кількісні характеристики системи. Логи містять інформацію про окремі події, помилки та операції. Трасування дозволяє простежити проходження конкретного запиту через декілька сервісів.

Комплексний аналіз цих джерел дає змогу встановлювати зв’язки між окремими подіями. Наприклад, збільшення часу відповіді може бути пов’язане не з самим вебсервером, а з перевантаженням бази даних або іншого мікросервісу.

Таким чином, Observability забезпечує більш глибоке розуміння поведінки розподіленої системи та створює інформаційну основу для подальшої автоматизації управління.

Управління навантаженням та автоматичне масштабування. Моніторинг повинен бути пов’язаний із механізмами управління ресурсами. Одним із найбільш поширених методів є балансування навантаження, яке забезпечує розподіл запитів між декількома серверами або екземплярами сервісів.

Використання балансувальника дозволяє уникати ситуацій, коли один сервер перевантажений, а інші використовуються недостатньо. Найпростішим алгоритмом є Round Robin. Більш складні алгоритми можуть враховувати поточне завантаження вузлів, кількість активних з’єднань та час відповіді.

Важливим механізмом сучасних хмарних платформ є автоматичне масштабування. Його метою є динамічна адаптація кількості або потужності ресурсів відповідно до навантаження.

Горизонтальне масштабування можна представити як:

$$N(t)=f(L(t)), \quad (1)$$

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

де N(t) - кількість активних екземплярів у момент часу t, а L(t) - поточне навантаження на систему.

Наприклад, якщо середня завантаженість процесора перевищує встановлений поріг протягом певного інтервалу часу, система автоматично додає нові екземпляри сервісу. Коли навантаження знижується нижче заданого порога, зайві екземпляри видаляються.

Горизонтальне масштабування забезпечує підвищення пропускної здатності, відмовостійкості та гнучкості хмарної інфраструктури. Його ефективність залежить від механізмів балансування навантаження, часу запуску нових екземплярів і коректного налаштування порогів масштабування.

Вертикальне масштабування передбачає зміну ресурсних характеристик окремого вузла: збільшення або зменшення кількості процесорних ядер, обсягу оперативної пам'яті, дискового простору чи інших обчислювальних ресурсів (scale-up / scale-down). На відміну від горизонтального масштабування, за якого змінюється кількість екземплярів сервісу, вертикальне масштабування передбачає зміну потужності вже наявного вузла:

$$R(t)= (C(t), M(t), D(t)), \quad (2)$$

де C(t) - обсяг обчислювальних ресурсів процесора, M(t) - обсяг оперативної пам'яті, D(t) - доступний дисковий простір у момент часу t.

Наприклад, у разі зростання навантаження на базу даних хмарна платформа може збільшити обсяг оперативної пам'яті та кількість процесорних ядер віртуальної машини, забезпечуючи швидше опрацювання запитів без додавання нових вузлів.

Основними перевагами вертикального масштабування є простота реалізації та відсутність необхідності розподіляти навантаження між додатковими екземплярами. Водночас його застосування обмежується максимально доступними характеристиками обладнання або віртуальної машини, а зміна конфігурації в окремих середовищах може вимагати перезапуску сервісу чи тимчасового припинення його роботи.

Основною проблемою реактивного автомасштабування є затримка між зміною навантаження та зміною ресурсів. Тому перспективним є використання прогнозування.

Прогнозування навантаження методами машинного навчання. Навантаження на IT-інфраструктуру часто має виражену часову структуру. Можуть спостерігатися добові, тижневі та сезонні закономірності. Використання історичних даних дозволяє побудувати модель прогнозування майбутнього навантаження.

Для прогнозування можуть застосовуватися статистичні методи, регресійні моделі, дерева рішень, ансамблеві алгоритми та нейронні мережі.

Одним із простих методів є ковзне середнє. Основна ідея методу полягає в тому, що прогнозне значення навантаження визначається як середнє арифметичне за останніми \(n\) інтервалами спостереження.

$$x_{t+1} = \frac{1}{n}\sum_{i=0}^{n-1} x_{t-1}, \quad (3)$$

де x_t - фактичне навантаження в момент часу t;

x_t - розмір вікна усереднення;

x_{t+1} - прогноз навантаження на наступний інтервал.

Такий підхід ефективний для згладжування випадкових коливань, але не завжди здатний врахувати складні нелінійні залежності.

Для складніших часових рядів можуть застосовуватися рекурентні нейронні мережі, зокрема LSTM, а також сучасні моделі глибокого навчання.

Основною перевагою прогнозування є можливість заздалегідь підготувати необхідні ресурси. Якщо модель прогнозує збільшення навантаження через певний проміжок часу, система може виконати масштабування до фактичного виникнення піку.

Виявлення аномалій. Виявлення аномалій є важливою складовою інтелектуального моніторингу. Аномалією може вважатися спостереження або послідовність спостережень, які істотно відрізняються від нормального режиму функціонування системи.

Для виявлення аномалій можуть використовуватися статистичні методи, кластеризація, Isolation Forest, методи аналізу часових рядів та нейронні мережі.

Простим статистичним показником є стандартизоване відхилення:

$$z = x - \mu / \sigma, \quad (4)$$

де x - поточне значення показника; μ - середнє значення показника за вибраний період; σ - стандартне відхилення.

При значному відхиленні від нормального діапазону система може сформувати повідомлення про потенційну аномалію.

Більш складні моделі дозволяють одночасно аналізувати декілька параметрів. Наприклад, одночасне збільшення CPU, мережевого трафіку та часу відповіді може бути більш інформативним, ніж аналіз кожного параметра окремо.

Концептуальна модель інтелектуального управління. На основі проведеного аналізу доцільно розглядати моніторинг і управління як замкнений контур (рис. 1).

На першому етапі здійснюється збір метрик, логів і трасувань. Далі дані очищуються, агрегуються та приводяться до формату, придатного для аналітичних моделей.

На етапі аналізу визначається поточний стан інфраструктури. Паралельно модель прогнозування оцінює майбутню зміну навантаження, а модуль виявлення аномалій визначає нетипові режими роботи.

На основі отриманих результатів формується керуюче рішення. Воно може передбачати збільшення кількості екземплярів сервісу, зміну ресурсних параметрів, перерозподіл запитів або, навпаки, зменшення кількості ресурсів.

Після виконання керуючої дії система повторно аналізує телеметрію та оцінює результат. Таким чином формується замкнений адаптивний контур управління. Безперервний цикл забезпечує стабільну роботу, високу продуктивність та ефективне використання ресурсів хмарної платформи.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

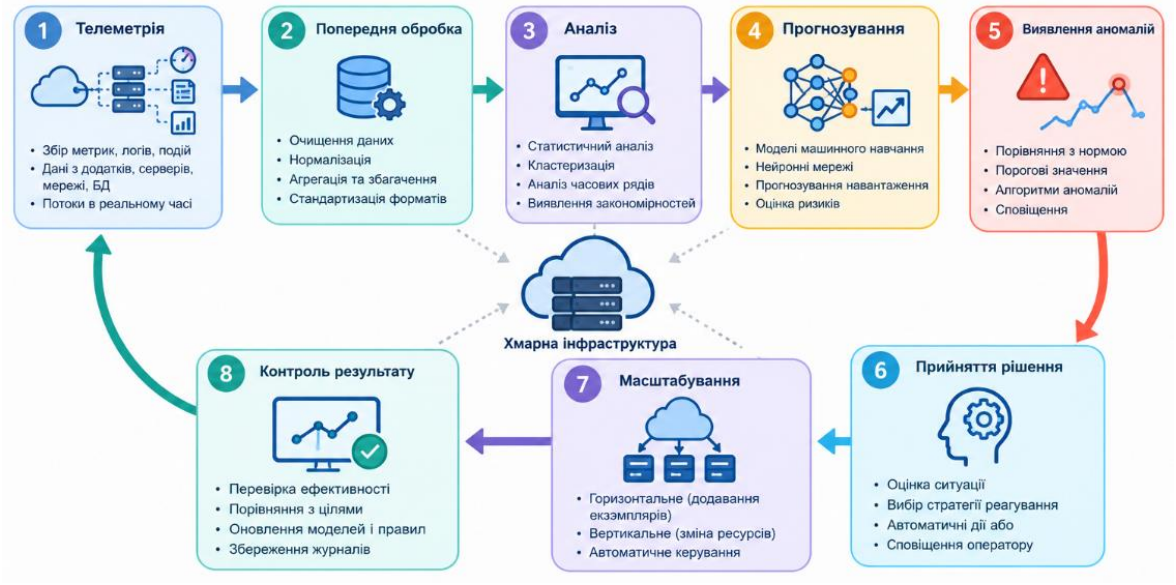


Рисунок 1 – Замкнений контур моніторигу і управління в хмарний інфраструктурі

Порівняльний аналіз підходів пердставленов таблиці 1.

Таблиця 1

Порівняння сучасних підходів до моніторингу та управління

Підхід	Принцип	Переваги	Обмеження
Пороговий моніторинг	Контроль заданих значень	Простота	Реактивність
Observability	Аналіз метрик, логів і трасувань	Комплексне бачення системи	Великий обсяг даних
Балансування	Розподіл запитів	Підвищення доступності	Не прогнозує навантаження
Автомасштабування	Динамічна зміна ресурсів	Еластичність	Можлива затримка
Прогнозування	Оцінювання майбутнього стану	Проактивність	Потреба в історичних даних
Виявлення аномалій	Пошук нетипової поведінки	Раннє виявлення проблем	Хибні спрацювання
AIOps	Інтелектуальна автоматизація	Висока адаптивність	Складність реалізації

Порівняння показує, що зазначені підходи не слід розглядати як взаємовиключні. Найбільш перспективним є їх поєднання в єдиній системі управління.

AIOps та перспективи розвитку. Подальший розвиток IT-інфраструктур пов’язаний із переходом від автоматизованого моніторингу до інтелектуального управління. Одним із перспективних напрямів є AIOps, що передбачає використання методів штучного інтелекту та машинного навчання для автоматизації операційних процесів.

AIOps-система може аналізувати великі обсяги телеметричних даних, автоматично виявляти аномалії, класифікувати інциденти, прогнозувати навантаження та формувати рішення щодо зміни конфігурації інфраструктури.

Перспективним є використання підкріплювального навчання для вибору оптимальної стратегії управління. У такій системі алгоритм може навчатися на основі результатів попередніх рішень.

Цільова функція може одночасно враховувати продуктивність, вартість ресурсів і ризик деградації сервісу:

$$J = \alpha P + \beta C + \gamma R, \quad (5)$$

де P - показник продуктивності інфраструктури, C - вартість використання обчислювальних ресурсів, R - ризик деградації сервісу, α, β, γ - вагові коефіцієнти, що визначають відносну важливість відповідних критеріїв.

Отже, завдання інтелектуального управління полягає не просто у максимізації доступних ресурсів, а в пошуку оптимального балансу між якістю обслуговування, продуктивністю та вартістю.

Водночас автоматизація породжує додаткові ризики. Помилки прогнозної моделі або некоректні керуючі рішення можуть призвести до необґрунтованого збільшення витрат або деградації сервісів. Тому перспективні системи повинні містити механізми контролю, обмеження допустимих дій, оцінювання впевненості прогнозу та можливість втручання оператора.

Окремого значення набуває пояснюваність моделей. Адміністратор повинен мати можливість встановити, чому певний стан було класифіковано як аномальний і чому система прийняла рішення про масштабування.

Обговорення результатів. Проведений аналіз показує, що кожен із розглянутих підходів вирішує окрему частину задачі управління IT-інфраструктурою. Традиційний моніторинг забезпечує контроль базових параметрів, Observability розширює можливості діагностики, балансування забезпечує рівномірний розподіл запитів, а автомасштабування адаптує ресурси до поточних потреб.

Водночас використання прогнозування дозволяє усунути один із основних недоліків реактивного управління - затримку між зміною навантаження та відповідною реакцією системи.

Інтеграція прогнозування та виявлення аномалій створює можливість переходу до більш адаптивної моделі управління. У цьому випадку система може враховувати як поточний стан інфраструктури, так і прогнозовану динаміку її параметрів.

Важливим результатом аналізу є висновок про доцільність побудови багаторівневої системи управління, у якій базові механізми моніторингу доповнюються методами інтелектуального аналізу. Такий підхід дозволяє зберегти простоту та надійність традиційних механізмів, одночасно розширюючи їх прогнозними та адаптивними можливостями.

Висновки. В роботі проведено аналіз сучасних підходів до моніторингу та управління навантаженням в IT-інфраструктурах. Встановлено, що традиційний пороговий моніторинг є ефективним базовим засобом контролю, але має обмежені можливості щодо прогнозування та адаптації до динамічної зміни навантаження.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Досліджено концепцію Observability, яка забезпечує комплексний аналіз метрик, логів і трасувань та дозволяє підвищити ефективність діагностики розподілених систем. Проаналізовано балансування навантаження та механізми горизонтального і вертикального масштабування, які забезпечують динамічну адаптацію ресурсів.

Встановлено, що застосування методів машинного навчання дозволяє перейти від реактивного до проактивного управління шляхом прогнозування майбутнього навантаження та виявлення аномальної поведінки.

Запропоновано концептуальну послідовність інтелектуального управління: **збір телеметрії → попередня обробка → аналіз → прогнозування → виявлення аномалій → прийняття рішення → масштабування → контроль результату.**

Перспективним напрямом подальших досліджень є розроблення та експериментальна перевірка інтелектуальної системи управління навантаженням із використанням моделей прогнозування часових рядів, алгоритмів виявлення аномалій та механізмів автоматичного масштабування. Особливу увагу доцільно приділити оцінюванню точності прогнозів, зменшенню хибних спрацювань, пояснюваності прийнятих рішень та оптимізації співвідношення між продуктивністю і вартістю використання ресурсів.

ЛІТЕРАТУРА/ REFERENCES

1. Beyer B., Jones C., Petoff J., Murphy N. R. Site Reliability Engineering: How Google Runs Production Systems. Sebastopol: O'Reilly Media, 2016. 552 p. URL: https://repo.darmajaya.ac.id/4636/1/Site%20Reliability%20Engineering_%20How%20Google%20Runs%20Production%20Systems%20%28%20PDFDrive%20%29.pdf

2. Beyer B., Murphy N. R., Rensin D. K., Kawahara K., Thorne S. The Site Reliability Workbook: Practical Ways to Implement SRE. Sebastopol : O'Reilly Media, 2018. 512 p. URL: https://lrita.github.io/images/posts/com/the-site-reliability-workbook-next18.pdf

3. Burns B., Beda J., Hightower K. Kubernetes: Up and Running: Dive into the Future of Infrastructure. 3rd ed. Sebastopol : O'Reilly Media, 2022. 416 p.

4. Kleppmann M. Designing Data-Intensive Applications: The Big Ideas Behind Reliable, Scalable, and Maintainable Systems. Sebastopol : O'Reilly Media, 2017. 616 p.

5. Han J., Kamber M., Pei J. Data Mining: Concepts and Techniques. 3rd ed. Waltham : Morgan Kaufmann, 2011. 744 p.
URL: https://myweb.sabanciuniv.edu/rdehkharghani/files/2016/02/The-Morgan-Kaufmann-Series-in-Data-Management-Systems-Jiawei-Han-Micheline-Kamber-Jian-Pei-Data-Mining.-Concepts-and-Techniques-3rd-Edition-Morgan-Kaufmann-2011.pdf

6. Géron A. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow. 3rd ed. Sebastopol : O'Reilly Media, 2022. 861 p.
URL: https://32931414.s21i.faiusr.com/61/ABUIABA9GAAg1-qMuwYo-LWn0Ac.pdf

7. Hwang K., Fox G.C., Dongarra J.J. Distributed and Cloud Computing: From Parallel Processing to the Internet of Things. Waltham: Morgan Kaufmann, 2012. 672 p. URL: https://cutepooji.wordpress.com/wp-content/uploads/2017/01/distributed-and-cloud-computing-from-parallel-processing-to-the-internet-of-things.pdf

Received 02.07.2026.
Accepted 02.09.2026.
Published 30.10.2026

Analysis of modern approaches to monitoring and load management in IT infrastructures

Abstract. The article analyzes modern approaches to monitoring and load management in IT infrastructures. Traditional resource monitoring methods, Observability, load balancing, autoscaling, containerization, and machine learning methods for workload forecasting and anomaly detection are considered. The features of reactive, proactive, and intelligent resource management are analyzed. A conceptual model of intelligent load management is proposed, integrating telemetry data collection, preprocessing, workload forecasting, anomaly detection, and automated resource scaling. A comparative analysis of the main approaches is performed according to the level of automation, adaptability, forecasting capabilities, and data requirements. The feasibility of transitioning from reactive to proactive IT infrastructure management is substantiated, and the prospects of using machine learning and AIOps technologies are identified.

It has been established that the application of machine learning methods enables a transition from reactive to proactive management by forecasting future load and detecting anomalous behavior.

A conceptual sequence for intelligent management is proposed: telemetry collection → preprocessing → analysis → forecasting → anomaly detection → decision-making → scaling → outcome monitoring.

A promising avenue for future research involves the development and experimental validation of an intelligent load management system utilizing time-series forecasting models, anomaly detection algorithms, and automated scaling mechanisms. Particular attention should be paid to evaluating forecast accuracy, reducing false positives, ensuring the explainability of decisions, and optimizing the balance between performance and resource costs.

Keywords: IT infrastructure, monitoring, load management, Observability, cloud technologies, autoscaling, machine learning, forecasting, anomaly detection, AIOps.

Островська Катерина Юріївна – к.т.н., доцент, доцент кафедри інформаційних технологій і систем Українського державного університету науки і технологій.
ORCID: https://orcid.org/0000-0002-9375-4121
Самуськов Олександр – аспірант кафедри інформаційних технологій і систем Українського державного університету науки і технологій.
ORCID: https://orcid.org/0009-0004-0336-9353

Ostrovska Kateryna Yuriivna – Ph.D., Associate Professor of the Department of Information Technology and Systems of Ukrainian State University of Science and Technology.
ORCID: https://orcid.org/0000-0002-9375-4121
Samuskov Oleksandr – postgraduate student of the Department of Information Technology and Systems of Ukrainian State University of Science and Technology.
ORCID: https://orcid.org/0009-0004-0336-9353

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

DOI 10.34185/1562-9945-5-166-2026-27
УДК 621.867:62-52

А.І. Аніщенко, М.М. Нестеренко, В.В. Блажко, Т.М. Нестеренко, Д.Ю.Бугров

АВТОМАТИЗАЦІЯ ТЕХНОЛОГІЧНОГО ОБЛАДНАННЯ ДЛЯ ВИРОБНИЦТВА ДЕКОРАТИВНИХ ЕЛЕМЕНТІВ МОЩЕННЯ

Анотація. У статті розглянуто сучасні технології виробництва декоративних елементів мощення та проаналізовано особливості автоматизації технологічного обладнання, що застосовується для їх виготовлення. Проведено аналіз технології вібропресування, як найбільш ефективного способу формування дрібноштучних бетонних виробів. Обґрунтовано структуру автоматизованої технологічної лінії, яка охоплює процеси транспортування сировини, дозування компонентів, приготування бетонної суміші, вібропресування, транспортування готових виробів і контролю технологічних параметрів. Розглянуто застосування програмованих логічних контролерів, цифрових датчиків, частотно-регульованих електроприводів і SCADA-систем для керування виробничими процесами. Визначено переваги комплексної автоматизації, зокрема підвищення продуктивності, покращення якості бетонних виробів, зниження енерговитрат, мінімізацію впливу людського фактору та забезпечення можливості інтеграції виробництва до цифрових систем відповідно до концепції Industry 4.0. Окреслено перспективи подальшого розвитку автоматизованих систем керування із застосуванням технологій Industrial Internet of Things (IIoT), цифрових двійників та прогнозного технічного обслуговування.

Ключові слова: декоративні елементи мощення; вібропресування; бетонні вироби; автоматизація виробництва; технологічна лінія; програмований логічний контролер; PLC; SCADA; Industry 4.0; Industrial Internet of Things (IIoT).

Постановка проблеми. Одним із пріоритетних напрямів розвитку сучасної будівельної індустрії є підвищення ефективності виробництва бетонних виробів, що широко застосовуються під час благоустрою територій, будівництва транспортної інфраструктури та об'єктів міського господарства. До найбільш поширених виробів належать тротуарна плитка, бутовий камінь, декоративні елементи мощення та інші будівельні конструкції, які повинні відповідати високим вимогам щодо міцності, довговічності, морозостійкості, геометричної точності та декоративних властивостей [1–6].

Зростання обсягів будівництва, підвищення вимог до якості продукції та необхідність скорочення виробничих витрат зумовлюють потребу у впровадженні сучасних автоматизованих технологічних комплексів [1, 8]. Використання традиційних виробничих схем із високою часткою ручної праці не дозволяє забезпечити стабільність технологічних параметрів, що негативно впливає на якість бетонної суміші, продуктивність обладнання та економічні показники підприємства.

© Аніщенко А.І., Нестеренко М.М., Блажко В.В., Нестеренко Т.М., Бугров Д.Ю., 2026
Copyright for this paper by its authors. Use permitted under License CC BY 4.0.
ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Одним із найбільш ефективних способів виготовлення бетонних виробів є технологія вібропресування, яка поєднує механічне ущільнення бетонної суміші із вібраційним впливом [1–5]. Такий метод забезпечує формування щільної структури матеріалу, зменшення пористості виробів, підвищення їх механічної міцності та довговічності. Разом із тим ефективність процесу значною мірою визначається рівнем автоматизації виробничої лінії, точністю дозування компонентів, стабільністю режимів змішування та керування технологічним обладнанням [7–10].

У сучасних умовах розвитку концепції Industry 4.0 автоматизація виробництва вже не обмежується лише використанням окремих засобів керування [11–15]. Все більшого поширення набувають інтегровані автоматизовані системи, що поєднують програмовані логічні контролери (PLC), цифрові датчики технологічних параметрів, частотні електроприводи, системи диспетчеризації SCADA та засоби дистанційного моніторингу виробничих процесів [11–15]. Це дозволяє забезпечити безперервний контроль роботи обладнання, своєчасне виявлення відхилень технологічних параметрів та підвищити надійність функціонування виробничих комплексів.

Таким чином, розроблення автоматизованого технологічного обладнання для виробництва декоративних елементів мощення є актуальним науково-практичним завданням, вирішення якого сприятиме підвищенню продуктивності підприємств, покращенню якості готової продукції, зниженню енерговитрат та впровадженню сучасних цифрових технологій у виробництво будівельних матеріалів.

Аналіз останніх публікацій і досліджень. Питання вдосконалення технології виробництва бетонних виробів широко висвітлені у працях сучасних вітчизняних і зарубіжних дослідників, присвячених удосконаленню складу бетонних сумішей, технології вібропресування та забезпеченню довговічності бетонних виробів [1–6].

Сучасні дослідження свідчать, що найбільш ефективною технологією виготовлення тротуарної плитки, бутового каменю та інших дрібноштучних бетонних виробів є метод напівсухого вібропресування. Застосування комбінованої дії статичного навантаження та вібрації дозволяє отримати бетон із мінімальною пористістю, високою міцністю та точними геометричними параметрами [1–6].

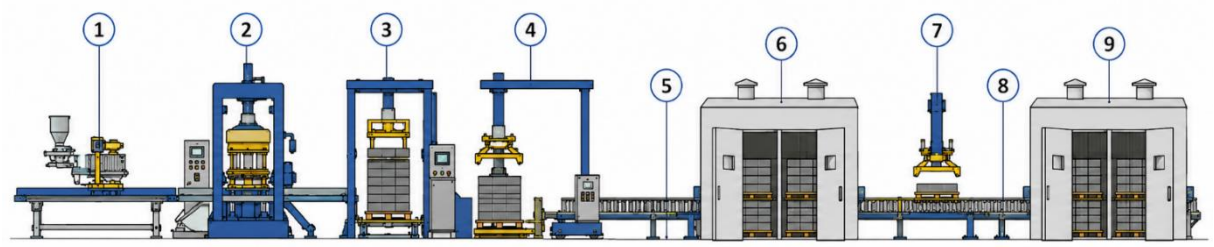
Важливим напрямом розвитку технологічного обладнання є автоматизація процесів дозування сировинних компонентів, приготування бетонної суміші, транспортування матеріалів, формування виробів та контролю технологічних параметрів [7 – 13]. Для цього широко застосовуються програмовані логічні контролери, електронні вагові дозатори, частотно-керовані електроприводи, датчики рівня, температури, вологості та системи диспетчерського керування виробництвом [7–15].

Незважаючи на значну кількість досліджень, питання комплексної автоматизації технологічних ліній виробництва декоративних елементів мощення потребують подальшого розвитку. Особливо актуальними залишаються задачі інтеграції сучасних цифрових систем керування із технологічним обладнанням, підвищення енергоефективності виробничих процесів, оптимізації режимів роботи вібропресів та впровадження інтелектуальних систем моніторингу технічного стану обладнання.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Мета статті. Метою статті є аналіз сучасних технологій виробництва декоративних елементів мощення та обґрунтування застосування автоматизованого технологічного обладнання для підвищення продуктивності виробництва, забезпечення стабільної якості бетонних виробів, зниження енерговитрат і мінімізації впливу людського фактора. Для досягнення поставленої мети виконано аналіз технології вібропресування, досліджено структуру автоматизованої виробничої лінії, розглянуто сучасні засоби автоматизації технологічних процесів та визначено перспективні напрями їх використання у виробництві бетонних виробів.

Виклад основного матеріалу. Одним із найбільш ефективних способів виготовлення дрібноштучних бетонних виробів є технологія вібропресування, яка поєднує дію вібраційних коливань і статичного пресування (рис. 1). Такий спосіб забезпечує рівномірне ущільнення напівсухої бетонної суміші, зменшення її пористості та формування виробів із високими фізико-механічними характеристиками. У порівнянні з технологією вібролиття вібропресування характеризується більшою продуктивністю, меншою тривалістю виробничого циклу, стабільною якістю готової продукції та можливістю комплексної автоматизації технологічного процесу [1–6].



1 – устаткування для приймання розподілення бетонної суміші;
2 – пост вібропресування бетону; 3 – автомат-пакетувальник; 4 – автомат для перекладання пакетів; 5 – конвейєр камери першої стадії ТВО; 6 – камера першої стадії ТВО;
7 – автозакват; 8 – конвейєр камери другої стадії ТВО; 9 – камера другої стадії ТВО
Рисунок 1 – Технологічна схема виробництва

Аналіз сучасних виробництв показав, що автоматизація процесу повинна охоплювати всі стадії технологічного циклу: приймання і транспортування сировини, дозування компонентів, приготування бетонної суміші, формування виробів, транспортування піддонів, тверднення продукції та контроль параметрів роботи обладнання. Комплексна автоматизація дозволяє мінімізувати вплив людського фактору, підвищити продуктивність технологічної лінії та забезпечити стабільність технологічних режимів [8, 11–14].

Запропонована технологічна схема виробництва декоративних елементів мощення передбачає автоматизоване виконання всіх основних виробничих операцій. Структуру технологічної лінії наведено на рис. 2. Вона включає системи приймання та накопичення сировини, транспортне обладнання, вузли дозування компонентів, бетонозмішувач, вібропрес, обладнання для транспортування сформованих виробів і ділянку їх тверднення. Така побудова виробництва забезпечує безперервність технологічного процесу та ефективне використання обладнання.



1 – склад інертних матеріалів; 2 – дозування компонентів; 3 – змішування бетонної суміші; 4 – подача суміші у вібропрес; 5 – вібропресування та формування виробів; 6 – транспортування піддонів; 7 – твердіння виробів; 8 – пакетування та відвантаження

Рисунок 2 – Автоматизована лінія виробництва декоративних елементів мощення методом вібпропресування

На першому етапі виробництва здійснюється автоматизоване приймання інертних матеріалів. Пісок та гранвідсів надходять до приймальних бункерів, звідки за допомогою стрічкових живильників і конвеєрів транспортуються до витратних бункерів. Цемент подається в силос, а потім через шнековий транспортер надходить до вагового дозатора. Автоматизація транспортних операцій забезпечує рівномірну подачу матеріалів, скорочує втрати сировини та дозволяє підтримувати необхідний режим роботи технологічної лінії [7, 8, 11].

Однією з найважливіших операцій є дозування компонентів бетонної суміші. Саме точність дозування визначає фізико-механічні властивості готової продукції. Для реалізації даного процесу використовуються електронні вагові дозатори, які забезпечують автоматичний контроль маси цементу, заповнювачів і води.

Система керування здійснює безперервне порівняння фактичних параметрів із заданими значеннями та автоматично коригує роботу дозувального обладнання [5–8].

Після завершення дозування всі компоненти надходять до одновального бетонозмішувача. На цьому етапі особливе значення має автоматичне керування процесом змішування. Програмований контролер забезпечує підтримання необхідної тривалості циклу, послідовності подачі компонентів і кількості води відповідно до встановленої рецептури. Використання автоматичного регулювання дозволяє отримувати бетонну суміш зі стабільними технологічними характеристиками незалежно від умов експлуатації обладнання [2–4, 7, 8].

Після завершення змішування бетонна суміш автоматично транспортується до бункера вібропреса. Формування виробів здійснюється методом вібропресування, який поєднує дію статичного навантаження та високочастотних вібрацій. У процесі ущільнення частинки бетонної суміші займають більш компактне положення, що сприяє зменшенню пористості та підвищенню щільності виробів. Завдяки цьому готова продукція характеризується високою міцністю, морозостійкістю, водонепроникністю та зносостійкістю [1–6].

Однією з головних переваг запропонованої технології є використання автоматизованої системи керування вібропресом (рис. 3).

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

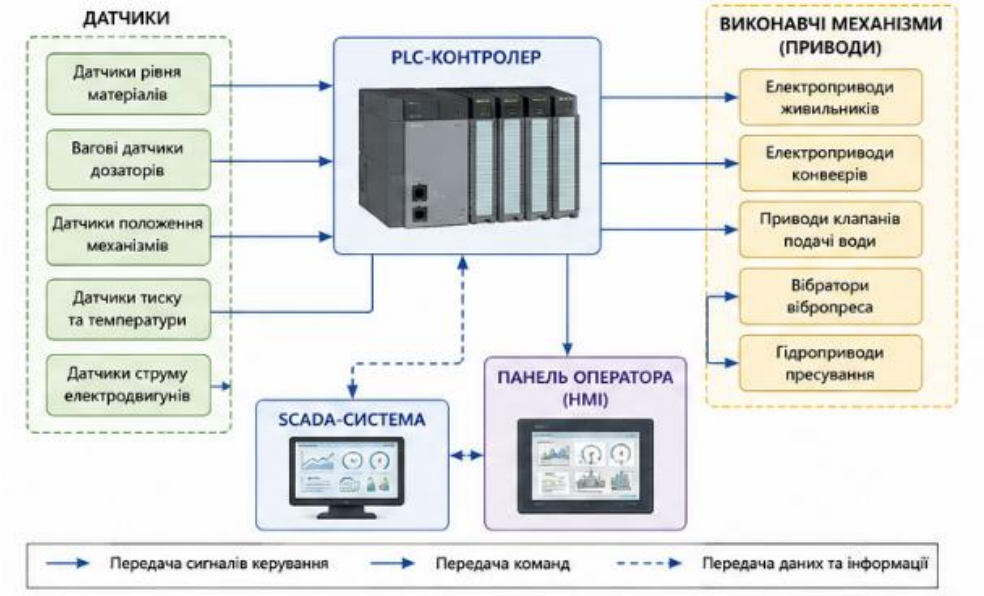


Рисунок 3 – Функціональна схема автоматизації виробництва декоративних елементів мощення

Робота пресового обладнання координується програмованим логічним контролером, який забезпечує синхронізацію роботи електроприводів, виконавчих механізмів і транспортного обладнання. У процесі роботи контролюються тривалість циклу формування, параметри вібрації, положення пуансона, стан прес-форми та транспортування сформованих виробів на наступні технологічні операції [7–10].

Після завершення формування вироби автоматично переміщуються на піддонах до ділянки тверднення. Транспортні механізми забезпечують безперервне переміщення продукції без участі оператора, що дозволяє скоротити тривалість виробничого циклу та зменшити ймовірність механічних пошкоджень свіжоформованих виробів. Після набору необхідної міцності готова продукція автоматично передається на склад готової продукції для подальшого відвантаження споживачам [1, 8].

Важливою складовою автоматизованої виробничої лінії є система моніторингу технологічних параметрів (рис. 4). Для контролю роботи обладнання застосовуються датчики рівня матеріалів, вагові датчики, датчики положення виконавчих механізмів, температури, тиску та електричних параметрів приводів [7-10]. Отримана інформація надходить до програмованого логічного контролера, де здійснюється її оброблення, аналіз і формування команд керування. Такий підхід дозволяє оперативно реагувати на відхилення технологічних параметрів і підтримувати оптимальні режими роботи обладнання [7-10].



Рисунок 4 – Схема автоматизації та керування

Сучасні автоматизовані виробничі комплекси дедалі частіше оснащуються SCADA-системами, які забезпечують диспетчерський контроль технологічного процесу в режимі реального часу [9-11]. Оператор отримує інформацію про стан усіх технологічних вузлів, продуктивність обладнання, витрати матеріалів, кількість виробленої продукції та можливі аварійні ситуації. Архівування технологічних параметрів дозволяє проводити аналіз роботи обладнання, оцінювати його ефективність та своєчасно планувати технічне обслуговування [9-11].

Проведений аналіз сучасного обладнання показав, що впровадження комплексної автоматизації забезпечує підвищення продуктивності виробничих ліній, покращення стабільності технологічних процесів, скорочення енерговитрат і зменшення кількості браку. Автоматизовані системи керування забезпечують точне виконання технологічних операцій, мінімізують вплив людського фактора та створюють умови для інтеграції виробництва в сучасні цифрові виробничі системи відповідно до концепції Industry 4.0 [11–15].

Проведений аналіз сучасного обладнання для виробництва декоративних елементів мощення показав, що більшість виробників застосовує автоматизовані вібропреси з програмованими системами керування, які забезпечують стабільність технологічного процесу та високу якість готової продукції. Найбільшого поширення набули вібропреси, обладнані електрогідравлічними або електромеханічними приводами, електронними системами дозування та цифровими засобами контролю параметрів роботи [1, 7–11].

Одним із перспективних напрямів розвитку такого обладнання є використання програмованих логічних контролерів (PLC), які забезпечують централізоване керування всіма технологічними операціями виробничої лінії [7–10]. На відміну від традиційних релейно-контакторних схем, PLC-системи характеризуються високою швидкодією, можливістю зміни алгоритмів керування без модернізації апаратної частини, простотою інтеграції з цифровими датчиками та інформаційними системами підприємства [7–10].

Використання PLC-контролера дозволяє реалізувати автоматичне керування всіма основними етапами виробництва: дозуванням компонентів бетонної суміші, процесом змішування, транспортуванням матеріалів, роботою вібропреса, подачею піддонів, переміщенням готової продукції та контролем аварійних режимів (рис. 5). Крім того, програмне забезпечення забезпечує можливість оперативного коригування параметрів технологічного процесу відповідно до рецептури бетонної суміші та типу продукції, що виготовляється [7–11].

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

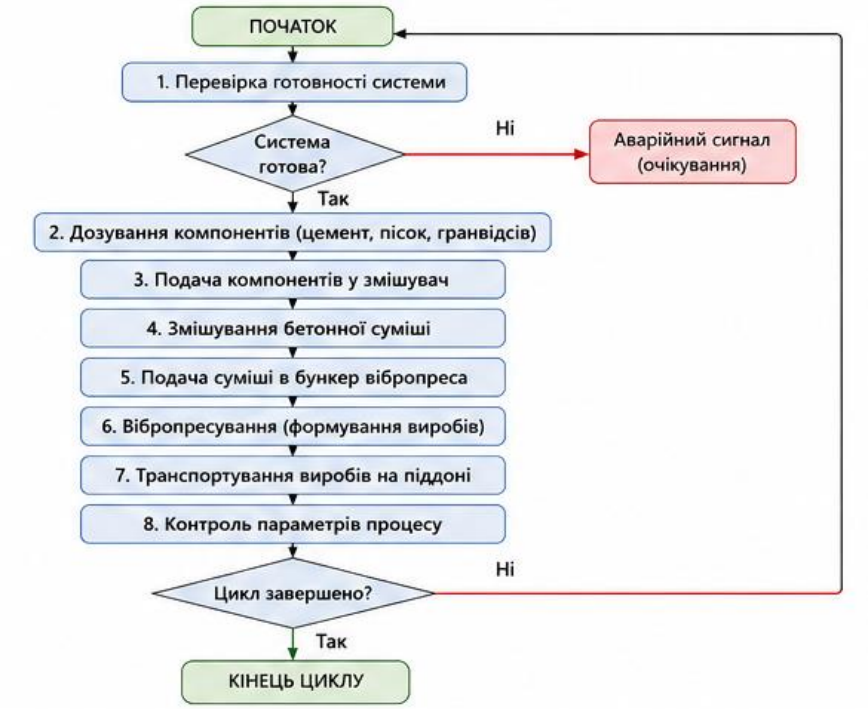


Рисунок 5 – Алгоритм роботи PLC (технологічний цикл)

Важливим елементом автоматизованої системи є використання частотно-регульованих електроприводів. Регулювання швидкості обертання електродвигунів дозволяє оптимізувати режими роботи транспортного обладнання, змішувача та допоміжних механізмів, зменшити пускові струми, знизити динамічні навантаження та підвищити енергоефективність виробництва. Крім економії електроенергії це забезпечує збільшення ресурсу роботи механічного обладнання та скорочення витрат на його технічне обслуговування [8, 11, 13].

Під час дослідження встановлено, що автоматизація процесу дозування є одним із найважливіших факторів забезпечення якості бетонної суміші. Застосування вагових дозаторів із цифровими тензометричними датчиками дозволяє підтримувати високу точність дозування цементу, заповнювачів та води, що безпосередньо впливає на однорідність суміші, міцність виробів та зменшення кількості браку [2–8].

Особливу увагу приділено автоматизації процесу вібропресування. Аналіз технологічного циклу показав, що підтримання постійних параметрів частоти та амплітуди коливань, часу ущільнення і величини пресового зусилля забезпечує формування виробів із мінімальною пористістю та високою щільністю структури бетону. Використання автоматичного керування режимами вібропресування дозволяє стабілізувати якість готової продукції незалежно від зміни зовнішніх умов або людського фактора [1–8].

Виконаний аналіз також показав доцільність інтеграції автоматизованої системи керування із сучасними SCADA-системами диспетчеризації (рис. 6). Такий підхід забезпечує централізований контроль усіх технологічних процесів, накопичення архівних даних, відображення параметрів роботи обладнання у реальному часі та оперативне інформування персоналу про виникнення аварійних ситуацій. Це значно підвищує ефек-

тивність експлуатації технологічної лінії та створює передумови для переходу до цифрового виробництва [9–13].

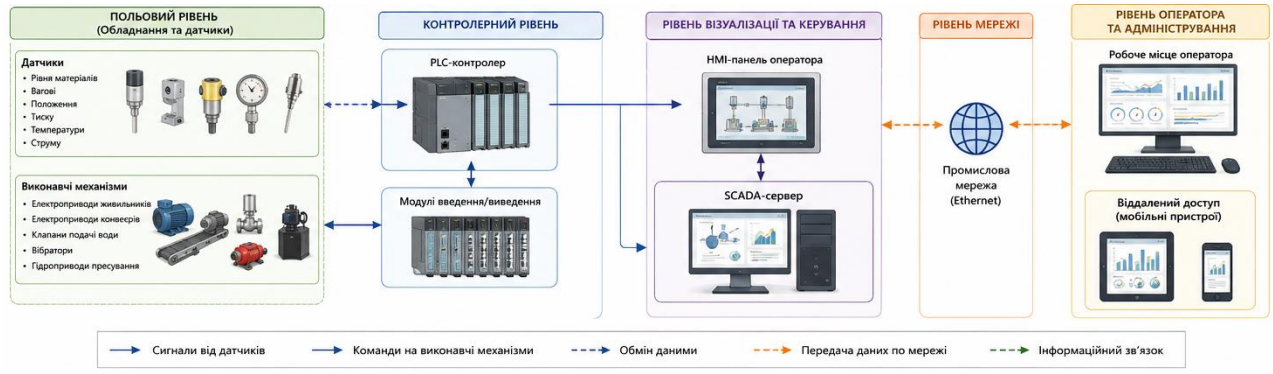


Рисунок 6 – Структура системи керування (PLC–SCADA)

Окремим напрямом розвитку автоматизації є використання технологій дистанційного моніторингу та промислового Інтернету речей (Industrial Internet of Things, IIoT). Передавання інформації від датчиків до серверів збору даних дозволяє здійснювати безперервний аналіз технічного стану обладнання, прогнозувати можливі відмови окремих вузлів та своєчасно виконувати профілактичне технічне обслуговування. Реалізація концепції Predictive Maintenance забезпечує скорочення простоїв обладнання та підвищення його експлуатаційної надійності [11–15].

Висновки. На основі аналізу сучасних технологій виготовлення декоративних елементів мощення обґрунтовано доцільність застосування вібропресування як способу формування бетонних виробів, що поєднує високу продуктивність, точність геометричних параметрів і забезпечення необхідних фізико-механічних властивостей продукції. Обґрунтовано структуру автоматизованої технологічної лінії, яка охоплює транспортування та дозування сировини, приготування бетонної суміші, формування й переміщення виробів, а також контроль технологічних параметрів із використанням програмованих логічних контролерів, цифрових датчиків, частотно-регульованих електроприводів і SCADA-систем. Визначено, що інтеграція зазначених засобів автоматизації створює передумови для підвищення точності керування виробничими процесами, стабілізації якості продукції, зниження матеріало- й енергоємності виробництва та скорочення простоїв обладнання. Подальші дослідження доцільно спрямувати на розроблення та експериментальну верифікацію інтелектуальної системи керування із застосуванням промислового Інтернету речей, цифрових двійників і алгоритмів машинного навчання для оптимізації технологічних режимів та прогнозування технічного стану обладнання.

ЛІТЕРАТУРА

1. Reichenbach S., Kromoser B. State of practice of automation in precast concrete production. Journal of Building Engineering. 2021. Vol. 43. Article 102527.
DOI: 10.1016/j.jobe.2021.102527. URL:
https://www.sciencedirect.com/science/article/pii/S2352710221003843?utm_sourc%D0%B5

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

2. Zhao Z., Liu Y., Lu Y. et al. Prediction of properties of recycled aggregate concrete using machine learning models: A critical review. Journal of Building Engineering. 2024. Vol. 90. Article 109516. DOI: 10.1016/j.jobe.2024.109516.
URL: https://researchportal.ulisboa.pt/en/publications/prediction-of-properties-of-recycled-aggregate-concrete-using-mac/?utm_source
3. Antunes A., Costa H., Do Carmo R., Júlio E. A comprehensive review of sustainable use of construction and demolition waste as recycled aggregates in the production of concrete – Properties, mix design and on-site applications. Construction and Building Materials. 2025. Vol. 482. Article 141733. DOI: 10.1016/j.conbuildmat.2025.141733.
URL: https://researchportal.ulisboa.pt/en/publications/a-comprehensive-review-of-sustainable-use-of-construction-and-dem/
4. Recycled aggregates from construction and demolition waste towards an application on structural concrete: A review. Journal of Building Engineering. 2022. Vol. 52. Article 104452. DOI: 10.1016/j.jobe.2022.104452.
URL: https://www.sciencedirect.com/science/article/abs/pii/S235271022200465X
5. Golewski G. L. An assessment of microstructure and durability of concrete modified with mineral additives. Materials. 2021. Vol. 14(23). Article 7348.
6. EN 206:2013+A2:2021. Concrete – Specification, performance, production and conformity. Brussels: European Committee for Standardization, 2021.
7. Bolton W. Programmable Logic Controllers. 7th ed. Oxford : Newnes, 2021. 376 p.
8. Groover M. P. Automation, Production Systems, and Computer-Integrated Manufacturing. 5th ed. Hoboken : Pearson, 2020. 816 p.
9. Boyer S. A. SCADA: Supervisory Control and Data Acquisition. 5th ed. Research Triangle Park : ISA, 2020. 384 p.
10. IEC 61131-3:2013. Programmable Controllers – Part 3: Programming Languages. Geneva: International Electrotechnical Commission, 2013.
11. Yadav G., Paul K. Architecture and Security of SCADA Systems: A Review. arXiv. 2020. arXiv:2001.02925. URL: https://arxiv.org/abs/2001.02925?utm_source
12. Derigent W., Cardin O., Trentesaux D. Industry 4.0: Contributions of holonic manufacturing control architectures and future challenges. arXiv. 2020. arXiv:2002.04525.
URL: https://arxiv.org/abs/2002.04525?utm_source
13. Javaid M., Haleem A., Singh R.P., Khan S. Industry 4.0 technologies and their applications in manufacturing industries. Sustainable Operations and Computers. 2022. Vol. 3. P. 96–103.
14. International Federation of Robotics. World Robotics 2024: Industrial Robots. Frankfurt : IFR, 2024.
15. Li Z., Radlińska A. Artificial Intelligence in Concrete Materials: A Scientometric View. arXiv. 2022. arXiv:2209.09636. URL: https://arxiv.org/abs/2209.09636?utm_source

REFERENCES

1. Reichenbach S., Kromoser B. State of practice of automation in precast concrete production. Journal of Building Engineering. 2021. Vol. 43. Article 102527.
DOI: 10.1016/j.jobe.2021.102527.

URL:
https://www.sciencedirect.com/science/article/pii/S2352710221003843?utm_sourc%D0%B5
2. Zhao Z., Liu Y., Lu Y. et al. Prediction of properties of recycled aggregate concrete using machine learning models: A critical review. Journal of Building Engineering. 2024. Vol. 90. Article 109516. DOI: 10.1016/j.jobe.2024.109516.
URL: https://researchportal.ulisboa.pt/en/publications/prediction-of-properties-of-recycled-aggregate-concrete-using-mac/?utm_source.
3. Antunes A., Costa H., Do Carmo R., Júlio E. A comprehensive review of sustainable use of construction and demolition waste as recycled aggregates in the production of concrete – Properties, mix design and on-site applications. Construction and Building Materials. 2025. Vol. 482. Article 141733. DOI: 10.1016/j.conbuildmat.2025.141733.
URL: https://researchportal.ulisboa.pt/en/publications/a-comprehensive-review-of-sustainable-use-of-construction-and-dem/.
4. Recycled aggregates from construction and demolition waste towards an application on structural concrete: A review. Journal of Building Engineering. 2022. Vol. 52. Article 104452. DOI: 10.1016/j.jobe.2022.104452.
URL: https://www.sciencedirect.com/science/article/abs/pii/S235271022200465X .
5. Golewski G. L. An assessment of microstructure and durability of concrete modified with mineral additives. Materials. 2021. Vol. 14(23). Article 7348.
6. EN 206:2013+A2:2021. Concrete – Specification, performance, production and conformity. Brussels: European Committee for Standardization, 2021.
7. Bolton W. Programmable Logic Controllers. 7th ed. Oxford : Newnes, 2021. 376 p.
8. Groover M. P. Automation, Production Systems, and Computer-Integrated Manufacturing. 5th ed. Hoboken: Pearson, 2020. 816 p.
9. Boyer S. A. SCADA: Supervisory Control and Data Acquisition. 5th ed. Research Triangle Park: ISA, 2020. 384 p.
10. IEC 61131-3:2013. Programmable Controllers – Part 3: Programming Languages. Geneva : International Electrotechnical Commission, 2013.
11. Yadav G., Paul K. Architecture and Security of SCADA Systems: A Review. arXiv. 2020. arXiv:2001.02925. URL: https://arxiv.org/abs/2001.02925?utm_source .
12. Derigent W., Cardin O., Trentesaux D. Industry 4.0: Contributions of holonic manufacturing control architectures and future challenges. arXiv. 2020. arXiv:2002.04525.
URL: https://arxiv.org/abs/2002.04525?utm_source (дата звернення: 25.07.2025 р.)
13. Javaid M., Haleem A., Singh R. P., Khan S. Industry 4.0 technologies and their applications in manufacturing industries. Sustainable Operations and Computers. 2022. Vol. 3. P. 96–103.
14. International Federation of Robotics. World Robotics 2024: Industrial Robots. Frankfurt : IFR, 2024.
15. Li Z., Radlińska A. Artificial Intelligence in Concrete Materials: A Scientometric View. arXiv. 2022. arXiv:2209.09636. URL: https://arxiv.org/abs/2209.09636?utm_source.

Received 06.07.2026.
Accepted 09.09.2026.
Published 30.10.2026

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Automation of technological equipment for the production of decorative paving elements

The paper examines modern technologies for manufacturing decorative paving elements and analyzes the automation of technological equipment used in their production. Vibropressing technology is considered as the most efficient method for manufacturing small-sized concrete products. The structure of an automated production line, including raw material handling, component dosing, concrete mixing, vibropressing, product transportation, curing, and process monitoring, is substantiated. The application of programmable logic controllers (PLC), digital sensors, variable frequency drives, and SCADA systems for process control is analyzed. The advantages of integrated automation are identified, including increased productivity, improved product quality, reduced energy consumption, minimized human influence, and integration into digital manufacturing systems in accordance with the Industry 4.0 concept. Prospects for further development are associated with the implementation of Industrial Internet of Things (IIoT) technologies, digital twins, and predictive maintenance systems.

Keywords: decorative paving elements; vibropressing; concrete products; production automation; technological line; programmable logic controller; PLC; SCADA; Industry 4.0; Industrial Internet of Things (IIoT).

Аніщенко Анна Ігорівна – к.т.н., доцент, доцент кафедри Автоматизації та комп'ютерно-інтегрованих технологій, Харківський національний університет міського господарства ім. О.М. Бекетова, Україна.
ORCID: https://orcid.org/0000-0002-5649-9379

Нестеренко Микола Миколайович – к.т.н., доцент, доцент кафедри галузевого машинобудування та мехатроніки, Національний університет «Полтавська політехніка імені Юрія Кондратюка», Україна.
ORCID: https://orcid.org/0000-0002-4073-1233

Блажко Володимир Володимирович - кандидат технічних наук, доцент, доцент кафедри Автоматизації та комп'ютерно-інтегрованих технологій, Харківський національний університет міського господарства ім. О.М. Бекетова, Україна.
ORCID: https://orcid.org/0000-0002-3411-0385

Нестеренко Тетяна Миколаївна - доцент кафедри нафтогазової інженерії та технологій, Національний університет «Полтавська політехніка імені Юрія Кондратюка», Україна.
ORCID: https://orcid.org/0000-0002-2387-8575

Дмитро Юрійович Бугров - викладач кафедри автоматики, електроніки та телекомунікацій, Національний університет «Полтавська політехніка імені Юрія Кондратюка», Україна.
ORCID: https://orcid.org/0009-0000-4571-5147

Anishchenko Anna - PhD (Tech), Associate Professor, Associate Professor of the Department of Automation and Computer-Integrated Technologies, O. M. Beketov National University of Urban Economy in Kharkiv.
ORCID: https://orcid.org/0000-0002-5649-9379

Nesterenko Mykola - PhD (Tech), Associate Professor, Associate Professor of the Department of industrial mechanical engineering and mechatronics National University «Yuri Kondratyuk Poltava Polytechnic», Ukraine.
ORCID: https://orcid.org/0000-0002-4073-1233
Blazko Vladimir - PhD (Tech), Associate Professor, Associate Professor of the Department of Automation and Computer-Integrated Technologies, O. M. Beketov National University of Urban Economy in Kharkiv.
ORCID: https://orcid.org/0000-0002-3411-0385
Nesterenko Tetiana - PhD (Tech), Associate Professor, Associate Professor of the Department of industrial mechanical engineering and mechatronics, National University «Yuri Kondratyuk Poltava Polytechnic».
ORCID: https://orcid.org/0000-0002-2387-8575
Buhrov Dmytro - PhD (Tech) Department of Automation, Electronics and Telecommunications National University «Yuri Kondratyuk Poltava Polytechnic», Ukraine.
ORCID: https://orcid.org/0009-0000-4571-5147

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

«Системні технології» 5 (166) 2026 Том 2 «System technologies»
DOI 10.34185/1562-9945-5-166-2026-28
УДК: 004.056.5

С.Б. Бурченко

ЕКСПЕРИМЕНТАЛЬНА АГЕНТНА СИСТЕМА КІБЕРЗАХИСТУ НА ЦИФРОВОМУ ДВІЙНИКУ ЕНЕРГЕТИЧНОГО ОБ'ЄКТА ТА КІЛЬКІСНИЙ МЕТОД ВИЗНАЧЕННЯ ДОПУСТИМОГО РІВНЯ АВТОНОМНОСТІ

Анотація. Посилення взаємозв'язку між інформаційними мережами та системами управління на енергетичних об'єктах підвищує вимоги до швидкості виявлення кіберзагроз, водночас створюючи ризик того, що автоматизовані заходи реагування можуть негативно вплинути на технологічний процес. У статті запропоновано концептуальну архітектуру агентної системи кіберзахисту, що функціонує на базі цифрового двійника енергетичного об'єкта, а також кількісний метод визначення допустимого рівня автономності захисних дій. Запропонована архітектура об'єднує підсистеми збору та нормалізації телеметрії, граф активів, агентів моніторингу й планування, а також детермінований механізм перевірки та виконання політик безпеки. Цифровий двійник слугує контрольованим ізольованим середовищем для моделювання станів системи та сценаріїв атак, а також для валідації рішень перед їх застосуванням у реальній (операційній) мережі. Метод враховує критичність активів, масштаб потенційних кіберфізичних наслідків, граничні значення похибок (на етапах виявлення, атрибуції, планування та виконання), якість телеметрії, а також розбіжність між цифровим двійником і фізичною системою. Для дозволених дій розраховується безрозмірний «індекс ризикового тиску», на основі якого визначається рівень автономності — від A0 до A4. Детерміновані обмеження безпеки мають пріоритет: команди, заборонені політикою або відсутні у «білому списку», блокуються незалежно від значення індексу; крім того, автономне виконання дій забороняється у разі критичної втрати телеметрії або виявлення неідентифікованої атаки. Система передбачає підтвердження дій оператором, суворе розмежування операційного та тестового середовищ, використання гістерезису та ведення журналу аудиту. Наукова новизна підходу полягає в поєднанні формалізованих правил делегування дій агенту з перевіркою обмежень безпеки та врахуванням похибок цифрового двійника — у межах єдиної системи політик. Наведено ілюстративні сценарії та план подальшої валідації, що включає порівняння з ручним реагуванням і системами, які базуються на жорстких правилах, а також перевірку можливості скасування безпечних дій та оцінювання невизначеності. Стаття має концептуально-методологічний характер: наведені параметри та порогові значення є ілюстративними; експериментальні результати не розглядаються. Питання практичної доцільності, безпеки та впливу на час реагування потребують уточнення після впровадження цифрового двійника та проведення незалежного тестування в ізольованому середовищі.

© Бурченко С.Б., 2026
Copyright for this paper by its authors. Use permitted under License CC BY 4.0.
ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Ключові слова: Agentic AI, цифровий двійник, кібербезпека, OT security, енергетика, автономне реагування, кіберфізичний ризик, IEC 62443.

1. Вступ. Енергетична інфраструктура дедалі частіше об'єднує корпоративні IT мережі, диспетчерські системи та польові контролери. У такому середовищі швидке виявлення аномалій є необхідним, але автономна дія штучного інтелекту може сама створити небезпечний технологічний стан. Тому практичне питання полягає не лише у точності моделі, а й у тому, який клас дій вона має право виконувати щодо конкретного активу [1–3, 17, 18].

Мета роботи — запропонувати архітектуру агентного кіберзахисту й відтворюване правило обмеження автономності за критичністю активу, помилками агентного контуру, наслідками та похибкою цифрового двійника. Стаття є методологічною пропозицією: експеримент і валідацію ще не проведено; числові сценарії є лише умовними прикладами [7, 11, 12].

2. Пов'язані стандарти та обмеження. Архітектуру узгоджено з принципом сегментації зон і каналів IEC 62443 та з підходом NIST до виявлення, реагування і відновлення. Модель не отримує прямого привілейованого каналу до пристроїв безпеки. Захисні SIS, аварійне відключення, релейний захист і будь-які команди, що змінюють уставки або стан силового обладнання, виключено з автономного контуру. [1–3, 7, 10]

Передбачено використання цифрового двійника як ізольованого середовища для перевірки політик. Планується моделювати топологію, телеметрію та зв'язки між кіберподіями і технологічними станами; відповідність моделі фізичному процесу потребує валідації [11, 17].

3. Огляд сучасних досліджень і наукова прогалина. Огляди 2024–2025 років підтверджують, що цифрові двійники дедалі частіше використовують як середовище для моделювання атак, прогнозування стану та перевірки захисних політик [11, 12]. Для енергетики це поєднання особливо важливе: кіберінцидент може впливати не лише на конфіденційність або доступність даних, а й на безпечність технологічного процесу, безперервність електропостачання та ресурс обладнання [17, 18].

Паралельно формується напрям agentic AI для операційної підтримки захисту. Праці з критичної інфраструктури наголошують на потенціалі LLM-агентів для кореляції доказів, автоматизованої аналітики та оцінювання сценаріїв [12]. Однак нові огляди загроз до AI-агентів демонструють, що автономний контур сам розширює площину атаки: prompt injection через неочищені журнали, компрометація інструментів, tool abuse, витік секретів, отруєння контексту та помилкова взаємодія агентів [13, 19, 20].

Роботи з human-in-the-loop навчання та пояснюваного AI у промислових кіберфізичних системах підтримують принцип делегування дій за умов перевірюваності, зрозумілого обґрунтування й можливості скасування [14, 15]. Окремі дослідження часових аномалій створюють методичну основу для оцінювання помилки детекції у технологічній телеметрії [16]. Разом ці праці пояснюють, чому точність детектора не може бути єдиною підставою для автономного реагування.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Аналіз використаних джерел виявляє потребу в подальшому дослідженні правила делегування дій, яке одночасно враховує критичність активу, кіберфізичний наслідок, помилки агентного контуру та розбіжність цифрового двійника зі стендом. Запропонований індекс є дослідницькою гіпотезою; повну новизну відносно всієї літератури необхідно підтвердити систематичним оглядом.

4. Експериментальна архітектура. Запропонована архітектура складається з п’яти логічних шарів: збирання телеметрії, графа активів, агента спостереження, агента планування та виконавця політики. Планується використовувати журнали HMI й інженерної станції, NetFlow, події автентифікації, ознаки OPC UA та Modbus TCP і технологічні теги. Виконавець перевіряє план до застосування дії. Архітектура є проєктом, її працездатність потребує експериментальної перевірки. [3, 11, 12]

Запропонована експериментальна архітектура описана в таблиця 1.

Таблиця 1

Шар	Функція	Приклад контрольного обмеження
Телеметрія	Нормалізація IT та OT подій	Часова синхронізація та контроль повноти
Граф активів	Зв’язок ідентичності, вузла, зони та процесу	Версіонування конфігурації
Агент спостереження	Кореляція аномалій і гіпотез	Не ініціює жодних команд
Агент планування	Вибір оборонної дії та обґрунтування	План оцінюється детермінованою політикою
Виконавець політики	Ізоляція, блокування, ескалація	Заборона запису до SIS та PLC у A0 A3



Рисунок 1 - Архітектура експериментальної Agentic AI системи кіберзахисту

5. Кількісний метод визначення автономності. Для кожної запропонованої дії a щодо активу i визначимо нормовані параметри: C_i — критичність активу, e_m — оці-

нена імовірність помилки моделі в релевантному класі сценаріїв, S_i(a) — масштаб найгіршого правдоподібного кіберфізичного наслідку. Усі величини лежать у діапазоні від 0 до 1. Критичність C_i задає власник процесу з урахуванням ролі активу, резервування та часу відновлення. Значення e_m обчислюють на відкладеній валідаційній вибірці й збільшують калібрувальним запасом. S_i(a) оцінює мультидисциплінарна група за шкалою людей, довкілля, безперервності постачання та обладнання [7, 14, 15].

$$R_i(a) = 1 - [1 - C_i] \cdot [1 - e_m] \cdot [1 - S_i(a)] \quad (1)$$

R_i(a) — безрозмірний евристичний індекс, а не ймовірність збитку. Пороги та шкали є проектними припущеннями. Для ухвалення рішення використовується лише узгоджена процедура розділу 11: жорсткі заборони, перевірка якості даних, обчислення R_safe, пороги, обмеження зони та гістерезис.

Перелік запропонованих дії та контроля наведено у таблиця 2.

Таблиця 2

Рівень	**Дозволений тип дії**	**Людський контроль**
A0	Лише спостереження, журналювання, збагачення контексту	Не потрібен для збору даних
A1	Рекомендація оператору, створення інциденту	Підтвердження перед впливом
A2	Зворотні IT дії: блокування токена, карантин робочої станції	Автоматично в межах політики; повідомлення оператору
A3	Обмежена мережна ізоляція OT з попередньо схваленого набору	Черговий оператор має право негайного скасування
A4	Автономне виконання низькоризикових зворотних дій у тестовій зоні	Постійний аудит; заборонено для виробничих критичних зон

6. Проєкт експерименту та ілюстративні сценарії. Запропоновано створити модель умовної підстанційної ділянки з DMZ, HMI, інженерною станцією, двома PLC, historian і технологічними тегами. Плануються ручна базова процедура, rule-based SOAR та запропонована політика з однаковими вхідними даними. Наведені нижче параметри — умовні приклади; вимірювання часу, реалізацію та прогони ще не виконано [3, 17, 18].

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)



Рисунок 2 - Порядок визначення допустимої автономності

Запропоновані сценарії для експерименту наведені у таблиця 3

Таблиця 3

Сценарій	**C**	**e**	**S**	**R^+**	**Умовний рівень**
Фішинг інженерної станції	0,65	0,08	0,35	0,791	A1
Підміна команд до PLC	0,90	0,12	0,85	0,987	A0: заборона
Компрометація HMI сесії	0,70	0,10	0,45	0,852	A1
Сканування OT сегмента	0,35	0,06	0,20	0,511	A2*

** За умови $F \leq R^+$, відсутності штрафів, дозволеної зворотної IT дії та виконання всіх safety gates. Приклади не є результатами експерименту.*

У майбутньому експерименті для кожного сценарію слід зафіксувати початковий стан, подію, часові мітки, дію політики, підтвердження оператора, наслідок і можливість скасування. Час локалізації та частку коректних ескалацій буде визначено лише після проведення вимірювань.

7. Обговорення. Метод корисний тим, що робить межу автономності перевірюваною: кожна дія має простежуваний набір вхідних оцінок, розрахунок і правило відсікання. Однак e_m не є сталою характеристикою моделі. Вона залежить від дрейфу даних, зміни конфігурації, неповноти журналів та поведінки супротивника. Тому значен-

ня має переглядатися після істотних змін у мережі, моделі або процесі, а також при погіршенні якості спостережень [7, 14, 15].

Запропонована формула не замінює аналіз небезпек. Вона є політичним інструментом для вибору обсягу делегування дій агенту. У реальному розгортанні необхідні незалежна валідація цифрового двійника, Red Team перевірка, двоканальний журнал аудиту, аварійне скасування дій та погодження з відповідальними за функціональну безпеку [1–3, 7, 10].

8. Невизначеність, калібрування та аналіз чутливості

Оцінки похибок слід визначати окремо для кожного етапу й умовного набору даних. Наприклад, для частоти пропусків атак p_FN за незалежних бінарних спостережень і апріорного розподілу Beta(1,1), після FN пропусків та TP правильних виявлень апостеріорний розподіл дорівнює Beta(FN+1, TP+1). Для false positive знаменником є FP+TN; для помилок атрибуції, планування та виконання потрібні власні визначення і вибірки. За кластерних або залежних даних припущення незалежності необхідно перевіряти [7, 16].

Для ілюстративного аналізу чутливості використовуємо базовий індекс R^{+}..

Калібрування прогнозних імовірностей можна оцінювати за Brier score та ECE на окремому наборі, вказуючи правила групування і розміри вибірки. Параметри δ_cal, δ_obs та δ_twin є запропонованими запасами, а не стандартизованими ймовірностями; їх значення й спосіб калібрування слід визначити до експерименту. За критичної втрати телеметрії дозвіл на дію не видається [7, 11].

Джерела невизначеності, операційні індикатори та вплив на рішення наведено у таблиця 4.

Таблиця 4

Джерело невизначеності	**Операційний індикатор**	**Вплив на рішення**
Статистична похибка моделі	Верхня 95% квантиль Beta розподілу	Заміна e_m на $E_{0.95}$(e_m)
Некалібровані оцінки	ECE > τ_ECE	Додавання δ_cal
Пропуски телеметрії	Частка відсутніх критичних подій	Додавання δ_obs або пониження до A1
Дрейф конфігурації	Зміна графа активів або версії PLC	Примусова переоцінка C і S
Помилка цифрового двійника	Розходження з тестовим стендом	Зменшення максимального рівня автономності

Нижче наведено аналітичні приклади чутливості індексу; це не результати Монте-Карло. За C=0,60, e=0,12, S=0,40 маємо R^{+}=0,7888 ≈ 0,789. Якщо F ≤ 0,789 і жодна жорстка заборона не спрацювала, R_safe=0,789 і попередній рівень — A1. Розрахунки

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

округлено до трьох знаків; повний аналіз розподілів заплановано як окремий експеримент.

Аналітичні приклади чутливості індексу наведено у таблиця 5.

Таблиця 5

Варіація параметра	**R^+**	**Наслідок для політики**
C: 0,60 → 0,80 при E=0,12, S=0,40	0,789 → 0,894	Жорсткий поріг: максимум A1
E: 0,12 → 0,25 при C=0,60, S=0,40	0,789 → 0,820	Пониження або збереження низького рівня
S: 0,40 → 0,70 при C=0,60, E=0,12	0,789 → 0,894	Жорсткий поріг: максимум A1
δ_obs: 0,00 → 0,15	0,789 → 0,825	A1 за наведених умов; критична втрата телеметрії блокує дію

9. Протокол валідації та критерії приймання. Валідація має проводитися послідовно у чотирьох середовищах: offline набір подій, цифровий двійник, ізольований hardware in the loop стенд і пілотний режим спостереження без виконання дій. Перехід між етапами можливий тільки після документованого виконання критеріїв приймання. Жоден позитивний результат симуляції сам по собі не надає права на дію у виробничому середовищі[3, 7, 11, 14].

Перелік етапів валідації, мінімальних доказів та критеріїв переходу між етапами наведено у таблиця 6.

Таблиця 6

Етап	**Мінімальний доказ**	**Критерій переходу**
Offline валідація	Матриця помилок за класами атак, ECE, Brier score	Верхня межа помилки в межах політики
Цифровий двійник	Відтворювані атаки та регресійні сценарії	Нуль спостережених заборонених дій у всіх тестах; звіт про обсяг і статистичну межу
HIL стенд	Перевірка часових затримок і безпечного скасування	Скасування дії не порушує технологічний інваріант
Shadow mode	Рішення агента порівнюються з рішенням SOC	Погоджена точність ескалації та повний аудит
Обмежене впровадження	Лише A1 A2 для низькоризикових активів	Періодичний перегляд ризику та журнал винятків

Технологічні інваріанти визначаються до тесту разом з власником процесу: наприклад, заборона зміни командного каналу до PLC, неперевищення затримки телемет-

рії та збереження аварійного маршруту оператора. Кожен план агента зберігається з версією моделі, значеннями C, e_eff, S, R^+, обраним рівнем A, доказами вхідних подій і результатом виконання. Це забезпечує відтворюваність і можливість незалежного аудиту. [1, 3, 7]

10. Обмеження моделі та управління ризиком. Мультиплікативна форма є евристикою, а не оцінкою ймовірності шкоди. Для C, e, S ∈ [0,1] індекс R^+ не менший за C або S, тому $\max(R^+, C, S, F)$ математично еквівалентний $\max(R^+, F)$. Для уникнення подвійного обліку у базовій політиці fidelity gap враховується окремо через F; δ_twin у e_eff не додається [7, 11].

Окремим класом є adversarial та novel attacks, для яких історична помилка моделі не є репрезентативною. Для них діє default deny: агент може виконувати A0 або A1, доки новий клас не пройде повний цикл валідації. Рекомендовано щоквартальний перегляд калібрування, повторне моделювання після зміни топології та незалежну перевірку policy enforcement point [12, 13, 19, 20].

11. Ризик-обмежена автономність і safety gates. Політика застосовується в такому порядку. (1) Для дії, забороненої або відсутньої у білому списку, G(a)=0: виконання блокується, система переходить до A0 і створює повідомлення оператору. (2) У разі втрати критичної телеметрії або невідомого класу атаки автономне виконання блокується; можлива рекомендація A1. (3) Лише для дозволених дій розраховується R_safe і застосовується таблиця порогів. Доступ до SIS, релейного захисту, аварійного вимкнення та зміни уставок або програми PLC заборонено автономному контуру [1–3, 7, 13]

Визначимо e_det, e_attr, e_plan та e_exec як частоти помилок виявлення, атрибуції, планування і виконання на окремих релевантних вибірках. За припущення незалежності формула $1-\prod(1-e_j)$ описує ймовірність принаймні однієї помилки; незалежність не гарантує консервативності при залежних етапах. Без такого припущення може застосовуватися верхня межа об'єднання $\min(1, \Sigma\, e_j)$, якщо e_j оцінюють імовірності відповідних подій в одному визначеному операційному розподілі. Для невизначених параметрів використовують одночасні верхні межі з документованим рівнем покриття [7, 13, 16].

$$e_eff = \min(1, U_det + U_attr + U_plan + U_exec + \delta_cal + \delta_obs) \quad (2)$$

тут U_j — одночасні верхні межі частот помилок за визначених припущень і спільного рівня покриття; F_i(a) — нормована розбіжність двійника зі стендом, оцінена для відповідної дії. Межі, шкали та пороги є проектними й потребують калібрування. Для ОТ дій використовується $R_safe = \max(R^+, F_i(a))$; це індекс для вибору політики, а не кількісна гарантія безпеки [7, 11].

$$R^+ = 1 - (1-C_i)(1-e_eff)(1-S_i(a)); \quad R_safe = \max(R^+, F_i(a)) \quad (3)$$

Умови виконання або блокування автономних дій наведено у таблиця 7.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Таблиця 7

Умова	Дозволений рівень	Правило виконання
G(a)=0 або R_safe ≥ 0,90	A0	A0: виконання заблоковано; повідомлення оператору
0,55 ≤ R_safe < 0,90	A1	Рекомендація; вплив після підтвердження оператора
0,30 ≤ R_safe < 0,55	A2	Автоматична зворотна IT дія у затвердженій зоні
0,15 ≤ R_safe < 0,30	A3	Лише попередньо схвалена сегментація з негайним скасуванням
R_safe < 0,15 і тестова зона	A4	A4 у тестовій зоні; поза нею максимум A3

Для підвищення автономності R_safe має залишатися нижче відповідного порога щонайменше на h=0,05 протягом визначеного оператором вікна; зниження рівня застосовується негайно. Тривалість вікна підлягає калібруванню. Якщо рівень A4 недоступний поза тестовою зоною, діє верхня межа A3; для критичних виробничих зон застосовуються додаткові затверджені заборони [13, 14].

12. Посилений експериментальний протокол. Оцінювати слід повний агентний цикл, а не лише точність детектора: якість спостереження, точність контексту, безпечність плану, результат виконання, можливість скасування та повноту аудиту. Кожен прогін повинен фіксувати ідентифікатор сценарію, seed, версії моделі й графа активів, значення C, S, F, e_eff, R_safe, спрацьовані safety gates, дії оператора та наслідок. Це забезпечує відтворюваність, порівнюваність і незалежний розбір інцидентів [7, 12–16].

Перелік груп перевірок, сценаріїв та метрик оцінки дій та їх наслідків наведено у таблиця 8.

Кількість прогонів визначають до тесту за цільовою точністю, частотою подій та припущеннями про незалежність; 30 прогонів не є достатньою універсальною гарантією. Для порівняння потрібно фіксувати план вибірки, seed, базові лінії, критерії виключення, медіану, міжквартильний розмах, інтервал невизначеності та величину ефекту. За нуля порушень у n незалежних однаково розподілених тестах наближена верхня одностороння 95% межа дорівнює 3/n; це не доводить нульовий ризик. [7, 11, 14]

Safety case має бути окремим додатком або підрозділом: твердження про допустимість конкретної автономної дії, докази з тестів, припущення, технологічні інваріанти, залишкові ризики й власник кожного контролю. Практичні правила узгоджуються з IEC 62443, NIST CSF 2.0, NIST SP 800-82 Rev. 3 та NIST AI RMF; їх не слід трактувати як заміну вимог функціональної безпеки або юридичної оцінки для конкретної юрисдикції. [1–3, 7, 10, 15]

У такий спосіб експериментальний дизайн пов’язує чотири сучасні напрями: цифровий двійник як контрольоване середовище [11], захист критичної інфраструктури із

застосуванням LLM та агентів [12], стійкість самих агентів до ворожих впливів [13, 20] і людиноцентричне пояснюване ухвалення рішень [14, 15]. Ці джерела обґрунтовують вибір сценаріїв, safety gates, метрик помилки та обов'язкового аудиту.

Таблиця 8

Група перевірок	**Мінімальні сценарії**	**Метрики**
Кіберзахист	Фішинг, lateral movement, підміна команд, OT scanning	TTD, TTContain, precision, recall, false-negative rate
Безпечність агента	Небезпечний план, tool abuse, prompt injection через лог	Частка заблокованих заборонених дій, повнота доказів
Стійкість даних	Пропуски, десинхронізація, telemetry poisoning, дрейф топології	Зміна e_eff, стабільність рівня A, час fail-safe
Кіберфізичний контур	Каскадна відмова, скасування дії, HIL затримка	Порушення інваріантів, успішність rollback, latency
Людський контроль	Порівняння з SOC та rule-based SOAR у shadow mode	Час підтвердження, частка погоджених рекомендацій, workload

13. Висновки. Запропоновано концепцію ризик-обмеженої автономності для агентного кіберзахисту та план її перевірки на цифровому двійнику енергетичного об'єкта. Обмеження безпеки мають пріоритет над евристичним індексом; шкали, пороги й припущення потребують калібрування та незалежної валідації. Без реалізації стенду й експериментальних даних висновок про прискорення реагування або безпечність системи зробити неможливо.

ЛІТЕРАТУРА/REFERENCES

1. IEC 62443-3-3:2013. Industrial communication networks - Network and system security - Part 3-3: System security requirements and security levels.

2. NIST. The NIST Cybersecurity Framework 2.0. NIST CSWP 29, 2024.

3. NIST. Guide to Operational Technology Security. SP 800-82 Rev. 3, 2023.
DOI: 10.6028/NIST.SP.800-82r3.

4. MITRE. ATT&CK for ICS: Knowledge Base of Adversary Tactics and Techniques in Industrial Control Systems, 2025.

5. Cárdenas A. A., Amin S., Sastry S. Research challenges for the security of control systems. Proceedings of HotSec, 2008.

6. Kritzinger W. et al. Digital Twin in manufacturing: A categorical literature review and classification. IFAC-PapersOnLine, 2018.

7. National Institute of Standards and Technology. AI Risk Management Framework 1.0. NIST AI 100-1, 2023.

8. NIST. Secure Software Development Framework SSDF Version 1.1. SP 800-218, 2022.

9. OWASP. Top 10 for Large Language Model Applications. Guidance for LLM application security, 2025.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

10. ISO/IEC 23894:2023. Information technology - Artificial intelligence - Guidance on risk management.
11. Homaei M., Mogollón Gutiérrez Ó., Sancho Núñez J. C., Ávila M. A review of digital twins and their application in cybersecurity based on artificial intelligence. Artificial Intelligence Review, 2024. DOI: 10.1007/s10462-024-10805-3.
12. Yigit Y., Ferrag M. A., Ghanem M. C., Sarker I. H. Generative AI and LLMs for Critical Infrastructure Protection: Evaluation Benchmarks, Agentic AI, Challenges, and Opportunities. Sensors, 2025. DOI: 10.3390/s25061666.
13. Deng Z., Guo Y., Han C., Ma W. AI Agents Under Threat: A Survey of Key Security Challenges and Future Pathways. ACM Computing Surveys, 2025. DOI: 10.1145/3716628.
14. Retzlaff C. O., Das S., Wayllace C., Mousavi P. Human-in-the-Loop Reinforcement Learning: A Survey and Position on Requirements, Challenges, and Opportunities. Journal of Artificial Intelligence Research, 2024. DOI: 10.1613/jair.1.15348.
15. Moosavi S., Farajzadeh-Zanjani M., Razavi-Far R., Palade V. Explainable AI in Manufacturing and Industrial Cyber-Physical Systems: A Survey. Electronics, 2024.
DOI: 10.3390/electronics13173497.
16. Darban Z. Z., Webb G. I., Pan S., Aggarwal C. C. Deep Learning for Time Series Anomaly Detection: A Survey. ACM Computing Surveys, 2024. DOI: 10.1145/3691338.
17. Abdelkader S. M., Amissah J., Kinga S., Mugerwa G. Securing modern power systems: Implementing comprehensive strategies to enhance resilience and reliability against cyberattacks. Results in Engineering, 2024. DOI: 10.1016/j.rineng.2024.102647.
18. Alomari M. A., Al-Andoli M. N., Ghaleb M., Thabit R. Security of Smart Grid: Cybersecurity Issues, Potential Cyberattacks, Major Incidents, and Future Directions. Energies, 2025. DOI: 10.3390/en18010141.
19. Bilal M., Crowcroft J., Wang R., Xu X. Large language models for agentic NetOps and AIOps: Architectures, evaluation, and safety. Computer Science Review, 2026. DOI: 10.1016/j.cosrev.2026.101075.
20. Anbiaee Z., Rabbani M., Mirani M., Piya G. Security threat modeling for emerging AI-agent protocols: A comparative analysis of MCP, A2A, agora, and ANP. Journal of Information Security and Applications, 2026. DOI: 10.1016/j.jisa.2026.104645.

Received 10.07.2026.
Accepted 14.09.2026.
Published 30.10.2026

Experimental agentic ai cyber-defense system on a digital twin of an energy facility and a quantitative method for determining the permissible level of autonomy

Abstract. The increasing interconnection between information networks and control systems in energy facilities raises requirements for the speed of cyber-threat detection, while simultaneously creating a risk that automated responses may adversely affect the technological process. This article aims to develop a conceptual architecture for an agent-based cyber-defense system operating on a digital twin of an energy facility, along with a quantitative method for determining the permissible level of autonomy for defensive actions. The proposed architecture integrates telemetry collection and normalization, an asset graph, a monitoring agent, a planning agent, and a deterministic mechanism for policy verification and execution.

The digital twin serves as a controlled, isolated environment for modeling system states and attack scenarios, as well as for validating decisions before they are applied to the operational network. The method accounts for asset criticality, the scale of potential cyber-physical consequences, upper bounds on errors (in detection, attribution, planning, and execution), telemetry quality, and the divergence between the digital twin and the physical system. For permitted actions, a dimensionless "risk pressure index" is calculated, which is then used to determine an autonomy level ranging from A0 to A4. Deterministic security constraints take precedence: commands prohibited by policy or absent from the allowlist are blocked, regardless of the index value; furthermore, autonomous execution is disallowed in the event of a critical loss of telemetry or an unidentified attack. The framework incorporates operator confirmation, strict separation between operational and test environments, hysteresis, and an audit log. The scientific novelty of the proposed approach lies in combining a formalized rule for delegating actions to an agent with the verification of safety constraints and the consideration of digital twin error—all within a unified policy framework. Illustrative scenarios and a plan for future validation are presented; the latter encompasses comparisons against manual response and rule-based systems, verification of safe action revocation, and uncertainty assessment. The article is conceptual and methodological in nature: the parameters and thresholds are illustrative, and no experimental results are claimed. Practical viability, safety, and the impact on response time remain to be established following the implementation of the digital twin and independent testing in an isolated environment.

Keywords: agentic AI, digital twin, cybersecurity, operational technology security, energy infrastructure, autonomous response, cyber-physical risk, IEC 62443.

Бурченко Сергій – аспірант кафедри кібербезпеки інформаційних систем, мереж і технологій, навчально-науковий інститут комп'ютерних наук та штучного інтелекту, Харківський національний університет імені В.Н. Каразіна.
ORCID: https://orcid.org/0009-0009-4815-5507

Burchenko Serhii – PhD student at the Department of Cybersecurity of Information Systems, Networks and Technologies, Educational and Scientific Institute of Computer Sciences and Artificial Intelligence, V. N. Karazin Kharkiv National University.
ORCID: https://orcid.org/0009-0009-4815-5507

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

ЗМІСТ

Желдак Т.А., Владико О.Б.
Генетичні алгоритми та методи оптимального керування підприємством 3

Гончаров О.А., Філоненко М.М., Васильєва Л.В., Омельчук А.А., Крігер К.О.
Архітектура та оптимізація рекомендаційного модуля електронної бібліотеки закладу вищої освіти в умовах обмежених обчислювальних ресурсів 21

Бондаренко О.В., Бубліков А.В.
Аналіз сучасних підходів до енергоефективного планування використання приміщень будівель 33

Зародюк А.В., Ліхоузова Т.А.
Генерація ефектів реверберації для комп'ютерних ігор 49

Мартиненко А.А.
Інтелектуальні системи підтримки прийняття рішень для ідентифікації культурних цінностей: сучасні підходи та перспективи розвитку 60

Федорчук І., Шкурат О.
Часово-просторове прогнозування цифрових зображень з неоднорідними часовими інтервалами: огляд глибоких нейромережевих підходів 72

Палагута К.О., Савон О.Є.
Інформаційна технологія підтримки навчальних траєкторій на основі рекурентних нейронних мереж і ситуаційно-сценарного підходу 87

CONTENTS

Zheldak T.A., Vladiko O.B.
Genetic algorithms and methods of optimal enterprise management 3

Honcharov O.A., Filonenko M.M., Vasylieva L.V., Omelchuk A.A., Kriher K.O.
Architecture and optimization of the recommendation module of a higher education institution's electronic library under limited computational resources 21

Bondarenko O.V., Bublikov A.V.
Analysis of modern approaches to energy-efficient planning of building space use 33

Zarodiuk A.V., Likhouzova T.A.
Generating reverb effects for computer games 49

Martynenko A.A.
Intelligent decision support systems for identification of cultural values: modern approaches and development prospects 60

Fedorchuk I., Shkurat O.
Spatio-temporal forecasting of digital images with irregular time intervals: a review of deep neural network approaches 72

Palahuta K.O., Savon O.Ye.
Information technology for supporting learning trajectories based on recurrent neural networks and a situation-scenario approach 87

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Рибальченко О.Г., Гриценко А.М., Білашенко С.В., Сорокін В.С.
Проєктування архітектури та моделі даних інформаційно-аналітичної системи моніторингу водних ресурсів 101

Удовик І.М.
Методологія створення інтелектуальних інформаційних технологій розпізнавання та обробки багатоспектральних аерокосмічних зображень 114

Кравчук А.А., Онай М.В.
Проєктування архітектури програмної системи виявлення та блокування HTTP LR-DDoS атак на платформі Kubernetes 129

Бєляков Р.О.
Порівняльний аналіз процесів прогнозної синхронізації та вибору вузлів розподіленої антенної системи на основі виміряних характеристик радіоканалів 152

Островська К.Ю., Самуськов О.Д.
Аналіз сучасних підходів до моніторингу та управління навантаженням в IT-інфраструктурах 175

Аніщенко А.І., Нестеренко М.М., Блажко В.В., Нестеренко Т.М., Бугров Д.Ю.
Автоматизація технологічного обладнання для виробництва декоративних елементів мощення 185

Rybalchenko O., Hrytsenko A., Bilashenko S., Sorokin V.
Designing the architecture and data model of an information and analytical system for water resources monitoring 101

Udovyk I.M.
Methodology for creating intelligent information technologies for recognition and processing of multispectral aerospace images 114

Kravchuk A., Onai M.
Designing the architecture of a software system for detecting and blocking HTTP LR-DDoS attacks on the Kubernetes platform 129

Bieliakov R.
Comparative analysis of predictive synchronization and node selection processes in a distributed antenna system based on measured radio channel characteristics 152

Ostrovska K.Yu., Samuskov O.
Analysis of modern approaches to monitoring and load management in IT infrastructures 175

Anishchenko A.I., Nesterenko M.M., Blazhko V.V., Nesterenko T.M., Buhrov D.Yu.
Automation of technological equipment for the production of decorative paving elements cepstral domain 185

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Бурченко С.Б.
Експериментальна агентна система кіберзахисту на цифровому двійнику енергетичного об’єкта та кількісний метод визначення допустимого рівня автономності 197

Burchenko S.
Experimental agentic ai cyber-defense system on a digital twin of an energy facility and a quantitative method for determining the permissible level of autonomy 197

РЕФЕРАТИ

УДК 004.4: 622.2

Желдак Т.А., Владико О.Б. **Генетичні алгоритми та методи оптимального керування підприємством** // Системні технології. Регіональний міжвузівський збірник наукових праць. – Випуск 5(166). – Том 2. – Дніпро, 2026. – С.3 – 20.

Представлено результати дослідження, спрямоване на оптимізацію розподілу капітальних інвестицій промислового підприємства із застосуванням генетичного алгоритму. Розроблено математичну модель, у якій частки інвестицій розподілені на п’ять підсистем – ресурси, обладнання, технології, економіка та організація, які визначаються шляхом еволюційного пошуку оптимального вектору фінансування. Функцію пристосованості задано у вигляді чистої приведеної вартості (NPV), що дозволяє безпосередньо враховувати вартість грошей.

Бібл. 11, іл. 6, табл. 2.

УДК 004.738.5:004.65

Гончаров О.А., Філоненко М.М., Васильєва Л.В., Омельчук А.А., Крігер К.О. **Архітектура та оптимізація рекомендаційного модуля електронної бібліотеки закладу вищої освіти в умовах обмежених обчислювальних ресурсів** // Системні технології. Регіональний міжвузівський збірник наукових праць. – Випуск 5(166). – Том 2. – Дніпро, 2026. – С.21 – 32.

Робота присвячена розробці та архітектурній оптимізації рекомендаційного модуля для електронних бібліотек закладів вищої освіти, що функціонують в умовах обмежених обчислювальних ресурсів і не мають можливості використовувати ресурсомісткі нейромережеві підходи. Застосовано метод колаборативної фільтрації з інтеграцією експоненційної функції часового згасання.

Бібл. 12.

УДК 681.5:697.1

Бондаренко О.В., Бубліков А.В. **Аналіз сучасних підходів до енергоефективного планування використання приміщень будівель** // Системні технології. Регіональний міжвузівський збірник наукових праць. – Випуск 5(166). – Том 2. – Дніпро, 2026. – С.33 – 48.

Підвищення енергоефективності будівель потребує узгодження між роботою систем опалення, вентиляції та кондиціонування повітря (HVAC) і просторово-часовою організацією використання приміщень. У цій статті систематизовано сучасні підходи до виявлення та прогнозування присутності людей, енергетичного моделювання будівель, прогностичного керування системами HVAC, планування використання приміщень з урахуванням енергоспоживання, а також оптимізації гнучких просторів.

Бібл. 27.

УДК 004.043

Зародюк А.В., Ліхоузова Т.А. **Генерація ефектів реверберації для комп'ютерних ігор** // Системні технології. Регіональний міжвузівський збірник наукових праць. – Випуск 5(166). – Том 2. – Дніпро, 2026. – С.49 – 59.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Розглянуто проблему генерації реалістичних звукових ефектів реверберації, що враховують динамічні зміни ігрового середовища. Метою цієї роботи є підвищення якості ігрового процесу, збільшення рівня занурення гравця та покращення реалістичності віртуальних просторів шляхом розробки ефективного методу процедурної генерації реверберaційних ефектів, який буде забезпечувати реалістичну адаптацію звуку до змін середовища в реальному часі та оптимізувати використання апаратних ресурсів.

Бібл. 9, іл. 4, табл. 3.

УДК 004.8

Мартиненко А.А. **Інтелектуальні системи підтримки прийняття рішень для ідентифікації культурних цінностей: сучасні підходи та перспективи розвитку** // Системні технології. Регіональний міжвузівський збірник наукових праць. – Випуск 5(166). – Том 2. – Дніпро, 2026. – С.60 – 71.

У статті досліджено сучасні підходи до побудови інтелектуальних систем підтримки прийняття рішень для ідентифікації та класифікації культурних цінностей. Проаналізовано обмеження традиційних методів, зокрема k-NN, CNN та статичних моделей представлення графів знань. Обґрунтовано доцільність використання Graph Neural Networks, зокрема GAT і R-GCN, для врахування семантичних зв’язків та неповних даних. Запропоновано мультимодальну архітектуру, що поєднує Vision Transformer, BERT-подібні мовні моделі та графові нейронні мережі у спільному embedding-просторі. Розглянуто використання self-supervised learning, Explainable AI та сучасних функцій втрат. Визначено перспективи застосування LLM, автоматичного поповнення Knowledge Graph, цифрових музейних систем і federated learning.

Бібл. 19, іл. 1.

УДК 004.85:004.932:519.246.8

Федорчук І., Шкурат О. **Часово-просторове прогнозування цифрових зображень з неоднорідними часовими інтервалами: огляд глибоких нейромережевих підходів** // Системні технології. Регіональний міжвузівський збірник наукових праць. – Випуск 5(166). – Том 2. – Дніпро, 2026. – С.72 – 86.

У дослідженні оцінено чотири архітектури глибокого навчання (ConvLSTM, Transformer, Spatio-Temporal GAN та дифузійна модель) для прогнозування послідовностей зображень через нерегулярні часові інтервали. Тестування на наборах Moving MNIST таатмосферних кліматичних даних показало, що детерміновані моделі розмивають траєкторії жорстких об'єктів, тоді як GAN забезпечує високу структурну точність на обох наборах даних. Дифузійна модель чудово генерує траєкторії жорстких об'єктів, але не справляється з плавними атмосферними градієнтами. Запропоновано концепцію Transformer-GAN архітектури для забезпечення високоточного просторово-часового прогнозування цифрових зображень з неоднорідними часовими інтервалами.

Бібл. 15, іл. 10, табл. 2.

УДК 004.8:37.018.43

Палагута К.О., Савон О.Є. **Інформаційна технологія підтримки навчальних траєкторій на основі рекурентних нейронних мереж і ситуаційно-сценарного під-**

ходу // Системні технології. Регіональний міжвузівський збірник наукових праць. – Випуск 5(166). – Том 2. – Дніпро, 2026. – С.87 – 100.

У статті запропоновано інформаційну технологію підтримки навчальних траєкторій студентів у системах дистанційного навчання, що поєднує прогнозування на основі рекурентних нейронних мереж із ситуаційно-сценарним підходом. Розроблено багаторівневу архітектуру технології, яка забезпечує збір і підготовку освітніх даних, прогнозування навчальних результатів, ситуаційну інтерпретацію та формування сценаріїв педагогічної підтримки. RNN-модель використовується для прогнозування підсумкової оцінки та ймовірності успішного завершення курсу. Ситуаційно-сценарний модуль визначає стан навчальної траєкторії та відповідні педагогічні дії. Розроблено програмний прототип, інтегрований з LMS Moodle. Експериментальна апробація підтвердила працездатність запропонованої технології та можливість її використання для своєчасного виявлення ризиків, адаптивної підтримки студентів і прийняття обґрунтованих педагогічних рішень.

Бібл. 10, рис. 6.

УДК 004.65:502.51

Рибальченко О.Г., Гриценко А.М., Білашенко С.В., Сорокін В.С. **Проєктування архітектури та моделі даних інформаційно-аналітичної системи моніторингу водних ресурсів** // Системні технології. Регіональний міжвузівський збірник наукових праць. – Випуск 5(166). – Том 2. – Дніпро, 2026. – С.101 – 113.

Розроблено архітектуру реляційної бази даних інформаційно-аналітичної системи екологічного моніторингу. Збалансовано архітектурний конфлікт між необхідністю збереження безперервного аудиторського сліду водокористування та забезпеченням високої швидкодії багатовимірних аналітичних запитів. Запропоновано концептуальну модель на основі стандарту бітемпоральних таблиць. Для зниження ймовірності транзакційних блокувань застосовано архітектурний патерн CQRS. Фізичне розділення моделей запису та читання реалізовано через матеріалізовані представлення з використанням кластерних стовпчикових індексів. Запропонована структура запобігає критичній обчислювальній деградації та формує технологічне підґрунтя для використання системи як інструменту оперативного контролю водокористування.

Бібл. 10, іл. 4, табл. 5.

УДК 004.89:528.8

Удовик І.М. **Методологія створення інтелектуальних інформаційних технологій розпізнавання та обробки багатоспектральних аерокосмічних зображень** // Системні технології. Регіональний міжвузівський збірник наукових праць. – Випуск 5(166). – Том 2. – Дніпро, 2026. – С.114 – 128.

Розроблено методологію створення інтелектуальних інформаційних технологій розпізнавання та обробки багатоспектральних аерокосмічних зображень. Методологія поєднує синергетичну обробку різнорідних даних, цикл інтеграції даних, представлень і рішень Ψ–Ф–Λ, предметно залежний вибір конфігурації, мультиагентний і мультимодальний аналіз, наскрізний контроль якості й невизначеності, збереження походження

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

та адаптивну реконфігурацію. Структурно-функціональний аналіз підтверджує наявність механізмів повноти, узгодженості, модульності, адаптивності, простежуваності та інтерпретованості.

Бібл. 25, рис. 1, табл. 3.

УДК 004.056.5:004.738

Кравчук А.А., Онай М.В. **Проєктування архітектури програмної системи виявлення та блокування HTTP LR-DDoS атак на платформі Kubernetes** // Системні технології. Регіональний міжвузівський збірник наукових праць. – Випуск 5(166). – Том 2. – Дніпро, 2026. – С.129 – 151.

У статті запропоновано та реалізовано розподілену архітектуру захисту мікросервісів від низькошвидкісних HTTP DDoS-атак у середовищі Kubernetes без модифікації коду цільових застосунків. Рішення базується на асинхронному збиранні атрибутів трафіку прикладного рівня розширенням EnvoyFilter і мережевого рівня засобами eBPF та їхньої передачі до брокера Apache Kafka. У системі реалізовано три сценарії реагування на атаки з оперативним блокуванням зловмисного трафіку безпосередньо на вхідному шлюзі. Експерименти підтвердили високу продуктивність системи: додаткова затримка становить лише 2,933 мс, а час блокування на шлюзі – 5,674 мс. Порівняно з Coraza WAF система демонструє у 12,3 раза меншу затримку та вищу пропускну здатність: 1973 проти 473 запитів за секунду.

Бібл. 36, іл. 3, табл. 3.

УДК 621.396:004.89

Бєляков Р.О. **Порівняльний аналіз процесів прогнозної синхронізації та вибору вузлів розподіленої антенної системи на основі виміряних характеристик радіоканалів** // Системні технології. Регіональний міжвузівський збірник наукових праць. – Випуск 5(166). – Том 2. – Дніпро, 2026. – С.152 – 174.

У статті проведено порівняльний аналіз процесів прогнозної часової, частотної та фазової синхронізації і адаптивного вибору вузлів розподіленої антенної системи з одноранговою координацією. Розглядається система з 32 радіовузлів, які формують когерентний інформаційний сигнал для рухомого приймача в умовах втрати актуальності оцінок синхронізації через рух радіовузлів і приймача, нестабільність опорних генераторів радіовузлів і затримки службового обміну. Для проведення дослідження використано виміряні характеристики радіоканалів DICHASUS для 32 антенних елементів на частоті 1,272 ГГц, траєкторії руху БпЛА із набору DeepSense 6G та модель частотно-фазової нестабільності опорних генераторів радіовузлів, параметри якої визначаються на основі девіації Аллана.

Бібл. 22.

УДК 004.451:004.7

Островська К.Ю., Самуськов О.Д. **Аналіз сучасних підходів до моніторингу та управління навантаженням в IT-інфраструктурах** // Системні технології. Регіональний міжвузівський збірник наукових праць. – Випуск 5(166). – Том 2. – Дніпро, 2026. – С.175 – 185.

У статті проведено аналіз сучасних підходів до моніторингу та управління навантаженням в IT-інфраструктурах. Досліджено традиційні методи моніторингу ресурсів, концепцію Observability, балансування навантаження, автоматичне масштабування, контейнеризацію та застосування методів машинного навчання для прогнозування навантаження і виявлення аномалій. Проаналізовано особливості реактивного, проактивного та інтелектуального управління ресурсами. Запропоновано узагальнену концептуальну модель інтелектуального управління навантаженням, яка передбачає інтеграцію збору телеметричних даних, їх попередньої обробки, прогнозування навантаження, виявлення аномалій та автоматичного прийняття рішень щодо масштабування ресурсів. Проведено порівняльний аналіз основних підходів за критеріями автоматизації, адаптивності, прогнозних можливостей і вимог до даних. Обґрунтовано доцільність переходу від реактивного до проактивного управління IT-інфраструктурами та визначено перспективи використання технологій машинного навчання і AIOps.

Бібл. 7.

УДК 621.867:62-52

Аніщенко А.І., Нестеренко М.М., Блажко В.В., Нестеренко Т.М., Бугров Д.Ю. **Автоматизація технологічного обладнання для виробництва декоративних елементів мощення** // Системні технології. Регіональний міжвузівський збірник наукових праць. – Випуск 5(166). – Том 2. – Дніпро, 2026. – С.185 – 196.

Проведено аналіз сучасних технологій виробництва декоративних елементів мощення та особливостей автоматизації технологічного обладнання. Обґрунтовано доцільність застосування вібропресування й структуру автоматизованої лінії, що охоплює підготовку та дозування сировини, приготування бетонної суміші, формування і транспортування виробів. Розглянуто застосування програмованих логічних контролерів, цифрових датчиків, частотно-регульованих електроприводів і SCADA-систем для керування технологічними процесами та стабілізації якості продукції. Визначено перспективи вдосконалення систем керування із застосуванням промислового Інтернету речей, цифрових двійників і прогнозного технічного обслуговування.

Бібл. 15, іл. 6.

УДК: 004.056.5

Бурченко С.Б. **Експериментальна агентна система кіберзахисту на цифровому двійнику енергетичного об’єкта та кількісний метод визначення допустимого рівня автономності** // Системні технології. Регіональний міжвузівський збірник наукових праць. – Випуск 5(166). – Том 2. – Дніпро, 2026. – С.197 – 208.

У статті запропоновано концептуальну архітектуру агентної системи кіберзахисту, що функціонує на базі цифрового двійника енергетичного об’єкта, а також кількісний метод визначення допустимого рівня автономності захисних дій. Запропонована архітектура об’єднує підсистеми збору та нормалізації телеметрії, граф активів, агентів моніторингу й планування, а також детермінований механізм перевірки та виконання політик безпеки.

Бібл. 15.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

UDK 004.4: 622.2

Zheldak T.A., Vladiko O.B. **Genetic algorithms and methods of optimal enterprise management** // System technologies. № 5(166). – Part 2. – Dnipro, 2026.- P.3 – 20.

The results of a study aimed at optimizing the distribution of capital investments of an industrial enterprise using a genetic algorithm are presented. A mathematical model has been developed in which the shares of investments are distributed among five subsystems - resources, equipment, technology, economy and organization, which are determined by evolutionary search for the optimal financing vector. The adaptability function is given in the form of net present value (NPV), which allows directly considering the value of money.

Bib. 11, ill. 6, tab. 2.

UDC 004.738.5:004.65

Honcharov O.A., Filonenko M.M., Vasylieva L.V., Omelchuk A.A., Kriher K.O. **Architecture and optimization of the recommendation module of a higher education institution's electronic library under limited computational resources** // System technologies. № 5(166). – Part 2. – Dnipro, 2026.- P.21 – 32.

This paper focuses on the development and architectural optimization of a recommendation module for the electronic libraries of higher education institutions that operate under limited computational resources and cannot utilize resource-intensive neural network approaches. A collaborative filtering method is employed, incorporating an exponential time-decay function.

Refs. 12.

UDC 681.5:697.1

Bondarenko O.V., Bublikov A.V. **Analysis of modern approaches to energy-efficient planning of building space use** // System technologies. № 5(166). – Part 2. – Dnipro, 2026.- P.33 – 48.

Improving building energy efficiency requires coordination between HVAC operation and the spatial-temporal organisation of room use. This paper systematises recent approaches to occupancy sensing and prediction, building energy modelling, predictive HVAC control, energy-aware timetabling, and flexible-space optimisation.

Refs. 27.

UDC 004.043

Zarodiuk A.V., Likhouzova T.A. **Generating reverb effects for computer games** // System technologies. № 5(166). – Part 2. – Dnipro, 2026.- P.49 – 59.

The problem of generating realistic reverberation sound effects that take into account dynamic changes in the game environment is considered. The aim of this work is to improve the quality of the gameplay, increase the level of player immersion, and improve the realism of virtual spaces by developing an effective method for procedural generation of reverberation effects, which will provide realistic adaptation of sound to changes in the environment in real time and optimize the use of hardware resources.

Bibl. 9, pic. 4, table 3.

UDC 004.8

Martynenko A.A. **Intelligent decision support systems for identification of cultural values: modern approaches and development prospects** // System technologies. № 5(166). – Part 2. – Dnipro, 2026.- P.60 – 71.

The article examines modern approaches to developing intelligent decision support systems for the identification and classification of cultural heritage objects. The limitations of traditional methods, including k-NN, CNN, and static knowledge graph representation models, are analyzed. The feasibility of using Graph Neural Networks, particularly GAT and R-GCN, to account for semantic relationships and incomplete data is substantiated. A multimodal architecture combining Vision Transformer, BERT-based language models, and graph neural networks within a shared embedding space is proposed. The use of self-supervised learning, Explainable AI, and advanced loss functions is considered. Prospects for applying LLMs, automatic Knowledge Graph enrichment, digital museum systems, and federated learning are identified.

Ref. 19, fig. 1.

UDC 004.85:004.932:519.246.8

Fedorchuk I., Shkurat O. **Spatio-temporal forecasting of digital images with irregular time intervals: a review of deep neural network approaches** // System technologies. № 5(166). – Part 2. – Dnipro, 2026.- P.72 – 86.

This study evaluates four deep learning architectures – a Time-Aware ConvLSTM, Continuous Temporal Transformer, Spatio-Temporal GAN, and Continuous Diffusion Model – for forecasting digital image sequences at irregular time intervals. Tested on Moving MNIST and atmospheric climate records using a stochastic continuous-time masking protocol, the results demonstrate that the ConvLSTM and Transformer severely blur predictions of rigid-body trajectories. In contrast, the GAN achieves robust structural fidelity across both datasets. The CDM excels in modeling rigid-body trajectories but struggles significantly with smooth atmospheric gradients. A Transformer-GAN architecture is proposed to provide high-fidelity spatio-temporal forecasting of digital images with irregular time intervals.

Ref. 15, ill. 10, tab. 2.

UDC 004.8:37.018.43

Palahuta K.O., Savon O.Ye. **Information technology for supporting learning trajectories based on recurrent neural networks and a situation-scenario approach** // System technologies. № 5(166). – Part 2. – Dnipro, 2026.- P.87 – 100.

The article proposes an information technology for supporting students' learning trajectories in distance learning systems, combining forecasting based on recurrent neural networks with a situation-scenario approach. A multi-level architecture of the technology has been developed, providing the collection and preparation of educational data, prediction of learning outcomes, situational interpretation, and generation of pedagogical support scenarios. The RNN model is used to predict the final grade and the probability of successful course completion. The situation-scenario module determines the state of the learning trajectory and the corresponding pedagogical actions. A software prototype integrated with LMS Moodle has been

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

developed. Experimental validation confirmed the effectiveness of the proposed technology and demonstrated its potential for timely risk detection, adaptive student support, and informed pedagogical decision-making.

Ref.10. Fig. 6.

UDC 004.65:502.51

Rybalchenko O., Hrytsenko A., Bilashenko S., Sorokin V. **Designing the architecture and data model of an information and analytical system for water resources monitoring** // System technologies. № 5(166). – Part 2. – Dnipro, 2026.- P.101 – 113.

The architecture of a relational database for an environmental monitoring information and analytical system has been developed. An architectural trade-off between the necessity of maintaining a continuous water usage audit trail and ensuring high performance for multidimensional analytical queries has been balanced. A conceptual model based on the bitemporal table standard is proposed. To reduce the probability of transactional locks, the CQRS architectural pattern is applied. The physical separation of the write and read models is implemented via materialized views using clustered columnstore indexes. The proposed structure prevents critical computational degradation and establishes the technological foundation for utilizing the system as a tool for the operational control of water usage.

Ref. 10, fig. 4, tabl. 5.

UDC 004.89:528.8

Udovyk I.M. **Methodology for creating intelligent information technologies for recognition and processing of multispectral aerospace images** // System technologies. № 5(166). – Part 2. – Dnipro, 2026.- P.114 – 128.

A methodology has been developed for creating intelligent information technologies for recognition and processing of multispectral aerospace images. The methodology combines synergistic processing of heterogeneous data, the Ψ–Φ–Λ cycle of data, representation, and decision integration, task-dependent configuration selection, multiagent and multimodal analysis, end-to-end quality and uncertainty control, provenance preservation, and feedback-driven reconfiguration. Structural-functional analysis confirms mechanisms for completeness, consistency, modularity, adaptability, traceability, and interpretability.

Ref. 25, fig. 1, tab. 3.

UDC 004.056.5:004.738

Kravchuk A., Onai M. **Designing the architecture of a software system for detecting and blocking HTTP LR-DDoS attacks on the Kubernetes platform** // System technologies. № 5(166). – Part 2. – Dnipro, 2026.- P.129 – 151.

This paper proposes and implements a distributed architecture for protecting microservices against low-rate HTTP DDoS attacks in Kubernetes without modifying target application code. The solution relies on asynchronous collection of application-layer attributes via EnvoyFilter and network-layer attributes using eBPF, forwarding them to Apache Kafka. The system implements three attack mitigation scenarios with prompt traffic blocking directly at the ingress gateway. Experiments confirmed the system’s high performance: latency overhead is only 2.933 ms, and gateway-level blocking takes 5.674 ms. Compared to Coraza WAF, the

system demonstrates 12.3 times lower latency and higher throughput: 1,973 versus 473 requests per second.

Ref. 36, fig. 3, tabl. 3.

UDC 621.396:004.89

Bieliakov R. **Comparative analysis of predictive synchronization and node selection processes in a distributed antenna system based on measured radio channel characteristics** // System technologies. № 5(166). – Part 2. – Dnipro, 2026.- P.152 – 174.

This paper presents a comparative analysis of predictive time, frequency, and phase synchronization processes, as well as the adaptive selection of nodes within a distributed antenna system featuring peer-to-peer coordination. The study examines a system comprising 32 radio nodes that generate a coherent information signal for a mobile receiver. This scenario accounts for the degradation of synchronization estimates caused by the motion of both the radio nodes and the receiver, the instability of the nodes' reference oscillators, and delays in control signaling. The research utilizes measured radio channel characteristics (from the DICHASUS dataset) for 32 antenna elements at a frequency of 1.272 GHz, UAV trajectory data from the DeepSense 6G dataset, and a model of frequency-phase instability for the reference oscillators, with parameters derived from the Allan deviation.

Ref. 22.

UDC 004.451:004.7

Ostrovska K.Yu., Samuskov O. **Analysis of modern approaches to monitoring and load management in IT infrastructures** // System technologies. № 5(166). – Part 2. – Dnipro, 2026.- P.175 – 185.

The article analyzes modern approaches to monitoring and load management in IT infrastructures. Traditional resource monitoring methods, Observability, load balancing, autoscaling, containerization, and machine learning methods for workload forecasting and anomaly detection are considered. The features of reactive, proactive, and intelligent resource management are analyzed. A conceptual model of intelligent load management is proposed, integrating telemetry data collection, preprocessing, workload forecasting, anomaly detection, and automated resource scaling. A comparative analysis of the main approaches is performed according to the level of automation, adaptability, forecasting capabilities, and data requirements. The feasibility of transitioning from reactive to proactive IT infrastructure management is substantiated, and the prospects of using machine learning and AIOps technologies are identified.

Ref. 7.

UDC 621.867:62-52

Anishchenko A.I., Nesterenko M.M., Blazhko V.V., Nesterenko T.M., Buhrov D.Yu. **Automation of technological equipment for the production of decorative paving elements** // System technologies. № 5(166). – Part 2. – Dnipro, 2026.- P.185 – 196.

Modern technologies for producing decorative paving elements and approaches to automating the associated technological equipment are analyzed. The suitability of vibropressing is substantiated, along with the structure of an automated production line encompassing

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

raw material preparation and batching, concrete mixing, product forming, and conveying. The application of programmable logic controllers, digital sensors, variable-frequency drives, and SCADA systems for process control and consistent product quality is examined. Prospects for improving control systems through the use of the Industrial Internet of Things, digital twins, and predictive maintenance are outlined.

Ref. 15. Fig. 6.

UDC004.056.5

Burchenko S. **Experimental agentic ai cyber-defense system on a digital twin of an energy facility and a quantitative method for determining the permissible level of autonomy** // System technologies. № 5(166). – Part 2. – Dnipro, 2026.- P.197 – 208.

This article aims to develop a conceptual architecture for an agent-based cyber-defense system operating on a digital twin of an energy facility, along with a quantitative method for determining the permissible level of autonomy for defensive actions.

Ref. 15.

Системні технології

ЗБІРНИК НАУКОВИХ ПРАЦЬ

Випуск 5 (166)
Том 2

Головний редактор: к.т.н., доц. Т.В. Селівьорстова

Відповідальний секретар та технічний редактор: к.т.н., доц. К.Ю. Островська

Здано до набору 19.10.2026. Підписано до друку 30.10.2026.
Формат 60х84 1/16. Друк - різограф. Папір типограф.
Умов. друк арк. – 15,785. Обл.– видавн. арк. – 13,813.
Тираж 300 прим. Замовл. – 05/26

Український державний університет науки і технологій,
ННІ «Дніпровський металургійний інститут»,
кафедра Інформаційних технологій та систем: ІВК «Системні технології»
49600, Дніпро, а/с 493
http://journals.nmetau.edu.ua/index.php/st
Свідоцтво про державну реєстрацію друкованого засобу масової інформації:
Серія КВ № 8684 від 23 квітня 2004 рік

Редакційна колегія

Селівьорстова Тетяна Віталіївна ***(головний редактор)*** доцент, кандидат технічних наук	Український державний університет науки і технологій, ННІ «Дніпровський металургійний інститут», Україна
Гнатушенко Вікторія Володимирівна ***(заступник головного редактора)*** професор, доктор технічних наук	Український державний університет науки і технологій, ННІ «Дніпровський металургійний інститут», Україна
Островська Катерина Юріївна ***(відповідальний секретар)*** доцент, кандидат технічних наук	Український державний університет науки і технологій, ННІ «Дніпровський металургійний інститут», Україна

Члени редакційної колегії

Бердник Михайло Геннадійович професор, доктор технічних наук	Національний технічний університет «Дніпровська політехніка», Україна
Гвоздева Ірина Маратівна професор, доктор технічних наук	Національний університет «Одеська морська академія», Україна
Гнатушенко Володимир Володимирович професор, доктор технічних наук	Національний технічний університет «Дніпровська політехніка», Україна
Редчиць Дмитро Олександрович професор, доктор технічних наук	Інститут транспортних систем і технологій НАНУ, Україна
Svyetlichnyy Dmytro Ph.D, Professor, Full Professor	Akademia Górniczo-Hutnicza im. Stanisława Staszica, Poland
Heipke Christian Dr.-Ing., Professor	Leibniz University Hannover, Germany
Шинкаренко Віктор Іванович професор, доктор технічних наук	Український державний університет науки і технологій, ННІ «Дніпровський інститут інфраструктури і транспорту», Україна
Shults Roman DSc, PhD, Associate Professor	Interdisciplinary Research Center for Aviation and Space Exploration, Kingdom of Saudi Arabia