

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ТАВРІЙСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ В.І. ВЕРНАДСЬКОГО**

Журнал заснований у 1918 році

**ВЧЕНІ ЗАПИСКИ
ТАВРІЙСЬКОГО НАЦІОНАЛЬНОГО УНІВЕРСИТЕТУ
ІМЕНІ В.І. ВЕРНАДСЬКОГО**

Серія: Технічні науки

Том 36 (75) № 4 2025

Частина 2



Видавничий дім
«Гельветика»
2025

Головний редактор:

Кисельов Володимир Борисович – доктор технічних наук, професор, директор Навчально-наукового інституту муніципального управління та міського господарства Таврійського національного університету імені В.І. Вернадського.

Члени редакційної колегії:

Медведєв Микола Георгійович (відповідальний секретар) – доктор технічних наук, професор, завідувач кафедри загальноінженерних дисциплін та теплоенергетики Таврійського національного університету імені В.І. Вернадського;

Бронін Сергій Вадимович – кандидат технічних наук, доцент, доцент кафедри інформаційних систем та технологій Київського національного університету імені Тараса Шевченка;

Домніч Володимир Іванович – кандидат технічних наук, професор, завідувач кафедри автоматизованого управління технологічними процесами Таврійського національного університету імені В.І. Вернадського;

Дехтяр Анатолій Соломонович – доктор технічних наук, професор, професор кафедри архітектурних конструкцій Національної академії образотворчого мистецтва і архітектури;

Дичко Аліна Олегівна – доктор технічних наук, професор, професор кафедри загальноінженерних дисциплін та теплоенергетики Таврійського національного університету імені В.І. Вернадського;

Дубко Валерій Олексійович – доктор фізико-математичних наук, професор, професор кафедри вищої математики Київського національного університету технологій та дизайну;

Єремєєв Ігор Семенович – доктор технічних наук, професор, професор кафедри автоматизованого управління технологічними процесами Таврійського національного університету імені В.І. Вернадського;

Лисенко Олександр Іванович – доктор технічних наук, професор, професор кафедри телекомунікацій Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»

Огородник Станіслав Станіславович – доктор технічних наук, старший науковий співробітник, професор кафедри загальноінженерних дисциплін та теплоенергетики Таврійського національного університету імені В.І. Вернадського;

Сегай Олександр Михайлович – кандидат технічних наук, доцент, доцент кафедри загальноінженерних дисциплін та теплоенергетики Таврійського національного університету імені В.І. Вернадського;

Чумаченко Сергій Миколайович – доктор технічних наук, старший науковий співробітник, завідувач кафедри інформаційних систем Національного університету харчових технологій;

Цомко Олена – доктор філософії по спеціальності «Безпека і управління інформацією», відділення комп'ютерної інженерії, Інститут Міжнародної освіти, Університет Донгсо, Республіка Корея.

Статті у виданні перевірені на наявність плагіату за допомогою програмного забезпечення StrikePlagiarism.com від польської компанії Plagiat.pl.

**Рекомендовано до друку та поширення через мережу Internet
Вченою радою Таврійського національного університету імені В.І. Вернадського
(протокол № 1 від 29 серпня 2025 року)**

Науковий журнал «Вчені записки ТНУ імені В.І. Вернадського. Серія: Технічні науки»
zareєстровано відповідно до Рішення Національної ради України
з питань телебачення і радіомовлення № 1136 від 11.04.2024 року.
Ідентифікатор медіа: R30-04000.

Суб'єкт у сфері друкованих медіа – Таврійський національний університет імені В.І. Вернадського
(просп. Берестейський, 10, м. Київ, 01135, e-mail: crimea.tnu@gmail.com, тел.: +38 (044) 529-05-16).

Мови видання: українська, англійська, польська, німецька,
французька, словацька, румунська, італійська.

Періодичність: 6 разів на рік.

Журнал включено до Переліку наукових фахових видань України (категорія «Б») з технічних наук
(спеціальності: G1 Хімічні технології та інженерія; G4 Енерговиробництво (за спеціалізацією); G5 Електроніка,
електронні комунікації, приладобудування та радіотехніка) відповідно до Наказу МОН України
від 17.03.2020 № 409 (додаток 1), F2 Інженерія програмного забезпечення; F6 Інформаційні системи і технології;
F7 Комп'ютерна інженерія; G7 Автоматизація, комп'ютерно-інтегровані технології та робототехніка;
D3 Менеджмент; G11 Машинобудування (за спеціалізаціями); J5 Морський та внутрішній водний транспорт;
J6 Авіаційний транспорт; J7 Залізничний транспорт; J8 Автомобільний транспорт
відповідно до Наказу МОН України від 02.07.2020 № 886 (додаток 4)

Журнал включено до міжнародної наукометричної бази Index Copernicus International
(Республіка Польща)

Сторінка журналу: www.tech.vernadskyjournals.in.ua

ISSN 2663-5941 (Print)

ISSN 2663-595X (Online)

© Таврійський національний університет ім. В.І. Вернадського, 2025

ЗМІСТ

ІНФОРМАТИКА, ОБЧИСЛЮВАЛЬНА ТЕХНІКА ТА АВТОМАТИЗАЦІЯ

Алтухов О.В., Васильєва Л.В., Алтухов В.О.

КЛАСИФІКАЦІЯ МЕДИЧНИХ ЗОБРАЖЕНЬ З ВИКОРИСТАННЯМ НЕЙРОННИХ МЕРЕЖ
НА БАЗІ ФРЕЙМВОРКУ TENSORFLOW.....1

Баглай О.П.

ПРОГРАМНО-ІМІТАЦІЙНА ОЦІНКА ВПЛИВУ ПРИРОДНИХ ФАКТОРІВ НА ЯКІСТЬ
ЗВАРЮВАННЯ.....10

Бачкала Б.О., Тимошин Ю.А.

МОДЕЛІ ЗАБЕЗПЕЧЕННЯ ВИСОКОЇ ДОСТУПНОСТІ ТА ВІДМОВСТІЙКОСТІ
РОЗПОДІЛЕНИХ СИСТЕМ ЦИФРОВОЇ ІНФРАСТРУКТУРИ.....16

Біщак Д.С., Петрик М.Р.

КОРЕЛЯЦІЯ МІЖ ГРАФОМОТОРНОЮ АКТИВНІСТЮ ТА ЕЕГ У ПАЦІЄНТІВ ІЗ ТРЕМОРОМ
ПРИ ХВОРОБІ ПАРКІНСОНА.....25

Большакова О.О., Кошкін В.К., Арамян А.М., Бандура В.М., Єганов О.Ю.

СУЧАСНІ ПІДХОДИ ДО ПОБУДОВИ ЕФЕКТИВНИХ ТА БЕЗПЕЧНИХ КОМП'ЮТЕРНИХ
СИСТЕМ І МЕРЕЖ.....31

Бугаєва Л.М., Абрамова А.О., Билбас Д.О.

РОЗРОБКА АВТОМАТИЗОВАНОГО ПРОЦЕСУ АБСОРБЦІЙНОГО ОЧИЩЕННЯ ГАЗУ
ВІД CO₂ РОЗЧИНОМ МЕТИЛДИЕТАНОЛАМІНУ.....41

Burchak P.V., Oleshchenko L.M.

MACHINE LEARNING SOFTWARE FRAMEWORK FOR PREDICTING AND OPTIMIZING
WEB APPLICATION PERFORMANCE.....48

Веселовська Г.В., Плахотнюк М.Ф.

СПОСІБ ПІДВИЩЕННЯ ПРОДУКТИВНОСТІ МЕТОДІВ УПРАВЛІННЯ В СКЛАДНИХ
КОМП'ЮТЕРНО-ОРІЄНТОВАНИХ СИСТЕМАХ НАВЧАННЯ НА ОСНОВІ МЕРЕЖНИХ
ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ.....56

Вишневий С.В., Журба А.В.

СИНТЕЗ ВІДЕОПОСЛІДОВНОСТЕЙ, ЩО МІСТЯТЬ ОДНОРІДНУ ТЕКСТУРУ
ІЗ СЕПАРАБЕЛЬНОЮ АВТОКОРЕЛЯЦІЙНОЮ ФУНКЦІЄЮ.....63

Волк М.О., Бугрій А.М., Ковтун Є.І., Самойлов І.А., Брестовицький Р.М., Соробей Б.В.

МЕТОДИ РОЗПОДІЛЕННЯ ОБЧИСЛЮВАЛЬНИХ ЗАДАЧ У ГЕТЕРОГЕННИХ
КОМП'ЮТЕРНИХ СИСТЕМАХ.....73

Гончаров Д.С.

ІНТЕЛЕКТУАЛЬНІ МОДЕЛІ ТА МЕТОДИ МОНІТОРИНГУ СТАНУ ЗДОРОВ'Я ЛЮДИНИ
НА ОСНОВІ ДАНИХ СЕНСОРНИХ МЕРЕЖ.....80

Горбаченко С.А., Овчинников М.Ю.

ЗАХИСТ ІНТЕЛЕКТУАЛЬНОЇ ВЛАСНОСТІ У СФЕРІ ЦИФРОВОГО КОНТЕНТУ:
ВИКЛИКИ ТА РІШЕННЯ ДЛЯ РЕКЛАМНОГО МЕНЕДЖМЕНТУ.....87

Грезіна О.М., Гурін Є.П.

КІБЕРБЕЗПЕКА МОДУЛІВ JAVASCRIPT-ЗАСТОСУНКІВ.....97

Дадиверін В.В., Бісікало О.В.

АНАЛІЗ ТА ІМПЛЕМЕНТАЦІЯ ІДЕЇ НАВЧАННЯ ВЕЛИКИХ МОВНИХ МОДЕЛЕЙ
ЗА АНАЛОГІЄЮ З ДІТЯЧИМ КОГНІТИВНИМ РОЗВИТКОМ.....103

Дука А.І.

DETECTION OF SVD-BASED HIDDEN DATA IN WEB APPLICATIONS USING
WALSH-HADAMARD TRANSFORMANT ANALYSIS.....110

Добровольський Ю.Г., Дячук Р.Л.

ДОСЛІДЖЕННЯ ГЕНЕРАЦІЇ ВИПАДКОВОЇ ПОСЛІДОВНОСТІ ЧИСЕЛ НА ОСНОВІ
ТЕМНОВОГО СТРУМ ФОТОДІОДА ПРИ РІЗНИХ ТЕМПЕРАТУРАХ.....119

Дубко В.О. ЗАСТОСУВАННЯ ІНДИКАТОРНИХ ПРОЦЕСІВ ПРИ МОДЕЛЮВАННІ ТА ДОСЛІДЖЕНІ ПОКАЗНИКІВ СТОХАСТИЧНИХ СИСТЕМ ІЗ ЗМІННОЮ СТРУКТУРОЮ.....	127
Здольник М.О., Степанець О.В. МУЛЬТИАГЕНТНИЙ ПІДХІД ДО АРХІТЕКТУРИ ЛОКАЛЬНОЇ СИСТЕМИ АВТОМАТИЧНОГО КЕРУВАННЯ ОБ'ЄКТАМИ ТЕПЛОПОСТАЧАННЯ ПРОМИСЛОВОГО ТА ЖИТЛОВОГО ПРИЗНАЧЕННЯ.....	133
Зіноватна С.Л., Білий Д.О., Оніщенко Т.В. ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ВЗАЇМОДІЇ ПЛАНУВАННЯ МЕНЮ ТА ВИКОРИСТАННЯ СКЛАДСЬКИХ ЗАПАСІВ.....	142
Зінченко Д.Д., Маріяш Ю.І. КООРДИНОВАНЕ КЕРУВАННЯ ЧАСТОТОЮ CPU ТА GPU ДЛЯ ПІДВИЩЕННЯ ПРОДУКТИВНОСТІ ТА КОМФОРТУ КОРИСТУВАЧА МОБІЛЬНОГО ПРИСТРОЮ.....	148
Піп М.О., Oleshchenko L.M. SOFTWARE-DRIVEN DATA ANALYSIS OF PROPAGANDA NARRATIVES AND SOCIAL MEDIA ENGAGEMENT.....	155
Козуб Г.О., Артеменко О.І., Осадчук С.І. ПАРАЛЕЛЬНЕ ПРОГРАМУВАННЯ В МОБІЛЬНИХ ІГРАХ: МЕТОДИКИ ОПТИМІЗАЦІЇ ГРАФІКИ ТА ІГРОВОЇ ЛОГІКИ.....	163
Комлева Н.О. РОЗШИРЕННЯ БАЗИСУ UML ДЛЯ ПІДТРИМКИ НЕДОСКОНАЛИХ ТА ЗМІНЮВАНИХ ВИМОГ У ПРОЄКТУВАННІ ІНТЕЛЕКТУАЛЬНИХ СИСТЕМ.....	170
Крамар А.В., Кравчук Я.Я., Осадчук С.І. АДАПТИВНЕ НАВЧАННЯ ЧАТ-БОТІВ НА ОСНОВІ ДИНАМІЧНОГО АНАЛІЗУ ПОВЕДІНКОВИХ ПАТЕРНІВ КОРИСТУВАЧІВ.....	177
Кримська А.О., Дідок В.Ю., Столярчук В.В. ОПТИМІЗАЦІЯ ЕНЕРГОЕФЕКТИВНОСТІ В ІНЖЕНЕРНИХ СИСТЕМАХ ПРОМИСЛОВИХ ОБ'ЄКТІВ.....	186
Латипов О.В., Єфименко М.В., Миронова Н.О., Молочков Д.Є., Куляба-Харитоновна Т.І. РОЗРОБКА СИСТЕМИ РОЗПІЗНАВАННЯ ПОЛОЖЕННЯ ОБ'ЄКТІВ У РОБОЧІЙ ЗОНІ ПРОМИСЛОВОГО РОБОТА.....	195
Мовчан К.О. КОМПЛЕКСНИЙ ПІДХІД ДО ЗАБЕЗПЕЧЕННЯ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ В ГЕТЕРОГЕННИХ ЕКОСИСТЕМАХ ІНТЕРНЕТУ РЕЧЕЙ.....	203
Нікітченко М.І. АРХІТЕКТУРА ТА РЕАЛІЗАЦІЯ ПЛАГІНА ІНКРЕМЕНТАЛЬНОЇ ВАЛІДАЦІЇ UML-МЕТАМОДЕЛЕЙ З ДВОМА ПОДАННЯМИ.....	208
Оніщенко Т.В., Оніщенко І.О. ПОРІВНЯЛЬНИЙ АНАЛІЗ ОНЛАЙН-ІНСТРУМЕНТІВ КОНВЕРТАЦІЇ СИСТЕМ ЧИСЛЕННЯ ДЛЯ ВИКОРИСТАННЯ В ОСВІТНЬОМУ КОНТЕКСТІ.....	217
Пахомова В.М., Іванченко Д.В. ДОСЛІДЖЕННЯ СТРУКТУРИ МЕРЕЖІ MPLS НА ЗАЛІЗНИЧНОМУ ТРАНСПОРТІ З ВИКОРИСТАННЯМ ІНТЕЛЕКТУАЛЬНОГО ЗАСОБУ.....	226
Петросян А.Р. МЕТОД ЛОКАЛІЗАЦІЇ ОБ'ЄКТА У ПРОСТОРІ ІЗ ЗАСТОСУВАННЯМ ГЕНЕТИЧНО-ОПТИМІЗОВАНОГО ФІЛЬТРА ЧАСТИНОК.....	233
Правосєров А.О. МОДЕЛЮВАННЯ ПРОЦЕСУ ЗГОРЯННЯ БІОМАСИ В КОТЕЛЬНОЇ УСТАНОВЦІ З МЕТОЮ ПОКРАЩЕННЯ ЕКОЛОГІЧНИХ ХАРАКТЕРИСТИК....	240
Прокопович-Ткаченко Д.І., Хрушков Б.С., Білан М.В., Черкаський О.В., Деркач Я.О. БРОУНІВСЬКИЙ ПІДХІД ДАЙСОНА ТА ЗАСТОСУВАННЯ ТЕОРІЇ ВИПАДКОВИХ МАТРИЦЬ ДЛЯ МОДЕЛЮВАННЯ РИЗИКІВ У КІБЕРФІЗИЧНИХ ТА НЕЙРОМЕРЕЖЕВИХ СИСТЕМАХ....	249

Rolik O.I., Ulianytska K.O., Amons O.A., Boiko O.V., Tsymbal S.I. MODERN INTERPRETATION FOR DATABASE DEVELOPMENT OF THE FIFTH NORMAL FORM.....	257
Рудницький В.М., Бабенко В.Г., Рудницький С.В., Короткий Т.К., Ковтюх В.А. ОСОБЛИВОСТІ ГРУП НЕСИМЕТРИЧНИХ СЕТ-ОПЕРАЦІЙ СИНТЕЗОВАНИХ З ТОЧНІСТЮ ДО ПЕРЕСТАНОВКИ ПЕРШОГО ОПЕРАНДА.....	265
Сичков В.С. ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ ДЛЯ АВТОМАТИЗОВАНОГО ПРОФЕСІЙНОГО НАСТРОЮВАННЯ АКУСТИЧНИХ СТРУННИХ МУЗИЧНИХ ІНСТРУМЕНТІВ: ЕКСПЕРТНИЙ ПОГЛЯД.....	273
Стаценко Д.В., Стаценко В.В., Мельник Г.В., Калашник В.Ю. ЗАСТОСУВАННЯ ШТУЧНОГО ІНТЕЛЕКТУ І ХМАРНИХ БАЗ ДАНИХ: НАПРЯМИ РОЗВИТКУ ТА ПРИКЛАДНІ МОЖЛИВОСТІ.....	279
Стецюк Ю.В., Савенко О.С. ЧАСТКОВО-ЦЕНТРАЛІЗОВАНА СИСТЕМА БЕЗПЕКИ МЕРЕЖЕВИХ ОС З ДИНАМІЧНОЮ ПЕРЕДАЧЕЮ КЕРУВАННЯ МІЖ ЇЇ ВУЗЛАМИ.....	286
Стоцький І.В., Устинов Д.А., Налісний Г.І., Костенко В.П. НЕЙРОМЕРЕЖЕВІ МОДЕЛІ ПРОГНОЗУВАННЯ ЛОГІСТИЧНИХ ПОТРЕБ В УМОВАХ БОЙОВИХ ДІЙ.....	300
Фомін О.С. МОДЕЛІ МАШИННОГО НАВЧАННЯ ДЛЯ РОЗПІЗНАВАННЯ ТИПІВ БПЛА ЗА СПЕКТРАЛЬНИМИ ХАРАКТЕРИСТИКАМИ РАДІОСИГНАЛІВ.....	306
Чепурна О.Є., Черевко Є.В., Лобода Ю.Г., Кухаренко С.В. ЗАСТОСУВАННЯ МАТЕМАТИЧНОЇ СТАТИСТИКИ ДЛЯ АНАЛІЗУ ВЕЛИКИХ ДАНИХ В АЛГОРИТМАХ ОБРОБКИ ІНФОРМАЦІЇ.....	313
Щербина Ю.В., Логінова Н.І., Манаков С.Ю. АНАЛІЗ СУЧАСНИХ СПОСОБІВ КОМП'ЮТЕРНОГО МОДЕЛЮВАННЯ СТОХАСТИЧНИХ ПРОЦЕСІВ.....	321
Яковенко В.О., Прокопович-Ткаченко Д.І., Ульяновська Ю.В. ЕНЕРГЕТИЧНО-ІНФОРМАЦІЙНА МОДЕЛЬ ДЕГРАДАЦІЇ ТЕЛЕКОМУНІКАЦІЙНИХ ВУЗЛІВ ПІД ВПЛИВОМ ГІБРИДНИХ КІБЕРАТАК.....	328
АКТУАЛЬНІ ПИТАННЯ ТЕХНІЧНИХ НАУК	
Вергелес А.В., Юрчак І.Ю. ЛОКАЛЬНА СИСТЕМА ВИЯВЛЕННЯ СЛОВА-ТРИГЕРА НА БАЗІ ESP32 ДЛЯ ГОЛОСОВОГО КЕРУВАННЯ БЕЗ ВИКОРИСТАННЯ ХМАРНИХ СЕРВІСІВ.....	339
Волохов Д.Г., Муляревич О.В. ГІБРИДНА МОДЕЛЬ ПРОГНОЗУВАННЯ ФІНАНСОВИХ РИНКІВ ІЗ ВРАХУВАННЯМ ТОНАЛЬНОСТІ НОВИННИХ ПОВІДОМЛЕНЬ.....	344
ВІДОМОСТІ ПРО АВТОРІВ.....	350

CONTENTS

INFORMATICS, COMPUTER ENGINEERING AND AUTOMATION

Altukhov O.V., Vasylieva L.V., Altukhov V.O.

CLASSIFICATION OF MEDICAL IMAGES USING NEURAL NETWORKS
BASED ON THE TENSORFLOW FRAMEWORK.....1

Bahlai O.P.

PROGRAM-SIMULATION ASSESSMENT OF THE INFLUENCE
OF NATURAL FACTORS ON THE QUALITY OF WELDING.....10

Bachkala B.O., Tymoshyn Yu.A.

MODELS FOR ENSURING HIGH AVAILABILITY AND FAULT
TOLERANCE IN DISTRIBUTED DIGITAL INFRASTRUCTURE SYSTEMS.....16

Bishchak D.S., Petryk M.R.

CORRELATION BETWEEN GRAPHOMOTOR ACTIVITY AND EEG IN PATIENTS
WITH TREMOR IN PARKINSON'S DISEASE.....25

Bolshakova O.O., Koshkin V.K., Aramian A.K., Bandura V.M., Yehanov O.Yu.

MODERN APPROACHES TO BUILDING EFFICIENT AND SECURE COMPUTER SYSTEMS
AND NETWORKS.....31

Bugaieva L.M., Abramova A.O., Bilbas D.O.

DEVELOPMENT OF AN AUTOMATED PROCESS FOR ABSORPTION PURIFICATION
OF CO₂ GAS WITH METHYLDIETHANOLAMINE SOLUTION.....41

Burchak P.V., Oleshchenko L.M.

MACHINE LEARNING SOFTWARE FRAMEWORK FOR PREDICTING AND OPTIMIZING
WEB APPLICATION PERFORMANCE.....48

Veselovska H.V., Plakhotniuk M.F.

THE WAY OF INCREASING THE PRODUCTIVITY OF CONTROL METHODS IN COMPLEX
COMPUTER-ORIENTED TRAINING SYSTEMS BASED ON NETWORK INFORMATION
TECHNOLOGIES.....56

Vyshnevyyi S.V., Zhurba A.V.

SYNTHESIS OF VIDEO SEQUENCES CONTAINING A HOMOGENEOUS TEXTURE
WITH A SEPARABLE AUTOCORRELATION FUNCTION.....63

Volk M.O., Buhrii A.M., Kovtun Ye.I., Samoilo V.I., Brestovyt'skyi R.M., Sorobey B.V.

METHODS OF DISTRIBUTION OF COMPUTATIONAL TASKS IN HETEROGENEOUS
COMPUTER SYSTEMS.....73

Honcharov D.S.

INTELLIGENT MODELS AND METHODS OF MONITORING HUMAN HEALTH
BASED ON SENSOR NETWORK DATA.....80

Gorbachenko S.A., Ovchinnikov M.Yu.

INTELLECTUAL PROPERTY PROTECTION IN DIGITAL CONTENT: CHALLENGES
AND SOLUTIONS FOR ADVERTISING MANAGEMENT.....87

Hrezina O.M., Hurin Ye.P.

CYBERSECURITY OF JAVASCRIPT APPLICATION MODULES.....97

Dadyverin V.V., Bisikalo O.V.

ANALYSIS AND IMPLEMENTATION OF IDEAS FOR TEACHING LARGE LANGUAGE
MODELS BY ANALOGY WITH CHILDREN'S COGNITIVE DEVELOPMENT.....103

Dyka A.I.

DETECTION OF SVD-BASED HIDDEN DATA IN WEB APPLICATIONS
USING WALSH-HADAMARD TRANSFORMANT ANALYSIS.....110

Dobrovolsky Yu.G., Diachuk R.L.

STUDY OF RANDOM NUMBER SEQUENCE GENERATION BASED ON DARK CURRENT
OF A PHOTODIODE AT DIFFERENT TEMPERATURES.....119

Doobko V.O. APPLICATION OF INDICATOR PROCESSES IN MODELING AND RESEARCH OF INDICATORS OF STOCHASTIC SYSTEMS WITH VARIABLE STRUCTURE.....	127
Zdolnyk M.O., Stepanets O.V. MULTI-AGENT APPROACH TO THE ARCHITECTURE OF LOCAL AUTOMATIC CONTROL SYSTEMS FOR INDUSTRIAL AND RESIDENTIAL HEAT SUPPLY FACILITIES.....	133
Zinovatna S.L., Bilyi D.O., Onishchenko T.V. INFORMATION TECHNOLOGY INTERACTION OF MENU PLANNING AND USING OF WAREHOUSE INVENTORIES.....	142
Zinchenko D.D., Mariiash Yu.I. COORDINATED CONTROL OF CPU AND GPU FREQUENCIES TO ENHANCE PERFORMANCE AND USER COMFORT IN MOBILE DEVICES.....	148
Ilin M.O., Oleshchenko L.M. SOFTWARE-DRIVEN DATA ANALYSIS OF PROPAGANDA NARRATIVES AND SOCIAL MEDIA ENGAGEMENT.....	155
Kozub H.O., Artemenko O.I., Osadchuk S.I. PARALLEL PROGRAMMING IN MOBILE GAMES: TECHNIQUES FOR OPTIMIZING GRAPHICS AND GAME LOGIC.....	163
Komleva N.O. EXTENDING THE UML FOUNDATION TO SUPPORT IMPERFECT AND EVOLVING REQUIREMENTS IN THE DESIGN OF INTELLIGENT SYSTEMS.....	170
Kramar A.V., Kravchuk Ya.Ya., Osadchuk S.I. ADAPTIVE LEARNING OF CHATBOTS BASED ON DYNAMIC ANALYSIS OF USER BEHAVIORAL PATTERNS.....	177
Krymska A.O., Didok V.Yu., Stoliarchuk V.V. OPTIMIZATION OF ENERGY EFFICIENCY IN ENGINEERING SYSTEMS OF INDUSTRIAL FACILITIES.....	186
Latypov O.V., Yefymenko M.V., Myronova N.O., Molochkov D.Ye., Kulyaba-Kharytonova T.I. DEVELOPMENT OF A SYSTEM FOR RECOGNISING THE POSITION OF OBJECTS IN THE WORKING AREA OF AN INDUSTRIAL ROBOT.....	195
Movchan K.O. A COMPREHENSIVE APPROACH TO ENSURING INFORMATION SECURITY IN HETEROGENEOUS INTERNET OF THINGS ECOSYSTEMS.....	203
Nikitchenko M.I. ARCHITECTURE AND IMPLEMENTATION OF A PLUGIN FOR INCREMENTAL VALIDATION OF UML METAMODELS WITH TWO VIEWS.....	208
Onishchenko T.V., Onishchenko I.O. COMPARATIVE ANALYSIS OF ONLINE NUMBER SYSTEM CONVERSION TOOLS FOR USE IN AN EDUCATIONAL CONTEXT.....	217
Pakhomova V.M., Ivanchenko D.V. RESEARCH OF MPLS NETWORK STRUCTURE IN RAILWAY TRANSPORT USING INTELLECTUAL PROPERTY.....	226
Petrosian A.R. OBJECT LOCALIZATION METHOD IN SPACE USING A GENETICALLY-OPTIMIZED PARTICLE FILTER.....	233
Pravovierov A.O. MODELING OF THE BIOMASS COMBUSTION PROCESS IN A BOILER PLANT TO IMPROVE ENVIRONMENTAL CHARACTERISTICS.....	240
Prokopovych-Tkachenko D.I., Khrushkov B.S., Bilan M.V., Cherkaskyi O.V., Derkach Ya.O. DYSON'S BROWNIAN APPROACH AND APPLICATION OF RANDOM MATRIX THEORY TO RISKS MODELING IN CYBERPHYSICAL AND NEURAL NETWORK SYSTEMS.....	249

Rolik O.I., Ulianytska K.O., Amons O.A., Boiko O.V., Tsymbal S.I. MODERN INTERPRETATION FOR DATABASE DEVELOPMENT OF THE FIFTH NORMAL FORM.....	257
Rudnytskyi V.M., Babenko V.H., Rudnytskyi S.V., Korotkyi T.K., Kovtiukh V.A. PARTICULAR FEATURES OF GROUPS OF ASYMMETRIC SET OPERATIONS SYNTHESIZED WITH AN ACCURACY OF FIRST OPERAND PERMUTATION.....	265
Sychkov V.S. SOFTWARE FOR AUTOMATED PROFESSIONAL TUNING OF ACOUSTIC STRING MUSICAL INSTRUMENTS: AN EXPERT VIEW.....	273
Statsenko D.V., Statsenko V.V., Melnyk G.V., Kalashnik V.Yu. APPLICATION OF ARTIFICIAL INTELLIGENCE AND CLOUD DATABASES: DIRECTIONS OF DEVELOPMENT AND APPLICATION OPPORTUNITIES.....	279
Stetsyuk Yu.V., Savenko O.S. PARTIALLY CENTRALIZED NETWORK OS SECURITY SYSTEM WITH DYNAMIC CONTROL TRANSFER BETWEEN ITS NODES.....	286
Stotskyi I.V., Ustynov D.A., Nalisnyi H.I., Kostenko V.P. NEURAL NETWORK MODELS FOR FORECASTING LOGISTICS NEEDS IN COMBAT CONDITIONS.....	300
Fomin O.S. MACHINE LEARNING MODELS FOR UAV TYPE RECOGNITION BASED ON THE SPECTRAL CHARACTERISTICS OF RADIO SIGNALS.....	306
Chepurna O.Ye., Cherevko Ye.V., Loboda Yu.G., Kukharenyk S.V. APPLICATION OF MATHEMATICAL STATISTICS FOR BIG DATA ANALYSIS IN INFORMATION PROCESSING ALGORITHMS.....	313
Shcherbyna Yu.V., Loginova N.I., Manakov S.Yu. ANALYSIS OF MODERN METHODS OF COMPUTER MODELING OF STOCHASTIC PROCESSES.....	321
Yakovenko V.O., Prokopovich-Tkachenko D.I., Ulyanovska Yu.V. ENERGY-INFORMATION MODEL OF TELECOMMUNICATION NODE DEGRADATION UNDER HYBRID CYBERATTACKS.....	328
CURRENT ISSUES OF TECHNICAL SCIENCES	
Verheles A.V., Yurchak I.Yu. LOCAL TRIGGER WORD DETECTION SYSTEM BASED ON ESP32 FOR VOICE CONTROL WITHOUT CLOUD SERVICES.....	339
Volokhov D.H., Muliarevych O.V. HYBRID MODEL FOR FINANCIAL MARKET FORECASTING WITH SENTIMENT-AWARE NEWS INTEGRATION.....	344
INFORMATION ABOUT AUTHORS.....	350

ІНФОРМАТИКА, ОБЧИСЛЮВАЛЬНА ТЕХНІКА ТА АВТОМАТИЗАЦІЯ

УДК 004.8

DOI <https://doi.org/10.32782/2663-5941/2025.4.2/01>

Алтухов О.В.

Донбаська державна машинобудівна академія

Васильєва Л.В.

Київський національний економічний університет імені Вадима Гетьмана

Алтухов В.О.

Донбаська державна машинобудівна академія

КЛАСИФІКАЦІЯ МЕДИЧНИХ ЗОБРАЖЕНЬ З ВИКОРИСТАННЯМ НЕЙРОННИХ МЕРЕЖ НА БАЗІ ФРЕЙМВОРКУ TENSORFLOW

У статті розглянуто підхід до класифікації медичних зображень за допомогою нейронних мереж, реалізованих на базі фреймворку TensorFlow. Проаналізовано методи та моделі розпізнавання, класифікації та попередньої обробки зображень. Доведено доцільність використання методів штучного інтелекту під час аналізу медичних зображень, отриманих різними методами, для діагностування та попередження захворювань на ранній стадії розвитку, що забезпечує своєчасне лікування та дозволяє розширити межі використання автоматизованої обробки медичних даних.

Розглянуто використання згорткових нейронних мереж (ЗНМ) глибокого навчання для аналізу та класифікації зображень. Наведено переваги використання ЗНМ у порівнянні з класичними методами класифікації зображень. Проаналізовані сучасні дослідження в даній галузі. Розглянуто використання згорткових нейронних мереж з різними за розмірами наборами даних (датасетами) для навчання.

Для навчання нейронної мережі обрано набір зображень (датасет) рентгенографічних зображень здорових та хворих пневмонією легень людини. Обраний датасет розділений на частини для навчання, тестування та валідації нейронної мережі.

Запропоновано архітектуру програмного забезпечення для автоматизованої обробки медичних даних. Розроблено діаграми прецедентів та класів системи класифікації медичних зображень з використанням нейронної мережі. Розроблено програмне забезпечення на мові Python з використанням фреймворку TensorFlow для навчання нейронної мережі, яке дозволяє змінювати параметри дослідження для пошуку оптимальних параметрів навчання нейронної мережі. Розроблено програмне забезпечення для класифікації зображень рентгенографії для діагностування діагнозу пневмонія.

Проведені емпіричні дослідження зі зміною параметрів нейронної мережі (кількість епох та кількість кроків у епоху) для оптимізації параметрів навчання моделі та визначення її точності. У ході досліджень встановлено, що максимальна точність класифікації при навчання досягається з кількістю епох 20. Проведене дослідження впливу кількості кроків навчання на епоху показало, що максимальна точність навчання досягається при 326 кроках за епоху.

Ключові слова: медичні зображення, класифікація, нейронна мережа, TensorFlow, Python, глибоке навчання.

Постановка проблеми. З розвитком цифрових технологій і методів машинного навчання зростає роль автоматизованих систем в області медичної діагностики. Одним із перспективних напрямків

є застосування методів глибокого навчання, зокрема згорткових нейронних мереж (ЗНМ), для класифікації медичних зображень. Класифікація дозволяє точно і швидко розпізнавати патології

на основі рентгенівських знімків, мамограм, МРТ та інших типів медичних візуалізацій. Це сприяє зменшенню навантаження на медичних фахівців, підвищує точність діагностики і дозволяє раннє виявлення захворювань.

Саме тому актуальними є створення гнучких, ефективних і адаптивних рішень, що дозволяють масштабовано впроваджувати інтелектуальні системи в медичну практику.

Для вирішення цього завдання виникає потреба в аналізі сучасних методів класифікації зображень, вивченні принципів побудови математичної моделі класифікатора, а також створенні програмного забезпечення для автоматичної класифікації медичних зображень з використанням нейронних мереж на базі фреймворку TensorFlow. Такий підхід має як наукову, так і прикладну цінність, відкриваючи перспективи для створення систем підтримки прийняття рішень у галузі охорони здоров'я.

Аналіз останніх досліджень і публікацій. Класифікація зображень – одна з ключових задач комп'ютерного зору, яка передбачає автоматичне визначення класу об'єкта, що зображений на цифровому знімку. У випадку медичних зображень це завдання має особливу важливість, оскільки правильне розпізнавання патологій може напряму впливати на якість і своєчасність діагностики.

Класифікація медичних зображень за допомогою нейронних мереж широко використовується у комп'ютерній діагностиці, пропонуючи високу точність та ефективність порівняно з традиційними методами. В [1] підкреслюється, що традиційні методи досягли своєї межі продуктивності, а їх використання вимагає великих витрат часу та зусиль на вилучення та вибір ознак класифікації. Згорткові нейронні мережі є домінуючим підходом, причому останні досягнення зосереджені на гібридних моделях, ансамблях та адаптаціях для невеликих наборів даних та різноманітних модальностей. ЗНМ перевершують традиційне машинне навчання, автоматично витягуючи відповідні ознаки, зменшуючи потребу в ручному проєктуванні ознак [2, 3, 4]. Згорткові нейронні мережі можуть використовуватися і для класифікації цифрових зображень за відповідними категоріями, такими як рентген, КТ, МРТ [5]. Зі зростанням доступності даних медичної візуалізації, досягнення в галузі штучного інтелекту дозволили використовувати ЗНМ для автоматизованої класифікації медичних зображень, допомагаючи клініцистам точно діагностувати захворювання з меншими витратами часу та зусиль [3].

В роботі [6] доводиться, що різні архітектури ЗНМ вивчають різні рівні семантичного представлення зображень, і таким чином ансамбль різних архітектур згорткових нейронних мереж дозволить витягувати ознаки вищої якості. Трансферне навчання та доповнення даних особливо ефективні для невеликих або незбалансованих медичних наборів даних, покращуючи точність класифікації та узагальнюваність [1, 4, 6]. Ансамблеві та гібридні моделі, наприклад, поєднання ЗНМ з локальними бінарними шаблонами або використання кількох архітектур ЗНМ, додатково підвищують продуктивність, особливо у складних або мультимодальних завданнях [6, 7, 8].

У статтях [7, 8] автори представляють рішення, адаптовані до аналізу малих вибірок, що демонструють збалансовану взаємодію між обчислювальною ефективністю та точністю отриманих результатів.

Також розглядаються багатомасштабні рекурентні нейронні мережі (MS-RNN), які стали відомим інструментом для категорії медичних зображень, оскільки вони можуть допомогти клініцистам діагностувати захворювання за медичними зображеннями з більшою точністю. Вони складаються з кількох шарів глибокої згорткової нейронної мережі та рекурентної нейронної мережі (RNN), які можна навчити систематизувати та сприймати медичні фотографії для правильної класифікації [9]. Використовуючи RNN, дослідники можуть навчати глибоку мережу контрольованим способом без необхідності ручної сегментації фотографій [10].

Медичні дослідження генерують різноманітні типи даних, включаючи зображення, отримані за допомогою спеціалізованих методів візуалізації, текстові описи діагнозів та числові набори діагностичних показників. Медичні зображення мають специфічними характеристиками, що відрізняють їх від загальнодоступних наборів даних (наприклад, CheXpert (Stanford AIMI) [11], MedPix [12]). Такі зображення, отримані за допомогою рентгенівських апаратів, КТ, МРТ, УЗД, ЕЕГ та ЕКГ, мають чітку діагностичну спрямованість на виявлення патологій в онкології, неврології, кардіології, дерматології тощо. Вони можуть містити неочевидні візуальні аномалії, які, попри їхню низьку помітність для неозброєного ока, є критично важливими для встановлення медичного діагнозу. Критичними аспектами при підготовці наборів даних (датасет) є баланс між класами та точність ручної анотації зображень. Ці фактори безпосередньо визначають здатність

нейронної мережі до адекватного узагальнення на нових, раніше не бачених вхідних прикладах. Ретельна передобробка даних, включаючи процеси нормалізації, аугментації та усунення шумів, є обов'язковою передумовою для успішного навчання моделі.

Аналіз та обробка різних типів медичних даних застосовуються у навчальному процесі для підготовки бакалаврів та магістрів з біоінженерії при розробці мультидисциплінарних навчальних програм [13, 14]. Вивчення підходів та алгоритмів обробки вхідних даних дозволяє коректно проектувати предметну область та розробляти програмне забезпечення.

Постановка завдання. Метою дослідження є подальший розвиток моделей нейронних мереж для класифікації медичних зображень та допомоги у встановленні діагнозу в автономних системах.

Виклад основного матеріалу. У рамках даного дослідження класифікації медичних зображень було використано датасет Chest X-Ray Images (Pneumonia) [15]. Цей набір даних містить 5 863 зображення, розподілені між двома категоріями (класами): «NORMAL» (для здорових пацієнтів) та «PNEUMONIA» (для пацієнтів з пневмонією). Загальний обсяг датасету становить 1 179 МБ даних у форматі JPEG. Датасет розділений на три групи: навчальну (train), тестову (test) та валідаційну (validate). Кожна група включає зображення обох класів. Оскільки датасет є бінарним, мітки класів безпосередньо представлені назвами директорій, що містять відповідні зображення.

Конкретний розподіл зображень за класами та групами виглядає наступним чином:

- Група train для навчання нейронної мережі включає 1 341 зображення категорії «NORMAL» та 3 875 зображень категорії «PNEUMONIA».

- Група test для оцінки навченої моделі містить 243 зображення категорії «NORMAL» та 390 зображень категорії «PNEUMONIA».

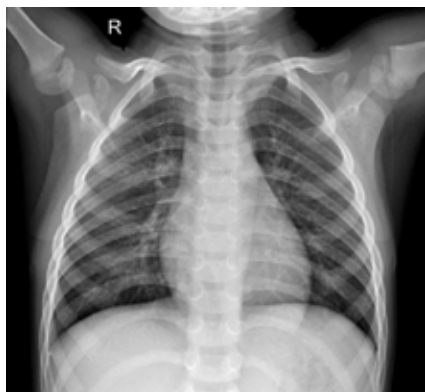
- Група validate для перевірки налаштувань моделі перед фінальним тестуванням складається з 8 зображень категорії «NORMAL» та 8 зображень категорії «PNEUMONIA».

Обраний датасет є достатнім для навчання нейронної мережі з вчителем методом глибокого навчання, таким чином є можливість використовувати моделі нейронних мереж як з одним шаром, так із багатьма. На рис. 1 наведені типові зображення здорових легень (рис. 1 а) та хворих легень людини (рис. 1 б).

Основою побудови системи автоматизованої класифікації зображень є формалізація процесу як математичної задачі. У найпростішому варіанті задача класифікації зображень формулюється як знаходження функції $f: X \rightarrow Y$, де X – простір вхідних зображень, а Y – множина міток класів. Метою є знаходження такої функції f , яка мінімізує функцію втрат $L(f(x), y)$, де $y \in Y$ – правильна мітка для прикладу $x \in X$.

Для розв'язання цієї задачі було обрано згорткову нейронну мережу, яка складається з кількох згорткових шарів, шарів активації (ReLU), пулінгу та повнозв'язних шарів. Така архітектура дозволяє моделі виділяти релевантні ознаки зображень, поступово переходячи від низькорівневих (лінії, кути) до високорівневих (структурні особливості). Навчання мережі виконується методом зворотного поширення помилки з використанням оптимізатора Adam і функції втрат – категоріальної крос-ентропії.

З програмної точки зору було реалізовано програмний комплекс мовою Python з використанням фреймворку TensorFlow. Архітектура системи



а)



б)

Рис. 1. Зображення здорових легень (а) та хворих легень (б) людини

описана за допомогою діаграм UML. Діаграма класів предметної області представлена на рис. 2, діаграма прецедентів – на рис. 3.

Основні модулі:

- med.py – навчання моделі на основі датасету;

- med_test.py – класифікація нових зображень з виведенням результатів.

У логічній моделі передбачено чіткий поділ між процесом попередньої обробки, навчанням та тестуванням. Застосовано об'єктно-орієнтований підхід із забезпеченням масштабованості та роз-

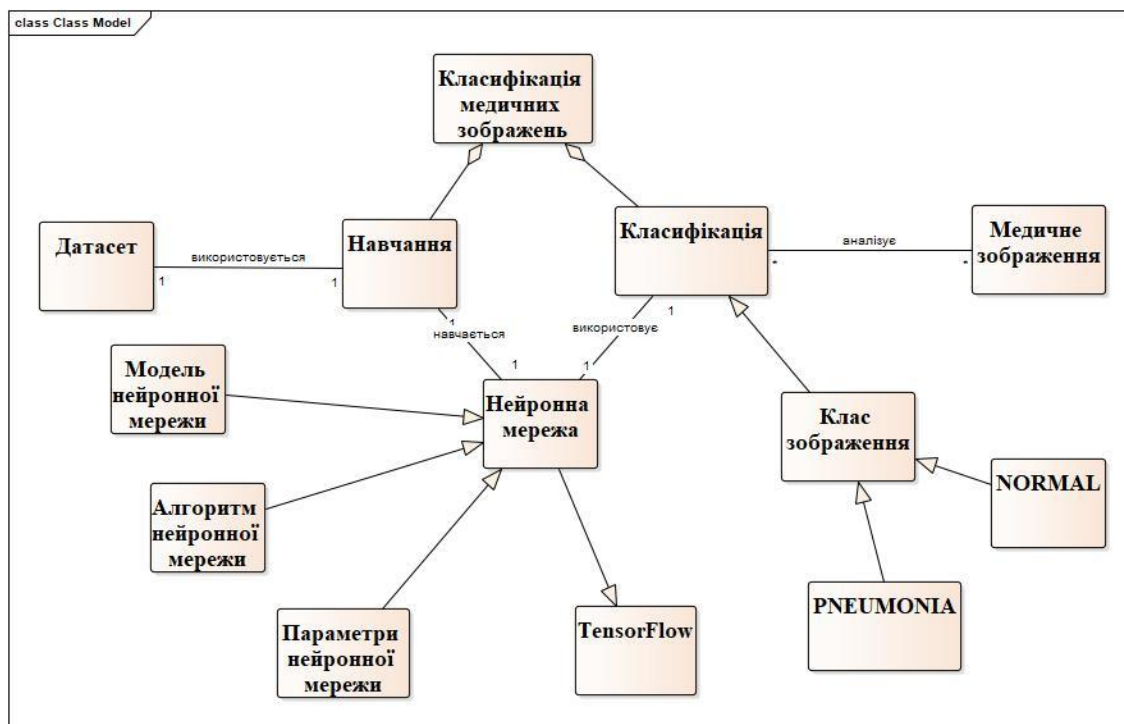


Рис. 2. Діаграма класів предметної області «Класифікація медичних зображень»

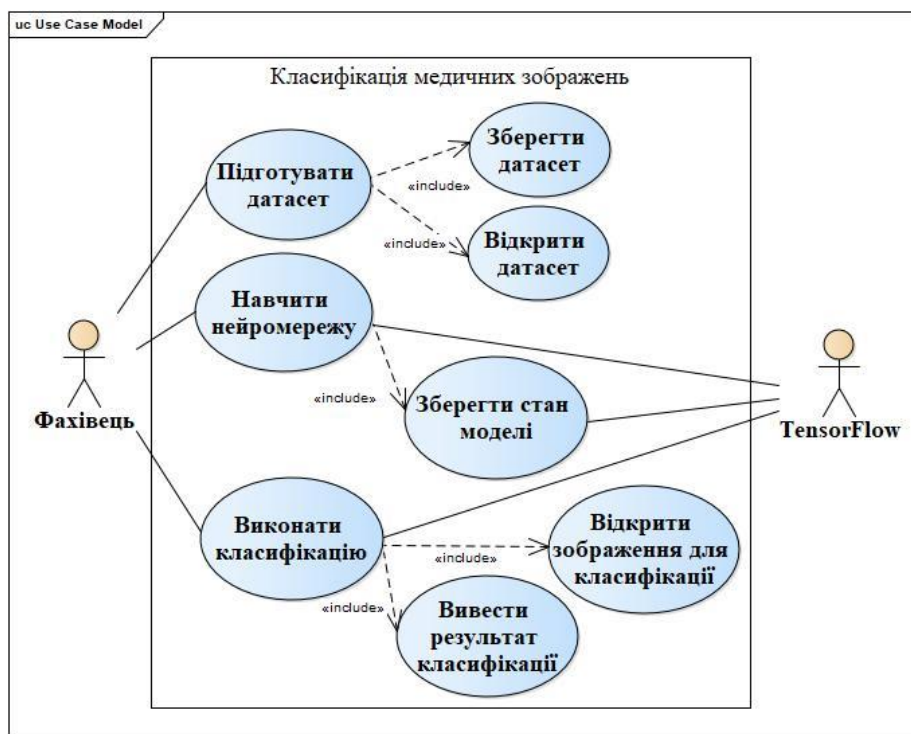


Рис. 3. Діаграма прецедентів предметної області «Класифікація медичних зображень»

ширюваності коду. Така архітектура забезпечує надійність, гнучкість і зручність подальшого доопрацювання системи з можливістю інтеграції нових типів медичних зображень або додаткових модулів аналітики.

Реалізація системи класифікації медичних зображень була виконана у вигляді модульного програмного забезпечення, що складається з двох основних компонентів: модуля навчання (`med.py`) та модуля класифікації (`med_test.py`). Обидва компоненти реалізовано мовою програмування Python із використанням фреймворку TensorFlow та бібліотек Keras, NumPy, OpenCV.

Для навчання та класифікації медичних зображень двох класів обрано модель згорткової нейронної мережі з послідовним порядком шарів (Sequential) (табл. 1).

Така архітектура мережі використовується з різними параметрами навчання (кількість епох, кількість кроків за епоху) для пошуку оптимальних параметрів. Обрана архітектура нейронної мережі створюється за допомогою фреймворку TensorFlow мовою програмування Python.

Модуль `med.py` призначений для навчання згорткової нейронної мережі на основі підготовленого датасету. Він включає завантаження та розмітку зображень, їхню нормалізацію, аугментацію (зсув, обертання, віддзеркалення, масштабування) та побудову архітектури нейронної мережі. Мережа складається з декількох згорткових шарів (Convolution2D), шарів активації ReLU, шарів максимального пулінгу (MaxPooling2D), шару згортання (Flatten) і завершального

повнозв'язаного шару із функцією Softmax для двокласової класифікації.

Навчання відбувається за допомогою оптимізатора Adam, з використанням функції втрат `categorical_crossentropy`. Для запобігання перенавчанню використано механізм `early stopping` і збереження найкращої моделі на валідаційній вибірці (ModelCheckpoint). Програма виводить графіки точності та функції втрат по епохах.

Модуль `med_test.py` виконує завантаження збереженої моделі та здійснює класифікацію зображень, поданих на вхід (одного або групи зображень). Результати класифікації (ймовірність кожного класу) виводяться на екран або записуються у файл. Також реалізовано візуалізацію вхідного зображення з прогнозованою міткою.

Завдяки модульній структурі програмний комплекс може бути адаптований до інших типів медичних зображень або розширений до багатокласової класифікації.

Для перевірки ефективності запропонованої системи класифікації було проведено низку експериментів, спрямованих на вивчення впливу різних параметрів навчання нейронної мережі на точність класифікації медичних зображень.

Навчання здійснювалось на підмножині медичного датасету, сформованого за принципом бінарної класифікації (наявність/відсутність патології). Для тренування мережі використовувались 80% зображень, решта 20% – для валідації та тестування. Було експериментально визначено, що оптимальні результати досягаються при навчанні мережі протягом 20 епох із кількістю кроків у кожній епосі, що дорівнює 342. Застосування таких

Таблиця 1

Опис шарів нейронної мережі

№	Назва шару	Параметри шару	Опис
1	conv2d (Conv2D)	(None, 148, 148, 32)	згортковий шар 1
2	activation (Activation)	(None, 148, 148, 32)	шар активації
3	max_pooling2d (MaxPooling2D)	(None, 74, 74, 32)	об'єднуючий шар
4	conv2d_1 (Conv2D)	(None, 72, 72, 32)	згортковий шар 2
5	activation_1 (Activation)	(None, 72, 72, 32)	шар активації
6	max_pooling2d_1 (MaxPooling2D)	(None, 36, 36, 32)	об'єднуючий шар
7	conv2d_2 (Conv2D)	(None, 34, 34, 64)	згортковий шар 3
8	activation_2 (Activation)	(None, 34, 34, 64)	шар активації
9	max_pooling2d_2 (MaxPooling2D)	(None, 17, 17, 64)	об'єднуючий шар
10	flatten (Flatten)	(None, 18496)	вхідний шар
11	dense (Dense)	(None, 64)	повнозв'язний шар
12	activation_3 (Activation)	(None, 64)	шар активації
13	dropout (Dropout)	(None, 64)	вихідний шар
14	dense_1 (Dense)	(None, 1)	повнозв'язний шар
15	activation_4 (Activation)	(None, 1)	шар активації

параметрів дозволило отримати високу точність класифікації без ознак перенавчання.

Протягом експериментів вивчалися графіки функції втрат (loss) та точності (accuracy) на тренувальному та валідаційному наборах. На етапах перевірки було досягнуто точності понад 94%, що підтверджує ефективність використаної архітектури мережі. Усі результати були виведені у вигляді таблиць та графіків, що дозволяють порівнювати вплив різних гіперпараметрів.

Крім того, протестовано роботу модуля класифікації (med_test.py) на нових зображеннях, відсутніх у навчальному наборі (рис. 4). Система змогла успішно ідентифікувати зображення з високою впевненістю, що свідчить про добру здатність до узагальнення.

Експерименти підтверджують, що запропонована модель може бути використана як основа для впровадження систем підтримки прийняття рішень у медичних установах, зокрема для попереднього автоматичного аналізу діагностичних зображень.

Для пошуку оптимальних параметрів навчання нейронної мережі було проведено навчання з різними параметрами та виконано оцінювання точності класифікації медичних зображень.

Перший параметр, який змінювався при навчанні мережі, це кількість епох, тобто кількість проходів всього набору даних скрізь всі шари мережі. Досліджувались такі значення EPOCHS: 5, 10, 20, 25. Результати досліджень наведені у таблиці 2.

Аналіз результатів навчання мережі з різними значеннями кількості епох показує, що опти-

мальним значенням є 20 епох. Для кожної моделі мережі та датасету кількість епох підбирається окремо. Залежність точності класифікації від значення EPOCHS наведена графіку (рис. 5).

Таблиця 2
Залежність точності класифікації від значення EPOCHS

Кількість епох (EPOCHS)	Точність, %	Час навчання, хв.
5	82.21	7
10	81.25	11
20	87.66	23
25	86.54	28

Дослідження кількості кроків у кожній епохі було проведено зі значеннями 100, 200 та 326. Значення кількості кроків 326 було отримано згідно з рекомендаціями для навчання мереж подібних моделей.

Значення точності мережі в залежності від кількості кроків у епоху наведені у табл. 3. Збільшення кількості кроків навчання у кожну епоху до значень 400 та 500 призводило помилок при навчанні мережі. Таким чином оптимальним значенням кількості кроків навчання для кожної епохи можна вважати 326.

Таблиця 3
Точність мережі в залежності від кількості кроків у епоху (20 епох)

Кількість кроків у епоху	Точність, %	Час навчання, хв.
100	81.25	15
200	86.06	19
326	87.66	23

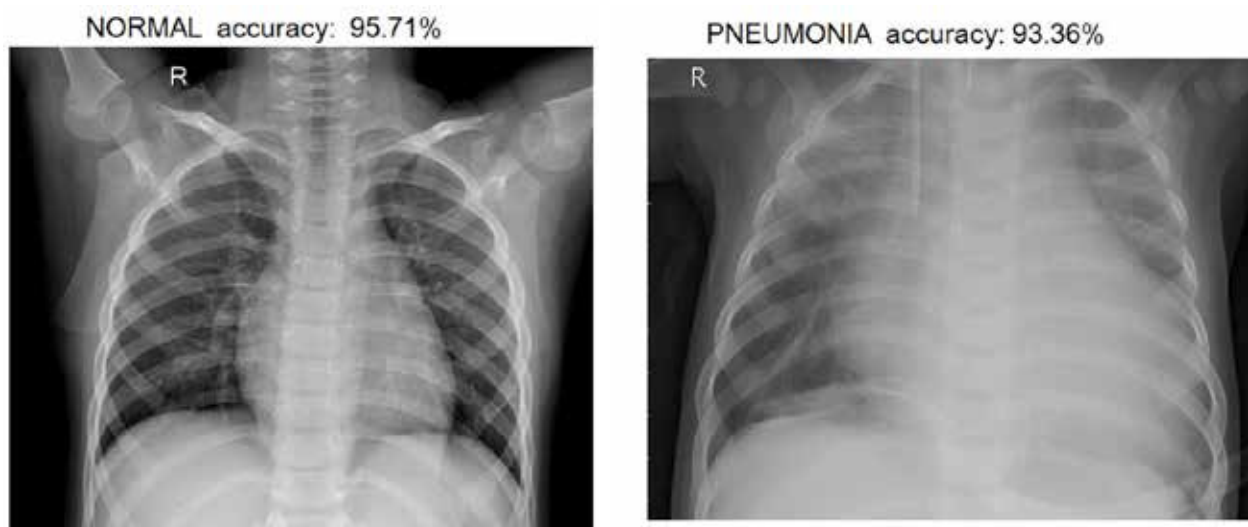


Рис. 4. Результат класифікації медичних зображень

```

nb_train_samples = 5217
batch_size = 16
...
steps_per_epoch=nb_train_samples // batch_size

```

Залежність точності класифікації нейронною мережею від кількості епох навчання

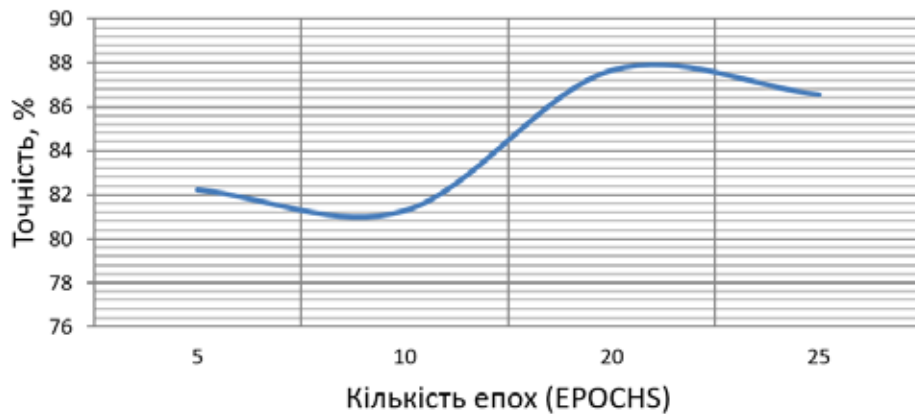


Рис. 5. Залежність точності класифікації нейронною мережею від значення кількості епох

Точність класифікації нейронною мережею в залежності від кількості кроків у епоху (20 епох)

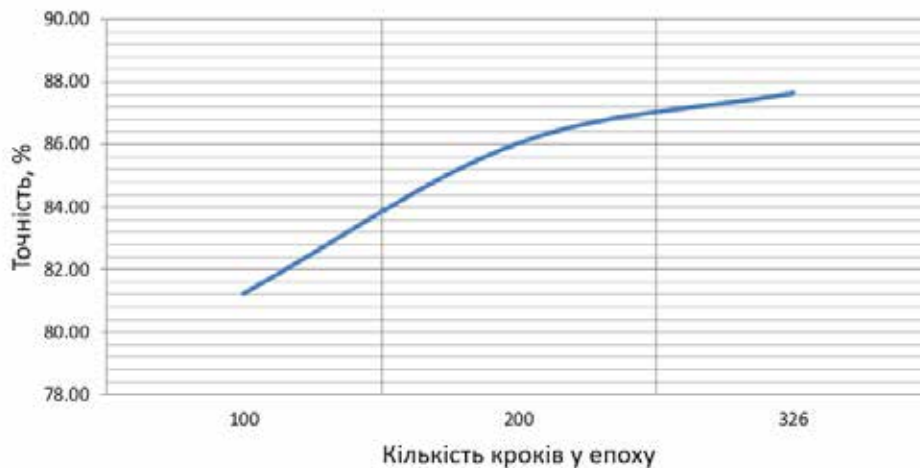


Рис. 6. Точність класифікації нейронною мережею в залежності від кількості кроків у епоху (20 епох)

Залежність точності мережі від кількості кроків у епоху наведена на рис. 6.

Висновки. У статті представлено комплексне дослідження процесу класифікації медичних зображень з використанням методів глибокого навчання на базі фреймворку TensorFlow. Здійснено аналіз сучасних підходів до автоматизованої обробки зображень, зокрема

згорткових нейронних мереж, які довели свою високу ефективність у задачах медичної діагностики.

Було розроблено програмний комплекс, що включає модулі для навчання та тестування нейронної мережі, а також реалізовано архітектурну модель системи з урахуванням вимог масштабованості, гнучкості та зручності модифікації.

Проведено навчання та тестування моделі на базі медичного датасету, що дало змогу досягти високих показників точності – понад 94%, а також виявити оптимальні гіперпараметри навчання (20 епох, 326 кроки на епоху).

Результати експериментів підтверджують доцільність та ефективність застосування штучних нейронних мереж у задачах бінарної класифікації медичних зображень. Запропоноване

рішення може стати основою для впровадження інтелектуальних систем підтримки прийняття рішень у клінічну практику.

У перспективі дослідження передбачає розширення системи до багатокласової класифікації, застосування трансферного навчання та адаптацію до інших типів діагностичних зображень, що відкриває широкі можливості для її використання в різних напрямках медицини.

Список літератури:

1. Yadav S., Jadhav S. Deep convolutional neural network based medical image classification for disease diagnosis. *Journal of Big Data*. 2019. Vol. 6. № 113. DOI: 10.1186/s40537-019-0276-2.
2. Li Q., Cai W., Wang X., Zhou Y., Feng D., Chen M. Medical image classification with convolutional neural network. *13th International Conference on Control Automation Robotics & Vision (ICARCV)*. 2014. P. 844–848. DOI: 10.1109/ICARCV.2014.7064414.
3. Pillai N., Manimala S., Rongali A., Gughani A., Kumar S., Sriramakrishnan G. Automated Classification of Medical Images Using Convolutional Neural Networks. *15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*. 2024. P. 1–6. DOI: 10.1109/ICCCNT61001.2024.10725642.
4. Zhang J., Xie Y., Wu Q., Xia Y. Medical image classification using synergic deep learning. *Medical Image Analysis*. 2019. Vol. 54, P. 10–19. DOI: 10.1016/j.media.2019.02.010.
5. Ghodeswar U., Borkar A., Bagde A. Classification of Different Medical Images Using Neural Network Approach. *Indian Journal Of Science And Technology*. 2022. Vol. 15, № 46, P. 2555–2561. DOI: 10.17485/ijst/v15i46.1160.
6. Kumar A., Kim J., Lyndon D., Fulham M., Feng D. An Ensemble of Fine-Tuned Convolutional Neural Networks for Medical Image Classification. *IEEE Journal of Biomedical and Health Informatics*. 2017. Vol. 31, № 1, P. 31–40. DOI: 10.1109/JBHI.2016.2635663.
7. Huang Z., Zhu X., Ding M., Zhang X. Medical Image Classification Using a Light-Weighted Hybrid Neural Network Based on PCANet and DenseNet. *IEEE Access*. 2020. Vol. 8, P. 24697–24712. DOI: 10.1109/ACCESS.2020.2971225.
8. Rocha M., Landini G., Florindo J. Medical image classification using a combination of features from convolutional neural networks. *Multimedia Tools and Applications*. 2022. Vol. 82, P. 19299–19322. DOI: 10.1007/s11042-022-14206-y.
9. Agarwal P., Nachappa M., Gautam C. Multi-Scale Recurrent Neural Networks for Medical Image Classification. *2024 International Conference on Optimization Computing and Wireless Communication (ICOCWC)*. 2024. P. 1–6. DOI: 10.1109/ICOCWC60930.2024.10470694.
10. Singh U.K., Kavitha R., Saraswat P. Recurrent Neural Networks for Improved Medical Image Classification. *2024 International Conference on Optimization Computing and Wireless Communication (ICOCWC)*. 2024. P. 1–6. DOI: 10.1109/ICOCWC60930.2024.10470908.
11. Stanford Center for Artificial Intelligence in Medicine and Imaging. Datasets. URL: <https://aimi.stanford.edu/shared-datasets> (дата звернення 22.07.2025).
12. The National Library of Medicine. MedPix®. URL: <https://medpix.nlm.nih.gov/home> (дата звернення 22.07.2025).
13. Teaching and subjects on bio-medical engineering. Approaches and experiences from the BIOART-project / Peter Arras and David Luengo (Eds.). Leuven (Belgium), 2021. 391 p.
14. Vasylieva L.V., Mikhieienko D.Y., Getman I.A. Integration of laboratory equipment in remote learning environments. *Proceedings of the 6th International Workshop on Augmented Reality in Education (AREdu 2023)*. 2024. Vol. 3844, pp. 35–44. URL: <https://ceur-ws.org/Vol-3844/paper07.pdf> (дата звернення 22.07.2025).
15. Kermany D.S., et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell*. 2018. Vol. 172 № 5. P. 1122–1131. URL: [https://www.cell.com/cell/fulltext/S0092-8674\(18\)30154-5](https://www.cell.com/cell/fulltext/S0092-8674(18)30154-5) (дата звернення 22.07.2025).

Altukhov O.V., Vasylieva L.V., Altukhov V.O. CLASSIFICATION OF MEDICAL IMAGES USING NEURAL NETWORKS BASED ON THE TENSORFLOW FRAMEWORK

The article considers an approach to medical image classification using neural networks implemented based on the TensorFlow framework. Methods and models of image recognition, classification, and preprocessing are analyzed. The feasibility of using artificial intelligence methods when analyzing medical images obtained by

various methods is proven for diagnosing and preventing diseases at an early stage of development, ensuring timely treatment and expanding the scope of automated medical data processing.

The use of deep learning convolutional neural networks (CNNs) for image analysis and classification is considered. The advantages of using CNNs compared to classical image classification methods are presented. Modern research in this field is analyzed. The use of convolutional neural networks with different sizes of data sets (datasets) for training is considered.

A set of images (dataset) of X-ray images of healthy and pneumonia-infected human lungs was selected for training the neural network. The selected dataset is divided into parts for training, testing, and validation of the neural network.

The software architecture for automated medical data processing is proposed. The diagrams of precedents and classes of the medical image classification system using a neural network are developed. The software in Python was developed using the TensorFlow framework for neural network training, which allows for the changing of study parameters to find the optimal neural network training parameters. The software was developed to classify radiography images and diagnose pneumonia.

Empirical studies were conducted by changing the neural network parameters (number of epochs and number of steps per epoch) to optimize the model training parameters and determine its accuracy. The research found that the maximum classification accuracy during training is achieved with the number of epochs of 20.

The study of the influence of the number of training steps per epoch showed that the maximum training accuracy is achieved with 326 steps per epoch.

Key words: *medical images, classification, neural network, TensorFlow, Python, deep learning.*

Дата надходження статті: 23.07.2025

Дата прийняття статті: 28.07.2025

Опубліковано: 27.10.2025

Баглай О.П.

Український державний університет залізничного транспорту

ПРОГРАМНО-ІМІТАЦІЙНА ОЦІНКА ВПЛИВУ ПРИРОДНИХ ФАКТОРІВ НА ЯКІСТЬ ЗВАРЮВАННЯ

У статті представлено дослідження впливу географічних і кліматичних чинників на якість зварювання металевих конструкцій шляхом створення програмно-імітаційної моделі. Метою роботи є розробка інструменту, що дозволяє кількісно оцінити залежність між природними умовами (зокрема температурою та висотою над рівнем моря) і технологічними параметрами зварювання – середньою напругою, силою струму та імовірністю дефектів у зварних з'єднаннях. Модель побудована на основі відкритих кліматичних, географічних і виробничих даних, що дало змогу здійснити імітаційний аналіз впливу зовнішнього середовища на стабільність та повторюваність параметрів зварного процесу в різних умовах.

У рамках дослідження виконано статистичну обробку даних та побудовано графічні залежності, які відображають варіації параметрів зварювання під впливом температури і висоти. Виявлено клас-тери і тенденції: зокрема, зони температур 0–10°C і висоти 3500–4500 м характеризуються помірними змінами напруги та дефектності, що вказує на потенційний вплив кліматичних умов на якість формування зварного шва. Аналіз графіків свідчить про наявність зв'язку між зовнішнім середовищем і технологічними характеристиками, що може бути враховано при плануванні монтажних робіт у складних географічних зонах.

Запропонована модель підтверджує ефективність застосування інформаційних технологій для вивчення багатofакторного впливу природних умов на виробничі процеси. Результати можуть бути використані для розробки адаптивних рекомендацій щодо зварювання в різних регіонах, оптимізації технологічних режимів, вибору обладнання та підвищення надійності конструкцій. Модель також створює підґрунтя для подальших досліджень із залученням розширених емпіричних даних, розробки превентивних сценаріїв та інтеграції в системи управління якістю монтажу в умовах географічної варіативності.

Ключові слова: зварювання металевих конструкцій, географічні умови, кліматичні фактори, температурний вплив, висота рельєфу, статистичний аналіз, монтажні роботи, якість зварювання, програмне моделювання, стійкість конструкцій.

Постановка проблеми. Якість зварювання металевих конструкцій є критично важливою для безпеки, довговічності та надійності інфраструктурних об'єктів, особливо в умовах впливу зовнішніх факторів. Зварні шви схильні до виникнення залишкових напружень, утворення дефектів, зниження механічної міцності та втомної стійкості через умови навколишнього середовища під час монтажу. У контексті глобальних кліматичних змін зростає актуальність вивчення впливу географічного розташування, включаючи температуру, вологість, атмосферний тиск, висоту над рівнем моря, а також екстремальні погодні явища, на процеси зварювання та монтажу.

У дослідженні Михай А. Попеску (Mihai A. Popescu) та співавторів, яке присвячене вимірюванню залишкових напружень у зварних

з'єднаннях трубчастих сталевих веж вітрових турбін, показано, що залишкові напруження після зварювання мають значний негативний вплив на втомну міцність конструкцій, особливо в умовах циклічного навантаження [23]. Зі свого боку, у роботі Leila Kamalian та колег, яка аналізує вплив кліматично зумовлених екстремальних погодних подій на затримки залізничного транспорту, доведено, що такі явища можуть спричиняти критичні збої в інфраструктурі. Автори підкреслюють важливість прогнозування ризиків за допомогою штучного інтелекту та аналізу великих даних [24]. У сфері зварювання, яка часто виконується в умовах відкритого середовища, подібні ризики залишаються недостатньо дослідженими.

Дослідження Марсель Рік (Marcel Rieck) та співавторів присвячене впливу зовнішнього

середовища на процес підводного зварювання та показує, що навіть зміна середовища з атмосферного на підводне значно впливає на стабільність дуги та характеристики зварного шва, що підтверджує високу чутливість зварювального процесу до зовнішніх умов [26]. У роботі Guido Di Bella та колег, яка стосується аналізу життєвого циклу та ризиків фрикційного зварювання алюмінієвих та сталевих елементів, наголошено на необхідності врахування географічного середовища в оцінках сталості та ефективності з'єднань [25].

З огляду на це, виникає потреба у розробці ІТ-рішень, які дозволяють аналізувати ризики зварювання в різних географічних зонах на основі кліматичних та виробничих даних. Метою дослідження є створення програмної моделі, яка дозволяє оцінити потенційний вплив кліматичних і географічних чинників (зокрема температури та висоти рельєфу) на якість зварювання металевих конструкцій, виражену у вигляді середньої напруги та ймовірності виникнення дефектів у зварних швах. Запропонований підхід дозволяє не лише формалізувати вплив природних умов, а й виявити критичні зони ризику виникнення дефектів, що є особливо важливим у контексті індустріалізації та переходу до стійкої інженерії.

Аналіз останніх досліджень і публікацій. Дослідження, присвячені географічним і кліматичним аспектам впливу на якість зварних з'єднань, підкреслюють важливість урахування зовнішніх природних чинників.

Наукова праця Сун (Sun) та співавторів розглядає атмосферну корозію зварних з'єднань типу Т з різномірних сталей та вдосконалення моделі прогнозування корозії [5]. Автори показали, що навіть незначні відхилення у складі зварного шва можуть призвести до помітного прискорення корозійного руйнування, що є важливим аспектом для оцінки довговічності конструкцій у різних географічних умовах.

Вахіді (Wahidi) із колегами у систематичному огляді присвяченому роботизованим технологіям зварювання в морських конструкціях акцентують увагу на важливості адаптації технологій зварювання до складних умов середовища, таких як висока вологість і солоність [6]. Це підкреслює необхідність врахування кліматичних факторів під час проектування та монтажу металевих конструкцій.

Чень Пенчжао (Chen Pengchao) досліджує новітні методи моніторингу та оцінки стану нафтогазопроводів, включаючи вплив зовнішніх факторів і дефектів у зварних швах [7]. Автор висвітлює досягнення в області інспекційних технологій, що забезпечують високу точність визна-

чення залишкової міцності та прогнозування довговічності конструкцій, зокрема в умовах складної географії та клімату.

Лукач (Lukacs) та співавтори вивчали підвищення цілісності пошкоджених трубопроводів за допомогою композитних армуючих систем [8]. Вони підтверджують, що навіть у випадках суттєвих дефектів у зварних з'єднаннях використання таких систем дозволяє продовжити експлуатаційний ресурс конструкцій, що важливо для підвищення безпеки в умовах різної географії.

Фекаду (Fekadu) зі співавторами у своєму дослідженні геологічної оцінки ділянки дамби підкреслюють, що геологічні та кліматичні умови можуть істотно впливати на якість і довговічність монтажних робіт [9]. Вони показали, що врахування геологічного середовища є важливим для запобігання аварійних ситуацій, що також має значення для зварювальних конструкцій, зокрема при використанні матеріалів, які зварюються в складних географічних умовах.

Підсумовуючи наведений огляд літератури, слід відзначити, що існує об'єктивна необхідність у подальших дослідженнях, які комплексно враховують географічні та кліматичні фактори у процесах проектування, монтажу і контролю якості зварних конструкцій. Такий висновок підкреслює актуальність проведеного дослідження та обґрунтовує наукову доцільність розробки програмних засобів для моделювання впливу природних умов на зварювальні процеси.

Постановка завдання. Метою статті є створення програмної моделі для оцінки впливу географічних і кліматичних чинників на якість зварювання металевих конструкцій.

Виклад основного матеріалу. Дослідження було проведено з використанням імітаційного моделювання впливу географічних умов на якість зварювання металевих конструкцій. Основною метою дослідження було оцінити потенційний зв'язок між кліматичними та географічними умовами та технологічними параметрами процесу зварювання (зокрема, середньою напругою, струмом і силою під час зварювання), а також кількістю дефектів, які виникають у результаті монтажу металевих конструкцій.

Для виконання дослідження було використано декілька відкритих датасетів.

Датасет «Resistance Spot Welding Dataset» містить інформацію про зварювальні процеси, включаючи часові ряди напруги, струму та сили, а також бінарну класифікацію дефектів (0 – без дефекту, 1 – дефектний зразок). Усього у цьому датасеті міститься

1 976 зразків. Напруга, струм і сила представлені у вигляді нормалізованих часових рядів (1000 точок у кожному), що дозволяє отримати середні значення цих параметрів для кожного зразка [10].

Вибірка GlobalLandTemperaturesByCountry із датасету «Climate Change: Earth Surface Temperature Data» містить історичні кліматичні дані (середньомісячні температури) для різних країн світу, включаючи Україну [11]. Ці дані були згруповані за місяцями та використані для аналізу сезонних коливань температури, які можуть впливати на процеси зварювання та формування дефектів.

Додатково було використано вибірку mountain_ranges.csv із датасету «World Geographic and Climatic Dataset», який містить інформацію про гірські хребти світу, включаючи їхню середню висоту [12]. Цей показник було застосовано як імітаційний географічний фактор для дослідження можливого впливу висоти над рівнем моря на технологічні параметри зварювання.

Завдання дослідження полягало у побудові програмної моделі, яка дозволяє поєднати кліматичні, географічні та технологічні дані з метою виявлення можливих закономірностей. З огляду на відсутність реальних даних, що одночасно містять усі ці показники, дослідження носить імітаційний характер. Було виконано статистичну оцінку та побудовано численні графіки залежностей, щоб візуально виявити можливі тренди між параметрами зварювання та географічно-кліматичними умовами.

Дані зварювання було попередньо очищено від нечислових значень та оброблено для отримання середніх значень по кожному зразку. Кліматичні дані для України було згруповано за місяцями та середньомісячні температури було використано для генерації імітаційної температурної вибірки. Географічні дані було оброблено для отримання середньої висоти гір як постійної характеристики.

Імітаційні значення температури та висоти було випадковим чином зіставлено з даними про зварювання. Це дозволило побудувати зведену таблицю, яка включала такі стовпці: середня напруга, струм, сила, дефектність (0/1), температура та висота гір.

У ході дослідження було розроблено програмну модель, яка дозволяє оцінити вплив географічних та кліматичних факторів на технологічні параметри процесу зварювання, зокрема середню напругу, струм і силу, а також на ймовірність виникнення дефектів.

Технологічно рішення реалізовано мовою програмування Python із застосуванням таких бібліотек: pandas – для роботи з табличними даними, numpy – для числових обчислень, matplotlib – для

побудови графіків та їхнього оформлення. Програма побудована у вигляді лінійного сценарію, у якому кожен етап обробки даних логічно впливає з попереднього. Вона не передбачає паралельної обробки даних або інтеграції зовнішніх API, що забезпечує простоту та стабільність виконання.

Основними конструкціями програми є циклічні та групувальні операції (наприклад, усереднення часових рядів або обчислення ковзного середнього), а також операції фільтрації даних за умовами (наприклад, відбір кліматичних даних лише для України). Імітаційне зіставлення кліматичних та географічних параметрів здійснюється за допомогою генерації випадкових вибірок, що дозволяє моделювати можливі зв'язки між показниками у реальному середовищі. Схематично логіку роботи показано на рис. 1.

Для візуалізації отриманих результатів були побудовані два основних типи графіків: точкові діаграми (scatter plot), що ілюструють загальний розподіл значень та можливі кореляційні зв'язки між показниками, а також лінійні графіки з використанням ковзного середнього (rolling mean), що дозволяють згладити шумові коливання та виявити загальні тенденції. Результати візуа-



Рис. 1. Логіка роботи програми моделювання зварювання та аналізу впливу географічних умов

лізації наведені на трьох графіках (рис. 2), що демонструють імітаційні залежності між технологічними показниками процесу зварювання та географічно-кліматичними умовами.

На першому графіку представлено розподіл кількості дефектів (0 – відсутність дефекту, 1 – наявність дефекту) залежно від температури середовища. Помітно, що дефекти розподілені досить рівномірно в межах усього діапазону температур, що свідчить про відсутність явно вираженого лінійного зв'язку між температурою та дефектністю. Водночас незначні кластери точок при окремих температурних діапазонах (близько 0°C та 10°C) можуть вказувати на локальні аномалії, що варто розглядати як потенційні зони ризику при реальних умовах експлуатації.

Другий графік ілюструє залежність середньої напруги зварювального процесу від температури. Розподіл точок носить статистично рівномірний характер, що підтверджує відсутність жорсткої лінійної залежності між температурою та середньою напругою. Проте можна виокремити слабкі тенденції у вигляді легких скупчень точок у межах температури від 5 до 15°C. Це може вказувати на можливий вплив кліматичних умов на стабільність процесу зварювання, оскільки саме у цьому діапазоні можуть виникати сприятливі умови для зварювання металевих конструкцій.

Третій графік демонструє залежність середньої напруги від висоти гірського рельєфу. Розподіл точок має виражений концентрований характер, утворюючи центральне «ядро» навколо середньої висоти гір. Така форма розподілу свідчить про те, що у межах середніх висот напруга коливається

в обмежених межах, що може бути інтерпретовано як ознака стабільності процесу при певних географічних умовах. Незважаючи на загальну відсутність чіткої залежності, локальні варіації напруги поблизу середніх висот вказують на потенційну чутливість процесу зварювання до змін атмосферного тиску чи кисневого складу повітря.

Додатково для демонстрації принципів візуалізації та залежностей, було збудовано три умовні графіки залежності вартості від маси, що за форматом і структурою відповідають класичним прикладам лінійних залежностей у наукових дослідженнях. На рис. 3 представлено результати дослідження взаємозв'язку між технологічними параметрами процесу зварювання металевих конструкцій та географічно-кліматичними чинниками. Графіки відображають лінійні залежності дефектності та середньої напруги від температури довкілля, а також від висоти гірського рельєфу.

Перший графік демонструє коливання кількості дефектів у зварювальних швах залежно від температури. Спостерігається чітке зростання кількості дефектів у діапазоні температур від -5°C до 5°C, що може вказувати на підвищену ймовірність виникнення дефектів за умов середньої температури. У межах температурних екстремумів (нижче -10°C та вище 15°C) кількість дефектів залишається стабільною, що дозволяє зробити висновок про потенційно безпечні умови зварювання у цих кліматичних зонах.

Другий графік ілюструє залежність середньої напруги процесу зварювання від температури середовища. У діапазоні від -10°C до 0°C напруга демонструє відносно стабільний рівень, після

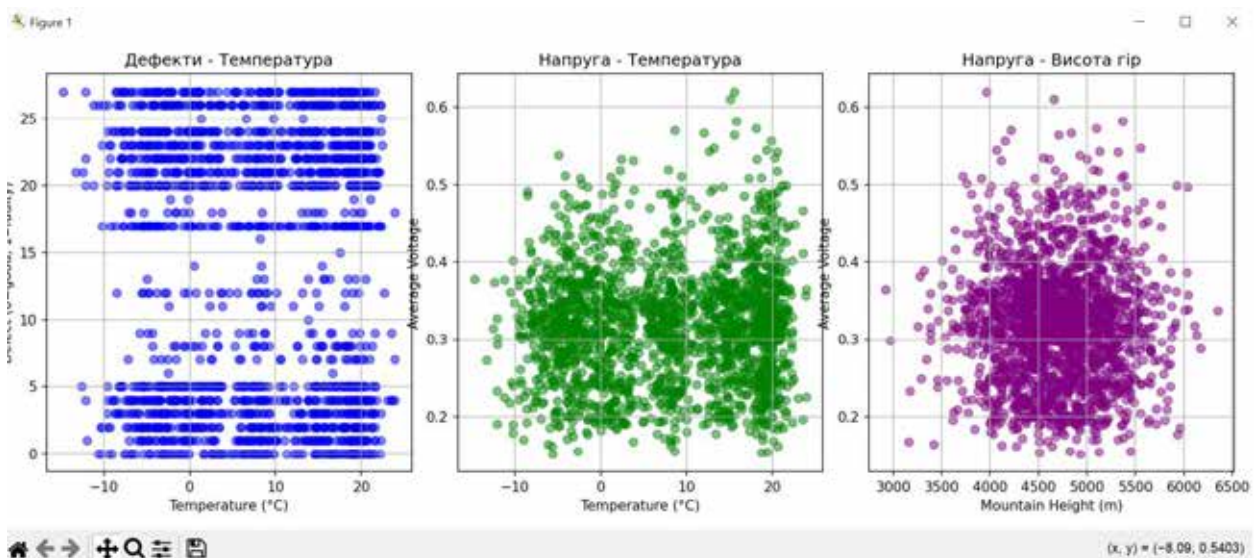


Рис. 2. Залежності між дефектністю, середньою напругою зварювання та географічно-кліматичними показниками (температура, висота гір)

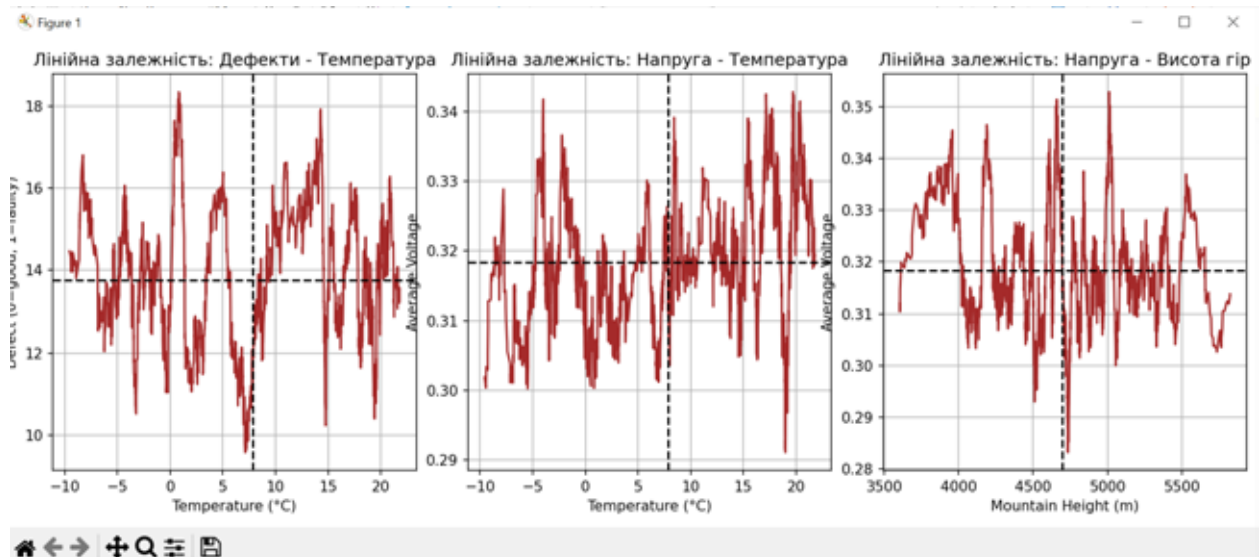


Рис. 3. Залежність технологічних параметрів процесу зварювання від температури та висоти гірської місцевості

чого у температурному діапазоні 0–15°C простежується збільшення середньої напруги. Це може бути пов'язано з покращенням умов зварювання та зниженням теплових втрат при температурі, близькій до середньої кліматичної.

Третій графік відображає вплив висоти гірської місцевості на середню напругу процесу зварювання. У межах висот від 3500 м до 4500 м напруга залишається на помірному рівні, що свідчить про стабільність процесу зварювання за умов середньої висоти. Проте, у висотному діапазоні 4500–5500 м спостерігається зростання середньої напруги, що може вказувати на адаптаційні зміни у параметрах зварювального процесу під впливом зниженого атмосферного тиску та зміненої концентрації кисню.

Загалом, аналіз отриманих графіків дозволяє зробити висновок про наявність залежності між географічно-кліматичними умовами та параметрами процесу зварювання. Найбільш чітко ця залежність проявляється у межах середніх температур та висот, що є важливою інформацією для формування практичних рекомендацій щодо підготовки зварювального обладнання та проведення монтажних робіт у різних географічних зонах.

Висновки. Проведене дослідження засвідчило, що технологічні параметри зварювання металевих конструкцій можуть демонструвати

варіабельність залежно від географічних і кліматичних умов. Під час аналізу було встановлено, що температурні коливання та висотні характеристики регіонів здатні чинити вплив на розподіл напруги та кількість дефектів у зварних швах. Виявлені локальні кластери і тенденції свідчать про можливий вплив природно-географічних чинників на стабільність процесу зварювання.

Результати дослідження також підкреслюють потребу у розробці адаптивних технологічних рішень для зварювання, здатних враховувати специфічні умови монтажу у різних регіонах. Створена програмна модель та проведена статистична оцінка демонструють потенціал інформаційних технологій у виявленні таких зв'язків, що є актуальним у контексті підвищення надійності зварних конструкцій в умовах зростаючої географічної різноманітності. Отримані результати можуть слугувати основою для формування рекомендацій з урахуванням факторів клімату та висотної зони під час організації монтажних робіт.

У подальших дослідженнях доцільно зосередитися на розширенні масиву емпіричних даних та створенні більш реалістичних моделей, що дозволять комплексно оцінювати вплив географічних умов на якість зварювання у широкому спектрі виробничих і монтажних практик.

Список літератури:

1. Kamalian L., Li H., Ching-Pong Poo M., Bras A., Adolf K.Y. Ng, Yang Z. Analysis of the impact of climate-driven extreme weather events (EWEs) on the UK train delays: A data-driven BN approach. *Reliability Engineering & System Safety*. 2025. № 262. ISSN 0951-8320. URL: <https://doi.org/10.1016/j.res.2025.111189>.
2. Di Bella G., Alderucci T., Salmeri F., Cucinotta F. Integrating the sustainability aspects into the risk analysis for the manufacturing of dissimilar aluminium/steel friction stir welded single lap joints used in marine

applications through a Life Cycle Assessment. *Sustainable Futures*. 2022. № 4. ISSN 2666-1888. URL: <https://doi.org/10.1016/j.sft.2022.100101>.

3. Rieck M., Koglin J., Emde B., Hermsdorf J. Influence of NIR-laser radiation on the underwater arc welding process with flux-cored wire. *Procedia CIRP*. 2024. № 124. C. 378-382. ISSN 2212-8271. URL: <https://doi.org/10.1016/j.procir.2024.08.137>.

4. Popescu M. A., Braga D. F. O., Vaz M. A. P., Melo F. J. M. Q. Residual stress measurements and weld characterization in wind turbine support structures. *Procedia Structural Integrity*. 2022. № 42. C. 1626-1633. ISSN 2452-3216. URL: <https://doi.org/10.1016/j.prostr.2022.12.205>.

5. Sun S., Liu C., Ferraz G., Karabulut B., Verhoeven B., Ruan X., Dewil R., Rossi B. Atmospheric corrosion of welded dissimilar T-joints made of steel and improvements in the corrosion prediction model. *Engineering Failure Analysis*. 2025. № 170. ISSN 1350-6307. URL: <https://doi.org/10.1016/j.engfailanal.2025.109273>.

6. Wahidi S. I., Oterkus S., Oterkus E. Robotic welding techniques in marine structures and production processes: A systematic literature review. *Marine Structures*. 2024. № 95. ISSN 0951-8339. URL: <https://doi.org/10.1016/j.marstruc.2024.103608>.

7. Pengchao C. Advancements and future outlook of safety monitoring, inspection and assessment technologies for oil and gas pipeline networks. *Journal of Pipeline Science and Engineering*. 2025. ISSN 2667-1433. URL: <https://doi.org/10.1016/j.jpse.2025.100267>.

8. Lukács J., Koncsik Z., Chován P. Integrity increasing of damaged transporting pipelines using fiber reinforced polymer composite wrap systems. *Engineering Failure Analysis*. 2022. № 137. ISSN 1350-6307. URL: <https://doi.org/10.1016/j.engfailanal.2022.106284>.

9. Fekadu A., Chemedi Y. C., Karuppannan S., Tadele Chala E. Engineering geological evaluation of Bowa Dayole dam site, North Shewa Zone, Oromia, Ethiopia. *Heliyon*. 2025. № 11 (1). e41385. ISSN 2405-8440. URL: <https://doi.org/10.1016/j.heliyon.2024.e41385>.

10. The dataset contains data about resistance spot welding used for the publication: «Fault Prediction in Resistance Spot Welding: A Comparison of Machine Learning Approaches». *GitHub – smartdatapolito/resistance_spot_welding_dataset*. URL: https://github.com/smartdatapolito/resistance_spot_welding_dataset.

11. Climate Change: Earth Surface Temperature Data. *Kaggle*. URL: <https://surli.cc/gmrfrpx>

12. World Geographic and Climatic Dataset. *Kaggle*. URL: <https://surl.li/nakitf>

Bahlai O.P. PROGRAM-SIMULATION ASSESSMENT OF THE INFLUENCE OF NATURAL FACTORS ON THE QUALITY OF WELDING

The article presents a study of the influence of geographical and climatic factors on the quality of welding of metal structures by creating a program-simulation model. The purpose of the work is to develop a tool that allows you to quantify the relationship between natural conditions (in particular, temperature and altitude above sea level) and technological parameters of welding – average voltage, current strength and the probability of defects in welded joints. The model is built on the basis of open climatic, geographical and production data, which made it possible to carry out a simulation analysis of the influence of the external environment on the stability and repeatability of the parameters of the welded process under different conditions.

Within the framework of the study, statistical data processing was performed and graphical dependencies were built, which reflect variations in welding parameters under the influence of temperature and altitude. Clusters and trends have been identified: in particular, temperature zones of 0–10°C and altitudes of 3500–4500 m are characterized by moderate changes in stress and defects, which indicates a potential impact of climatic conditions on the quality of weld formation. The analysis of the graphs indicates the presence of a connection between the external environment and technological characteristics, which can be taken into account when planning installation work in complex geographical areas.

The proposed model confirms the effectiveness of the use of information technologies to study the multifactorial influence of natural conditions on production processes. The results can be used to develop adaptive recommendations for welding in different regions, optimize process modes, select equipment and improve structural reliability. The model also provides the basis for further research involving advanced empirical data, development of preventive scenarios and integration into installation quality management systems in conditions of geographical variability.

Key words: welding of metal structures, geographical conditions, climatic factors, temperature influence, relief height, statistical analysis, installation works, welding quality, software modeling, stability of structures.

Дата надходження статті: 14.07.2025

Дата прийняття статті: 31.07.2025

Опубліковано: 27.10.2025

Бачкала Б.О.

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»

Тимошин Ю.А.

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»

МОДЕЛІ ЗАБЕЗПЕЧЕННЯ ВИСОКОЇ ДОСТУПНОСТІ ТА ВІДМОВОСТІЙКОСТІ РОЗПОДІЛЕНИХ СИСТЕМ ЦИФРОВОЇ ІНФРАСТРУКТУРИ

З розвитком інформаційних технологій та стрімким зростанням обсягу даних, сучасні організації все частіше надають перевагу використанню хмарних обчислень для підтримки цифрової інфраструктури. Це сприяє поширенню складних розподілених комп'ютерних систем, побудованих на основі хмарних, мультихмарних, туманних та периферійних рішень. Такі системи зазвичай функціонують в умовах можливих технічних (мережевих, апаратних та програмних збоїв), і зовнішніх впливів, а отже мають залишатися працездатними у разі відмови як окремих компонентів, так і цілих вузлів таких систем. Сучасний бізнес висуває жорсткі вимоги до надійності таких систем, оскільки навіть мінімальний час простою може призвести до значних фінансових втрат та зниження довіри з боку клієнтів. Щоб організувати безвідмовну роботу хмарних розподілених систем необхідно реалізувати моделі автоматичного аналізу параметрів стану вузлів обробки, їх відновлення та забезпечити високу доступність за рахунок реплікації даних, архітектурних рішень та резервування компонентів.

У статті проведено аналіз моделей забезпечення відмовостійкості та високої доступності сучасних розподілених систем, зокрема хмарних, мультихмарних, туманних та периферійних систем Інтернету Речей (IoT). У статті визначені та проаналізовані сучасні методи забезпечення відмовостійкості та високої доступності для різних типів розподілених систем. Надано рекомендації щодо застосування таких моделей та методів залежно від специфіки середовища та вимог щодо якості обслуговування. У статті запропоновано модель адаптивного резервування даних для периферійних систем з урахуванням мережесих відмов на основі розподілу Вейбула. Викладено результати моделювання співвідношень цільового рівня надійності системи та інтервалу резервування. Основною мотивацією проведених досліджень є систематизація підходів до підвищення надійності розподілених систем та розробка стратегії резервування, адаптованої для периферійних систем з характерними мережевими відмовами компонентів.

Ключові слова: висока доступність, відмовостійкість, розподілені системи, хмарні обчислення, мультихмара, периферійні обчислення, IoT.

Постановка проблеми. Попри широкий спектр існуючих інформаційних технологій, нині не існує універсального рішення, здатного забезпечити безперервну роботу розподілених систем цифрової інфраструктури для обробки прикладних задач з урахуванням багатьох умов їх експлуатації та різних внутрішніх та зовнішніх факторів технічного, організаційно-управлінського та операційного характеру. Отже, актуальним є завдання аналізу моделей забезпечення високої доступності та відмовостійкості таких систем, а також розроблення рекомендацій щодо вибору оптимальних архітектурних рішень та операцій-

них характеристик, залежно від особливостей конкретного середовища, включаючи хмарні сервіси IaaS, та вимог щодо якості обслуговування. Для периферійних систем актуальною є проблема втрати даних та зниження продуктивності через відсутність можливості своєчасно копіювати чи реплікувати дані, а також через специфіку роботи таких систем у середовищах із обмеженим або нестабільним мережесим з'єднанням. Тому під час реалізації методів підвищення надійності периферійних систем необхідно враховувати як специфіку мережевої інфраструктури, так і особливості локальної обробки даних, резервування

каналів зв'язку, розміщення резервних ЦОД, ізоляції «зон доступності» та можливість автономної роботи периферійних пристроїв за умов періодичних мережових чи апаратних відмов.

Аналіз останніх досліджень і публікацій. Велика кількість наукових праць присвячена аналізу та класифікації підходів до забезпечення безперервності роботи ІТ-сервісів [1, с. 76-86]. Забезпечення високої доступності зазвичай не потребує повного дублювання всіх компонентів, але включає резервування критичних вузлів, кешування важливих даних, механізми автоматичного відновлення та балансування навантаження. В свою чергу, відмовостійкість фокусується на повному уникненні простою за будь-яких одиничних відмов – тобто система продовжує працювати без перерви навіть у разі відмови одного чи кількох компонентів. У роботі [2] проаналізовані підходи до забезпечення надійності та високої доступності в середовищах хмарних обчислень, які відкривають широкі можливості для підвищення надійності та забезпечення безперервності бізнес-процесів завдяки своїй гнучкості, масштабованості та можливості оперативного відновлення роботи сервісів після відмов, що дає змогу зменшити ризики і витрати на підтримку ІТ-інфраструктури.

У статті [3, с. 1-14] запропоновано новий метод адаптивного управління хмарними ресурсами (FT-ERM) з вбудованим механізмом відмовостійкості для забезпечення високої доступності серверів та віртуальних машин (ВМ), що передбачає проактивне резервування додаткових ресурсів для попередження аварійних ситуацій шляхом завчасної міграції навантаження, показано підвищення загальної доступності серверів на 34.47% та значне зменшення надлишкових міграцій віртуальних машин порівняно з традиційними підходами. Однак динамічні характеристики хмарного середовища підвищують ймовірність виникнення додаткових помилок та відмов. У дослідженні [4, с. 75920-75940] представлено нову гібридну модель відмовостійкого планування та балансування навантаження (HFSLM). Це рішення дозволило усунути проблему відмов та нерівномірності розподілу задач між віртуальними машинами і скоротити час виконання задач на 8% порівняно з базовими алгоритмами планування.

У дослідженні [5] запропоновано нову платформу FogROS2-FT для підвищення якості обслуговування та відмовостійкості за рахунок мультихмарного розгортання обчислювальних модулів. Платформа автоматично реплікує незалежні роботизовані сервіси, направляє запити до цих

реплік і повертає назад першу отриману відповідь. Авторами роботи [6, с. 10501-10511] запропоновано новий підхід для підвищення відмовостійкості промислових хмарних платформ під назвою TOLERANCER у вузлах систем Індустріального IoT (ІІoT). Запропонована система здатна ідентифікувати відмови програмного чи апаратного рівня на крайових обчислювальних серверах та автоматично перерозподіляти навантаження між працездатними вузлами, гарантуючи мінімальний час простою. У статті [7] представлено метод динамічного балансування навантаження (FTSM), який віддає пріоритет обробці чутливих до затримок запитів у середовищі периферійних та туманних обчислень, враховує час виконання, критичність задач та поточне навантаження на вузлах для оптимального перенесення процесів у випадку відмов.

Технологія розподіленої обробки даних та додатків з використанням динамічно інтерпретованих метамоделей, а також використання такого прикладного середовища для міжкорпоративної обробки бізнес-об'єктів, що описані у дослідженнях [8, с. 128-138; 9, с. 137-150], дозволяють зменшити вектор параметрів управління для систем розподіленої обробки, адаптувати і покращити балансування навантаження на вузли, збільшити надійність ІТ-систем.

Питанням забезпечення відмовостійкості хмарних систем на основі оцінок метрик надійності присвячена робота [10, с. 29-31], в якій досліджено вплив метрик RPO (Recovery Point Objective) на потенційні втрати даних та RTO (Recovery Time Objective) на допустимий час простою в середовищі хмарних сервісів моделі IaaS.

В роботі [11, с. 158-161] запропоновано модель оцінки параметрів обробки інформації в розподілених системах з динамічною структурою, які реалізують неоднорідні розподілені бази даних, що дозволяє формалізувати структуру керування обробкою запитів.

Робота [12, с. 1-8] присвячена процесу моделювання аналізу ситуацій та прийняття рішень щодо функціонування вузлів розподілених комп'ютерних систем (PKC) на основі спеціального механізму стохастичних RA-мереж, що дозволяє моделювати потік даних з різними, у тому числі суттєво різними інтенсивностями, що особливо важливо для аналізу ситуацій та прийняття рішень у динаміці керування параметрами вузлів PKC.

У роботі [13, с. 161-166] розглянуто методи підвищення ефективності виявлення загроз у роз-

поділених базах даних за допомогою системи моніторингу подій. Зазначено, що система працює на основі моделі моніторингу подій різнорідних нереляційних розподілених баз даних.

У статті [14, с. 320-325] описана складна адаптивна система, яка базується на нейронній мережі та інтелектуальних агентах. Розроблена система дозволяє ідентифікувати та класифікувати атаки та запобігати їм в режимі реального часу. В роботі [15, с. 195-203] показано, що автоматичне масштабування є ключовим механізмом адаптивного управління ресурсами, що дозволяє системі своєчасно реагувати на зміни навантаження. Своєчасному виявленню відхилень, їх діагностиці, реконфігурації та відновленню під час аварійних ситуацій присвячена робота [16, с. 104-111].

В роботі [17, с. 84-91] показано, що мультихмарна архітектура дозволяє організаціям використовувати неінтегровані послуги кількох публічних хмарних платформ. Створення ефективного мультихмарного середовища потребує системного підходу до вибору оптимальних платформ і рішень – цьому питанню присвячено дослідження [18, с. 32-39] [19, с. 113-121] [20, с. 51-62], включаючи географічно розподілені архітектури [21, с. 39-40].

Для забезпечення високої доступності у туманних обчисленнях застосовуються проактивні та реактивні стратегії управління відмовами, чому присвячена робота [22, с. 415-422]. В роботі [23, с. 49-52] показано як реплікація даних та сервісів забезпечує безперервність роботи при виході з ладу окремих компонентів периферійних вузлів.

Публікації [24, с. 217-245; 25, с. 233-246] присвячені порівнянню ймовірнісних моделей для опису процесу відмови та як вони відповідають експоненціальному, Вейбуллівському, нормальному та логнормальному розподілам ймовірностей для задач надійності, включаючи моделювання для реальних об'єктів.

В роботі [26, с. 1520-1532] обговорюються питання віртуалізації мережеских функцій, що забезпечує ефективне планування хмарних ресурсів шляхом віртуалізації мережеских сервісів та програм у алгоритми та відповідне ПЗ, яке працює на стандартних серверах. Провайдер повинен забезпечувати доступність сервісів мережі паралельних віртуалізованих мережеских функцій (ВМФ).

Вище описане підтверджує, що основними факторами, які впливають на стійкість та надійність розподілених систем обробки є технічні компоненти інфраструктури ЦОД, їх конфігурація,

ізоляція сегментів у просторі та працездатність LAN- та інтернет-з'єднань, ефективне балансування потоків даних всередині вузлів та між ними, спеціалізоване ПЗ для управління розподіленими системами обробки та моделі відновлення, впровадження оптимальних і адаптивних моделей управління з резервуванням ресурсів і проактивними заходами виявлення та запобігання нетипових та аварійних ситуацій для забезпечення відмовостійкості.

Постановка завдання. Цілі пропонованого дослідження полягають у тому, щоб:

(I) розробити математичну модель адаптивної стратегії резервування для периферійних систем. Побудувати аналітичні залежності узгодження цільового рівня надійності системи із інтервалом резервування з урахуванням ймовірності мережеских відмов;

(II) побудувати графіки навантаження для гібридного хмарного середовища з метою визначення параметричних діапазонів стійкої високонадійної роботи з різними законами відмов компонентів периферійної системи;

(III) надати рекомендації щодо вибору конкретних моделей залежно від специфіки середовища та вимог щодо якості обслуговування для забезпечення відмовостійкості та високої доступності розподілених систем.

Виклад основного матеріалу. Висока доступність та відмовостійкість у хмарних середовищах реалізується повним або частковим резервуванням всіх компонентів інфраструктури. Використання географічно розподілених центрів обробки даних (ЦОД) або зон доступності дозволяє захистити IT-сервіси від зовнішніх факторів (стихійних лих, перебоїв з електроенергією). Балансування навантаження є ще одним ключовим механізмом забезпечення високої доступності у хмарній інфраструктурі. Цей підхід дозволяє рівномірно розподіляти запити та інші типи навантаження між декількома фізичними або віртуальними серверами, забезпечуючи стабільність та високу продуктивність системи. Балансування може здійснюватися за допомогою програмних або апаратних засобів. У разі відмови одного із серверів балансувальник навантаження автоматично виключає його з процесу обробки та перенаправляє трафік на інші справні вузли. Важливою складовою забезпечення високої доступності у хмарі є також реалізація механізмів аварійного відновлення, які дозволяють мінімізувати показники RTO (цільовий час відновлення – допустимий час простою системи) та RPO (цільова точка

відновлення – допустимий обсяг втрати даних) при плануванні стратегії резервного копіювання. В результаті поєднання цих підходів сучасні хмарні інфраструктури досягають рівня доступності 99.99% («чотири дев'ятки»), що є стандартом для критичних бізнес-застосунків.

Порівняно з мультирегіональними рішеннями у межах однієї хмари, мультихмарний підхід забезпечує вищий рівень відмовостійкості та доступності. Цей підхід усуває залежність від одного постачальника: глобальний збій однієї платформи не спричинить повної зупинки системи. Крім того, різні хмари можуть виконувати різні завдання. Наприклад, одна платформа може обробляти високонавантажені транзакції, а інша – забезпечувати резервне зберігання даних чи аналітику, що підвищує ефективність і стійкість системи. Одним з базових механізмів забезпечення високої доступності в мультихмарних середовищах є реплікація даних між хмарами. На відміну від класичних хмарних середовищ, де реплікація обмежена дата-центрами однієї хмари, мультихмарна реплікація підтримує декілька копій даних на різних платформах. Такий підхід дозволяє підтримувати актуальність та цілісність даних у всіх хмарах одночасно, гарантуючи їх доступність навіть за умов виходу з ладу одного з постачальників. Наступним важливим компонентом високої доступності мультихмарних архітектур є автоматичне балансування навантаження між хмарними платформами. Для цього застосовуються спеціалізовані програмні рішення, такі як глобальні балансувальники навантаження та динамічні механізми розподілу трафіку. Балансування навантаження дозволяє рівномірно розподілити запити між декількома платформами, забезпечуючи оптимальне використання ресурсів та уникнення перевантаження окремих вузлів. Крім цього, балансувальники здатні виконувати автоматичне перенаправлення трафіку на резервні хмарні майданчики у разі виявлення відмов чи зниження продуктивності однієї з платформ, завдяки чому забезпечується мінімізація часу відновлення роботи сервісів. Механізм автоматичного відновлення між кількома хмарами передбачає моніторинг стану хмарних платформ і своєчасне перенесення навантаження на резервні ресурси іншої платформи у разі виявлення аварійної ситуації. Мультихмарні архітектури пропонують вищий ступінь диверсифікації ризиків, оскільки резервні компоненти розташовані у ізольованих та незалежних хмарних платформах.

У середовищі периферійних та туманних обчислень висока доступність та відмовостійкість досягається шляхом децентралізації обчислювальних ресурсів, наближенням їх до кінцевих пристроїв та зниженням затримок, що особливо важливо для систем IoT та Індустріального Інтернету речей (IIoT). Аналіз та фільтрація даних на границі зменшує витрати на зв'язок, хмарну обробку та зберігання даних. Архітектури на базі периферійних та хмарних обчислень дозволяють створювати географічно розподілені системи з низькою затримкою, визначенням місцезнаходження та підтримкою мобільності. Головною метою цього підходу є покращення якості надання послуг, а саме: зменшення затримок передачі даних, забезпечення локальної обробки даних у реальному часі та гарантування роботи сервісів навіть за нестабільного зв'язку. За рахунок географічного рознесення вузлів і наближення їх до кінцевих пристроїв досягається стійкість до регіональних збоїв – наприклад, якщо виходить з ладу центр обробки даних у певній локації, пристрої на периферії можуть тимчасово виконувати основні функції самостійно або переключатися на резервний центр в іншому регіоні.

Для забезпечення високої доступності у туманних обчисленнях застосовуються проактивні та реактивні стратегії управління відмовами, причому проактивні методи передбачають використання систем моніторингу, які збирають інформацію про стан вузлів мережі та аналізують її для виявлення потенційних проблем ще до їх виникнення. Для цього часто використовують методи статистичного аналізу, моделі машинного навчання або алгоритми штучного інтелекту, що дають змогу завчасно спрогнозувати відмову вузла або зниження його продуктивності. На основі цих прогнозів система може адаптивно перерозподіляти навантаження, мігрувати критично важливі сервіси на більш стабільні вузли або резервувати додаткові ресурси, мінімізуючи таким чином ймовірність простою та забезпечуючи високу доступність. Реактивні ж стратегії забезпечують оперативне реагування після фактичного виникнення відмови. Серед таких методів особливо ефективними є автоматичне переключення маршрутів передачі даних та швидке перенаштування роботи системи для мінімізації негативних наслідків. Ще одним напрямом забезпечення відмовостійкості периферійних та туманних обчислень є реплікація даних та сервісів, що забезпечує безперервність роботи при виході з ладу окремих компонентів. Це дозволяє

забезпечити гнучкість і високу надійність систем навіть за умов значних навантажень або частих змін у структурі мережі.

У таблиці 1 наведено узагальнені рекомендації щодо вибору моделей забезпечення відмовостійкості розподілених систем.

Для розглянутих моделей та обчислювальних середовищ спільною рисою є використання механізмів резервування ресурсів та даних для забезпечення високої надійності розподілених систем. При виборі стратегії резервування для периферійних та туманних середовищ необхідно враховувати випадковий характер відмови мережевих компонентів та вузлів системи. Далі пропонується математична модель адаптивного резервування даних на прикладі периферійних систем.

Для врахування ймовірності відмов компонентів периферійної системи обрано дво-параметричний розподіл Вейбула. Цей розподіл описує різні режими відмов: при значенні параметра форми розподілу $k < 1$ інтенсивність відмов зменшується з часом; при $k = 1$ інтенсивність відмов є постійною; при $k > 1$ інтенсивність відмов зростає з часом. Параметр масштабу розподілу (λ) визначає характерний час, протягом якого більшість компонентів системи зазнає відмов. Формально функція надійності визначає ймовірність безвідмовної роботи системи до часу t :

$$R(t) = \exp \left[- \left(\frac{t}{\lambda} \right)^k \right] \quad (1)$$

де $k > 0$ – параметр форми розподілу Вейбула; $\lambda > 0$ – параметр масштабу розподілу Вейбула.

Адаптивне резервування передбачає динамічну зміну інтервалів між резервними копіями даних відповідно до вимог надійності системи. Нехай після чергового резервного копіювання минув час t , визначимо час Δt як додатковий інтервал часу

до наступного резервування. Введемо поняття відносної надійності на основі (1), що визначає ймовірність безвідмовної роботи системи на додатковому інтервалі Δt за умови, що система вже пропрацювала час t :

$$R(t, t + \Delta t) = \frac{R(t + \Delta t)}{R(t)} = \exp \left[- \left(\frac{t + \Delta t}{\lambda} \right)^k + \left(\frac{t}{\lambda} \right)^k \right] \quad (2)$$

Критерієм вибору оптимального інтервалу резервування є забезпечення відповідності ймовірності безвідмовної роботи системи цільовому рівню надійності R_{req} :

$$R(t, t + \Delta t) = R_{req} \quad (3)$$

Підставляючи вираз (2) у (3) отримаємо рівняння:

$$\exp \left[- \left(\frac{t + \Delta t}{\lambda} \right)^k + \left(\frac{t}{\lambda} \right)^k \right] = R_{req} \quad (4)$$

Застосувавши логарифмічне перетворення до рівняння (4) отримаємо вираз для визначення оптимального інтервалу резервування Δt :

$$\Delta t = \lambda \left[\left(\frac{t}{\lambda} \right)^k - \ln(R_{req}) \right]^{\frac{1}{k}} - t \quad (5)$$

Формула (5) дозволяє визначити оптимальний інтервал резервування на основі поточного часу роботи системи та цільового рівня надійності.

Розглянемо приклад застосування формули (5) для систем класу Industrial IoT (IIoT) з наступними параметрами:

$$k = 1.8; \lambda = 6 \text{ років}; R_{req} = 0.999$$

Початковий інтервал резервування визначимо за умови $t = 0$:

$$\Delta t = 6 \times \left[-\ln(0.999) \right]^{\frac{1}{1.8}} \approx 0.13 \text{ років} \approx 47 \text{ днів}$$

Таблиця 1

Моделі забезпечення відмовостійкості розподілених систем

Середовище	Ключові характеристики	Рекомендовані моделі
Хмара	Надлишковість ресурсів; багатозонне і багаторегіональне резервування; централізоване керування та автоматичне перемикавання при збоях.	FT-ERM, HFSLM
Мультихмара	Відсутність єдиної точки відмови; автоматичний перерозподіл навантаження між різними хмарами; інтегрована міжхмарна синхронізація.	FogROS2-FT
Периферія	Децентралізація обчислень; розподіл задач між сусідніми вузлами при відмові; низькі затримки обробки даних.	TOLERANCER
Туманні обчислення	Проактивне виявлення відмов; автоматичний перехід на альтернативні вузли у разі відмов.	FTSM

Визначимо інтервал резервування через 1 рік безвідмовної роботи:

$$\Delta t = 6 \times \left[\left(\frac{1}{6} \right)^{\frac{1}{1.8}} - \ln(0.999) \right]^{-1} \approx 0.014 \text{ років} \approx 5 \text{ днів}$$

Цей приклад демонструє адаптивність запропонованої моделі: з часом експлуатації системи у режимі зношування ($k > 1$) інтервал резервування зменшується з 47 днів до 5 днів, забезпечуючи незмінний рівень надійності.

Для аналізу впливу параметрів розподілу Вейбула на вибір інтервалу резервування було проведено моделювання засобами мови програмування Python. На графіках рис. 1–2 показані залежності цільового рівня надійності системи від інтервалу резервування за різних значень параметрів масштабу та форми розподілу Вейбула.

При малих значеннях середнього часу до відмови (рис. 1) спостерігається швидке зниження надійності системи зі збільшенням інтервалу резервування, особливо для систем у режимі раних відмов. Збільшення середнього часу до відмови (рис. 2) дозволяє підтримувати високий рівень надійності при більш тривалих інтервалах резервування. Для обох графіків характерною є наявність спільної точки перетину кривих, яка відповідає однаковому значенню надійності для різних режимів відмов у момент, коли інтервал резервування дорівнює середньому часу до відмови системи.

На рис. 3 графічно визначено параметричні діапазони інтервалів резервування, що забезпечують стійку високонадійну роботу периферійних систем для режимів раних відмов, випадкових відмов та зношування.

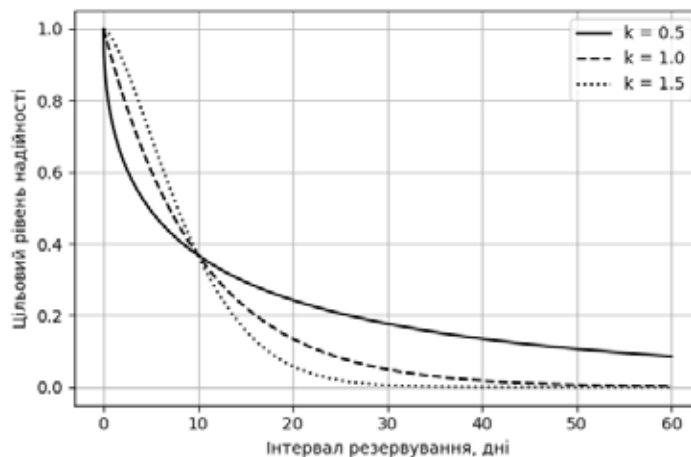


Рис. 1. Залежність цільового рівня надійності від інтервалу резервування за різних режимів відмов із середнім часом до відмови (λ) 10 днів

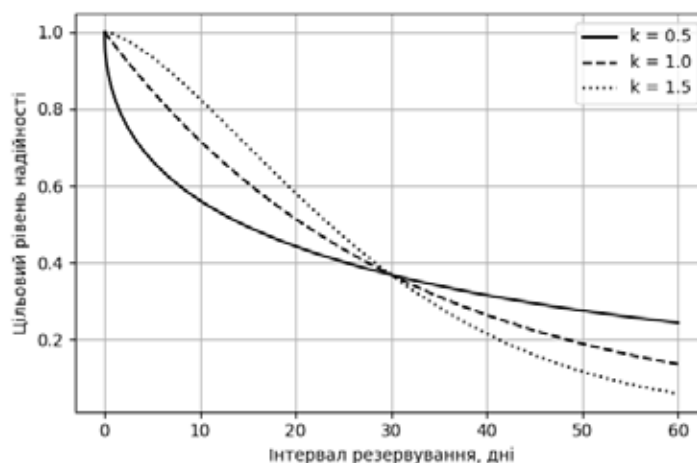


Рис. 2. Залежність цільового рівня надійності від інтервалу резервування за різних режимів відмов із середнім часом до відмови (λ) 30 днів

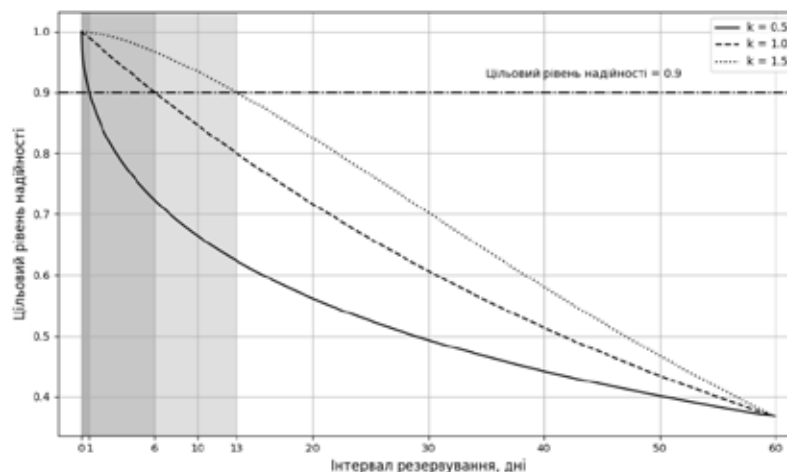


Рис. 3. Порівняння інтервалів резервування за різних режимів відмов із середнім часом до відмови (λ) 60 днів

На графіку (рис. 3) зафарбованими ділянками показано діапазони інтервалів резервування, що забезпечують цільовий рівень надійності для різних режимів відмов. Рівень надійності (0.9) було обрано відповідно до типових галузевих вимог до допустимих втрат даних у таких системах та для ілюстрації впливу різних режимів відмов. Порівнюючи значення інтервалів резервування при зменшенні ($k = 0.5$) та збільшенні ($k = 1.5$) інтенсивності відмов, стає очевидним, що системи з ранніми відмовами потребують значно коротших інтервалів резервування (до 24 годин) для підтримки цільового рівня надійності порівняно із системами, які функціонують у режимі зношування, де допустимі інтервали резервування можуть досягати 13 днів. Наведені графоаналітичні залежності можуть бути використані інженерами для оцінки допустимих меж інтервалів резервування залежно від характеру зміни інтенсивності відмов та для вибору стратегій резервування.

Висновки. У статті запропоновано та обґрунтовано комплекс моделей забезпечення високої доступності та відмовостійкості розподілених систем цифрової інфраструктури з акцентом на хмарні, мультихмарні, периферійні та туманні обчислення. Розглянуто специфіку реалізації

механізмів надійності у кожному з цих середовищ, виявлено їх переваги та недоліки. Аналіз літературних джерел і практичних підходів свідчить, що ефективність таких систем досягається шляхом впровадження комплексних архітектурних рішень – резервування, реплікації даних, автоматичного масштабування, балансування навантаження та проактивного моніторингу стану компонентів.

Розроблена математична модель адаптивного резервування дозволяє дослідникам та інженерам контролювати ризики втрати даних через динамічне керування інтервалом резервування відповідно до вимог надійності у периферійних середовищах. Надані графоаналітичні залежності дозволяють враховувати вплив мережових відмов на чисельні співвідношення частоти резервування із цільовим рівнем надійності периферійних систем.

Впровадження обґрунтованих підходів до забезпечення високої доступності та відмовостійкості є необхідною умовою для стабільного функціонування сучасної цифрової інфраструктури та підтримки критично важливих бізнес-процесів. Поєднання комплексних архітектурних рішень з адаптивними методами управління є ефективною стратегією для забезпечення надійності реальних розподілених систем.

Список літератури:

1. Panda J., Das S., Pattnaik D. Resilient IT Infrastructure: Strategies for Minimizing Downtime and Ensuring Business Continuity. *Journal of Humanities and Social Sciences Research*. 2024. Vol. 6 (S). P. 76–86. DOI: 10.37534/bp.jhssr.2024.v6.nS.id1255.p76
2. Mesbahi M. R., Rahmani A. M., Hosseinzadeh M. Reliability and high availability in cloud computing environments: a reference roadmap. *Human-centric Computing and Information Sciences*. 2018. Vol. 8, Article 20. DOI: 10.1186/s13673-018-0143-8

3. Saxena D., Gupta I., Singh A. K., Lee C.-N. A Fault Tolerant Elastic Resource Management Framework Towards High Availability of Cloud Services. *IEEE Transactions on Network and Service Management*. 2022. P. 1–14. DOI: 10.1109/TNSM.2022.3170379
4. Mushtaq S. U., Sheikh S., Idrees S. M. Next-Gen Cloud Efficiency: Fault-Tolerant Task Scheduling With Neighboring Reservations for Improved Resource Utilization. *IEEE Access*. 2024. Vol. 12. P. 75920–75938. DOI: 10.1109/ACCESS.2024.3404643
5. Chen K., Hari K., Chung T., Wang M., Tian N., Juette C., Ichnowski J., Ren L., Kubiawicz J., Stoica I., Goldberg K. FogROS2-FT: Fault Tolerant Cloud Robotics. arXiv preprint arXiv:2412.05408 [cs.RO], 2024. DOI: 10.48550/arXiv.2412.05408
6. Al-Dulaimy A., Sicari C., Papadopoulos A. V., Galletta A., Villari M., Ashjaei M. TOLERANCER: A Fault Tolerance Approach for Cloud Manufacturing Environments. *IEEE Transactions on Industrial Informatics*. 2023. Vol. 19, no. 10. P. 10501–10511. DOI: 10.1109/ETFA52439.2022.9921606
7. Alarifi A., Abdelsamie F., Amoon M. A fault-tolerant aware scheduling method for fog-cloud environments. *PLOS ONE*. 2019. Vol. 14, No. 10. e0223902. DOI: 10.1371/journal.pone.0223902
8. Тимошин Ю. А. Технологія розподіленої обробки даних і додатків з використанням динамічно інтерпретованих метамodelей. *Адаптивні системи автоматичного управління*. 2014. № 24. С. 128–138. DOI: 10.20535/1560-8956.24.2014.38199.
9. Стенін О. А., Тимошин Ю. А., Шемседінов Т. Г. Динамічне прикладне середовище з інтерпретацією метамodelей для міжкорпоративної обробки бізнес-об'єктів. *Адаптивні системи автоматичного управління*. 2016. № 28. С. 137–150. DOI: 10.20535/1560-8956.28.2016.82401.
10. Тимошин Ю., Бачкала Б. Проблеми забезпечення відмовостійкості хмарних систем на основі оцінок метрик надійності. *Матеріали XIII Міжнародної науково-практичної конференції з інформаційних систем та технологій «Infocom Advanced Solutions – 2025»*. Київ, 21–22 травня 2025. С. 29–31. URL: https://ist.kpi.ua/wp-content/uploads/2025/05/infocom-advanced-solutions-2025_compressed-3.pdf (дата звернення: 06.06.2025).
11. Mukhin V., Kornaga Y., Bondarenko V., Zavgorodnii V., Herasymenko O., Sholokhov O. Mathematical Model for Heterogeneous Databases Parameters Estimation in Distributed Systems with Dynamic Structure. *2020 IEEE 2nd International Conference on Advanced Trends in Information Theory (ATIT)*. Kyiv, Ukraine, 2020. P. 158–161. DOI: 10.1109/ATIT50783.2020.9349331.
12. Hu Z., Mukhin V., Loutskii H., Kornaga Y. Stochastic RA-Network for the Nodes Functioning Analysis in the Distributed Computer Systems. *International Journal of Computer Network and Information Security*. 2016. Vol. 8, No. 6. P. 1–8. URL: <https://sciup.org/15015169>
13. Корнага Я. І., Барабаш А. О. Математична модель моніторингу подій в системі управління розподіленими базами даних на основі сенсорів подій. *Інфокомунікаційні та комп'ютерні технології*. 2023. Т. 1, № 05. С. 161–166. DOI: 10.36994/2788-5518-2023-01-05-17.
14. Mukhin V., Kornaga Y., Steshyn V., Mostovoy Y. Adaptive Security System Based on Intelligent Agents for Distributed Computer Systems. *Proceedings of the 13th International Conference on Development and Application Systems (DAS 2016)*. Suceava, Romania, 2016. P. 320–325. DOI: 10.1109/DAAS.2016.7492595.
15. Bondi A. B. Characteristics of scalability and their impact on performance. *Proceedings of the 2nd International Workshop on Software and Performance*. 2000. P. 195–203. DOI: 10.1145/350391.350432.
16. Яскевич В. О., Ключко О. Ю. Методи підвищення відмовостійкості інтернет сервісів. *Кібербезпека: освіта, наука, техніка*. 2019. № 3 (3). С. 104–111. DOI: 10.28925/2663-4023.2019.3.104111
17. Касерес А., Глоба Л. С. Підхід до формування мультихмарного середовища на основі багатокритеріального аналізу та онтологічного моделювання. *Системи управління, навігації та зв'язку*. 2024. № 4. С. 84–91. DOI: 10.26906/SUNZ.2024.4.084
18. Гюргізова-Гай В. Ш., Кирюша Б. А. Хмарні послуги в корпоративній інфраструктурі. *Вчені записки ТНУ імені В. І. Вернадського. Серія: Технічні науки*. 2024. Т. 35 (74), № 6. С. 32–39. DOI: 10.32782/2663-5941/2024.6.2/05
19. Mankotia A. Comparing High Availability and Disaster Recovery in Multi-Cloud Environments. *International Journal of Computer Trends and Technology*. 2024. Vol. 72, No. 7. P. 113–121. DOI: 10.14445/22312803/IJCTT-V72I7P115
20. Yalate A. Demystifying Multi-Cloud Architecture: Foundational Concepts and Design Patterns. *European Journal of Computer Science and Information Technology*. 2025. Vol. 13, No. 24. P. 51–62. DOI: 10.37745/ejsit.2013/vol13n245162
21. Недоступ Д. М., Солом'яний М. В. Аналіз підходів забезпечення відмовостійкості архітектур Extended Cloud. *Матеріали VIII Міжнародної науково-технічної конференції «Проблеми електромагнітної сумісності перспективних безпроводових мереж зв'язку»*. Харків, 2022. С. 39–40. URL: <http://openarchive.nure.ua/handle/document/23319> (дата звернення: 06.06.2025).

22. Shah Y., Thakkar E., Bhavsar S. Fault Tolerance in Cloud and Fog Computing—A Holistic View. *Data Science and Intelligent Applications*. 2021. P. 415–422. DOI: 10.1007/978-981-15-4474-3_46
23. Agrawal A., Toshniwal D. Fault Tolerance in IoT: Techniques and Comparative Study. *Asian Journal For Convergence In Technology*. 2021. Vol. 7, No. 1. P. 49–52. DOI: 10.33130/AJCT.2021v07i01.011
24. Vardhan A. S. S., Verma A., Ogale J., Saket R. K., Galloway S. Modern aspects of probabilistic distributions for reliability evaluation of engineering systems. In: *Reliability Analysis of Modern Power Systems*. Wiley, 2024. P. 217–245. DOI: 10.1002/9781394226771.ch12
25. Ocansey S. A., Bikdash M. Temperature-Driven Reliability Analysis of Power Grid Failures: A Weibull Distribution Approach To Outage Prediction and Mitigation. *Journal of Statistical Theory and Applications*. 2025. Vol. 24, No. 1. P. 233–246. DOI: 10.1007/s44199-025-00109-y
26. Guo Z., Li J., Ramesh R. Scalable, adaptable, and fast estimation of transient downtime in virtual infrastructures using convex decomposition and sample path randomization. *INFORMS Journal on Computing*. 2020. Vol. 32, No. 2. P. 321–345. DOI: 10.1287/ijoc.2019.0888

Bachkala B.O., Tymoshyn Yu.A. MODELS FOR ENSURING HIGH AVAILABILITY AND FAULT TOLERANCE IN DISTRIBUTED DIGITAL INFRASTRUCTURE SYSTEMS

As information technologies advance and data volumes grow at a rapid pace, modern organizations increasingly favor the use of cloud computing to support their digital infrastructure. This trend has led to the spread of complex distributed computer systems built on cloud, multi-cloud, fog, and edge computing solutions. Such systems typically operate under conditions of potential technical failures (network, hardware, and software malfunctions) and external factors, and therefore must remain operational in the event of failures of individual components or even entire nodes. Modern businesses impose strict reliability requirements on these systems, since even minimal downtime can result in significant financial losses and a decline in customer trust. To achieve continuous operation of cloud-based distributed systems, it is necessary to implement models for automatic analysis of processing node status parameters and their recovery, and ensure high availability through data replication, architectural design, and component redundancy.

This article contains the analysis of models for ensuring fault tolerance and high availability in modern distributed systems, including cloud, multi-cloud, fog, and edge Internet of Things (IoT) systems. The article identifies and examines current methods for achieving fault tolerance and high availability for various types of distributed systems and provides recommendations for applying these models and methods based on the specifics of the environment and quality of service (QoS) requirements. In addition, the article proposes an adaptive data backup model for edge systems, which accounts for network failures using the Weibull distribution. The results of modeling the relationship between the target system's reliability level and the backup interval are presented. The main motivation of the conducted research is to systematize approaches to improving the reliability of distributed systems and to develop a backup strategy adapted for edge systems characterized by network-related component failures.

Key words: High availability, fault tolerance, distributed systems, cloud computing, multi-cloud, edge computing, IoT.

Дата надходження статті: 10.06.2025

Дата прийняття статті: 16.06.2025

Опубліковано: 27.10.2025

УДК 616.858-008.6:616.8-073.97:616.8-009.3
DOI <https://doi.org/10.32782/2663-5941/2025.4.2/04>

Біщак Д.С.

Тернопільський національний технічний
університет імені Івана Пулюя

Петрик М.Р.

Тернопільський національний технічний
університет імені Івана Пулюя

КОРЕЛЯЦІЯ МІЖ ГРАФОМОТОРНОЮ АКТИВНІСТЮ ТА ЕЕГ У ПАЦІЄНТІВ ІЗ ТРЕМОРОМ ПРИ ХВОРОБІ ПАРКІНСОНА

Прогресуюча нейродегенерація при хворобі Паркінсона (ХП) характеризується тремором спокою з частотою 4–6 Гц. Традиційні методи діагностики обмежені суб'єктивністю клінічної оцінки. Поєднання цифрових технологій створює передумови для об'єктивізації аналізу нейрофізіологічних основ тремору. Мета дослідження: визначення специфічних кореляційних зв'язків між показниками графомоторної активності та патернами електроенцефалографії (ЕЕГ) у пацієнтів із тремором при ХП для формування об'єктивного комбінованого нейрофізіологічного біомаркера. Дослідницька вибірка включала п'ять осіб із тремором різного походження (зокрема стадії 2–3 за шкалою Хен і Яра). Графомоторні параметри реєструвалися за допомогою планшету Huion KAMVAS Pro 16 graphics (250Гц), синхронно з 16-канальною ЕЕГ (500 Гц) під час виконання структурованих завдань. Запропоновано метрику «дельта-радіус» (ΔR) для квантифікації тремору. Здійснено спектральний аналіз, декомпозицію за частотними діапазонами та крос-кореляційний аналіз. Виявлено домінуючий пік ΔR у діапазоні 4–6 Гц, достовірне підвищення потужності ЕЕГ у тета-діапазоні (4–8 Гц) та зниження в альфа-діапазоні (8–13 Гц). Встановлено значущу кореляцію між амплітудою ΔR та потужністю тета-ритму в потиличних відведеннях ($r = 0,78, p < 0,01$) із часовим лагом 200–300 мс. Зафіксовано феномен «міжканальної синхронізації» – підвищення когерентності між усіма ЕЕГ-відведеннями в тета-діапазоні під час тремору з $0,35 \pm 0,08$ до $0,72 \pm 0,11$ ($p < 0,001$). Характер взаємозв'язку між ΔR та ЕЕГ демонстрував специфічність для ХП у порівнянні з іншими видами тремору. Вперше встановлено нейрофізіологічні кореляції між параметрами графомоторної активності та ЕЕГ-патернами при паркінсонічному треморі з часовим зсувом, що підтверджує причинно-наслідковий зв'язок. Феномен «міжканальної синхронізації» інтерпретується як прояв надмірної когерентності кортикальної активності. Запропонований комбінований аналіз може бути застосований для об'єктивної діагностики та моніторингу ефективності лікувальних інтервенцій при ХП.

Ключові слова: паркінсонічний тремор, дельта-радіус, електроенцефалографічні патерни, міжканальна синхронізація, кортикальна когерентність, нейрофізіологічний біомаркер.

Постановка проблеми. Хвороба Паркінсона (ХП) посідає друге місце за поширеністю серед нейродегенеративних захворювань, уражаючи близько 1 % населення старше 60 років [1]. Тремор у стані спокою з частотою 4–6 Гц є одним із кардинальних симптомів ХП, що спричиняє істотні функціональні обмеження [2]. Наявні інструменти оцінювання тяжкості ХП базуються здебільшого на суб'єктивному оцінюванні клініциста, що зумовлює варіабельність результатів, тому розроблення об'єктивних методів оцінювання симптомів становить актуальне завдання [3].

Механізми формування тремору на нейрофізіологічному рівні потребують подальшого вивчення. Відомо про дегенеративні зміни дофамінергічних нейронів чорної субстанції, що спричиняють дисбаланс у функціонуванні базальних гангліїв і таламо-кортикальних шляхів, проте взаємозв'язок між змінами електричної активності мозку та характеристиками тремору залишається недостатньо з'ясованим [4].

Електроенцефалографія (ЕЕГ) зарекомендувала себе як неінвазивний метод дослідження функціональної активності мозку з високою часовою роздільною здатністю [5]. Графомо-

торні тести із застосуванням цифрового планшета дають змогу кількісно вимірювати параметри руху при ХП [6], однак комплексні дослідження зв'язку між графомоторною активністю та електрофізіологічними характеристиками мозку при ХП є поодинокими.

Аналіз останніх досліджень і публікацій. Аналіз сучасної наукової літератури демонструє інтерес до дослідження взаємозв'язку між електрофізіологічними показниками мозку та моторними проявами ХП. Дослідження С. Фараші (S. Farashi) та співавторів [7] засвідчило можливість прогнозування епізодів тремору на основі аналізу ЕЕГ-сигналів із точністю до 73 %. У процесі вивчення графомоторних функцій Н. І. Ю (N. Y. Yu) та співавтори [8] виявили, що пацієнти з ХП демонструють виражене уповільнення під час виконання складних графомоторних завдань, а фаза декселяції при цільових рухах є суттєво подовженою.

Критичний огляд літератури дав змогу виявити декілька суттєвих прогалин у дослідженні взаємозв'язку між графомоторною активністю та ЕЕГ у пацієнтів із ХП, зокрема: відсутність комплексних досліджень, що одночасно фіксують та аналізують графомоторну активність і ЕЕГ-сигнали під час виконання стандартизованих рухових завдань; методологічна невизначеність щодо кількісної оцінки тремору під час графомоторних завдань, оскільки більшість дослідників використовує суб'єктивні шкали або загальні параметри руху, не пропонуючи чітких метрик для оцінювання тремору як відхилення від ідеальної траєкторії руху; брак досліджень, присвячених оцінці змін параметрів графомоторної активності та ЕЕГ-патернів під впливом медикаментозного лікування; недостатню увагу до частотно-специфічних кореляцій між ЕЕГ-активністю та характеристиками тремору; обмежене вивчення потенціалу комбінованого аналізу графомоторної активності та ЕЕГ як біомаркера для діагностики ХП [9–13].

Постановка завдання. Мета роботи: визначення кореляції між характеристиками тремору, зафіксованими під час графомоторної активності, та електроенцефалографічними патернами у пацієнтів із хворобою Паркінсона для розроблення комбінованого біомаркера діагностики та оцінки прогресування захворювання.

Для досягнення поставленої мети були сформульовані такі завдання: розробити метрику ΔR для кількісної оцінки тремору; провести одночасну реєстрацію графомоторної активності та ЕЕГ-

сигналів; виконати частотний і крос-кореляційний аналіз; дослідити феномен «резонансу каналів»; оцінити можливість використання комбінованого аналізу як біомаркера для діагностики ХП.

Виклад основного матеріалу. Обладнання та процедура дослідження графомоторної активності та ЕЕГ-сигналів у пацієнтів із ХП різного ступеня тяжкості наведені на рис. 1.

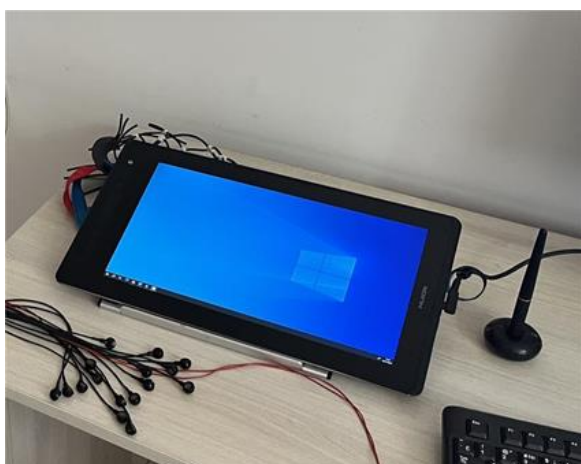
У дослідженні взяли участь п'ять осіб із тремором, зокрема три з ХП (стадії 2–3), а також двоє – з іншими його типами (табл. 1). Усі учасники були поінформовані про мету і методи дослідження та надали письмову згоду на участь. Дослідження було схвалено етичним комітетом і проводилося відповідно до принципів Гельсінської декларації.

Для реєстрації графомоторної активності використовували графічні планшети планшету Huion KAMVAS Pro 16 graphics (250Гц). Для фіксації ЕЕГ застосовувалася система ЕЕГ NEUROCOM, де розміщення електродів здійснено з акцентом на потиличній та тім'яній частках, а також ролі електродів REF та GND у навколоушних зонах, що пов'язане з участю задніх відділів мозку в управлінні рухом [14, 15].

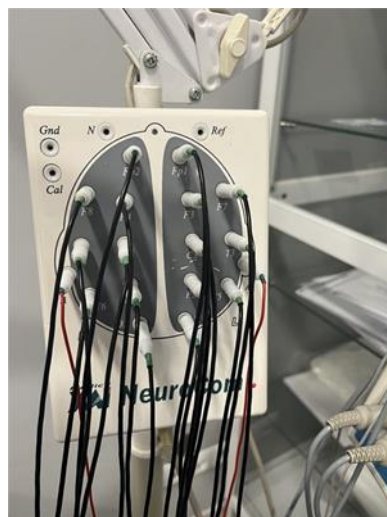
Процедура дослідження охоплювала підготовку учасника та виконання графомоторних завдань з одночасною реєстрацією ЕЕГ (малювання спіралі Архімеда). Кожне завдання тривало протягом приблизно до 2 хвилин, в залеженості від стану пацієнта із перервами між завданнями тривалістю 1 хвилину. Для оцінки ефективності медикаментозного лікування дослідження проводили до прийому дофамінергічних препаратів і через 40–60 хвилин після нього.

Для кількісного оцінювання тремору було розроблено метрику ΔR (радіус-відхилення), яку обчислювали як відхилення фактичної траєкторії від ідеальної: $\Delta R(i) = |R_{\text{факт}}(i) - R_{\text{ідеал}}(i)|$. Сигнал ΔR фільтрували для виокремлення компонентів, пов'язаних із тремором (3–7 Гц). ЕЕГ-сигнали обробляли з використанням фільтрації, декомпозиції за частотними діапазонами та аналізу когерентності. Для вивчення взаємозв'язку між графомоторною активністю та ЕЕГ-сигналами застосовували метод крос-кореляції.

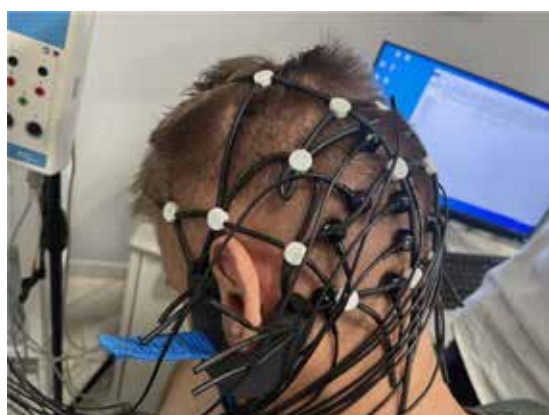
Результати дослідження. Аналіз ΔR -сигналів виявив характерні періодичні коливання з амплітудою, пропорційною інтенсивності тремору. Частотний аналіз показав наявність чіткого піку в діапазоні 4–6 Гц у всіх пацієнтів з ХП. У пацієнта № 4 з екстрапірамідним тремором частотний спектр демонстрував ширший діапазон піку (3–8 Гц), а пацієнт № 5 з хореїформними гіпер-



Графічні планшети для реєстрації графомоторної активності



Система NeuroPort



Спеціальні шапочки з електродами



Програмний інтерфейс, який відображає ЕЕГ-сигнали і параметри графомоторної активності в реальному часі

Рис. 1. Обладнання та процедура дослідження

Таблиця 1

Характеристики учасників дослідження

№	Стать	Вік, років	Діагноз	Примітки
1.	ж	55	Хвороба Паркінсона, стадія 3 за шкалою Хен і Яра	Сильний тремор (4–6 Гц) після зниження ефекту дії ліків
2.	ж	76	Хвороба Паркінсона, стадія 2 за шкалою Хен і Яра	Помірний тремор у правих кінцівках
3.	ч	73	Хвороба Паркінсона, стадія 3 за шкалою Хен і Яра	Після ліків проблеми із зором
4.	ч	32	Екстрапірамідний тремор неуточненого генезу	Порівняльна група
5.	ж	53	Наслідки менінгоенцефаліту з хореіформними гіперкінезами	Порівняльна група

кінезами мав нерегулярні, високоамплітудні коливання без вираженого частотного піку.

Аналіз ЕЕГ-сигналів виявив підвищення потужності в тета-діапазоні (4–8 Гц) та зниження в альфа-

діапазоні (8–13 Гц) у пацієнтів із ХП. Особливо важливим спостереженням було те, що посилення потужності в тета-діапазоні часто передувало посиленню тремору за даними ΔR на 200–300 мс.

Крос-кореляційний аналіз встановив високу кореляцію ($r = 0,78$, $p < 0,01$) між ΔR -сигналом і потужністю тета-ритму в потиличних відведеннях із часовим зсувом 200–300 мс, що свідчить про причинно-наслідковий зв'язок між змінами електричної активності мозку та проявами тремору. Детальні результати наведено в табл. 2.

Виявлено феномен «резонансу каналів» – підвищення когерентності між усіма ЕЕГ-відведеннями в тета-діапазоні під час епізодів посилення тремору від $0,35 \pm 0,08$ до $0,72 \pm 0,11$ ($p < 0,001$). Цей феномен спостерігався в усіх пацієнтів із ХП, але був найбільш вираженим у пацієнтів із тяжким тремором. Рівень когерентності корелював із клінічною оцінкою тяжкості тремору за шкалою Хен і Яра ($r = 0,85$, $p < 0,01$).

Виявлений часовий зсув між змінами електричної активності мозку та проявами тремору дає змогу припустити причинно-наслідковий зв'язок між цими явищами. Це узгоджується з патофізіологією ХП, згідно з якою порушення функції базальних гангліїв призводить до патологічної активності в таламо-кортикальних шляхах і, як наслідок, до тремору.

Феномен «резонансу каналів» можна інтерпретувати як прояв патологічної синхронізації активності різних ділянок кори головного мозку під час епізодів тремору. У нормі базальні ганглії забезпечують селективне гальмування та активацію кортикальних нейронів, що дає змогу виконувати тонкі координовані рухи. При ХП дегенерація дофамінергічних нейронів чорної субстанції порушує цю регуляцію.

Проведені дослідження продемонстрували можливість прогнозування епізодів тремору на основі аналізу ЕЕГ-сигналів, що узгоджується з виявленим нами часовим зсувом. Це надало

можливість описати особливості графомоторних функцій при ХП, які були доповнені аналізом кореляції з ЕЕГ-активністю.

Обмеження здійсненого дослідження були зумовлені невеликою кількістю учасників, використанням 8-канальної ЕЕГ-системи та наявністю рухових артефактів у ЕЕГ-сигналах. Проте результати демонструють потенціал комбінованого аналізу графомоторної активності та ЕЕГ для розробки нових діагностичних підходів.

Висновки.

1. Здійснене дослідження дало змогу встановити статистично значущий зв'язок між характеристиками тремору руки, виявленими під час графомоторної активності, та електроенцефалографічними патернами у пацієнтів із хворобою Паркінсона.

2. Розроблена метрика ΔR (радіус-відхилення) допомагає кількісно оцінювати тремор під час виконання графомоторних завдань та має високу чутливість до змін у стані пацієнта.

3. Частотний аналіз ΔR -сигналів свідчить про наявність чіткого піку в діапазоні 4–6 Гц у пацієнтів із хворобою Паркінсона, що відповідає типовій частоті паркінсонічного тремору.

4. У пацієнтів із хворобою Паркінсона спостерігається підвищення спектральної потужності ЕЕГ у тета-діапазоні (4–8 Гц) та зниження потужності в альфа-діапазоні (8–13 Гц) у порівнянні з порівняльною групою.

5. Виявлено високу кореляцію ($r = 0,78$, $p < 0,01$) між ΔR -сигналом і потужністю тета-ритму в потиличних відведеннях ЕЕГ з часовим зсувом 200–300 мс (ЕЕГ випереджає ΔR), що свідчить про причинно-наслідковий зв'язок між змінами електричної активності мозку та проявами тремору.

Таблиця 2

Результати крос-кореляційного аналізу між ΔR та ЕЕГ-сигналами

Частотний діапазон	Відведення	Коефіцієнт кореляції (r)	Часовий зсув (мс)	p-значення
Дельта (0,5–4 Гц)	O1	0,32	150	0,08
Дельта (0,5–4 Гц)	O2	0,29	140	0,09
Тета (4–8 Гц)	O1	0,78	210	<0,01
Тета (4–8 Гц)	O2	0,75	220	<0,01
Тета (4–8 Гц)	P3	0,65	260	<0,01
Тета (4–8 Гц)	P4	0,62	250	<0,01
Альфа (8–13 Гц)	O1	-0,45	180	0,02
Альфа (8–13 Гц)	O2	-0,42	190	0,03
Бета (13–30 Гц)	O1	0,21	90	0,14
Бета (13–30 Гц)	O2	0,19	80	0,17

6. Феномен «резонансу каналів» – підвищення когерентності між усіма ЕЕГ-відведеннями в тета-діапазоні під час епізодів посилення тремору – може бути інтерпретовано як прояв патологічної синхронізації активності різних ділянок кори головного мозку та використано як біомаркер хвороби Паркінсона.

7. Патерн кореляції між ΔR та ЕЕГ-сигналами відрізняється у пацієнтів із хворобою Паркінсона та пацієнтів з іншими видами тремору, що відкриває можливості для використання комбінованого

аналізу графомоторної активності та ЕЕГ у диференційній діагностиці.

Перспективи подальших досліджень полягають у збільшенні кількості учасників, використанні багатоканальних ЕЕГ-систем для більш повної оцінки просторового розподілу електричної активності мозку, а також у розробленні методів автоматичного виявлення характерних патернів кореляції між графомоторною активністю та ЕЕГ для клінічного застосування.

Список літератури:

1. Чиняк О., Дубенко О., Потапов О., Шульга А., Коцюба А. Хвороба Паркінсона – огляд сучасних методів лікування. *EUMJ*. 2023. Т. 11, № 1. С. 1–13. DOI: [https://doi.org/10.21272/eumj.2023;11\(1\):1-13](https://doi.org/10.21272/eumj.2023;11(1):1-13)
2. Петрик М., Біщак Д., Бачинський М., Бревус В., Чиж В., Михалик Д. Аналіз мимовільних рухів пацієнтів із симптомами тремору під впливом когнітивних впливів. *Прикладні питання математичного моделювання*. 2024. Т. 7, № 2. С. 150–165. DOI: <https://doi.org/10.32782/mathematical-modelling/2024-7-2-14>
3. Osiichuk I., Brevus V., Bishchak D., Mashtaliar Y., Mudryk I. Leveraging graphics tablet and JPen library to detect essential tremor. *CEUR Workshop Proceedings*. 2024. Vol. 3742. URL: <https://ceur-ws.org/Vol-3742/paper8.pdf> (date of access: 20.05.2025).
4. Zaiarnyi M. Multi-sensor analysis of cognitive signals for neurological disorders and diseases. *CEUR Workshop Proceedings*. 2024. Vol. 3742. URL: <https://ceur-ws.org/Vol-3742/paper21.pdf> (date of access: 20.05.2025).
5. Мудрик І. Я. Автоматизовані системи діагностування стану пацієнтів, хворих на есенціальний тремор : автореф. дис. ... д-ра філософії : 121 – Інженерія програмного забезпечення. Київ, 2021. 16 с. URL: <http://www.irbis-nbuv.gov.ua/aref/0821U102409> (дата звернення: 20.05.2025).
6. Petrik M. R., Gancarczyk T., Himič O. M. Methods of mathematical modeling and identification of complex processes and systems on the basis of high-performance calculations: (neuro-and nanoporous feedback cyber system, models with sparse structure data, parallel computations). Bielsko-Biała : Wydawnictwo Naukowe Akademii Techniczno-Humanistycznej, 2021. URL: <http://elartu.tntu.edu.ua/handle/lib/35999> (date of access: 20.05.2025).
7. Farashi S., Sarihi A., Ramezani M., Kashani H. Parkinson's disease tremor prediction using EEG data analysis – A preliminary and feasibility study. *BMC Neurology*. 2023. Vol. 23, № 1. P. 420. DOI: <https://doi.org/10.1186/s12883-023-03468-0>
8. Yu N. Y., Van Gemmert A. W. A., Chang S. H. Characterization of graphomotor functions in individuals with Parkinson's disease and essential tremor. *Behavior Research Methods*. 2017. Vol. 49, № 3. P. 913–922. DOI: <https://doi.org/10.3758/s13428-016-0752-y>
9. Brambilla C., Pirovano I., Mira R. M., Rizzo G., Scano A., Mastropietro A. Combined use of EMG and EEG techniques for neuromotor assessment in rehabilitative applications: A systematic review. *Sensors*. 2021. Vol. 21, № 21. P. 7014. DOI: <https://doi.org/10.3390/s21217014>
10. Wastensson G. Quantitative Methods for Evaluation of Tremor and Neuromotor Function. Gothenburg University, 2010. URL: <http://hdl.handle.net/2077/23133> (date of access: 20.05.2025).
11. Rothe J., Kattlun F. A., Kaufmann J., Uhlmann A., Wanderer S., Bluschke A., Roessner V. Effects of methylphenidate and physiotherapeutic treatment on graphomotor movements in children with ADHD. *European Child & Adolescent Psychiatry*. 2024. Vol. 33, № 1. P. 127–137. DOI: <https://doi.org/10.1007/s00787-023-02144-5>
12. Zhang X., Chen H., Tao L., Zhang X., Wang H., He W., Fang W. Combined multivariate pattern analysis with frequency-dependent intrinsic brain activity to identify essential tremor. *Neuroscience Letters*. 2022. Vol. 776. P. 136566. DOI: <https://doi.org/10.1016/j.neulet.2022.136566>
13. Góral-Pórola J., Mirski A., Knapik H., Pačalska M. Functional neuromarkers in Parkinson's disease (PD). *Acta Neuropsychologica*. 2021. Vol. 19. P. 147–168. URL: <https://actaneuropsychologica.com/article/148146/en>
14. Дегтяренко Т. В. Психофізіологія рухової діяльності : навч. посіб. для студ. закл. вищ. освіти 2024. 379. <http://dspace.pdpu.edu.ua/handle/123456789/17257>
15. Andersen R. A. From thought to action: The brain-machine interface in posterior parietal cortex. *Proceedings of the National Academy of Sciences*. 2019. Vol. 116. 52. P. 26274–26279. <https://doi.org/10.1073/pnas.1902276116>

Bishchak D.S., Petryk M.R. CORRELATION BETWEEN GRAPHOMOTOR ACTIVITY AND EEG IN PATIENTS WITH TREMOR IN PARKINSON'S DISEASE

Progressive neurodegeneration in Parkinson's disease (PD) is characterized by resting tremor with a frequency of 4–6 Hz. Conventional diagnostic methods are limited by the subjectivity of clinical assessment. The integration of digital technologies provides a foundation for the objective analysis of the neurophysiological basis of tremor. The aim of this study was to identify specific correlation patterns between graphomotor activity indicators and electroencephalographic (EEG) patterns in patients with tremor associated with PD to develop an objective combined neurophysiological biomarker. Materials and Methods. The study sample included five individuals: three patients with Parkinson's disease (Hoehn and Yahr stages 2–3) and two individuals with other types of tremor. Graphomotor parameters were recorded using Huion KAMVAS Pro 16 graphics tablet (250Hz), synchronous with 16-channel EEG (500Hz) during structured task performance. A novel metric, the “delta radius” (ΔR), was introduced to quantify tremor. Spectral analysis, frequency band decomposition, and cross-correlation analysis were performed. A dominant ΔR peak was identified in the 4–6 Hz range, along with a significant increase in EEG power within the theta band (4–8 Hz) and a decrease in the alpha band (8–13 Hz). A significant correlation was observed between ΔR amplitude and theta rhythm power in occipital leads ($r = 0.78$, $p < 0.01$) with a time lag of 200–300 ms. A phenomenon of “interchannel synchronization” was recorded, with an increase in EEG coherence across all channels in the theta range during tremor from 0.35 ± 0.08 to 0.72 ± 0.11 ($p < 0.001$). The relationship between ΔR and EEG demonstrated specificity to PD when compared to other types of tremor.

For the first time, neurophysiological correlates between graphomotor activity parameters and EEG patterns in parkinsonian tremor with temporal lag were established, confirming a causal relationship. The “interchannel synchronization” phenomenon is interpreted as a manifestation of excessive cortical coherence. The proposed combined analysis may be applied for objective diagnosis and monitoring of therapeutic interventions in PD.

Key words: parkinsonian tremor; delta radius, electroencephalographic patterns, interchannel synchronization, cortical coherence, neurophysiological biomarker.

Дата надходження статті: 14.07.2025

Дата прийняття статті: 25.07.2025

Опубліковано: 27.10.2025

Большакова О.О.

ПЗВО «Міжнародний класичний університет імені Пилипа Орлика»

Кошкін В.К.

ПЗВО «Міжнародний класичний університет імені Пилипа Орлика»

Арамян А.М.

ПЗВО «Міжнародний класичний університет імені Пилипа Орлика»

Бандура В.М.

ПЗВО «Міжнародний класичний університет імені Пилипа Орлика»

Єганов О.Ю.

ПЗВО «Міжнародний класичний університет імені Пилипа Орлика»

СУЧАСНІ ПІДХОДИ ДО ПОБУДОВИ ЕФЕКТИВНИХ ТА БЕЗПЕЧНИХ КОМП'ЮТЕРНИХ СИСТЕМ І МЕРЕЖ

У статті проаналізовано актуальні підходи до проєктування ефективних та безпечних комп'ютерних систем і мереж (КСМ) у контексті цифрової трансформації та зростання кіберзагроз. Показано, що збільшення обсягів даних, розвиток хмарної інфраструктури та підвищення складності ІТ-середовищ потребують запровадження архітектурних рішень, здатних забезпечити гнучкість, масштабованість та кіберстійкість. Розглянуто мікросервісну та сервісно-орієнтовану архітектуру як основу для побудови адаптивних КСМ, що легко масштабуються й оновлюються відповідно до змін бізнес-вимог.

Хмарні технології визначені як ключовий елемент сучасної ІТ-інфраструктури, проте підкреслено ризики, які вони породжують: збільшення поверхні атак, ускладнення захисту даних і потенційна втрата контролю. У зв'язку з цим обґрунтовано доцільність впровадження принципу *Security by Design*, що передбачає інтеграцію безпеки на всіх етапах розробки, а також застосування *Zero Trust Architecture (ZTA)*, яка забезпечує постійний контроль і перевірку доступу незалежно від джерела запиту.

Детально проаналізовано роль таких механізмів, як багатобакторна автентифікація (MFA) і наскрізне шифрування, як засобів підвищення кіберстійкості систем. Окрема увага приділена використанню штучного інтелекту (ШІ) в управлінні ресурсами, прогнозуванні збоїв і виявленні аномалій у режимі реального часу. ШІ розглянуто як інструмент підвищення динамічності систем захисту, зокрема завдяки адаптивному налаштуванню політик безпеки та аналізу поведінки користувачів. Водночас відзначено виклики, пов'язані із змагальними атаками, що вимагають стійких алгоритмів.

Результатом дослідження є запропонована інтегрована концепція побудови КСМ, яка синтезує можливості СОА, принципи ZTA та інтелектуальні засоби на базі ШІ. Такий підхід дозволяє створювати безпечні, адаптивні та ефективні системи для використання в державному, корпоративному й хмарному середовищах.

Ключові слова: мікросервісна архітектура, сервісно-орієнтована архітектура, кібербезпека, хмарні технології, штучний інтелект, *Zero Trust*.

Постановка проблеми. У сучасних умовах цифрової трансформації комп'ютерні системи та мережі (КСМ) є критично важливою інфраструктурою для державних і комерційних структур. Стрімкий розвиток цифрових технологій

ускладнив архітектуру КСМ, поставивши перед ними нові вимоги до продуктивності, масштабованості та безпеки, особливо в умовах використання хмарних технологій (AWS, Azure, Google Cloud). Хоча хмарні рішення забезпечують гнуч-

кість і економічність, вони одночасно створюють нові загрози – зростає поверхня атаки, виникають ризики втрати контролю над даними та проблеми з налаштуванням безпеки [1-2].

З огляду на це, інтеграція штучного інтелекту (ШІ) у КСМ набуває стратегічного значення. ШІ дозволяє автоматизувати управління ресурсами, виявляти аномалії у реальному часі, формувати адаптивні системи захисту на основі поведінкового аналізу (наприклад, UEBA), що підвищує рівень кіберзахисту в динамічних середовищах [3]. Водночас поєднання ШІ з хмарною інфраструктурою створює нові виклики, пов'язані з adversarial-атаками, проблемами конфіденційності та потребою у захисті гетерогенних і розподілених середовищ [4].

У відповідь на це виникає потреба у створенні інтегрованих архітектурних рішень, які поєднують хмарні сервіси, ШІ та засоби кіберзахисту в єдину масштабовану і керовану інфраструктуру [5]. Важливу роль у цьому відіграє концепція Zero Trust Architecture (ZTA), яка передбачає перевірку кожної дії незалежно від контексту. Ефективна кібербезпека має охоплювати як технічні (TLS, Post-Quantum Crypto, 2FA, SIEM), так і організаційні заходи [6]. Отже, науково-прикладна проблема полягає не лише у впровадженні окремих технологій, а у розробці цілісних, безпечних та адаптивних підходів до побудови КСМ, здатних відповідати вимогам сучасного цифрового середовища.

Аналіз останніх досліджень і публікацій. Стрімкий розвиток цифрових технологій, зокрема хмарних обчислень, IoT і гібридних мереж, кардинально трансформують архітектуру КСМ, водночас створюючи складні та швидкоплинні кіберзагрози. Це робить традиційні методи захисту застарілими та потребує нових гнучких підходів, серед яких особливу увагу привертає архітектура Zero Trust (ZTA). У поєднанні з технологіями ШІ, такими як машинне навчання та виявлення аномалій, ZTA дозволяє динамічно адаптувати політики безпеки до нових загроз, що особливо актуально в хмарних і розподілених середовищах.

Системи виявлення вторгнень (IDS), інтегровані з алгоритмами ШІ, є ключовими компонентами сучасної кібербезпеки [1]. Наприклад, у роботі Касонго С. (Kasongo S.) [7] описано IDS-платформу на основі рекурентних нейронних мереж (RNN), таких як LSTM і GRU, яка демонструє високу точність виявлення атак. Також застосування алгоритму XGBoost для вибору релевантних ознак дозволяє зменшити обчислювальне навантаження. Інше дослідження Маюра-

натан М., Сараванан С., Мутусенхіл Б. та Самудураї А. (Mayuranathan M., Saravanan S., Muthusenthil B., Samyudurai A.) [8] пропонує гібридний підхід EOS-IDS із використанням глибоких мереж Кронекера (DKNN) для точного виявлення загроз у хмарному середовищі. Робота Луо Х. (Luo X.) [9] презентує адаптивну IDS-модель на основі нечіткої логіки, яка забезпечує гнучке реагування на вторгнення в режимі реального часу.

Серед сучасних підходів особливо ефективним вважається виявлення аномалій на основі мережових потоків, що дозволяє відстежувати складні загрози, не помітні при аналізі окремих пакетів. У дослідженні Хефзіпах Дж., Валем Р., Шіла М. та Дханалакшмі Дж. (Nephzipah J., Vallem R., Sheela M. & Dhanalakshmi G.) [10] розроблено вдосконалену систему безпеки MMGT-ANN, засновану на штучній нейронній мережі з оптимізацією за теорією мінімаксних ігор, що забезпечує високу точність та швидке реагування на загрози.

Зростає популярність проактивних стратегій захисту, орієнтованих на передбачення загроз і динамічне оновлення політик безпеки. У роботі автора Вонг Й. (Wang Y.) [11] представлено інтелектуальну багаторівневу систему захисту в хмарі, яка об'єднує збір даних, аналіз загроз і адаптивний контроль доступу, що зменшує кількість помилкових спрацювань. Інше дослідження Пейді П. (Paidy P.) [12] досліджує адаптивне тестування безпеки в рамках DevSecOps, де інтеграція ШІ дозволяє автоматизувати тестування у реальному часі, реагуючи на нові вразливості. У роботі Зафера Ф. та Ахтара Ф. (Zafer F., & Akhtar F.) [13] запропоновано хмарне рішення з використанням ШІ для виявлення загроз у реальному часі, що дозволяє значно скоротити тривалість атак та операційні ризики.

Комбінування ZTA з ШІ також розглядається як революційний підхід до побудови сучасної безпеки КСМ. Дослідження Сайрай К. (Sairaj K.) [14] демонструє, що така архітектура перевершує традиційні стратегії за точністю виявлення загроз, швидкістю реакції та загальним рівнем захищеності системи.

Разом із тим, однією з ключових загроз для ШІ-систем залишаються змагальні атаки (adversarial attacks), які використовують майже непомітні зміни у вхідних даних для обману моделі. У роботі Вонг Й., Сан Т., Лі С., Юань Х. та ін. (Wang Y., Sun T., Li S., Yuan X. Et al.) [15] здійснено комплексний огляд таких атак і методів захисту, зокрема з використанням регуляризації, виявлення аномалій і підвищення стійкості

моделей. Автори класифікують наявні рішення та пропонують ефективні підходи для підвищення надійності ІІІ у системах безпеки.

Постановка завдання. Метою статті є комплексне дослідження сучасних підходів до побудови ефективних та безпечних комп'ютерних систем і мереж (КСМ) з урахуванням викликів цифрової трансформації. Зокрема, ставляться такі дослідницькі цілі:

- проаналізувати архітектурні рішення, що забезпечують гнучкість, масштабованість і надійність ІТ-інфраструктури, зокрема мікросервісну архітектуру (МСА) та сервісно-орієнтовану архітектуру (СОА);
- охарактеризувати роль хмарних технологій у забезпеченні масштабованості та економічної ефективності КСМ, а також виявити супутні ризики для кібербезпеки;
- дослідити сучасні підходи до кіберзахисту, зокрема застосування концепції Zero Trust Architecture, засобів шифрування, моніторингу та виявлення загроз;
- виявити переваги використання штучного інтелекту (ІІІ) для оптимізації, управління та забезпечення безпеки КСМ, включаючи виявлення аномалій, прогнозування інцидентів і адаптивне управління;
- розробити інтеграційну концепцію побудови КСМ на основі об'єднання принципів СОА, Zero Trust Security та можливостей ІІІ.

Виклад основного матеріалу. Побудова ефективних та безпечних КСМ вимагає комплексного підходу, що поєднує в собі технології, практики та постійну адаптацію до нових загроз. Основні напрямки включають архітектурні рішення, кібербезпеку, оптимізацію продуктивності та управління. Розглянемо кожен підхід окремо, де буде охарактеризовано архітектурні підходи, підходи щодо питань з кібербезпеки, підходи з використанням ІІІ, що розглядається для оптимізації та управління ефективністю, а також інтелектуального управління КСМ.

1. Архітектурні підходи.

Мікросервісна архітектура (МСА). Одним із ключових напрямів у побудові ефективних комп'ютерних систем є використання МСА. Цей підхід передбачає розбиття монолітного додатку на сукупність малих, незалежних сервісів, кожен з яких реалізує окрему бізнес-функцію та взаємодіє з іншими сервісами через стандартизовані API або повідомлення.

Перевагою підходу МСА є можливість незалежної розробки, тестування, розгортання та

масштабування окремих компонентів системи. Це значно підвищує гнучкість архітектури, дозволяє командам працювати автономно, сприяє оперативному оновленню сервісів та зменшує час виведення нових функцій на ринок. Крім того, підхід МСА забезпечує високу стійкість до збоїв, наприклад, у разі відмови одного з сервісів система в цілому продовжує функціонувати, обмежуючи область впливу інциденту. Такий підхід є критично важливим для розподілених систем, зокрема у хмарних обчисленнях, де забезпечення безперервності обслуговування користувачів має вирішальне значення.

Хмарні технології. ХТ є однією з важливої складових сучасної ІТ-інфраструктури, що забезпечує високу масштабованість, доступність та економічну ефективність. Залежно від потреб користувача чи організації використовуються а) публічні, б) приватні та в) гібридні хмарні інфраструктури, кожна з яких має свої переваги у контексті безпеки, контролю над даними та витрат.

На рис. 1 показано схему з основними принципами побудови архітектури хмарних обчислень, яка складається з моделей розгортання та основних послуг в залежності від користувацьких потреб. Цей рисунок візуалізує основні аспекти ХТ як фундаментальної складової сучасної ІТ-інфраструктури, що забезпечує високу масштабованість, доступність та економічну ефективність.

В центральній частині схеми відображено глобальний характер ХТ, що підкреслює їхню доступність з будь-якої точки світу. Збільшення та зменшення ресурсів, позначене розширювальними стрілками, ілюструє масштабованість. Елементи, що нагадують гроші, вказують на економічну ефективність, яка досягається завдяки моделі оплати за фактичне використання.

Рисунок розрізняє три основні моделі розгортання хмари, а саме:

– публічна хмара (Public Cloud) представлена хмарою, до якої підключається багато користувачів, символізуючи спільне використання ресурсів, яка є більш гнучкою та економічною, а також доцільною для некритичних додатків, розробки, тестування та швидкого масштабування;

– приватна хмара (Private Cloud) зображена більше як ізольована хмара в межах захищеного параметра, що підкреслює вищий рівень контролю, безпеки та відповідності регуляторним вимогам, оскільки інфраструктура виділена виключно для однієї організації;

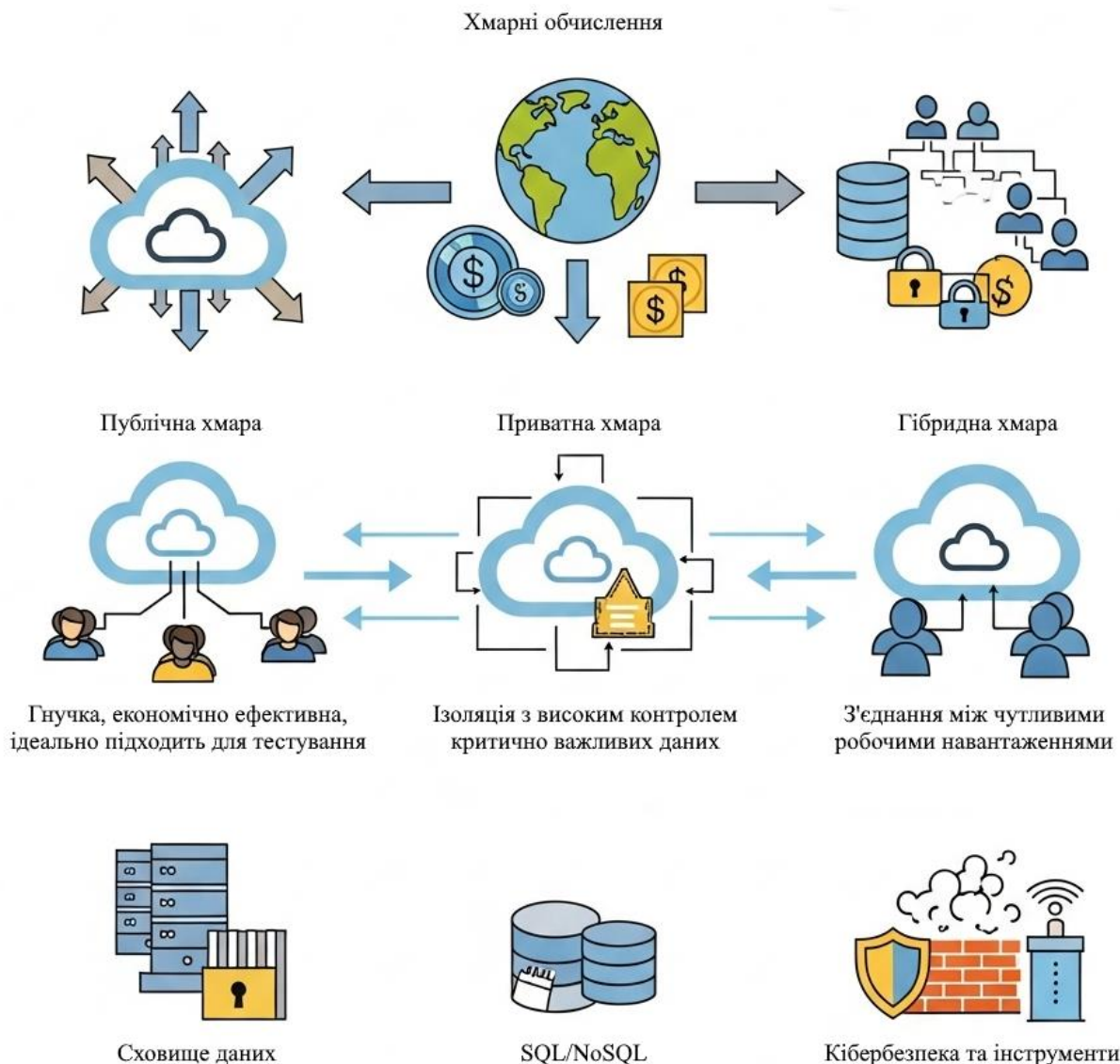


Рис. 1. Основні принципи архітектури хмарних обчислень з моделями розгортання та основними послугами

– гібридна хмара (Hybrid Cloud) представлена зв'язком між публічною та приватною хмарами, з потоками даних між ними, що демонструє її здатність поєднувати переваги обох підходів, дозволяючи розміщувати чутливі дані у приватній частині, а менш критичні навантаження – у публічній.

Сучасні хмарні провайдери надають широкий спектр сервісів, серед яких можна виділити: обчислювальні ресурси (compute) для розгортання і масштабування додатків; сховища даних для зберігання структурованої та неструктурованої інформації; керовані бази даних (SQL, NoSQL) з високою доступністю та надійністю; засоби та інструменти кібербезпеки, що містить шифрування, міжмережіві екрани, системи виявлення загроз, інструменти дотримання нормативів тощо.

Основні послуги провайдерів наведені в нижній частині схеми, які представляють ключові категорії сервісів, які надають хмарні провайдери. До них можна віднести:

- обчислювальні ресурси, які символізують собою іконки серверів, що вказують на віртуальні машини, контейнери та функції без серверів для розгортання та масштабування додатків;

- сховище даних, представлені у вигляді іконок сховищ або дисків, що відображають можливість для зберігання структурованої та неструктурованої інформації;

- керовані бази даних (SQL/NoSQL), позначені як бази даних, підкреслюючи високу доступність та надійність керованих рішень за допомогою SQL та NoSQL;

– засоби та інструменти кібербезпеки, відображені іконками щитів, замків та міжмережних екранів, що вказує на інтегровані можливості шифрування, моніторинг загроз та інструментів відповідності нормативним вимогам.

На основі рис. 1 можна зазначити, що ХТ є комплексним та гнучким рішенням, оскільки забезпечується необхідна інфраструктура та сервіси для побудови сучасних, ефективних та безпечних КСМ.

Сервісно-орієнтована архітектура (COA). COA є архітектурним стилем для розробки програмного забезпечення, який фокусується на створенні додатків з використанням слабо пов'язаних незалежних компонентів, що називаються сервісами. Кожен такий сервіс виконує певну бізнес-функцію і взаємодіє з іншими сервісами за допомогою стандартизованих протоколів.

Основні принципи COA пов'язані з: сервісами (які є самостійними, незалежними від програмних одиниць, які надають певну бізнес-функціональність, наприклад, сервіс з обробкою платежів); слабкою зв'язаністю (сервіси мінімально залежать один від одного, але зміна одного сервісу не повинна суттєво впливати на інші, що підвищує гнучкість системи); повторне використання (де однією з переваг є можливість повторного використання сервісів без написання

коду з нуля); стандартизовані інтерфейси (сервіси спілкуються через чітко визначені, стандартизовані інтерфейси, що дозволяє їм взаємодіяти незалежно від мови програмування чи платформи); абстракція (внутрішня реалізація сервісу прихована від його споживачів). До переваг COA можна віднести гнучкість та адаптивність, прискорення розробки, спрощення інтеграції та масштабованість.

Важливо розрізняти COA та MSA, оскільки MSA є більш суворою, специфічною реалізацією принципів COA. Мікросервіси ще більше «розбивають» додаток на дуже малі, незалежні сервіси, кожен з яких фокусується на одній функції та може розгортатися абсолютно незалежно. SOA ж є ширшою концепцією, яка може включати як великі, так і малі сервіси, і часто передбачає централізовані механізми комунікації. На рис. 2 показано порівняння MSA та COA.

2. Підходи щодо питань з кібербезпеки.

Безпека є невід'ємною складовою проектування, розгортання та експлуатації сучасних комп'ютерних систем і мереж. Ще на етапі розробки має бути реалізований принцип безпеки за замовчуванням (Security by Design), що передбачає інтеграцію захисних механізмів у кожен фазу життєвого циклу системи. Це охоплює проведення аналізу загроз, моделювання ризиків та



Рис. 2. Порівняння підходів побудови архітектури MSA та COA для КСМ

проектування захисту з урахуванням потенційних вразливостей з самого початку, а не як вторинний етап після реалізації функціональності.

Одним із базових підходів до забезпечення надійності є концепція нульової довіри (Zero Trust Architecture), що ґрунтується на принципі «нікому не довіряти, завжди перевіряти». У цій моделі кожен користувач, пристрій або сервіс незалежно від місця розташування повинен пройти сувору автентифікацію та авторизацію перед отриманням доступу до системних ресурсів. Такий підхід мінімізує ризики внутрішніх загроз і несанкціонованого доступу.

Важливе місце в системі кіберзахисту займає наскрізне шифрування даних як під час їх зберігання (на дисках), так і під час передачі мережею. Це дозволяє захистити конфіденційну інформацію навіть у випадку перехоплення чи несанкціонованого доступу до каналів зв'язку. Додатковим рівнем безпеки є багатофакторна автентифікація (БАФ), що передбачає використання кількох способів підтвердження особи (наприклад, поєднання пароля з кодом із SMS або біометричними даними), істотно ускладнюючи можливість злому облікових записів.

На рис. 3 показано архітектуру ZTA, яка на відміну від традиційних моделей, де ідентичність визначається на основі географічного положення або попередньо перевіреної аутентифікації, ZTA базується на припущенні, що порушення будуть, і їх необхідно контролювати.

З моменту первинної концептуалізації ZTA ця модель швидко поширилася в промисловості та урядових організаціях завдяки своїй здатності

адаптуватися до ризиків у нових децентралізованих та гібридних схемах обчислювальних систем. Основні елементи ZTA включають контроль ідентичності та доступу, сегментацію та моніторинг, що відображає поточні мінливі та зростаючі вимоги підприємств до роботи з хмарними системами, збільшення кількості віддалених працівників та способи підключення до систем за допомогою Інтернету речей.

3. Підходи з використанням ІІІ.

ІІІ є важливим для побудови як ефективних, так і безпечних КСМ, оскільки його застосування охоплює широкий спектр можливостей. Окрім того, ІІІ дозволяє вирішувати складні завдання, оскільки традиційні методи можуть бути надто трудомісткими або неможливими. Підходи використання ІІІ охоплюють собою наступне оптимізацію та управління ефективністю, покращення безпеки, а також інтелектуальне управління мережею.

ІІІ підвищує ефективність управління ІТ-ресурсами завдяки автоматизованому прогнозуванню потреб в обчислювальних можливостях (накопичувач, оперативна пам'ять, процесор тощо), що дозволяє динамічно масштабувати системи та оптимізувати використання інфраструктури, особливо в хмарних середовищах, знижуючи при цьому витрати. Крім того, алгоритми ІІІ аналізують мережевий трафік у реальному часі, виявляють перевантаження та охоплюють процес маршрутизації даних найоптимальнішими шляхами, що підвищує пропускну здатність та зменшує затримки. Завдяки прогнозованому обслуговуванню, яке базується на аналізі телеметрії обладнання та

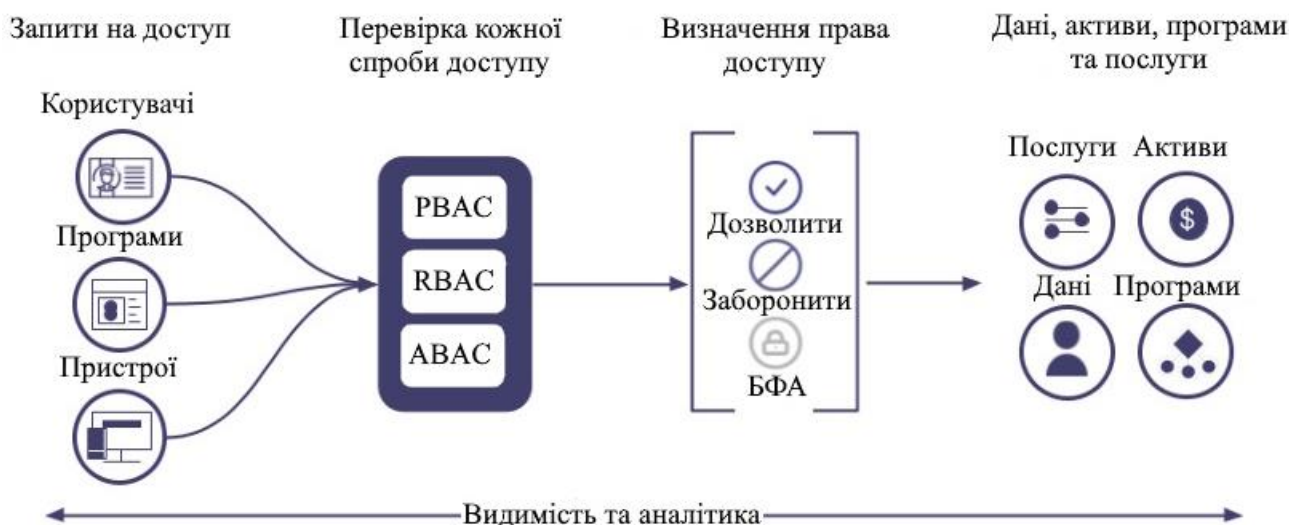


Рис. 3. Загальна архітектура Zero Trust

ПЗ, ШІ може передбачати потенційний збій в системі та проблеми з продуктивністю, що дозволяє здійснювати проактивне технічне обслуговування та мінімізувати час простою систем.

Штучний інтелект істотно покращує кібербезпеку завдяки здатності ефективно виявляти аномалії та нові загрози, виходячи з аналізу нормальної поведінки систем і мереж, що особливо важливо для боротьби з невідомими атаками, зокрема атаками нульового дня. Окрім виявлення, ШІ автоматизує реагування на інциденти, блокуючи шкідливий трафік, ізолюючи заражені пристрої або відкатуючи системи до безпечного стану. Завдяки аналізу великих обсягів журналів та подій безпеки, ШІ швидко корелює дані з різних джерел, виявляючи приховані закономірності та складні атаки, недоступні людині. Крім того, системи ШІ покращують автентифікацію, аналізуючи поведінкові патерни користувачів як додатковий фактор безпеки, та класифікують нові зразки шкідливого програмного забезпечення для швидкого виявлення їхніх характеристик. Важливою є і допомога ШІ в управлінні вразливостями, оскільки ШІ здатний розставити пріоритети та усунути їх, оцінюючи реальний ризик для конкретної інфраструктури.

Інтелектуальне управління мережею включає застосування ШІ для створення самостійно організованих мереж, які самостійно оптимізують та налаштовують свої параметри, значно знижуючи навантаження на адміністраторів, а також для прогнозного аналізу відмов шляхом вивчення історії збоїв і метрик продуктивності мережевого обладнання з метою своєчасного проактивного обслуговування. Таким чином, штучний інтелект виступає не просто додатковим інструментом, а є ключовим елементом сучасних методів побудови ефективних і безпечних комп'ютерних систем і мереж, що відкриває нові можливості для автоматизації, оптимізації та захисту.

4. Інтеграція COA та ZTS з використанням ШІ для створення загальної концепції побудови ефективних та безпечних КСМ.

Загальна концепція COA-ZTS на основі ШІ полягає в інтеграції принципів Zero Trust Security безпосередньо у сервісно-орієнтовані системи за допомогою штучного інтелекту, де ШІ забезпечує постійний моніторинг, аналіз поведінки користувачів і сервісів, автоматичне виявлення аномалій та адаптивне управління доступом у режимі реального часу.

Поєднання штучного інтелекту та архітектури Zero Trust може запропонувати новий погляд на

вирішення існуючих слабких місць як традиційних, так і інноваційних систем безпеки. В даному випадку інтеграція штучного інтелекту в принципи ZTA може допомогти організаціям створити контекстну безпеку, яка змінюється в режимі реального часу відповідно до нових загроз. Штучний інтелект покращує ZTA шляхом автоматизації функцій перевірки, визначення аномалій поведінки та встановлення нових політик на основі змінних ризикових середовищ.

На рис. 4 зображено концептуальну схему, яка ілюструє сучасні підходи до побудови ефективних та безпечних КСМ, який візуалізує взаємозв'язок між ключовими технологіями та архітектурними принципами, які забезпечують масштабованість, надійність та захищеність ІТ-інфраструктури.

Рисунок складається з двох основних секцій, що відображають архітектурні підходи та аспекти безпеки. Верхня секція з архітектурними підходами (архітектура мікросервісів) відображає схему з представленою мережею взаємопов'язаних, але окремих блоків, які з'єднані лініями й взаємодіють між собою. Такий архітектурний стиль показує, як великий монолітний блок розбивається на малі незалежні сервіси, де кожен сервіс виконує одну конкретну бізнес-функцію і може бути розроблений, розгорнутий або масштабований окремо, що забезпечує високу гнучкість, стійкість до збоїв (де відмова одного сервісу не впливає на всю систему) та швидкість розробки.

Хмарне середовище показує взаємодію та потоки даних, що передбачає використання обчислювальних ресурсів (серверів, сховищ, баз даних) від зовнішніх провайдерів через Інтернет, де забезпечується висока масштабованість, доступність та економічна ефективність, що робить зручним для розгортання мікросервісів.

Контейнеризація в цій концепції висувається в якості технології, яка дозволяє упакувати додаток та всі його залежності (бібліотеки, середовище виконання) в ізольовані, портативні контейнери, що забезпечує уніфіковане середовище виконання для додатка незалежно від базової інфраструктури, що значно спрощує розгортання, масштабування та управління. Варто зауважити, що саме мікросервіси можуть доволі часто розгортатися саме в контейнерах (Docker), а потім керуються за допомогою оркестраторів (Kubernetes).

Нижня секція охоплює питання безпеки, де основна увага буде зосереджена на методах забезпечення безпеки в поєднанні з можливостями ШІ. В даному контексті використовується Zero Trust Security (ZTS), що є одним з сучасних підходів

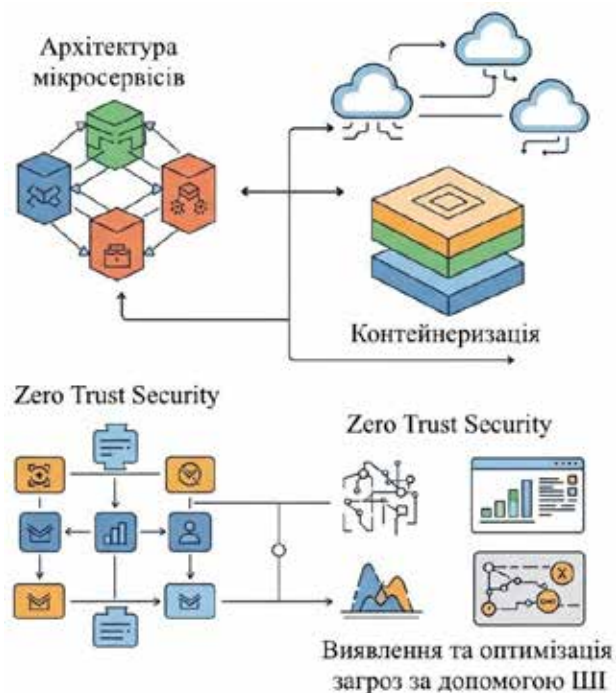


Рис. 4. Концептуальна схема сучасних підходів щодо побудови ефективних та безпечних КСМ

кібербезпеки, де кожен користувач, пристрій та додаток, незалежно від його розташування (навіть всередині мережі), має бути ретельно автентифікований та авторизований перед наданням доступу до будь-яких ресурсів. Діаграма показує складну мережу перевірок і потоків даних між користувачами, пристроями та системами, що ілюструє модель ZTS, де жодному елементу не довіряють за замовчуванням і вимагають постійної автентифікації та контролю.

ШІ є важливою складовою в сучасній кібербезпеці та оптимізації, оскільки системи на основі ШІ здатні аналізувати величезні обсяги даних у реальному часі, виявляти аномалії та невідомі раніше загрози, прогнозувати потенційні ризики та автоматично реагувати на інциденти безпеки. ШІ також використовується для оптимізації розподілу ресурсів, балансування навантаження та покращення загальної продуктивності системи. Друга частина ілюструє застосування штучного інтелекту (ШІ) для виявлення загроз і оптимізації безпекових рішень.

Далі розглянемо основні переваги та недоліки кожного з підходів до концепції COA-ZTA на основі ШІ в табл. 1.

Висновки. В роботі проаналізовано сучасні підходи до побудови ефективних та безпечних комп'ютерних систем і мереж (КСМ) в умовах стрімкої цифрової трансформації та зростаючих кіберзагроз.

Розглянуто та оцінено ключові архітектурні рішення, такі як мікросервісна архітектура (МСА) та сервісно-орієнтована архітектура (COA). Виявлено, що вони є критично важливими для забезпечення гнучкості, масштабованості та стійкості сучасних КСМ, дозволяючи системам адаптуватися до мінливих бізнес-вимог та технологічних викликів. Особливу увагу було приділено тому, як ці архітектури сприяють ефективному використанню хмарних технологій, надаючи високу доступність та економічну ефективність, при цьому враховуючи пов'язані з ними кіберризики.

Проведено аналіз сучасних стратегій кібербезпеки, з акцентом на принципах «безпека за замовчуванням» (Security by Design) та архітектури Zero Trust. Визначено, що ZTA є фундаментальним підходом для мінімізації ризиків несанкціонованого доступу та внутрішніх загроз шляхом постійної верифікації кожного користувача та сервісу. Також було вивчено роль таких механізмів, як багатофакторна автентифікація та наскрізне шифрування, у посиленні загального захисту системи.

Систематизовано та визначено ключові переваги використання ШІ для оптимізації функціонування та посилення кібербезпеки КСМ. З'ясовано, що ШІ може значно покращити автоматизацію управління ресурсами, прогнозування продуктивності та ефективне виявлення аномалій та нових кіберзагроз в реальному часі. Одночасно було

Характеристика переваг та недоліків концепції СОА-ZTA на основі ШІ

Підхід	Переваги	Недоліки
СОА (Сервісно-орієнтована архітектура)	1) Гнучка структура, яка дозволяє масштабування та повторне використання сервісів. 2) Полегшує інтеграцію між різними платформами. 3) Знижує залежність між компонентами системи.	1) Складність в управлінні великою кількістю сервісів. 2) Висока вартість впровадження та підтримки. 3) Вразливість до міжсервісних загроз без належної безпеки.
ZTA (Zero Trust Architecture)	1) Підвищений рівень безпеки завдяки постійній перевірці доступу. 2) Мінімізація ризиків внутрішніх та зовнішніх загроз. 3) Ефективна модель для хмарних і гібридних середовищ.	1) Висока складність впровадження в існуючу ІТ-інфраструктуру. 2) Потреба в постійному моніторингу та аналізі. 3) Можливе навантаження на продуктивність через багатоступеневу перевірку.
Штучний інтелект (ШІ)	1) Автоматизація виявлення загроз і реагування на них. 2) Оптимізація розподілу ресурсів та продуктивності систем. 3) Можливість прогнозування інцидентів у реальному часі.	1) Уразливість до змагальних атак (adversarial attacks). 2) Потреба у великих обсягах якісних даних для навчання моделей. 3) Високі обчислювальні витрати та складність реалізації.

розглянуто виклики, пов'язані з ШІ, такі як змагальні атаки, що підкреслює важливість розробки стійких та адаптивних ШІ-систем.

В якості основних результатів роботи розроблено та представлено інтегровану концепцію, яка поєднує СОА, архітектуру ZTA та принципи ZTS в поєднанні з ШІ. Обґрунтовано, що ця синергія дозволяє створювати високоадаптивні та захищені

системи, де ШІ посилює ZTA шляхом автоматизації верифікації, виявлення аномалій поведінки та динамічного налаштування політик безпеки відповідно до контексту та змінних ризиків. Таким чином, досягнуто мети створення комплексного підходу, який забезпечує покращену стійкість та ефективність комп'ютерних систем і мереж у сучасному складному цифровому ландшафті.

Список літератури:

1. Янко А.С. Система захисту комп'ютерної мережі. *Системи управління навігації та зв'язку*. 2022. № 2 (68). С. 91–94. URL: <https://doi.org/10.26906/SUNZ.2022.2.091> (дата звернення: 17.07.2025)
2. Кулаковська І. Ризики використання хмарних технологій. *Herald of Khmelnytskyi National University. Technical Sciences*. 2024. № 341(5). С. 45-52. URL: <https://doi.org/10.31891/2307-5732-2024-341-5-6> (дата звернення: 17.07.2025)
3. Марченко К. М., Оришак О. В., Марченко А. К., Мельник А. М. Ризики впровадження штучного інтелекту в комп'ютерні системи. *Центральноукраїнський науковий вісник. Технічні науки*. 2022. Вип. 5(36). Ч. 1. С. 119-124. URL: [https://doi.org/10.32515/2664-262X.2022.5\(36\).1.119-124](https://doi.org/10.32515/2664-262X.2022.5(36).1.119-124) (дата звернення: 17.07.2025)
4. Jaber T. A. Artificial intelligence in computer networks. *Periodicals of Engineering and Natural Sciences*. 2022. № 10(1). С. 309-322. URL: <https://doi.org/10.21533/pen.v10.i1.539> (дата звернення: 17.07.2025)
5. Hu Y., Kuang W., Qin Z., Li K., Zhang J., Gao Y., ... & Li K. Artificial intelligence security: Threats and countermeasures. *ACM Computing Surveys (CSUR)*. 2024. № 55(1). С. 1-36. URL: <https://doi.org/10.1145/3487890> (дата звернення: 16.07.2025)
6. Ahvanooey M. T., Mazurczyk W., Zhao J., Caviglione L., Choo K. R., Kilger M., ... & Misoczki R. Future of cyberspace: A critical review of standard security protocols in the post-quantum era. *Computer Science Review*. 2025. № 57. С. 100738. URL: <https://doi.org/10.1016/j.cosrev.2025.100738> (дата звернення: 15.07.2025)
7. Kasongo S. M. A deep learning technique for intrusion detection system using a Recurrent Neural Networks based framework. *Computer Communications*. 2023. № 199. С. 113-125. URL: <https://doi.org/10.1016/j.comcom.2022.12.010> (дата звернення: 16.07.2025)
8. Mayuranathan M., Saravanan S. K., Muthusenthil B., & Samyurai A. An efficient optimal security system for intrusion detection in cloud computing environment using hybrid deep learning technique. *Advances in Engineering Software*. 2022. № 173, С. 103236. URL: <https://doi.org/10.1016/j.advengsoft.2022.103236> (дата звернення: 16.07.2025)

9. Luo X. Model design artificial intelligence and research of adaptive network intrusion detection and defense system using fuzzy logic. *Journal of Intelligent & Fuzzy Systems*/ 2021. № 40(4). С. 8227-8235. URL: <https://doi.org/10.3233/JIFS-189645>_(дата звернення: 16.07.2025)
10. Hephzipah J. J., Vallem R. R., Sheela M. S., & Dhanalakshmi G. An efficient cyber security system based on flow-based anomaly detection using Artificial neural network. *Mesopotamian Journal of Cybersecurity*. 2023. С. 48-56. URL: <https://doi.org/10.58496/MJCS/2023/009>_(дата звернення: 15.07.2025)
11. Wang Y. Research on Intelligent Cybersecurity Protection System in Cloud Computing Environment. *Innovation in Science and Technology*. 2024. № 3(4). С. 71-78. URL: <https://www.paradigmppress.org/ist/article/view/1198>_(дата звернення: 16.07.2025)
12. Paidy P. Adaptive Application Security Testing with AI Automation. *International Journal of AI, BigData, Computational and Management Studies*. 2023. № 4(1). С. 55-63. URL: <https://doi.org/10.63282/3050-9416.IJAIBDCMS-V4I1P106>_(дата звернення: 16.07.2025)
13. Zafer F., Akhtar F. Enhancing Cyber Resilience with AI-Powered Security: Cloud-Based Real-Time Threat Detection, Autonomous Response, and Multi-Layered Adaptive Defense for Cloud Security. 2025. URL: <https://doi.org/10.13140/RG.2.2.14283.09766>_(дата звернення: 15.07.2025)
14. Sairaj Kommera. "Enhancing Zero Trust Architecture with AI-Driven Threat Intelligence in Cloud Environments". *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*. Vol. 11, no. 1. 2025. Pp. 1524-33. URL: <https://doi.org/10.32628/CSEIT251112163>_(дата звернення: 17.07.2025)
15. Wang Y., Sun T., Li S., Yuan X., Ni W., Hossain E., Poor H. V. Adversarial attacks and defenses in machine learning-empowered communication systems and networks: A contemporary survey. *IEEE Communications Surveys & Tutorials*. 2023. № 25(4). С. 2245-2298. URL: <https://doi.org/10.1109/COMST.2023.3319492>_(дата звернення: 17.07.2025)

Bolshakova O.O., Koshkin V.K., Aramian A.K., Bandura V.M., Yehanov O.Yu. MODERN APPROACHES TO BUILDING EFFICIENT AND SECURE COMPUTER SYSTEMS AND NETWORKS

The article analyzes current approaches to designing effective and secure computer systems and networks (CSN) in the context of digital transformation and the growth of cyber threats. It shows that the increase in data volumes, the development of cloud infrastructure, and the growing complexity of IT environments require the introduction of architectural solutions capable of ensuring flexibility, scalability, and cyber resilience. Microservice and service-oriented architecture are considered as the basis for building adaptive C&N that can be easily scaled and updated in line with changing business requirements. Cloud technologies are identified as a key element of modern IT infrastructure, but the risks they pose are highlighted: increased attack surface, data protection complications, and potential loss of control. In this regard, the feasibility of implementing the Security by Design principle, which provides for the integration of security at all stages of development, as well as the application of Zero Trust Architecture (ZTA), which ensures constant control and verification of access regardless of the source of the request, is justified. The role of mechanisms such as multi-factor authentication (MFA) and end-to-end encryption as means of increasing the cyber resilience of systems is analyzed in detail. Particular attention is paid to the use of artificial intelligence (AI) in resource management, failure prediction, and real-time anomaly detection. AI is considered as a tool for increasing the dynamism of protection systems, in particular through adaptive configuration of security policies and analysis of user behavior. At the same time, challenges related to competitive attacks that require robust algorithms are noted. The result of the study is a proposed integrated concept for building CSPs that synthesizes the capabilities of SOA, ZTA principles, and AI-based intelligent tools. This approach allows the creation of secure, adaptive, and effective systems for use in government, corporate, and cloud environments.

Key words: *microservice architecture, service-oriented architecture, cybersecurity, cloud technologies, artificial intelligence, Zero Trust.*

Дата надходження статті: 17.07.2025

Дата прийняття статті: 25.07.2025

Опубліковано: 27.10.2025

Бугаєва Л.М.Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»**Абрамова А.О.**Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»**Билбас Д.О.**Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»

РОЗРОБКА АВТОМАТИЗОВАНОГО ПРОЦЕСУ АБСОРБЦІЙНОГО ОЧИЩЕННЯ ГАЗУ ВІД CO₂ РОЗЧИНОМ МЕТИЛДИЕТАНОЛАМІНУ

У статті проведено дослідження процесу очищення газів від діоксиду вуглецю з використанням метилдиетаноламіну (MDEA). Авторами проведено аналіз найбільш вживаних в світі методів очищення промислових газів від CO₂, розглянуто їх переваги та недоліки. В першу чергу розглянуто найпоширеніші методи – хемосорбційні методи, які передбачають використання хімічних реагентів для поглинання CO₂. Крім них розглянуто фізичні методи очищення газу, що базуються на розчиненні CO₂ у спеціальних фізичних абсорбентах або на його уловлюванні за допомогою мембранних технологій. Мембранна сепарація передбачає використання напівпроникних мембран, що селективно пропускають CO₂, і має переваги у вигляді компактності, простоти експлуатації та низького енергоспоживання. Хімічна абсорбція є найпоширенішим методом очищення газів від CO₂ і базується на хімічних реакціях між CO₂ та реагентами, що утворюють розчинні сполуки. Особливу увагу авторами приділено абсорбційному методу на основі використання метилдиетаноламіну. Як показує аналіз, MDEA забезпечує високу ступінь очищення газу за мінімальних енергетичних витрат. Авторами пропонується варіант схеми автоматизації процесу, що включає контури для контролю таких параметрів як тиск, температура, концентрація CO₂, рівень рідини. Цей варіант схеми автоматизації забезпечений системою дистанційного керування насосами, регулюючими клапанами та іншими технологічними апаратами. В роботі досліджено математичну модель процесу абсорбції CO₂ у розчині MDEA, реалізовану в програмному середовищі MATLAB. Крім того, запропоновано проводити налаштування регуляторів різними методами (ПІД-регулятори, методи оптимізації) та обрано оптимальні значення параметрів процесу, що забезпечують його стабільність. Отримані результати досліджень основних показників якості керування підтверджують ефективність запропонованої системи автоматизації та її відповідність сучасним вимогам енергоефективності та екологічної безпеки.

Ключові слова: очищення газу, вуглекислий газ, метилдиетаноламін, автоматизація, контур контролю, математична модель, регулятор, оптимізація, MATLAB.

Постановка проблеми. У світлі глобальних кліматичних викликів і посилення екологічного контролю, очищення технічних газів від вуглекислого газу (CO₂) набуває особливого значення. Це актуально як для хімічної, так і для енергетичної промисловості. Сучасні технології газоочищення відіграють важливу роль у промисловості, особливо у контексті зменшення викидів парникових газів та збереження навколишнього середовища. Очищення від вуглекислого газу (CO₂) є критично

важливим процесом для хімічної, металургійної, енергетичної та інших галузей промисловості. Одним із найефективніших методів видалення CO₂ є хімічна абсорбція за допомогою розчину метилдиетаноламіну (MDEA), який широко використовується у промислових системах очищення газу. Метилдиетаноламін (MDEA) є третинним аміном, що має високу селективність до CO₂ при одночасному збереженні низької реакційної здатності до інших компонентів газової суміші, таких

як сірководень. Це дозволяє ефективно видаляти CO_2 , мінімізуючи втрати інших компонентів. Традиційно у промисловості використовують процес абсорбції CO_2 гарячим розчином поташу, який забезпечує високу ефективність очищення завдяки хімічним реакціям карбонатно-бікарбонатної рівноваги, що дозволяє уловлювати значну кількість CO_2 . Однак за кордоном усе більше набуває переваги метод очищення газу за допомогою метилдіетаноламіну (MDEA), так як є одним із найефективніших способів хімічної абсорбції. MDEA має високу селективність до CO_2 за умови низької здатності реагувати з іншими компонентами газової суміші, що призводить до зменшення побічних витрат. Перевагами використання MDEA є низька енергоємність регенерації розчину, висока стійкість до термічного та хімічного розкладу, зменшення корозійного впливу на обладнання, а також економічна та екологічна доцільність у порівнянні з традиційними амінами, такими як моноетаноламін (MEA) або діетаноламін (DEA).

Для підвищення ефективності процесу очищення газу необхідно автоматизоване керування, яке забезпечить стабільність параметрів, знизить витрати енергії та реагентів, а також підвищить загальну продуктивність системи. Автоматизація дозволяє підтримувати оптимальну концентрацію MDEA, регулювати температуру, тиск та витрату абсорбенту, а також контролювати баланс між швидкістю абсорбції CO_2 та витратами на регенерацію розчину.

В дослідженні пропонується ефективний підхід до автоматизації процесу очищення газу на основі хімічної абсорбції CO_2 розчином метилдіетаноламіну, який забезпечує високу ефективність видалення забруднень при знижених енергетичних витратах.

Аналіз останніх досліджень і публікацій. На сьогодні існує кілька методів очищення газів від CO_2 , які базуються на різних фізико-хімічних принципах. Найпоширенішими з них є хемосорбційні методи, які передбачають використання хімічних реагентів для поглинання CO_2 . Загальноприйнятими для видалення CO_2 із технологічних газів у виробництві аміаку використовуються методи, зокрема з використанням розчину моноетаноламіну (MEA) або гарячого розчину поташу (K_2CO_3), активованого діетаноламіном. Ці методи базуються на хімічних реакціях між CO_2 та абсорбентами, що дозволяє ефективно видаляти домішки з газових потоків [1].

Фізичні методи очищення газу базуються на розчиненні CO_2 у спеціальних фізичних абсорбентах або на його уловлюванні за допомогою мембранних технологій. До таких методів належить фізична абсорбція, що передбачає розчинення CO_2 у рідких розчинниках, таких як метанол, пропіленкарбонат. Перевагами цього методу є низькі енергетичні витрати та можливість регенерації абсорбенту без високотемпературного нагріву, проте ефективність значно залежить від тиску газу, а також обмежена розчинність CO_2 . Мембранна сепарація передбачає використання напівпроникних мембран, які селективно пропускають CO_2 , і має переваги завдяки компактності, простоти експлуатації та низького енергоспоживання. Однак вона має високу вартість мембран та обмежений ступінь очищення при низьких концентраціях CO_2 [1].

Хімічна абсорбція є найпоширенішим методом очищення газів від CO_2 і базується на хімічних реакціях між CO_2 та реагентами, що утворюють розчинні сполуки. Одним із найбільш використовуваних реагентів є моноетаноламін (MEA), що характеризується високою швидкістю реакції та здатністю поглинати значну кількість CO_2 . Однак цей метод має значні недоліки, серед яких висока енергоємність регенерації, корозійна активність та деградація розчину під час експлуатації. Альтернативою є гарячий розчин поташу (K_2CO_3), активований діетаноламіном, який потребує менше енергії для регенерації та є більш стабільним. Водночас цей метод є менш ефективним при низьких концентраціях CO_2 та потребує підвищених температур для забезпечення ефективної роботи [1].

У XXI столітті були запропоновані нові методи очищення газів від CO_2 , зокрема з використанням розчину метилдіетаноламіну (MDEA), активованого піперазином. Цей метод характеризується низькою енергомісткістю регенерації, меншою корозійною активністю та зниженою деградацією розчину порівняно з традиційними методами. Його основним недоліком є нижча швидкість реакції з CO_2 порівняно з MEA, через що для підвищення ефективності процесу застосовують активатори, наприклад, піперазин [2].

З метою підвищення ефективності очищення газу від CO_2 у деяких випадках застосовують комбіновані методи, такі як амінокарбонатні процеси, що поєднують використання амінів (наприклад, MDEA) та карбонатних розчинів для прискорення поглинання CO_2 . Також активно використовуються гібридні мембранно-абсорбційні техноло-

гії, які передбачають попереднє мембранне відділення CO_2 , а потім його доочищення хімічною абсорбцією [2].

Серед доступних методів очищення газів хімічна абсорбція виявилась найефективнішою, особливо у разі використання MDEA як абсорбенту. MDEA відзначається високою селективністю до CO_2 , низьким споживанням пари на регенерацію, а також зниженою корозійною активністю, що значно продовжує термін служби обладнання. Процес полягає у поглинанні CO_2 в абсорбері та його подальшому виділенні у регенераторі з використанням теплової енергії.

На рис. 1 представлено один з варіантів технологічної схеми очищення газу від CO_2 за допомогою метилдиетаноламіну [3, 4]. Важливими вузлами системи є абсорбер, регенератор, теплообмінники, насоси, конденсатори та система циркуляції розчину. Конденсат із процесу збирається у сепараторі конденсату (102-F) та відводиться за допомогою насоса конденсату (106-J). Частина використаної пари конденсується і повертається у вигляді конденсату пари, а для підтримки балансу в систему додається необхідна кількість додаткової води. Оптимізація процесу включає використання добавок, таких як піперазин (Pz), що покращує кінетику абсорбції, а також оптимі-

зацію складу розчинника (35–37% MDEA, 3–5% Pz, решта – вода) і вдосконалення теплообмінних апаратів для зменшення втрат тепла. Це забезпечує високу ефективність процесу очищення газу від CO_2 при низьких енергетичних витратах.

Постановка завдання. Метою статті є представити розроблену авторами автоматизовану систему керування процесом очищення газу від CO_2 за допомогою метилдиетаноламіну, яка дозволить підвищити ефективність процесу, знизити експлуатаційні витрати та забезпечити стабільність роботи установки.

У відповідності до поставленої мети авторами було вирішено такі задачі.

1. Дослідити технологічний процес очищення газу від CO_2 за допомогою MDEA та визначити його ключові параметри.
2. Проаналізувати існуючі методи автоматизації подібних процесів та обрати оптимальні підходи.
3. Розробити відповідні схеми автоматизації, а саме: технологічну, функціональну та принципову схеми. Забезпечити відповідне обладнання для автоматизації процесу.
4. Створити алгоритми автоматичного регулювання концентрації абсорбенту, температури та витрати газу.

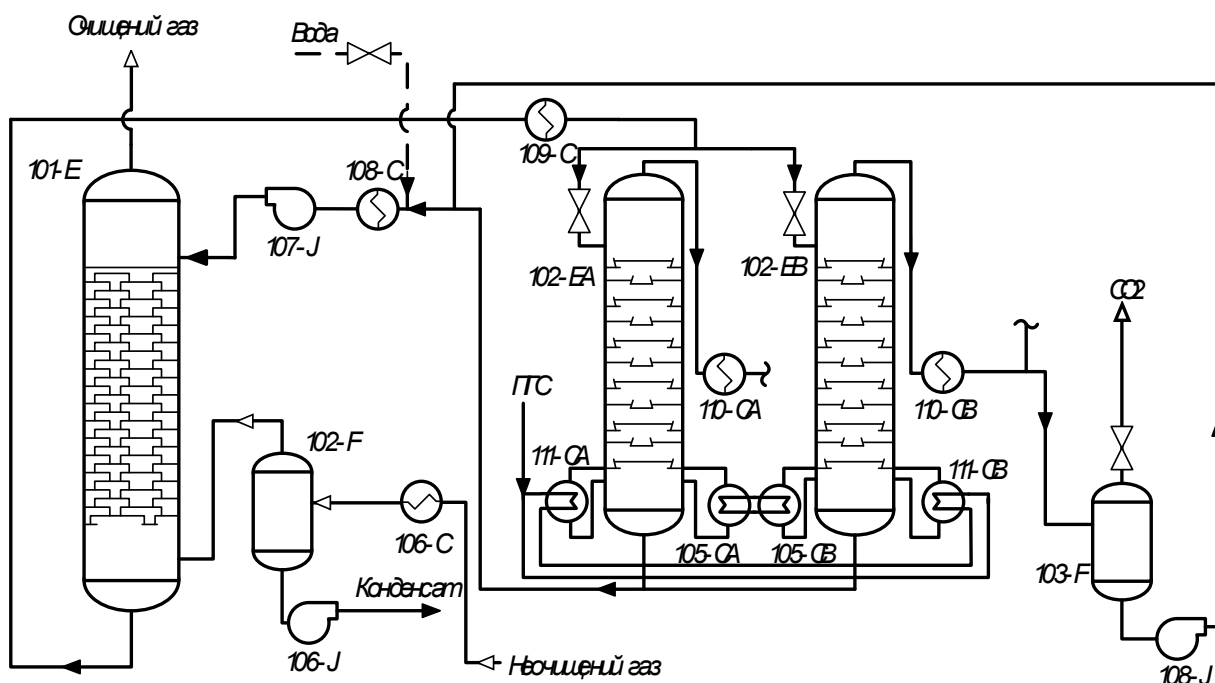


Рис. 1. Система видалення CO_2 на основі MDEA для аміачного заводу: 102-E – абсорбер, 102-F – сепаратор конденсату, 102-EA/102-EB – регенератор, 103-F – відстійник рефлюкс-конденсату, 105-CA/105-CB – нагрівник регенератора, 106-J – насос конденсату, 107-J – насос збідненого розчину, 108-C – охолоджувач збідненого розчину, 109-C – теплообмінник збагаченого розчину, 110-CA, 110-CB – конденсатор регенератора, 111-CA/111-CB – паровий нагрівач

5. Провести моделювання та аналіз ефективності впровадженої автоматизованої системи.

Виклад основного матеріалу. Для досягнення максимальної ефективності процесу очищення в дослідженні було запропоновано оптимальну систему автоматизації [3]. На підставі аналізу технологічної схеми сформовано вимоги до контролю наступних параметрів:

- тиску та температури в апаратах;
- витрати розчину та газу;
- рівня рідини;
- концентрації CO_2 у газі на вході та виході з абсорбера.

Вибір технічних засобів базувався на надійності, точності та здатності інтегруватися з програмованими логічними контролерами Siemens S7-1200. Авторами пропонуються для промисло-

вого впровадження контури регулювання температури на базі тиристорних регуляторів, витрат на базі частотних перетворювачів Danfoss FC 301), а також контроль вмісту CO_2 у газі з використанням газоаналізаторів та програмованих логічних контролерів (ПЛК).

Всі сигнали з датчиків мають передаватися на ПЛК, після чого оброблятися, відображатися на панелі оператора, і використовуватися для керування клапанами, насосами та обігрівачами. До того ж треба передбачити реалізацію системи аварійного захисту та сигналізації. Схема автоматизації процесу наведена на рис. 2.

У процесі дослідження також була розроблена принципова електрична схема керування електродвигунами насосів, які циркулюють розчин MDEA. Система передбачає живлення через

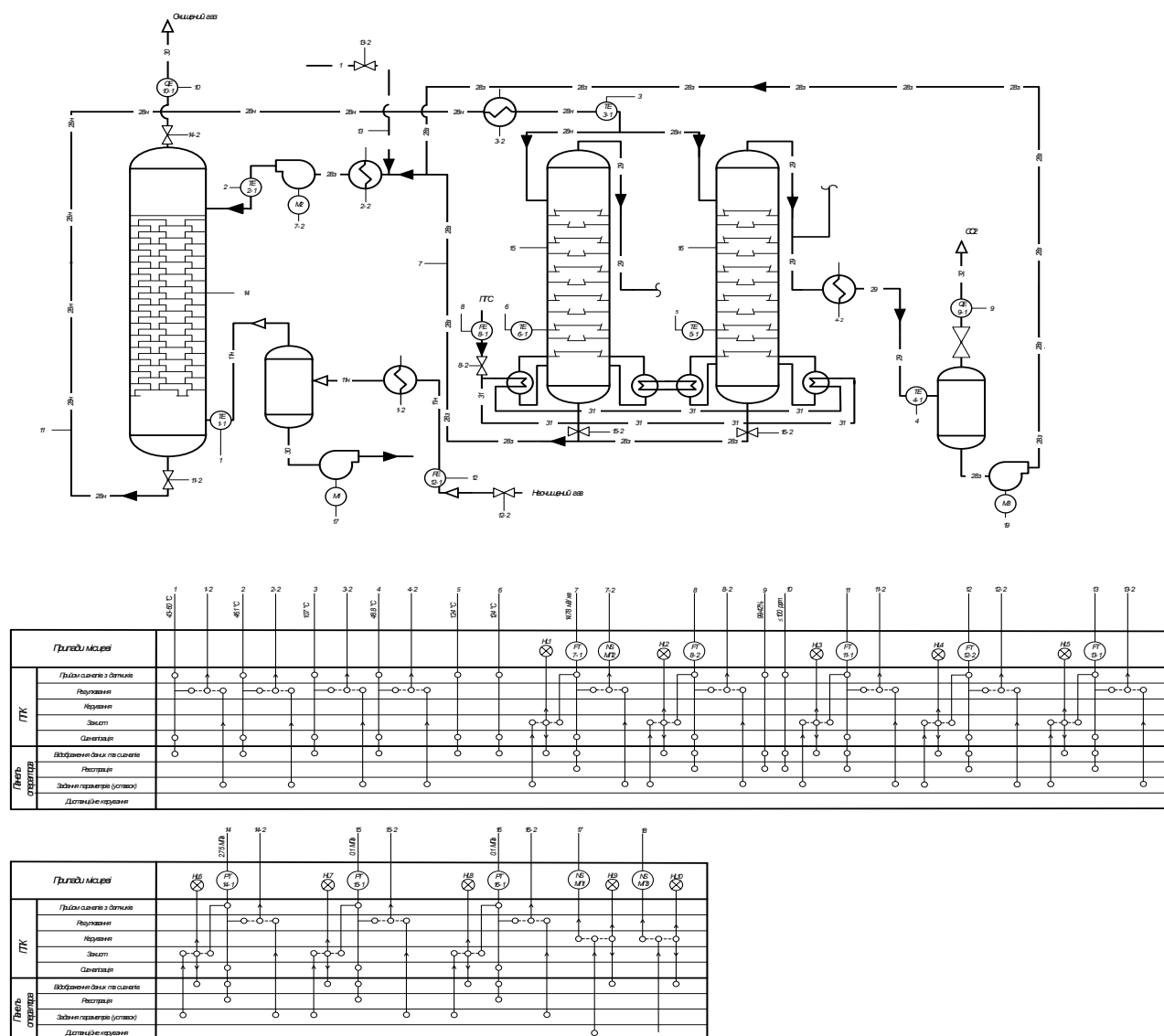


Рис. 2. Схема автоматизації процесу

частотні перетворювачі Danfoss з вбудованим захистом, а також релейну сигналізацію і аварійне відключення. Монтажно-комутаційну схему пропонується реалізувати з урахуванням клемних колодок і модулів аналогових/дискретних сигналів SM1234 та SM1223. Це дозволить оперативно підключати, змінювати або доповнювати конфігурацію схеми без суттєвих змін у програмному коді.

Для оцінки ефективності автоматизованої системи було створено математичну модель процесу абсорбції в середовищі MATLAB, яка враховує масо- та теплоперенесення, а також кінетику хімічної реакції [2]. Система рівнянь моделі набуває вигляду (1):

$$\begin{cases} \frac{dy}{dt} = \frac{G}{V_g}(Y_{CO2_{in}} - Y) - \frac{K_v S}{V_g}(Y - Y_{CO2_{out}}) \\ \frac{dx}{dt} = -\frac{L}{V_g}(X_{CO2_{in}} - X) - \frac{K_v S}{V_g}(X - X_{CO2_{out}}) \end{cases} \quad (1)$$

Рівняння енергетичного балансу для зміни температури T рідкої фази по висоті абсорбера:

$$\frac{dT}{dt} = \frac{K_T}{V_l}(T_{in} - T) - \frac{K_v S}{V_l * C_p}(Y - Y_{CO2_{out}})\Delta H. \quad (2)$$

Для лінеаризації рівняння приймаємо стаціонарні значення для вхідних параметрів і вводимо відхилення від їхніх стаціонарних значень. Таким

чином, замість значень, які змінюються, працюємо з відхиленнями від стану рівноваги. Рівняння після лінеаризації наведено нижче.

$$\frac{d\Delta X}{dt} = \frac{L_0}{V_L}(-\Delta X) + \frac{1}{V_L}(X_{in} - X_0)\Delta L + \frac{K_v S}{V_l}(\Delta Y - K_0 \Delta X), \quad (3)$$

$$\frac{d\Delta Y}{dt} = \frac{G}{V_G}(-\Delta Y_{CO2_{in}} - \Delta Y) - \frac{K_v S}{V_G}(\Delta Y - K_0 \Delta X). \quad (4)$$

Рівняння у просторі станів:

$$\begin{cases} \Delta \dot{X} = a_{11}\Delta X + a_{12}\Delta Y + b_1\Delta L \\ \Delta \dot{Y} = a_{21}\Delta X + a_{22}\Delta Y + b_2\Delta Y_{CO2_{in}} \end{cases} \quad (5)$$

Застосовуємо перетворення Лапласа для отриманого рівняння:

$$(sI - A)X(s) = B^*U(s) \rightarrow Y(s) = C(sI - A)^{-1}B^*U(s). \quad (6)$$

Далі отримаємо функцію за каналом «керування-вихід» яка має вигляд:

$$W(s) = \frac{0.0004021}{s^2 + 1.194e04s + 1.379e07}. \quad (7)$$

Побудована модель дозволила вивчити динаміку системи, а також провести ідентифікацію стійких і нестійких режимів. Додатково виконано синтез та налаштування ПД-регуляторів температури, витрати та концентрації CO_2 за допомогою автоматичних методів (Tuner, частотний підхід).

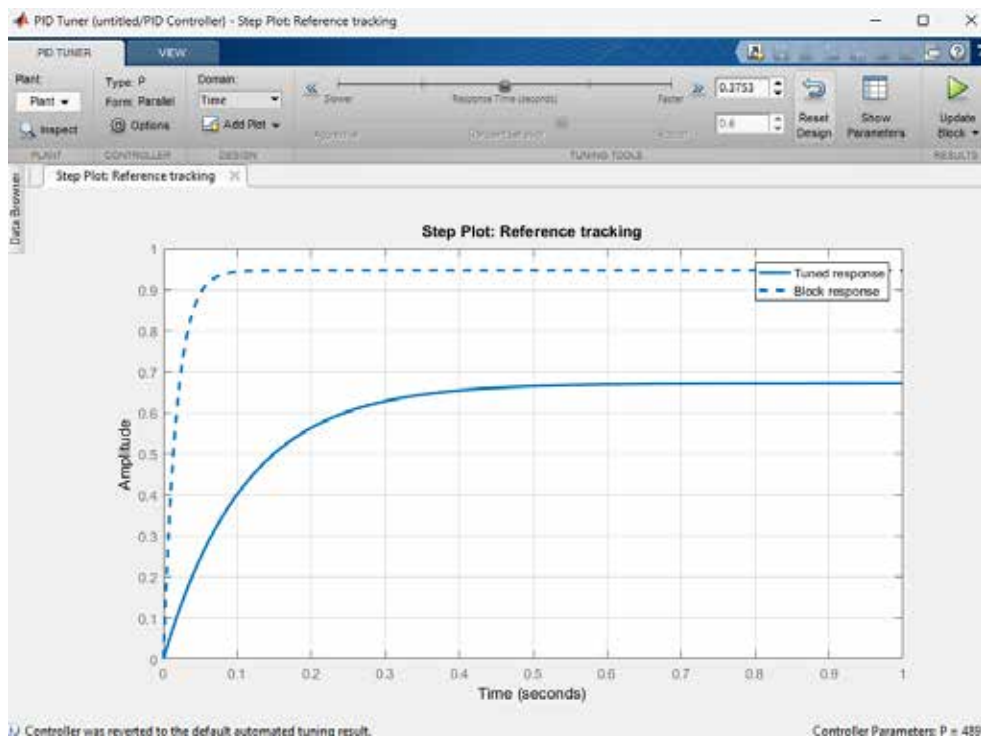


Рис. 3. Результати налаштування Р-регулятора в середовищі MATLAB PID Tuner

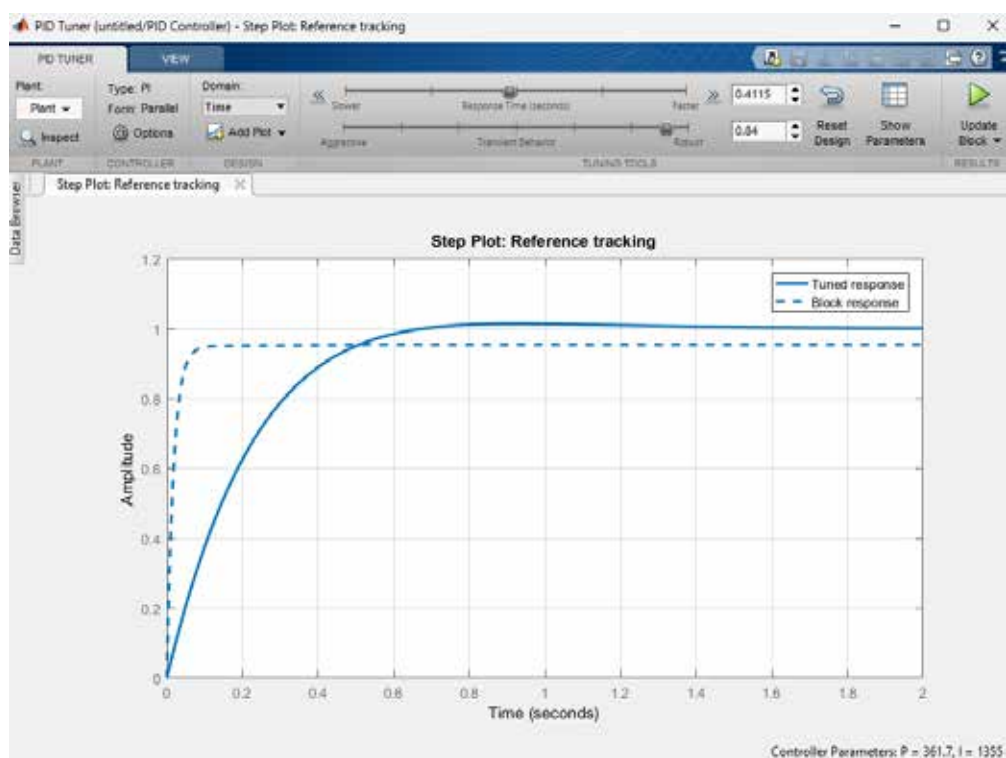


Рис. 4. Результати налаштування PI-регулятора в середовищі MATLAB PID Tuner

На основі побудованої моделі було вирішено задачу статичної оптимізації процесу очищення газу з метою мінімізації витрат MDEA без втрати ефективності вилучення CO_2 . Розв'язання задачі з використанням методу штрафних функцій і SQP показало можливість зменшення споживання розчину при збереженні рівня очищення на 99% і вище. Це дозволить не лише знизити експлуатаційні витрати, а й покращити екологічні показники підприємства.

Висновки. У представленому дослідженні пропонується повний інженерний цикл: від ана-

лізу та моделювання процесу очищення до розробки схем керування та оптимізації параметрів. Автоматизована система керування на базі ПЛК Siemens і частотних перетворювачів Danfoss дозволяє досягти стабільної, економічно ефективної та екологічно безпечної роботи установки. Запропоноване рішення застосовує інтеграцію сучасних підходів у галузі хімічної інженерії, автоматизації та системного моделювання, тому може бути основою для впровадження на реальних промислових об'єктах.

Список літератури:

1. Герасимов О. І. Теоретичні основи технологій захисту навколишнього середовища. Одеса, ТЕС, 2018. – 227 с.
2. Технологія зв'язаного азоту: курс лекцій: навч. посіб. для студ. спеціальності 161 «Хімічні технології та інженерія», ОПП «Хімічні технології неорганічних речовин та водоочищення» / А.Л. Концевой, С.А. Концевой ; КПІ ім. Ігоря Сікорського. – Електронні текстові данні (1 файл: 8,28 Мбайт). – Київ : КПІ ім. Ігоря Сікорського, 2021. – 293 с.
3. Бугаєва Л. М., Абрамова А. О., Билбас Д. О. Очищення технологічних газів від CO_2 на основі метилдіетаноламіну (MDEA) / Збірник наукових праць XI міжнародної конференції АКІТ-2025. Київ, 7 травня 2025 р. – Київ : КПІ ім. Ігоря Сікорського, Вид-во «Політехніка», 2025. – С. 34–35. ISBN 978-966-990-144-6.
4. Combs, G. & McGuire, MDEA Based CO_2 Removal System Process Simulation. URL: <https://www.chemengservices.com/co2-removal.html> (дата звернення: 06.05.2025).

**Bugaieva L.M., Abramova A.O., Bilbas D.O. DEVELOPMENT
OF AN AUTOMATED PROCESS FOR ABSORPTION PURIFICATION
OF CO₂ GAS WITH METHYLDIETHANOLAMINE SOLUTION**

The article studies the process of gas purification from carbon dioxide using methyldiethanolamine (MDEA). The authors analyze the most widely used methods of industrial gas purification from CO₂ in the world, consider their advantages and disadvantages. First of all, the most common methods are considered – chemisorption methods, which involve the use of chemical reagents for CO₂ absorption. In addition, physical methods of gas purification are considered, which are based on the dissolution of CO₂ in special physical absorbents or on its capture using membrane technologies. Membrane separation involves the use of semi-permeable membranes that selectively pass CO₂, and has the advantages of compactness, ease of operation and low energy consumption. Chemical absorption is the most common method of gas purification from CO₂ and is based on chemical reactions between CO₂ and reagents that form soluble compounds. The authors paid special attention to the absorption method based on the use of methyldiethanolamine. As the analysis shows, MDEA provides a high degree of gas purification with minimal energy costs. The authors propose a variant of the process automation scheme, which includes circuits for controlling such parameters as pressure, temperature, CO₂ concentration, liquid level. This variant of the automation scheme is equipped with a remote control system for pumps, control valves and other technological devices. The paper investigates a mathematical model of the CO₂ absorption process in MDEA solution, implemented in the MATLAB software environment. In addition, it is proposed to adjust the regulators using various methods (PID regulators, optimization methods) and select the optimal values of the process parameters that ensure its stability. The obtained results of studies of the main indicators of control quality confirm the effectiveness of the proposed automation system and its compliance with modern requirements for energy efficiency and environmental safety.

Key words: gas purification, carbon dioxide, methyldiethanolamine, automation, control loop, mathematical model, regulator, optimization, MATLAB.

Дата надходження статті: 07.06.2025

Дата прийняття статті: 17.06.2025

Опубліковано: 27.10.2025

Burchak P.V.

National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute”

Oleshchenko L.M.

National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute”

MACHINE LEARNING SOFTWARE FRAMEWORK FOR PREDICTING AND OPTIMIZING WEB APPLICATION PERFORMANCE

The article introduces a software framework for analyzing and predicting the performance of web applications using machine learning (ML) techniques. Instead of relying solely on traditional performance testing, an automated pipeline has been introduced that has collected, processed, and modeled key system metrics – such as server response time, memory usage, CPU load, and UI responsiveness – to proactively identify and address performance issues. The dataset used in the study has been gathered from real-world web applications and has included over 200,000 performance records, encompassing web page load time, layout recalculations, JavaScript memory usage, and network throughput.

Multiple ML algorithms have been integrated into the modeling pipeline, including linear regression (LR), polynomial regression, decision trees, and neural networks (NN), to compare their effectiveness in forecasting application performance. Experimental results have shown that LR has achieved the best overall performance ($MAE = 187.25$, $R^2 = 0.9357$), indicating a strong linear correlation in the data. Neural network also demonstrated high accuracy and advantages in modeling more dynamic, nonlinear behaviors. Proposed developed software framework has been implemented using Python technologies, such as Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn, and TensorFlow.

The research has resulted in a scalable, automated, and interpretable software solution for predictive diagnostics and performance optimization of web applications. Proposed software system has also generated categorized web applications performance reports with actionable recommendations. The research results have confirmed the feasibility and effectiveness of incorporating machine learning into modern web applications performance monitoring workflows.

Key words: *software framework, software performance monitoring, machine learning, predictive analytics, web systems optimization, regression models, neural network, Python, TensorFlow, web application performance forecasting.*

Formulation of the problem. In the modern digital landscape, the performance of web applications plays a crucial role in determining the success of software products. Fast load times, responsiveness, and system stability directly impact user satisfaction and a company's competitiveness. With the increasing number of users, the variety of devices, and the complexity of server-client interactions, evaluating and maintaining optimal performance levels has become a significant challenge.

Traditional methods, such as manual testing or standard automated scripts, often fall short when it comes to handling large volumes of data or adapting to dynamic conditions like network delays, traffic spikes, or shifts in user behavior. As web applications grow more sophisticated, performance becomes a

vital component of both user experience and business outcomes.

Research indicates that even a one-second delay in page loading can lead to a 7% drop in conversions, while more than half of mobile users abandon websites that take over three seconds to load. Search engines like Google factor in site speed when ranking content. Industry experience demonstrates that implementing optimizations – such as reducing HTTP requests, compressing resources, and applying caching techniques – can cut load times by 20–50%, leading to better retention and overall performance. Recently, machine learning (ML) has emerged as a promising solution, enabling predictive load analysis and adaptive resource management, thus enhancing performance without requiring major infrastructure upgrades.

Evaluating the performance of web applications remains a critical challenge in contemporary software engineering, as system efficiency directly impacts both user satisfaction and business outcomes. A wide range of techniques is employed in this domain, including analytical frameworks, empirical testing, and ML approaches. Conventional performance evaluation typically relies on metrics such as server response time, requests per second (RPS), and resource utilization (CPU, memory, and network bandwidth). Tools like Apache JMeter, New Relic, and Dynatrace provide valuable outcomes into system behavior, though they often lack capabilities for forecasting future loads or uncovering subtle performance trends. In response to these limitations, ML-based methods have gained popularity for their ability to analyze and predict application behavior more accurately.

Techniques such as linear and polynomial regression, decision trees, neural networks, and ensemble models have proven effective in detecting anomalies and forecasting performance issues. The research indicates that deep neural networks can outperform traditional models by 15–30% in prediction accuracy.

Another emerging area is the application of reinforcement learning for dynamic, real-time adjustment of system parameters, with some studies reporting up to 40% reduction in average response times.

Modern strategies for performance analysis are moving toward a hybrid approach that combines classical monitoring tools with intelligent ML models. Ongoing challenges include model interpretability, the need for extensive training datasets, and ensuring adaptability across various web application architectures.

This research focuses on creating a software-based solution for testing and evaluating the performance of web applications by integrating ML techniques, particularly regression models and neural networks.

Regression models help uncover key patterns and measure the influence of performance factors, while neural networks offer deeper outcomes into complex, nonlinear relationships in large datasets. This combined approach enables both real-time evaluation and performance forecasting under dynamic conditions.

The proposed method is designed to be scalable and efficient, capable of handling increased data volumes without compromising speed or accuracy.

Analysis of recent research and publications. The study in the article [1] demonstrated that among the existing approaches to regression testing of web applications, no single method consistently outperforms the others, as their effectiveness largely depends on various factors and deployment environments. The research emphasizes the importance of evaluat-

ing these methods not only by technical parameters but also by their cost-effectiveness. In the article [2], a performance evaluation methodology for web pages is proposed, leveraging ML algorithms to enhance the accuracy and efficiency of assessing web resource quality. A systematic literature review covering 2010–2024 identified 59 key factors affecting website performance. The proposed model integrates a wide range of metrics – such as usability, accessibility, content relevance, visual appeal, and technical characteristics – enabling it to overcome the limitations of traditional approaches. The study showed that SVM algorithms achieved the highest prediction accuracy (89%) after feature selection, compared to 87% without it. Random Forest models also saw a slight improvement, from 80% to 81%. The use of feature selection significantly improved model performance, highlighting the value of selecting influential predictors. This approach enables the automation of web performance evaluation and prediction of potential issues at the development stage, supporting more informed decision-making and setting new research benchmarks.

A systematic review in the article [3] presents an overview of current methods and strategies for regression testing of web applications, covering areas such as test prioritization, selection, and adaptation to dynamic environments. The review includes both classical and modern techniques, such as automated testing and change analysis tools. The results of research [4] indicate that optimizing non-functional performance attributes like speed, startup time, memory usage, and energy consumption is essential to improving user satisfaction with mobile applications. The authors also identify existing optimization techniques and research gaps for future exploration. According to the article [5], key design aspects – such as interface simplicity, navigability, content readability, feedback mechanisms, personalization, and aesthetic design – strongly influence users' perceptions of usability and utility, which in turn affects their engagement and satisfaction.

In the article [6], the authors examine performance optimization methods including caching, compression, CDNs, and tools like WebPageTest and YSlow, aimed at assisting developers in improving website speed and search engine visibility.

The study in the article [7] focuses on enhancing trace visualization for analyzing microservices performance by addressing the shortcomings of existing tools and introducing new techniques for better diagnostic efficiency.

The work [8] explores how Dynatrace monitoring data can be used to develop performance models for

Java EE applications, aiming to improve performance understanding and modeling precision.

The research [9] presents methods for optimizing the performance of web application state management, with a particular focus on improving responsiveness and reducing resource consumption. Through experimental evaluation of various state management strategies – such as client-side caching, minimizing state mutations, and asynchronous state updates – the authors demonstrate measurable gains in load time, memory usage, and CPU efficiency. The results confirm that careful architectural decisions in managing application state can significantly enhance overall performance, particularly in complex and dynamic web environments.

Across the reviewed literature, the main challenges in regression analysis of web application performance include the difficulty of adapting traditional testing approaches to dynamic, component-based architectures; minimizing test suite size while maintaining adequate coverage; prioritizing tests to detect critical defects more effectively; and automating the testing process – particularly for interactive elements like GUIs and JavaScript components.

Researchers also point to the high time and resource costs of large-scale regression testing and emphasize the need for tools that address both functional and non-functional aspects of performance. As application load or user numbers grow, predicting system behavior under future conditions becomes crucial. The nonlinear relationships between parameters such as requests per second, hardware resource usage, and average server response time complicate traditional analysis approaches, highlighting the need for more adaptive and flexible evaluation techniques that can account for these complex interdependencies.

Task statement. The goal of the research is to develop a performance evaluation software tool that automates analysis, detects system bottlenecks, and anticipates potential issues before they impact functionality. The key tasks include reviewing existing methods to identify their strengths and limitations; selecting a core set of performance metrics – such as response time, server latency, and resource usage – that provide meaningful outcomes; building and training ML models to generate predictive analytics; and developing a software framework that brings all these elements together into an integrated, usable solution.

The resulting software system must encompass every stage of performance analysis – from data gathering to classification and reporting – delivering a unified and scalable method. It also must provide actionable recommendations to improve perfor-

mance, including API optimization, asynchronous request handling, and reduction of data transfer volume and code complexity.

Outline of the main material of the research. To evaluate the performance of web applications, data containing over 200,000 real user records was selected for in-depth analysis of various performance aspects. The Chrome User Experience Report (CrUX) offers a valuable source of real-world performance data, gathered directly from users of the Chrome browser.

Unlike synthetic testing methods or controlled lab environments, CrUX reflects how websites actually perform under diverse real-life conditions, including different devices, network types, and geographic locations. This makes it an essential resource for studying web performance, enhancing user experience, and improving frontend efficiency.

The dataset features critical performance metrics such as First Contentful Paint (FCP), Largest Contentful Paint (LCP), First Input Delay (FID), Cumulative Layout Shift (CLS), and Interaction to Next Paint (INP). These metrics align closely with Google's Core Web Vitals, which serve as key indicators for assessing responsiveness and visual stability of web applications.

Since CrUX data is updated monthly, it enables continuous tracking of performance trends, identification of regressions, and assessment of the effects of UI or backend changes on user experience.

One of CrUX's major strengths lies in its ability to filter data by various dimensions, including device type (mobile, desktop, tablet), geographic region, and network quality (e.g., 3G, 4G). This flexibility allows for targeted analysis of specific user segments and helps identify performance issues affecting particular demographics. As the dataset is hosted on Google BigQuery, researchers can execute scalable SQL queries without the need for manual file handling, making it ideal for large-scale performance evaluations.

To analyze and model web application performance using Python, a variety of libraries and tools were utilized. Pandas was employed for convenient data manipulation in tabular format, while Numpy handled efficient numerical computations. For visualizing data and exploring inter-variable relationships, Matplotlib and Seaborn were applied. Regression models were developed using Scikit-learn, including *LinearRegression* for linear modeling, *PolynomialFeatures* for polynomial expansion, and *StandardScaler* for data normalization.

Accuracy evaluation was carried out using metrics such as mean absolute error (MAE), mean squared error (MSE), and the coefficient of determination

(R^2). To construct NN, TensorFlow was used – specifically, the Sequential API and Dense layers – to capture non-linear dependencies in the data and compare these findings to regression results.

TensorFlow proved beneficial due to its flexible architecture, which supports deployment on CPUs, GPUs, and TPUs, as well as mobile and embedded systems. This scalability makes it suitable for both small-scale experiments and complex production environments.

The research focused on a set of key performance indicators most relevant to evaluating web application performance. These included: *onloadtime* (total page load time including all resources), *dominteractive* (time at which the page becomes interactive), and *domcontentloadedeventend* (time after the HTML is fully parsed, before all media loads). Additional metrics involved *layoutcount* (number of layout recalculations), *jsheaptotalsize* (JavaScript memory usage), *img_initiated_transfer_size* and *css_initiated_transfer_size* (network sizes for images and CSS), and *numberofrequests* (total server requests triggered during load). To streamline processing and improve analysis efficiency, the dataset was reduced to include only the most relevant parameters. This helped focus the evaluation specifically on metrics that most significantly impact web application performance.

The selected data, stored in float64 format across 10,000 records, allowed for detailed dependency analysis and accurate regression modeling.

For model training and testing, the dataset was divided using an 80/20 split via the *train_test_split* function (with *random_state=42* for reproducibility). Training was conducted using a GPU-based setup (NVIDIA T4/P100), with 12 GB of RAM and 2 vCPUs, which significantly accelerated the learning process.

This research proposes a software-driven methodology for assessing web application performance through the use of ML techniques. It addresses both the structural components and predictive modeling of performance, aiming to detect inefficiencies and forecast issues before they negatively affect user experience.

The approach starts with selecting critical performance indicators such as response time, memory and CPU usage, network speed, full load time, time to interactivity, layout recalculations, JavaScript memory consumption, and the number of server requests. These metrics were chosen for their practical importance in evaluating real-world application behavior. To facilitate data acquisition, custom software agents and backend services were implemented, enabling continuous monitoring of application activity. After collection, the data was cleaned and prepared using

Python libraries to ensure quality and consistency for analysis. The modeling process incorporated ML algorithms – including linear and polynomial regression, decision tree, and neural network – to identify anomalies and predict potential performance drops.

This methodology offers a replicable and scalable framework that integrates automated data gathering, feature extraction, and predictive analysis. It provides a robust foundation for combining machine learning with advanced web performance monitoring systems.

To simplify the task and improve the efficiency of analysis, the dataset was reduced by limiting the number of rows and selecting only the most essential parameters. This approach minimized the volume of processed data while maintaining focus on the key metrics necessary for assessing the performance of web applications. The selected data is stored in the int64 numerical format, which allows for detailed examination of relationships between variables and enables the construction of regression models to predict web application performance:

```
< class 'pandas.core.frame.DataFrame '>
RangeIndex: 10000 entries, 0 to 9999
dtypes : float64(8)
memory usage: 625.1 KB
onloadtime dominteractive domcontentload-edeventend layoutcount \
count 10000.000000 10000.000000 10000.000000 10000.000000
mean 4849.606800 2165.309672 2345.924978 39.31970
standard 4577.348315 1961.989792 2084.977438 76.60846
min 0.000000 0.000000 0.000000 0.000000
25% 1820.750000 950.393750 1044.675000 7.000000
50% 3501.000000 1634.337500 1779.035000 20.00000
75% 6152.500000 2761.778750 3023.257500 44.00000
max 29919.000000 29423.770000 29423.805000 2566.00000
jsheaptotalsize img_initiated_transfer_size \
count 1.000000e+04 1.000000e+04
mean 1.422299e+07 7.054487e+05
std 1.344788e+07 2.328241e+06
min 1.843200e+06 0.000000e+00
25% 5.251072e+06 5.130000e+02
50% 1.041613e+07 7.473500e+04
75% 2.032845e+07 5.419240e+05
max 1.884570e+08 8.057756e+07
css_initiated_transfer_size numberofrequests
count 1.000000e+04 10000.000000 ...
```

The dataset was standardized and then underwent an initial correlation analysis to identify relationships between the variables (Fig. 1).

The data underwent normalization and preliminary correlation analysis. Clear correlations were found among key time-based metrics, such as *onloadtime*, *dominteractive*, and *domcontentloadedeventend*. A strong correlation was also observed between *jsheaptotalsize* and *onloadtime*, suggesting a link between JavaScript memory usage and loading time. Conversely, transfer sizes for images and CSS files showed weak correlations with performance metrics.

In this research, a combination of linear regression (LR), polynomial regression, decision tree, and neural network (NN) models was applied to evaluate web application performance. These models were chosen to capture a variety of complexities in the data.

Linear and polynomial regression serve as interpretable baselines, suitable for identifying straightforward relationships in performance metrics.

Decision trees offer a flexible, non-linear alternative capable of capturing feature interactions, while NN are ideal for modeling intricate, non-linear dependencies in high-dimensional data, offering robust performance in dynamic environments. This diverse set of models ensures a balance between interpretability and predictive power. For the linear

regression model, the LR class was used to fit the training data and generate predictions. The model achieved a high R^2 score of 0.936, explaining about 93.6% of the variance.

The resulting regression equation included key predictors with both positive and negative impacts on the target variable:

$$y = 2164.4 + 110.6 \cdot x_0 + 1885.6 \cdot x_1 + (-13.4) \cdot x_2 + (-9.6) \cdot x_3 + (-0.2) \cdot x_4 + (-80) \cdot x_5 + (-15.1) \cdot x_6, \quad (1)$$

where y represents the predicted outcome, while $x_0 \dots x_6$ are the predictor variables contributing to the model.

The variable x_1 had the strongest positive effect, while x_3 and x_6 contributed most significantly to decreases in the predicted value, providing valuable outcomes into performance influencers.

Polynomial regression was used to capture more complex, non-linear relationships by introducing second-degree polynomial features. The model achieved a MAE of 190.91, MSE of 222,434.98, and an R^2 of 0.9336, indicating strong performance.

A decision tree model was also tested, which, while capable of modeling non-linear patterns, underperformed with a lower R^2 of 0.8628 and higher error rates, possibly due to data noise or lack of tuning.

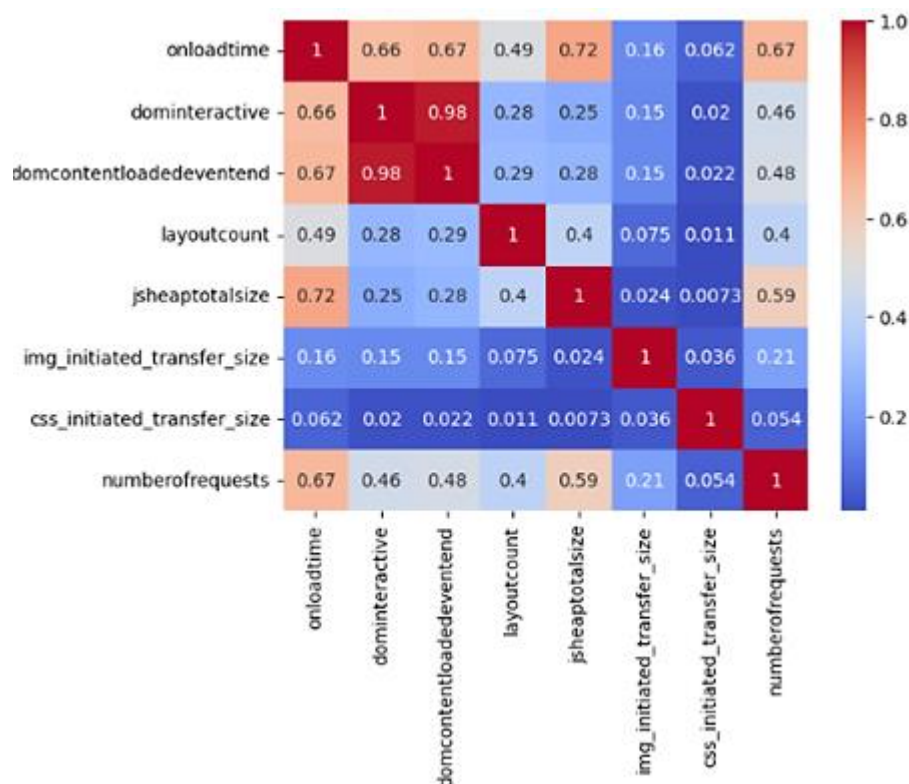


Fig. 1. The correlation between the variables

NN was designed with three dense layers (64, 32, and 1 neurons) using ReLU activation and trained for 50 epochs with the Adam optimizer. The model achieved a MAE of 191.47, MSE of 215,483.79, and an R^2 score of 0.9357, closely matching the performance of polynomial regression. These results suggest that while NN adds modeling power for complex data, polynomial regression offers similar accuracy with greater interpretability in this context.

The findings indicate that while NN offer a powerful and flexible modeling framework, they do not significantly outperform simpler regression models in this particular case. LR model demonstrated the best performance, achieving the lowest Mean Absolute Error (MAE = 187.25) and Mean Squared Error (MSE = 213,550.03), which confirms its reliability and consistency in predictions.

Proposed software architecture designed for evaluating web application performance. It outlines the full workflow – from initial data acquisition to analysis, reporting, and result visualization. The process starts with the Driver module, which communicates with the web application through the DevTools Protocol to gather performance-related data [10].

The gathered data is forwarded to the Data Analysis module, where regression techniques are applied to examine the key factors influencing web application performance. This step leads to the creation of an Analysis Report. Within this framework, regression models are utilized to assess how specific performance metrics affect the overall responsiveness of the application.

Through regression analysis, correlations between different parameters and page load efficiency are identified, and the influence of each metric on performance can be forecasted. Applying various regression approaches allows for comparison of prediction accuracy and the selection of the most suitable method for ongoing monitoring and optimization. After analysis, the refined data is directed to a Classification module, which organizes it for further interpretation. These categorized outputs – referred to as Artifacts – are sent to the Audit module, where they are transformed into JSON-formatted results.

Proposed software system compiles a comprehensive performance report covering several categories such as PWA, Performance, Accessibility, and Best Practices. It also considers multiple contextual factors including server and client-side data, user behavior, timestamps, and data source characteristics.

The analyzed results are delivered through a WebSocket connection and visualized in the Report Results Display module, enabling real-time tracking

and evaluation of web application performance. For the dataset and problem under research, LR proved to be the most suitable approach due to its simplicity, high efficiency, and strong accuracy. The table 1 summarizes the performance of four ML models used for web application performance prediction.

Table 1

Experimental research results

ML model	MAE	MSE	R^2
Linear regression	187.25	213550.03	0.9363
Polynomial regression	190.91	222434.98	0.9336
Decision tree	245.80	459650.08	0.8629
Neural network	191.48	215483.79	0.9357

LR model showed the best results with the lowest MAE (187.25) and MSE (213550.03), and the highest R^2 (0.9363), indicating high accuracy and stability. Polynomial regression and the NN achieved comparable results, slightly less accurate. The Decision Tree performed the worst, with the highest errors and the lowest R^2 , showing limited effectiveness for this task.

It effectively balances interpretability with predictive performance, making it an optimal choice for this analysis. The research developed a methodology for evaluating the performance of web applications, implemented as a software-based solution. This solution can be integrated into web applications to automatically analyze and monitor key performance indicators. The primary metrics selected include the total page load time (covering all resources such as images, stylesheets, and scripts), and the time at which the HTML document is fully loaded and parsed. Other resources, like images and multimedia content, may continue loading while the HTML is being processed. Additional performance indicators include the number of layout recalculations (which affect rendering efficiency), the total memory usage by JavaScript objects, the total size of images transferred over the network, the cumulative size of CSS files loaded during rendering, and the number of server requests initiated during the page load. To streamline the analysis process and enhance evaluation efficiency, the dataset was reduced by focusing solely on these critical parameters. This approach minimized the volume of data processed while maintaining emphasis on the most impactful performance metrics of the web application. The developed software solution unifies all key stages of web application performance assessment – data gathering, processing, classification, and reporting – into a cohesive and automated system. This comprehensive approach enhances the efficiency and accuracy of performance evaluation.

To boost web application performance based on the analytical outcomes, it is advised to minimize server requests by implementing data caching, consolidating multiple requests, and refining API interactions. These measures help decrease server load and enhance responsiveness. Incorporating asynchronous requests can further reduce delays and optimize resource usage. Effective monitoring of CPU and memory consumption is essential for timely scaling – either horizontally by adding more servers or vertically by upgrading existing ones.

Reducing data transfer volume through compression techniques, image and static resource optimization, and the use of Content Delivery Networks (CDNs) can significantly accelerate page load times. Optimizing code, particularly by simplifying SQL queries and computational algorithms, can also help avoid unnecessary resource strain. Unlike conventional methods that often depend on manual checks or isolated performance indicators, this solution leverages ML to automatically identify and predict performance bottlenecks. It delivers real-time, data-driven outcomes and offers practical recommendations such as caching improvements, asynchronous communication, API optimization, and dynamic scaling – resulting in a smarter and more proactive performance management strategy. A software architecture is proposed for intelligent analysis of JavaScript web application performance. It collects numerical data, applies regression analysis to assess key factors. The software offers recommendations to developers for optimizing performance and reducing server load.

Conclusions. Applying ML techniques to evaluate the performance of web applications enhances the precision of assessments and enables efficient processing of extensive datasets. Regression models are instrumental in uncovering key performance trends, while deep NN allow for forecasting outcomes under

dynamic and complex scenarios. The proposed software solution streamlines the entire evaluation process by automating data collection, analysis, reporting, and live monitoring. Data is collected through browser DevTools, analyzed using regression algorithms, and presented in real-time via WebSocket through a performance dashboard.

Optimization recommendations derived from this system include minimizing the number of server requests, utilizing asynchronous operations, implementing resource scaling, compressing transmitted data, and balancing loads across infrastructure. This comprehensive and predictive methodology not only increases the accuracy and speed of detecting performance bottlenecks but also supports timely backend and frontend optimization efforts, improving application responsiveness and scalability.

Future work will focus on enhancing neural network efficiency to speed up training processes and integrating ML with advanced performance monitoring tools for more insightful and automated diagnostics. Techniques such as transfer learning or dynamic learning rate adjustments can significantly shorten training durations. Coupling ML models with platforms like Google Lighthouse, New Relic, or Datadog could provide intelligent real-time feedback and adaptive system tuning. An additional direction involves developing hybrid models that merge statistical analysis with ML-based anomaly detection for deeper and more context-aware outcomes. For instance, integrating time-series analysis of server latency with NN capable of spotting unusual memory usage patterns could help flag early warning signs of overloads or memory leaks. Such preemptive detection facilitates proactive optimization, boosting both the reliability and performance of web systems while reducing operational costs by mitigating risks before they escalate.

Bibliography:

1. Anis Z. A systematic review on regression testing for web-based applications. *Journal of Software*. 2015. Vol. 10 (8). P. 971–990. DOI: 10.17706/jsw.10.8.971-990.
2. Ghattas M., Mora A. M., Odeh S. A novel approach for evaluating web page performance based on machine learning algorithms and optimization algorithms. *AI*. 2025. Vol. 6 (2). Article 19. DOI: 10.3390/ai6020019.
3. Thompson H. S. Improved methodology for longitudinal web analytics using Common Crawl. *Communications of the ACM*. 2024. Vol. 27 (6). P. 929–948. DOI: 10.1145/3614419.3644018.
4. Hort M., Kechagia M., Sarro F., Harman M. A survey of performance optimization for mobile applications. *IEEE Transactions on Software Engineering*. 2022. Vol. 48 (8). P. 2879–2904. DOI: 10.1109/TSE.2021.3071193.
5. Lun L., Zetian D., Hoe T. W., Juan X., Jiaxin D., Fulai W. Factors influencing user intentions on interactive websites: Insights from the technology acceptance model. *IEEE Access*. 2024. Vol. 12. P. 122735–122756. DOI: 10.1109/ACCESS.2024.3437418.
6. Shailesh S., Suresh P. V. A survey and analysis of techniques and tools for web performance optimization. *Journal of Information Organization*. 2018. Vol. 8 (2). P. 31–57. DOI: 10.6025/jio/2018/8/2/31-57.

7. Leone J., Traini L. Enhancing trace visualizations for microservices performance analysis. *Companion of the 2023 ACM/SPEC International Conference on Performance Engineering*. 2023. P. 283–287. DOI: 10.1145/3578245.3584729.
8. Willnecker F., Andreas B., Wolfgang G., Helmut K. Using Dynatrace monitoring data for generating performance models of Java EE applications. *Proceedings of the 6th ACM/SPEC International Conference on Performance Engineering*. 2015. P. 103–104. DOI: 10.1145/2668930.2688061.
9. Oleshchenko L., Burchak P. Web application state management performance optimization methods. *Lecture Notes on Data Engineering and Communications Technologies*. 2023. Vol. 181. P. 59–74. Springer, Cham. DOI: 10.1007/978-3-031-36118-0_6.
10. Oleshchenko L.M., Burchak P.V. Software system architecture development for intelligent analysis of web application performance metrics. *Вчені записки Таврійського національного університету імені В.І. Вернадського. Серія: Технічні науки*. Том 35 (74). № 4. 2024. С. 141–150. <https://doi.org/10.32782/2663-5941/2024.4/22>.

Бурчак П.В., Олещенко Л.М. ПРОГРАМНИЙ ФРЕЙМВОРК МАШИННОГО НАВЧАННЯ ДЛЯ ПРОГНОЗУВАННЯ ТА ОПТИМІЗАЦІЇ ПРОДУКТИВНОСТІ ВЕБЗАСТОСУНКІВ

У статті представлено програмний фреймворк для аналізу та прогнозування продуктивності вебзастосунків з використанням методів машинного навчання. Замість традиційного підходу до тестування продуктивності було запропоновано автоматизований конвеєр, який здійснює збір, обробку та моделювання ключових системних метрик – таких як час відповіді сервера, використання пам'яті, навантаження на процесор і чутливість інтерфейсу з метою проактивного виявлення та усунення проблем продуктивності. Для дослідження було використано набір даних із понад 200 000 реальних записів, що включають показники часу завантаження вебсторінки, перерахунку макета, споживання пам'яті JavaScript та пропускну здатність мережі.

У моделювальний конвеєр було інтегровано декілька алгоритмів машинного навчання, зокрема, лінійну регресію, поліноміальну регресію, дерева рішень та нейронну мережу для оцінювання їхньої ефективності у прогнозуванні продуктивності застосунків. За результатами експериментів найкращу точність показала лінійна регресія ($MAE = 187,25$, $R^2 = 0,9357$), що свідчить про сильний лінійний зв'язок у даних. Нейронна мережа також продемонструвала високу точність і переваги при моделюванні динамічних нелінійних залежностей. Розроблена програмна система реалізована з використанням технологій Python: Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn і TensorFlow.

У результаті проведеного дослідження було створено масштабоване, автоматизоване та інтерпретоване програмне рішення для прогнозової діагностики та оптимізації продуктивності вебзастосунків. Запропонована програмна система також дозволяє формувати категоризовані звіти з рекомендаціями щодо покращення продуктивності вебзастосунків. Отримані результати підтверджують ефективність і доцільність використання методів машинного навчання в сучасних системах моніторингу продуктивності вебзастосунків.

Ключові слова: програмний фреймворк, моніторинг продуктивності програмного забезпечення, машинне навчання, прогнозна аналітика, оптимізація вебсистем, регресійні моделі, нейронні мережі, Python, TensorFlow, прогнозування продуктивності вебзастосунків.

Дата надходження статті: 17.07.2025

Дата прийняття статті: 24.07.2025

Опубліковано: 27.10.2025

Веселовська Г.В.

Херсонський національний технічний університет

Плахотнюк М.Ф.

Херсонський національний технічний університет

СПОСІБ ПІДВИЩЕННЯ ПРОДУКТИВНОСТІ МЕТОДІВ УПРАВЛІННЯ В СКЛАДНИХ КОМП'ЮТЕРНО-ОРІЄНТОВАНИХ СИСТЕМАХ НАВЧАННЯ НА ОСНОВІ МЕРЕЖНИХ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

У статті досліджено актуальні аспекти подальшого вдосконалювання сучасних комп'ютерно-орієнтованих систем навчання, наділених властивістю посиленої складності. Основним спрямуванням дослідження став пошук способу досягнення можливості підвищення продуктивності методів управління у вищезазначених системах на засадах активнішого залучення мережних інформаційних технологій. На основі результатів здійсненого попереднього аналізу, в якості методів дослідження, було вибрано елементи методологічного апарату теорії інформації, системного аналізу, експертних систем штучного інтелекту. Для досягнення поставленої мети дослідження, було запропоноване впровадження такого способу, що полягає в інтеграції до певної частини контуру управління комп'ютерно-орієнтованої системи навчання, а саме, до ланцюга зворотного управління контролем знань, спеціальної ланки для забезпечення додаткового моніторингу та корекції (підвищення) мотивації здобувачів освіти до вивчення навчальних дисциплін і, відповідно, зростання ефективності освітніх процесів, на засадах активнішого застосування мережних інформаційних технологій. В якості базису для практичного здійснення зазначеної спеціальної ланки, було запропоновано впровадити модуль, спрямований на підтримку процесів вирішення завдань, котрі передбачають виконання інтенсифікованого цілеспрямованого пошуку інформації в мережі Internet з використанням сучасних мережних інформаційних технологій, відповідно до визначеної системи з восьми взаємно пов'язаних моделей-профілів характеристик пошуку. З метою підвищення загальної ефективності запропонованого підходу, вищевказані моделі було сформовано, як такі, що доповнюють одна одну та забезпечують взаємне посилення їхньої дії, будучи, при цьому, побудованими на засадах різних концепцій та технологій інформаційного пошуку, особливостей та складності практичної реалізації, механізмів впливу на створення потрібних мотиваційних установок. Для кожної з цих моделей, було визначено найдоцільніші технологічні умови та максимальні межі часових термінів їхнього застосування під час здійснення управління освітніми процесами на засадах складних комп'ютерно-орієнтованих систем навчання.

Ключові слова: комп'ютерно-орієнтована система, складна система, метод управління, мережна інформаційна технологія, навчання, продуктивність.

Постановка проблеми. В епоху стрімкого розвитку технологій, комп'ютерно-орієнтовані системи навчання стали міцним підґрунтям і прискорювачем постійного вдосконалювання підготовки фахівців, а методологічні засади організації, функціонування, розробки та практичного застосування зазначених систем набули фундаментального характеру. Сучасні комп'ютерно-орієнтовані системи навчання та методи управління в цих системах доволі часто мають підвищений рівень складності, що, з пли-

ном часу, безперервно зростає, вимагаючи модернізації відповідного методологічного апарату. Одним із актуальних напрямків підвищення продуктивності методів управління в складних комп'ютерно-орієнтованих системах навчання є активізація застосування мережних інформаційних технологій. Разом із тим, на даний час, указаний вище напрямок, незважаючи на потужність отриманих в його межах здобутків, не є достатньо опрацьованим і потребує проведення детальніших досліджень і розробок.

Аналіз останніх досліджень і публікацій.

Питання пошуку нових можливостей та вирішення наявних проблем активізації застосування цифрових і, зокрема, мережних інформаційних технологій, з метою підвищення ефективності організації та функціонування сучасних складних систем навчання, належать до категорії таких, що знаходяться у фокусі досліджень науковців і практиків. У контексті зазначеної предметної галузі досліджень і розробок, важливу роль відіграє розгляд як загальних її питань (поточного стану та перспектив, можливостей і переваг, викликів і проблем) [1-3], так і більш конкретизованих її аспектів, пов'язаних із особливостями практичної реалізації для закладів вищої освіти [4-5], при розв'язуванні певних категорій завдань і використанні визначеного інструментарію [6-10] тощо. Особливо слід відзначити перспективність завдань активізації впровадження до цифрових освітніх процесів Internet-технологій [11-13], що можуть надати численні потужні можливості для вдосконалювання. Одну з таких можливостей, обумовлену чинником підвищення мотивації навчання, що досягається на засадах активнішого використання мережних інформаційних технологій в перебігу виконання спеціальних систем завдань, було представлено на концептуальному рівні в авторській роботі [14]. Результати дослідження, подані до розгляду в цій науковій статті, є подальшою деталізацією концепції, викладеної у вищезазначеній роботі.

Постановка завдання. Метою статті є знаходження способу досягнення можливості підвищення продуктивності методів управління в сучасних комп'ютерно-орієнтованих системах навчання, наділених властивістю посиленої складності, на засадах активнішого залучення мережних інформаційних технологій Internet.

Виклад основного матеріалу. На сучасному етапі розвитку науково-технічного прогресу та тісно пов'язаного з ним удосконалювання соціально-економічних процесів, одне з чільних місць посідає спрямованість на максимально можливе підвищення рівня такого ключового чинника, як людський капітал, будь то індивідуальна або колективна форма його прояву. Зазначене є потужним стимулом для розвитку ринку праці та постійного підвищення його вимог до фахових компетентностей та результатів навчання, отримуваних здобувачами вищої освіти в підсумку опановування ними сучасних спеціальностей.

Одним із найсучасніших і найефективніших підходів до вдосконалювання процесів

і результатів освітньої діяльності є застосування комп'ютерно-орієнтованих систем навчання, що нині активно спираються на такі засоби, як: технології комп'ютерних мереж; методи та засоби інформаційних систем і технологій; можливості дистанційного інформаційно-комунікаційного інструментарію; системи та технології штучного інтелекту.

До підґрунтя успішності зазначеного вище підходу покладено цілу низку сучасних потужних технологічних можливостей, що спроможні підтримувати здійснення таких функцій, як: високоякісна візуалізація процесів викладання та вивчання (статична та динамічна; двовимірна та тривимірна; з використанням контенту з мультимедійними, гіпермедійними та інтерактивними властивостями); зручне користування неосяжними інформаційними ресурсами глобальної комп'ютерної мережі Internet, що мають навчально-методичний характер; комфортна дистанційна комунікація учасників освітнього процесу один із одним та з комп'ютерно-орієнтованою системою навчання; гнучка індивідуалізована та персоніфікована адаптація до особливостей користувачів комп'ютерно-орієнтованих систем навчання, а також детальне врахування специфіки цільових груп користувачів зазначених систем.

Особливо слід відзначити, що активне використання сучасних засобів і методів штучного інтелекту в комп'ютерно-орієнтованих системах навчання надає можливість суттєво підвищувати ефективність управління освітніми процесами, а також продуктивність викладання та опановування навчальних курсів, на засадах наданих цими засобами та методами потужних додаткових можливостей для виконання такої діяльності, як: пошук наявної інформації; здобування та вилучення знань; генерація нової інформації, знань та ілюстрованого контенту навчально-методичного призначення; збільшення ступеня гнучкості моделей адаптації до користувачів зазначених систем тощо.

Серед численних важливих переваг, вище-названі можливості сучасних комп'ютерно-орієнтованих систем навчання надають ґрунтовні передумови для суттєвого підвищення рівня мотивації здобувачів освіти (зокрема, міцного усвідомлення ними актуальності опановування навчальних дисциплін), як одного з найголовніших каталізаторів зростання ефективності освітніх процесів та якості їхніх результатів.

Згідно з даними багаторічних досліджень, одним із визначальних підґрунть для отримання

високого рівня вмотивованості здобувачів вищої освіти, при опановуванні ними певних навчальних курсів, є наявність у цих здобувачів позитивних вражень стосовно тематики та термінології, загального обсягу, активності та сучасності висвітлення ключових питань предметних галузей зазначених курсів, представлених у просторі інформаційних ресурсів всесвітньої комп'ютерної мережі Internet.

Разом із тим, приступаючи до вивчення певного навчального курсу, здобувачі освіти можуть не мати належних понять про ступінь актуальності його предметної галузі, а також попередньо не цікавитися представленістю визначальних понять цієї галузі в Internet. Відповідно, надзвичайно важливим є цілеспрямоване формування у користувачів навчальних курсів мережних комп'ютерно-орієнтованих систем навчання всебічних понять про подання предметних галузей цих курсів в Internet. При цьому, слід брати до уваги ті чинники якості подання предметних галузей навчальних курсів, що стосуються семантичної повноти та інтенсивності поновлення їхньої структури, змісту й обсягів відповідно до актуальних тенденцій сучасності.

На засадах проведеного дослідження, було прийняте рішення про доцільність і практичну обґрунтованість введення до розгляду такої моделі управління формуванням позитивної мотиваційної установки до вивчення навчальних курсів на засадах комп'ютерно-орієнтованих систем навчання, в межах якої передбачається практична реалізація та виконання здобувачами вищої освіти комплексу спеціалізованих завдань восьми базових типів, вирішення котрих надає можливість зазначеним здобувачам зрозуміти повноту та рівень представленості цих курсів у мережі Internet.

Завдання типу I передбачають пошук інформаційних ресурсів мережі Internet за запитом, що містять як повну назву, так і окремі ключові слова та словосполучення назви навчального курсу, розміщеного в комп'ютерно-орієнтованій сис-

темі навчання. Зазначені завдання, в першу чергу, доцільно виконувати в якості вправ для самостійної роботи студентів до першої лекції з дисципліни, дотримуючись заданої верхньої межі часового терміну їхнього здійснення (таблиця 1).

Завдання типу II передбачають Internet-пошук інформаційних ресурсів на підставі переліку визначальних тем дисципліни, що надається викладачем, і надалі може доповнюватися здобувачами вищої освіти, з урахуванням їхніх інформаційних потреб і запитів. Такі завдання доцільно виконувати, дотримуючись таких технологічних умов: а) на початку першого лабораторного (практичного, семінарського) заняття з дисципліни, одразу за вступними настановами, під безпосереднім керівництвом викладача та в активному діалозі з ним та іншими учасниками групи здобувачів вищої освіти, при дотриманні наданого викладачем переліку ключових запитів до пошукової системи, максимально інтенсивного темпу комунікації та певної верхньої межі часового терміну виконання завдань; б) в якості першого виконаного завдання для самостійної роботи студентів до першого лабораторного (практичного, семінарського) заняття з дисципліни, на засадах надання можливості самостійного вибору здобувачами освіти додаткових ключових запитів до пошукової системи, при дотриманні заданої верхньої межі часового терміну здійснення завдання (таблиця 1).

Завдання типу III передбачають пошук інформаційних ресурсів Internet за запитом, що містять теми, а також визначальні підтеми, ключові слова та словосполучення текстів лекцій. Завдання цього типу рекомендується пропонувати до кожної лекційної теми дисципліни, в якості вправ для самостійної роботи до лекцій, за умови дотримання певної верхньої межі часового терміну виконання завдань (таблиця 1).

Завдання типу IV передбачає отримання, на засадах виконання інформаційного Internet-пошуку, тих актуальних відомостей, що наочно розкривають питання ефективного застосування

Таблиця 1

Рекомендована верхня межа часового терміну виконання завдань

Тип завдання		Максимальний час виконання завдання, хв.	Тип завдання		Максимальний час виконання завдання, хв.
I		20	V		30
II	a	10	VI		40
	б	20	VII		20
III		20	VIII	a	40
IV		25		б	20

теоретичного базису навчального курсу на практиці. В цьому випадку, доцільно в першу чергу брати до уваги таку інформацію, що стосується: показових прикладів успішної діяльності (так званих історій успіху) підприємств, організацій та фахівців тієї предметної галузі, котру представлено чи асоційовано з навчальним курсом; можливостей подальшого підвищення ефективності роботи підприємств, організацій та їхніх працівників на засадах використання тих методів та інструментальних засобів, що належать до предметної галузі навчального курсу тощо. Завдання зазначеного типу рекомендовано пропонувати до кожної теми дисципліни, з метою їхнього виконання в ході самостійної роботи здобувачів освіти до лабораторних (практичних, семінарських) занять, при дотриманні певної верхньої межі часового терміну виконання завдань (див. таблицю 1).

Уведення до розгляду завдань типу V, пов'язаних із інформаційним пошуком в Internet, обумовлено ще одним ключовим напрямом підвищення мотивації здобувачів вищої освіти щодо активізації їхнього навчання. Зазначений напрям передбачає значне збільшення обсягу таких завдань практичного фахового спрямування, що разом утворюють наскрізні системи покрокового розв'язування задач, де формулювання кожної наступної задачі являє собою розширену, деталізовану та ускладнену версію постановки попередньої задачі, потребуючи, при цьому, пошуку додаткової інформації, потрібної для її розв'язування, в мережі Internet. Як свідчить практична діяльність, указаний тип завдань значно підвищує мотивацію здобувачів вищої освіти щодо посилення активності навчання, бо надає можливість цим здобувачам, дотримуючись визначених меж термінів опановування навчального курсу, виконувати повномасштабні завдання професійного спрямування, поетапно нарощуючи об'єм і складність як їхніх компонент, так і додаткової інформаційної підтримки цих компонент. Завдання такого типу рекомендується пропонувати здобувачам вищої освіти до кожної з тем дисципліни, розміщуючи їх у переліку завдань для виконання під час самостійної роботи до лабораторних (практичних, семінарських) занять, при дотриманні визначеної верхньої межі часового терміну їх опрацювання (див. таблицю 1).

Розглянемо особливості завдань типу VI, пов'язаних зі здійсненням пошуку інформації в мережі Internet. Підґрунтям актуальності цього типу завдань є той чинник, що значний вплив на зростання ступеня вмотивованості здобувачів

вищої освіти до вивчення навчальних курсів і відповідного покращання якості результатів їхнього навчання здійснює розв'язування спеціалізованих задач, спрямованих на здійснення, в ході підготовки до аудиторних занять, діяльності з самостійного формування полів знань базових навчальних текстів, асоційованих із цими курсами. Побудову полів знань на основі базових навчальних текстів, пов'язану з необхідністю проходження множини спеціальних дій, доцільно виконувати за допомогою представленого далі способу. Базовий навчальний текст обирають в якості такої інформації, що якнайповніше представляє семантику визначеної теми навчального курсу, шляхом фільтрування цієї інформації з масиву наявного актуального контенту інформаційних ресурсів Internet (за необхідності, інтегрують декілька текстів). Надалі базовий навчальний текст прочитують та аналізують, виділяючи з нього ключові слова та словосполучення. Наступне, детальніше читання базового навчального тексту, з його подальшим аналізом, надає можливість встановлення семантичних зв'язків між попередньо виділеними ключовими словами та словосполученнями, а також побудови відповідної таблиці чи графу, з метою наочної формалізації зазначених зв'язків. Під час останнього, ще більш детального, читання та аналізу базового навчального тексту, спираючись на інформацію, отриману під час здійснення попередніх етапів роботи над формуванням поля знань, складають стислий текст-анотацію, що відображає виділені раніше ключові слова, словосполучення та семантичні зв'язки між ними. Таким чином, за допомогою наведеного вище способу, на основі формування, вивчення, аналізу та модифікації базового навчального тексту, поетапно моделюється поле знань, що в підсумку подається в форматі стисненого анотованого тексту. Цей тип завдань варто використовувати під час самостійної роботи до лабораторних (практичних, семінарських) занять, при дотриманні заданого максимуму часового терміну їхнього виконання (див. таблицю 1).

Навчальні завдання наступного представленого до розгляду типу VII, що теж пов'язані з діяльністю інформаційного пошуку в Internet і надають можливість значного посилення мотиваційного рівня здобувачів вищої освіти та сприяння відповідному зростанню продуктивності освітніх процесів, передбачають складання словників термінів (тезаурусів). Застосування такого типу завдань є доцільним в ході здійснення самостійної роботи здобувачів вищої освіти до лекційних занять, при

дотриманні заданого максимуму часового терміну їхнього виконання (див. таблицю 1).

Перед розглядом особливостей завдань типу VIII, асоційованих із інформаційним пошуком у мережі Internet, зробимо наголос на надзвичайній важливості ролі активних методів навчання в сприянні зростанню його вмотивованості та відповідної ефективності. Виходячи з результатів практичних досліджень, посилену продуктивність освітніх процесів було продемонстровано на засадах застосування таких видів вищезазначених методів, де за організаторів, експертів та арбітрів обираються здобувачі вищої освіти, а викладачу відводиться роль постановника початкового навчального завдання, спостерігача та консультанта. Дотримуючись вищезначеного підходу, доцільно в першу чергу використовувати ті заходи, асоційовані з тематикою навчальних курсів, що втілюють у практику інтегровані моделі організації та практичного здійснення таких активних методів навчання, котрі передбачають роботу з експертами, як то: мозкові штурми; круглі столи; анкетування; опитування; інтерв'ювання; ділові ігри. Роль інформаційного пошуку в Інтернет при виконанні завдань зазначеного типу полягає в отриманні в такий спосіб учасниками діяльності, що передбачає використання активних методів навчання, потрібних інформаційних підґрунть для здійснення цієї діяльності. Робота з завданнями вказаного типу є доцільною в таких випадках: а) під час виконання самостійної роботи здобувачів вищої освіти до лабораторних (практичних, семінарських) занять (при дотриманні заданої максимальної часової межі – див. таблицю 1); б) власне на лабораторних заняттях (із визначеним обмеженням за часом – див. таблицю 1).

За базовий формат звітності за підсумками виконання інформаційного пошуку в межах кожного з завдань типів I-VIII, рекомендується взяти об'єктно-орієнтовану структуру з полями, що міститимуть таку інформацію: посилання на актуальний інформаційний Internet-ресурс; найменування зазначеного ресурсу; коротка (анотована) характеристика вмісту цього ресурсу.

Висновки. В роботі здійснене дослідження тих актуальних аспектів, що сформували підґрунтя для подальшого вдосконалювання сучасних комп'ютерно-орієнтованих систем навчання, котрим властива посилена складність. Дослідження було спрямоване на знаходження спо-

соби отримання таких передумов, що дозволять підвищити продуктивність методів управління у складних комп'ютерно-орієнтованих системах навчання на основі більш активного застосування мережних інформаційних технологій. Виходячи з попередньо проведеного аналізу, за основні методи дослідження, було взято методологічні елементи низки фундаментальних теорій, таких як теорія інформації, системний аналіз, теорія експертних систем штучного інтелекту. Для того, щоб здійснити поставлену мету дослідження, було запропоновано впровадити певний спосіб, що передбачає інтегрувати спеціальну ланку до такої частини контуру управління комп'ютерно-орієнтованої системи навчання, як ланцюг зворотного управління контролем знань. Особливістю призначення вказаної ланки стала підтримка додаткових функцій зі спостереження та подальшого корегування (посилення) рівня вмотивованості тих, які навчаються, до опанування навчальних курсів, із відповідним підвищенням продуктивності процесу освіти, на основі активізації використання мережних інформаційних технологій. За основу для втілення вказаної спеціальної ланки в практику, було запропоноване впровадження модулю, спрямованого на забезпечення підтримки процесів розв'язування задач, що висувують завдання здійснення інтенсивного цілеспрямованого пошуку інформації в Internet із застосуванням сучасних мережних інформаційних технологій, згідно з визначеною системою, що охоплює вісім взаємопов'язаних моделей-профілів характеристик пошуку. Для посилення результативності пропонованих підходів у цілому, вищезазначені моделі було побудовано таким чином, що вони являють собою доповнення одна одної та взаємно підвищують ефективність своєї дії при спільному застосуванні, будучи створеними на підґрунті різних концептуальних і технологічних способів здійснення пошуку інформації, специфіки та складності втілення в практичній діяльності, сутності впливу на формування необхідних мотиваційних установок. За кожною з указаних моделей, були визначені найбільш доцільні технологічні підстави та верхні межі часових інтервалів для їхнього використання в ході виконання управління освітніми процесами в тих середовищах, що надають складні комп'ютерно-орієнтовані системи навчання.

Список літератури:

1. Толмач М.С. Цифрові технології в освіті: можливості й тенденції застосування. *Цифрова платформа: інформаційні технології в соціокультурній сфері*. 2021. Т. 4, № 2. С. 159–171. URL: http://nbuv.gov.ua/UJRN/dpitsca_2021_4_2_6 (дата звернення: 10.09.2024).
2. Гуревич Р. Цифрові грамотність, компетентність, технології – точки дотику в освітньому процесі / Р. Гуревич, Л. Коношевський, О. Коношевський, Н. Опушко, М. Драчук. *Збірник наукових праць Національної академії Державної прикордонної служби України. Серія «Педагогічні науки»*. 2023. № 1. С. 64–86. URL: http://nbuv.gov.ua/UJRN/znpradpcpn_2023_1_7 (дата звернення: 12.09.2024).
3. Медведовська О.Г. Впровадження цифрових технологій в освітній процес в Україні. *Наукові записки Центральноукраїнського державного педагогічного університету імені Володимира Винниченка. Серія «Педагогічні науки»*. 2021. Вип. 198. С. 132–136. URL: http://nbuv.gov.ua/UJRN/Nz_p_2021_198_31 (дата звернення: 17.09.2024).
4. Шищенко І. Деякі аспекти впливу цифрових технологій на освітній процес закладів вищої освіти: огляд проблем та викликів. *Освіта. Інноватика. Практика*. 2022. Т. 10, № 5. С. 42–47. URL: http://nbuv.gov.ua/UJRN/edinpr_2022_10_5_8 (дата звернення: 19.09.2024).
5. Потюк І.Є. Використання цифрових технологій в навчальному середовищі закладів вищої освіти: офлайн та онлайн формати. *Наукові записки Національного університету «Острозька академія»*. 2021. Вип. 11. С. 219–221. URL: http://nbuv.gov.ua/UJRN/Nznuoaf_2021_11_56 (дата звернення: 20.09.2024).
6. Садовий М.І., Соменко Д.В., Трифонова О.М., Доброван К.М. Розвиток техніко-технологічної компетентності під час виконання навчального проекту з використанням цифрових ресурсів. *Наукові записки Центральноукраїнського державного педагогічного університету імені Володимира Винниченка. Серія «Проблеми природничо-математичної, технологічної та професійної освіти»*. 2023. Вип. 1. С. 41–47. URL: http://nbuv.gov.ua/UJRN/nzprpmtpo_2023_1_8 (дата звернення: 24.09.2024).
7. Козир М.В., Іващенко А.В. Цифрові технології як інструмент консультативної діяльності в освіті дорослих. *Педагогічна освіта: теорія і практика. Психологія. Педагогіка*. 2023. № 39. С. 73–79. URL: http://nbuv.gov.ua/UJRN/Potip_2023_39_13 (дата звернення: 26.09.2024).
8. Мар'єнко М.В., Сухих А.С. Особливості організації змішаного навчання з використанням цифрових технологій. *Освітній дискурс*. 2021. Вип. 32. С. 45–52. URL: http://nbuv.gov.ua/UJRN/eddcsp_2021_32_7 (дата звернення: 27.09.2024).
9. Козяр М.М., Тимошук О.С. Інформаційно-цифрові технології доповненої та віртуальної реальності в освіті: огляд технічних і дидактичних можливостей. *Інноваційна педагогіка*. 2022. Вип. 45. С. 179–183. URL: http://nbuv.gov.ua/UJRN/innpred_2022_45_38 (дата звернення: 08.10.2024).
10. Романишин Ю.Л. Веборієнтовані віртуальні спільноти та комунікації в навчанні фахівців інформаційної сфери. *Інформаційні технології і засоби навчання*. 2021. Т. 85, № 5. С. 228–243. URL: http://nbuv.gov.ua/UJRN/ITZN_2021_85_5_17 (дата звернення: 10.10.2024).
11. Кузьменко О.Ю. Теоретичні аспекти запровадження інтернет-технологій в умовах цифровізації закладів вищої освіти. *Інноваційна педагогіка*. 2021. Вип. 40. С. 239–243. URL: http://nbuv.gov.ua/UJRN/innpred_2021_40_52 (дата звернення: 15.10.2024).
12. Семчук С. Впровадження інтернет-технологій в освітній процес закладів вищої освіти. *Психолого-педагогічні проблеми сучасної школи*. 2021. Вип. 1. С. 136–142. URL: http://nbuv.gov.ua/UJRN/Ppps_2021_1_18 (дата звернення: 17.10.2024).
13. Грицюк Ю.І., Дяк Т.П. Застосування інтернет-технологій для організації навчального процесу у вищих навчальних закладах. *Науковий вісник НЛТУ України*. 2021. Т. 31, № 1. С. 137–146. URL: http://nbuv.gov.ua/UJRN/nvnltu_2021_31_1_25 (дата звернення: 22.10.2024).
14. Веселовська Г.В. Удосконалення процесів викладання в сфері вищої освіти на засадах підвищення мотивації навчання шляхом використання мережних інформаційних технологій та спеціальних систем завдань. *Професійний розвиток, удосконалення загальних і професійних компетентностей викладачів вищої школи: матеріали Всеукраїнського науково-педагогічного підвищення кваліфікації (Кропивницький, 8 липня – 18 серпня 2024 р.)*. Львів–Торунь: Liha-Pres, 2024. 164 с. С. 17–21.

Veselovska H.V., Plakhotniuk M.F. THE WAY OF INCREASING THE PRODUCTIVITY OF CONTROL METHODS IN COMPLEX COMPUTER-ORIENTED TRAINING SYSTEMS BASED ON NETWORK INFORMATION TECHNOLOGIES

The actual aspects in further improvement of modern computer-oriented training systems that are endowed with the property of increased complexity have been investigated in article. The main direction of the research was to find the way to achieve the possibility to increase the productivity of control methods in the above systems on the basis of more active involvement of network information technologies. Based on the results

of the preliminary analysis, the elements of the methodological apparatus in the information theory, system analysis, and expert systems of artificial intelligence were selected as research methods. In order to achieve the set goal of the research, it was proposed to introduce such a way that is to integrate into the certain part of the control circuit of the computer-oriented training system, namely, into the reverse control chain of knowledge control, the special link to ensure additional monitoring and correction (increase) of motivation for education applicants to study disciplines and, accordingly, to increase the effectiveness of educational processes, on the basis of more active use of network information technologies. As a basis for practical implementation of the specified special link, it was proposed to introduce the module aimed at supporting the processes for solving problems that involve the performance of intensified purposeful search for information in the network Internet using modern network information technologies according to the defined system of eight mutually linked models-profiles of search characteristics. In order to increase the overall efficiency of the proposed approach, the above models have been formed as complementing each other and ensuring mutual strengthening of their action, being built on the basis of different concepts and technologies of information search, features and complexity of practical implementation, mechanisms of influence on creating the required motivational attitudes. For each of these models, the most appropriate technological conditions and maximum time limits for their application during control of educational processes on the basis of complex computer-oriented training systems were determined.

Key words: computer-oriented system, complex system, control method, network information technology, training, productivity.

Дата надходження статті: 19.06.2025

Дата прийняття статті: 26.06.2025

Опубліковано: 27.10.2025

Вишневий С.В.Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»**Журба А.В.**Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»

СИНТЕЗ ВІДЕОПОСЛІДОВНОСТЕЙ, ЩО МІСТЯТЬ ОДНОРІДНУ ТЕКСТУРУ ІЗ СЕПАРАБЕЛЬНОЮ АВТОКОРЕЛЯЦІЙНОЮ ФУНКЦІЄЮ

В статті розглянуто задачу синтезу відеопослідовностей, в яких всі кадри містять однорідну текстуру із однаковими статистичними характеристиками. Виконано стислий огляд сфер практичного застосування методів генерування текстурних зображень в контексті обробки відеоданих. Проведено дослідження алгоритму синтезу відеопослідовностей, який побудований на основі застосування авторегресійної моделі гаусівського випадкового поля із сепарабельною автокореляційною функцією, що враховує просторово-часову кореляцію відліків. В статті наведено розрахункові вирази, а також блок-схеми алгоритмів, які можуть бути використані для реалізації розглянутого підходу формування відеопослідовностей за допомогою відповідної мови програмування. Використання представленого алгоритму синтезу послідовності текстурних зображень потребує виконання попереднього визначення параметрів авторегресійної моделі шляхом аналізу навчальних даних, які містять однорідну текстуру, в ході якого мають бути оцінені коефіцієнти кореляції вздовж кожної із координат, а також виконано обрахунок середнього значення яскравості та середнього квадратичного відхилення значень пікселів текстури. Моделювання алгоритму реалізовано в програмному середовищі MATLAB. В статті представлено зразки синтезованих текстур, надано показники часових витрат створення відеофайлів для різних форматів кадру, приведено графіки автокореляційних функцій та гістограми отриманих зображень. З'ясовано, що кадри відеопослідовностей характеризуються анізотропною автокореляційною функцією, що в деяких випадках може бути розглянуто як недолік, для подолання якого необхідно ускладнення процедури синтезу текстури шляхом застосування моделі більш високого порядку. Результати роботи можуть бути використані в наукових дослідженнях та практичних задачах, в яких необхідно виконувати синтез відеопослідовностей, що містять текстуру.

Ключові слова: відеопослідовність, текстура, авторегресійна модель, гаусівське випадкове поле, алгоритм, цифрове зображення.

Постановка проблеми. Стрімкий розвиток електронної компонентної бази та значне зростання обчислювальної потужності мікропроцесорних пристроїв сприяло розвитку та широкому впровадженню апаратно-програмних комплексів та систем різного практично-прикладного призначення, що передбачають реєстрацію, передачу, обробку та збереження даних, які представлені у вигляді послідовності цифрових зображень, що застосовуються для вирішення цілого переліку різноманітних задач аналізу даних [1, 2]. Важливою відмінністю відеопослідовності від статичних одиночних зображень є можливість реєструвати дина-

міку зміни об'єкту інтересу. Відеопослідовності загалом характеризуються суттєвою надмірністю [3]. З метою зниження розміру відеоданих застосовуються алгоритми компресії [4–6]. Наприклад, в цифровому телебаченні широке поширення набув стандарт сімейства MPEG, використання якого дає можливість значно зменшити об'єм даних за допомогою використання алгоритму стиснення зображення із втратами [7, 8]. В деяких практичних задачах може виникати необхідність не використовувати стиснення даних задля забезпечення високої якості кадрів, однак при цьому відеодані характеризуються значно більшим розміром.

Серед властивостей об'єктів, які можуть бути представлені на кадрах відеопослідовності, текстура є однією із визначальних особливостей [9, 10]. Текстура може бути розглянута як псевдоперіодичний візерунок, який поширюється в межах деякої області на зображенні. Наявністю текстури можуть характеризуватися послідовності зображень, що отримуються в спеціалізованих медичних діагностичних комплексах, в системах моніторингу та спостереження за ділянками земної поверхні, текстуру також можуть мати об'єкти що містяться на кадрах телевізійних зображень тощо.

З-поміж широкого кола завдань, для вирішення яких використовуються методи цифрового оброблення зображень, своє застосування знаходить синтез текстур. До практичних задач, в яких виникає необхідність застосування алгоритмів синтезу текстур, можна віднести відновлення окремих пошкоджених ділянок в кадрі або видалення небажаних об'єктів [11–13]. Вирішення цього завдання може передбачати аналіз властивостей неушкодженої текстурної області, синтез штучної текстури, і заміщення пошкоджених або небажаних фрагментів в кадрах відеопослідовності згенерованими текстурними зразками. Алгоритми синтезу текстур можуть застосовуватися в задачах стиснення та передачі даних, зокрема, в каналах зв'язку з низькою пропускнуою здатністю замість передачі інформації про значення пікселів деякої текстурної ділянки кадру, можуть бути визначені параметри та характеристики текстури, які передаються на приймальну сторону по каналу зв'язку, а на приймальній стороні виконується синтез відповідних текстурних ділянок кадру [14–16]. Таким чином, відбувається фактичне заміщення реальної текстури синтезованим аналогом. Синтез текстурних відеопослідовностей може застосовуватися в задачах комп'ютерної графіки, в індустрії комп'ютерних ігор, в мультимедійних системах доповненої реальності [17]. Крім того, процедура синтезу текстур може бути задіяна в задачах дослідження алгоритмів та методів обробки цифрових зображень та відеопослідовностей. Застосування алгоритмів синтезу текстур дозволяє виконати генерування даних, які по своїм статистичним властивостям відповідають реальним даним. Отримані текстурні зображення можуть додатково спотворюватися відповідним видом завади із заданими параметрами, які узгоджуються із завадою, вплив якої передбачається в умовах функціонування деякої реальної технічної системи. Синтезовані «ідеальні» дані, а також дані, що спотворені завадою, можуть бути під-

дані обробці відповідним алгоритмом (фільтрації, сегментації тощо) з метою оцінки кількісних показників його роботи [18, 19]. Зміна параметрів моделі текстури дозволяє створити цілий набір різноманітних даних, які по своїм характеристикам співпадають із реальними даними. Це дає змогу проведення детального аналізу відповідного алгоритму обробки відеопослідовностей без необхідності використання великої колекції реальних даних, доступ до яких може вимагати додаткових витрат. Замість цього використовуються синтезовані дані. Таким чином, важливою складовою є вибір алгоритму синтезу текстур; визначення особливостей його практичної реалізації; оцінка часових витрат процесу синтезу текстурних відеопослідовностей; аналіз розміру отриманих відеофайлів тощо.

Аналіз останніх досліджень і публікацій.

Враховуючи велике різноманіття текстур, які можуть зустрічатися на цифрових зображеннях та відеопослідовностях, на поточний момент не існує єдиного підходу, який би задовольняв всі вимоги, що можуть висуватися до методів та алгоритмів синтезу статичних або динамічних текстур. Своє застосування на практиці знаходять алгоритми, які передбачають використання текстурного зразка, що представляє собою деякий блок, який відтворюється в межах області синтезу [20, 21]. Такий підхід може бути використаний для текстур, які містять псевдоперіодичний характер паттерну, однак в цьому випадку треба враховувати можливу необхідність застосування процедур усунення артефактів, які проявляються у вигляді швів між окремими блоками. Широке поширення набули моделі, які основані на використанні марковських випадкових полів [22, 23]. В якості недоліків таких моделей можна віднести обчислювальну складність їхньої практичної реалізації. Останнім часом все більшу популярність здобувають алгоритми синтезу текстур, що основані на використанні технологій машинного навчання та нейронних мереж, які знаходять своє застосування в різних сферах від комп'ютерної графіки до стиснення та відновлення зображень у відеопослідовностях та статичних цифрових знімках [24–27]. Дані алгоритми дозволяють отримувати текстури, які мають високий рівень реалістичності, однак при практичному використанні таких алгоритмів виникає необхідність проведення попереднього навчання нейронних мереж, що потребує значних обсягів навчальних даних, а також такі алгоритми можуть характеризуватися високими обчислювальними витратами. Для

деяких задач може виникати необхідність використання моделей синтезу текстурних зображень та відеопослідовностей, які характеризуються помірним рівнем складності своєї реалізації та не вимагають значних обчислювальних потужностей при їх використанні, а також не потребують проведення тривалої процедури навчання моделей. На основі використання авторегресійної моделі можуть будуватися алгоритми синтезу текстур для генерування двовимірних зображень [28, 29]. Авторегресійна модель може застосовуватися і для реалізації генерування відеопослідовностей. Параметри моделі можуть визначатися на основі попередньо оціненої автокореляційної функції, дисперсії та середнього значення яскравості початкової вибірки.

Постановка завдання. Метою статті є проведення аналізу алгоритму синтезу відеопослідовностей, що характеризуються наявністю однорідної текстури в межах сформованої множини кадрів, використовуючи для реалізації алгоритму авторегресійну модель генерування гаусівського випадкового поля, що характеризується сепарабельною автокореляційною функцією, та виконання оцінки результатів синтезу відеоданих.

Виклад основного матеріалу. В загальному вигляді відеопослідовність представимо у формі множини кадрів [7, 30]. Окремий кадр в свою чергу може бути розглянутий як двовимірний масив, де значення кожного елементу масиву буде відповідати рівню яскравості пікселя. Такий підхід буде адекватним опису послідовності, яка складається із зображень в градаціях сірого. При переході до кольорового зображення, використовуючи модель RGB, в загальному випадку кожен піксель характеризується значенням інтенсивності в червоному, зеленому та синьому каналах, таким чином це несуттєво ускладнює структуру представлення цифрового зображення. Будемо покладати, що окремий кадр є однорідним текстурним зображенням, причому статистичні характеристики та параметри текстури в межах відеопослідовності є незмінними.

Позначимо піксель відеопослідовності таким чином: $v(x,y,k)$, де v – піксель, значення якого відповідає рівню яскравості елементу зображення; x – координата, яка відповідає за номер стовпця, в якому розміщується піксель в кадрі ($x=1, \dots, N$); y – координата, яка відповідає за номер рядка, в якому розташований піксель в кадрі ($y=1, \dots, M$); k – координата, яка відповідає за номер кадру ($k=1, \dots, K$). З урахуванням прийнятих позначень відеопослідовність схематично представлено на рис. 1.

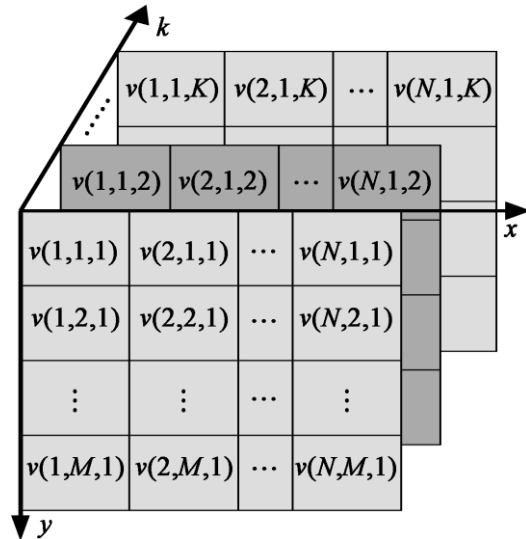


Рис. 1. Схематичне представлення відеопослідовності

В загальному випадку двовимірна авторегресійна модель передбачає проведення обчислення коефіцієнтів моделі та дисперсії шуму збудження, використовуючи для цього систему лінійних рівнянь розміри якої залежать від обраного порядку моделі (a,b) [31]. В роботі [32] приведено загальний вид авторегресійної моделі, яка враховує просторово-часову кореляцію відліків.

В даній роботі для реалізації алгоритму синтезу відеопослідовностей застосовується авторегресійна модель, яка конкретизована для випадку при якому для синтезу першого кадру ($k=1$) використовується модель порядку $(1,1)$, параметри якої мають зрозумілий фізичний смисл, а саме представляють собою коефіцієнти кореляції значень пікселів вздовж координати x та y . Такий підхід відповідає моделі із сепарабельною автокореляційною функцією виду $R(\Delta x, \Delta y) = \sigma_v^2 r_p^{|\Delta x|} r_c^{|\Delta y|}$, де r_p , r_c – коефіцієнти кореляції пікселів вздовж рядка і стовпця [33]. А при синтезі наступних кадрів ($k=2, \dots, K$) враховується також кореляція значень відповідних пікселів вздовж координати k . Отримані дані характеризуються сепарабельною автокореляційною функцією виду $R(\Delta x, \Delta y, \Delta k) = \sigma_v^2 r_p^{|\Delta x|} r_c^{|\Delta y|} r_k^{|\Delta k|}$, де r_k – коефіцієнт кореляції вздовж координати, що позначає номер кадру [34]. Розглянутий в роботі алгоритм дозволяє синтезувати кадри відеопослідовності, які характеризуються однаковістю своїх статистичних характеристик. В разі необхідності виконати симуляцію зміни властивостей сцени відеоряду (наприклад, моделюючи зміну сцени реєстрації, що отримується в деякій телевізійній системі), необхідно зада-

тися новими параметрами моделі і почати процес синтезу кадрів заново.

Загалом, процедура генерування відеопослідовності, передбачає виконання таких етапів: 1) задати коефіцієнти кореляції r_p , r_c , r_k . Визначити середнє квадратичне відхилення значень пікселів та середнє значення яскравості. Зазначені параметри можуть бути розраховані на основі аналізу деякої навчальної відеопослідовності; 2) задати розміри кадру, формат збереження зображення, кількість кадрів у відеопослідовності, та виконати синтез послідовності текстурних зображень; 3) на основі множини зображень сформулювати відеопослідовність та отримати відеофайл.

Для подальших викладок з метою компактного представлення записів, координату k , яка позначає номер кадру, представимо у вигляді нижнього індексу: $v_k(x,y)=v(x,y,k)$, $x=1,\dots,N$; $y=1,\dots,M$; $k=1,\dots,K$. Покладається що кадри відеопослідовності не є незалежними між собою. Перший кадр буде формуватися у вигляді одиночного зображення, при синтезі якого враховується кореляція вздовж координат x та y . При формуванні другого та наступних кадрів береться до уваги також коефіцієнт кореляції r_k .

Алгоритм № 1, який дозволяє синтезувати значення відліків першого кадру ($k=1$), передбачає виконання наступних кроків:

Крок 1: синтез першого пікселя в першому рядку зображення:

$$v_1(1,1) = \sigma_v \omega(1,1), \quad (1)$$

де σ_v – середнє квадратичне відхилення значень пікселів зображення; $\omega(1,1)$ – гаусівська некорельована випадкова величина із нульовим математичним сподіванням та дисперсією, що дорівнює 1.

Крок 2: синтез пікселів з 2-го по N -й, що розміщуються в першому рядку зображення ($x=2,\dots,N$):

$$v_1(x,1) = r_p v_1(x-1,1) + \sigma_v \sqrt{1-r_p^2} \omega(x,1). \quad (2)$$

Крок 3: синтез першого пікселя наступного рядка зображення ($y=2,\dots,M$):

$$v_1(1,y) = r_c v_1(1,y-1) + \sigma_v \sqrt{1-r_c^2} \omega(1,y). \quad (3)$$

Крок 4: синтез всіх наступних пікселів поточного рядка зображення ($x=2,\dots,N$):

$$v_1(x,y) = r_p v_1(x-1,y) + r_c v_1(x,y-1) - r_p r_c v_1(x-1,y-1) + \sigma_v \sqrt{1-r_p^2} \sqrt{1-r_c^2} \omega(x,y). \quad (4)$$

Блок-схема алгоритму приведена на рис. 2, де *randn* представляє собою генератор незалежних гаусовських випадкових величин із нульовим

математичним сподіванням та одиничною дисперсією (використовується в блок-схемі замість параметра $\omega(x,y)$ у виразах (1)–(4)).

Інший підхід реалізується за допомогою *Алгоритму № 2*, блок-схема якого приведена на рис. 3. Синтез значень пікселів зображення передбачає

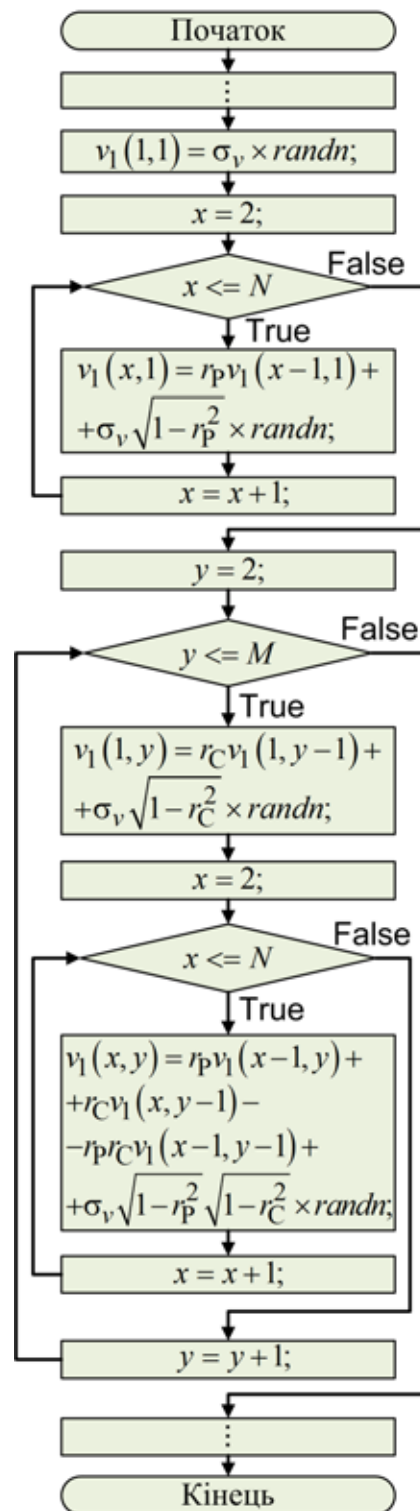


Рис. 2. Алгоритм №1

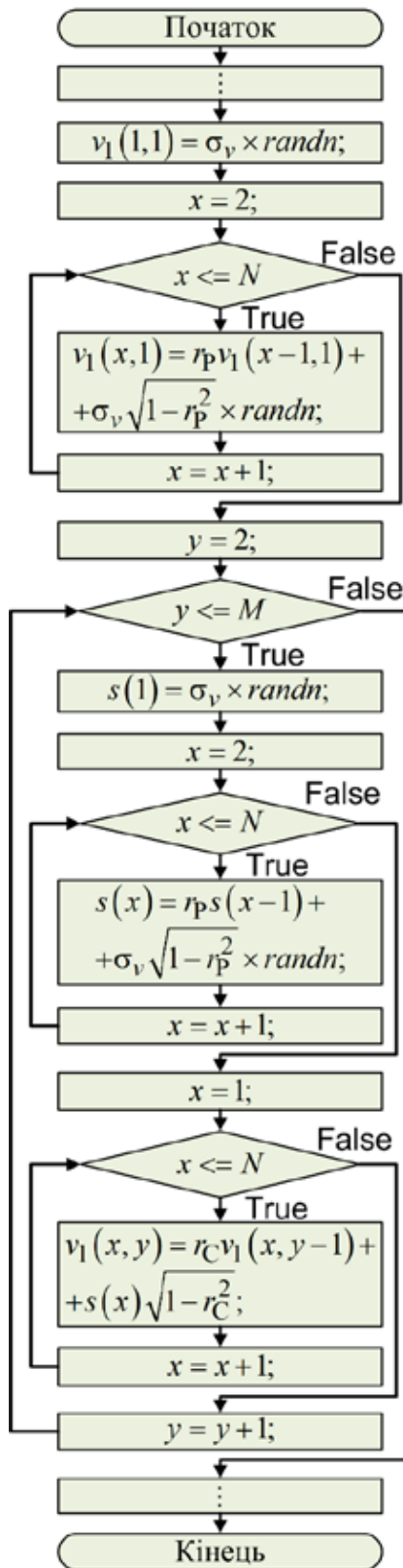


Рис. 3. Алгоритм № 2

залучення додаткового одновимірного масиву, який позначений як $s(x)$, $x = 1, \dots, N$. Одновимірний масив $s(x)$ виконує роль допоміжного рядка

зображення, значення елементів якого використовуються для синтезу пікселів, що розміщуються в рядках з 2-го по M -й в першому кадрі відеопослідовності. Значення елементів масиву $s(x)$ синтезуються кожного разу заново при генеруванні значень пікселів чергового рядка першого кадру, таким чином, елементи масиву $s(x)$ постійно оновлюються, причому процедура формування елементів масиву $s(x)$ відповідає синтезу пікселів першого рядка в першому кадрі відеопослідовності. Значення статистичних параметрів (коефіцієнт кореляції та середнє квадратичне відхилення), які використовуються у виразі обрахунку значень елементів масиву $s(x)$, повинні обиратися такими як і для формування значень пікселів першого рядка першого кадру відеопослідовності. Загалом, результати генерування значень пікселів зображення, які формуються за допомогою Алгоритму № 1 або Алгоритму № 2, при умові обрання однакових вхідних параметрів для кожного із алгоритмів, повинні володіти однаковими статистичними характеристиками. Нижче в статті проводиться дослідження швидкості роботи обох алгоритмів при їх застосуванні для синтезу множини однорідних текстурних зображень, на основі яких формується відеопослідовність.

Після синтезу першого зображення, що виступає першим кадром ($k=1$), процедура формування наступного кадру передбачає використання додаткового двовимірного масиву для зберігання додаткового зображення, яке синтезується на кожному кроці генерування кадрів з 2-го по K -й, застосовуючи для цього обраний раніше алгоритм – Алгоритм № 1 або Алгоритм № 2, використовуючи ті самі значення параметрів, які поклалися при генеруванні зображення першого кадру. Якщо позначити додаткове зображення як $q(x, y)$, $x=1, \dots, N$, $y=1, \dots, M$, тоді пікселі кадру $v_k(x, y)$, $k=2, \dots, K$ будуть формуватися таким чином:

$$v_k(x, y) = r_k v_{k-1}(x, y) + q(x, y) \sqrt{1 - r_k^2}. \quad (5)$$

Сама процедура синтезу кадрів відеопослідовності може бути представлена за допомогою блок-схеми алгоритму, яка приведена на рис. 4.

Дослідження алгоритмів проводилося в MATLAB. В ході синтезу множини зображень з метою їх використання для формування відеопослідовності не обов'язково застосовувати трьохвимірний масив $v_k(x, y)$, де відповідні виміри позначають номери стовпця та рядку зображення та номер кадру згідно із прийнятими позначеннями на рис. 1. Натомість при моделюванні було використано 4 двовимірних масиви: перший

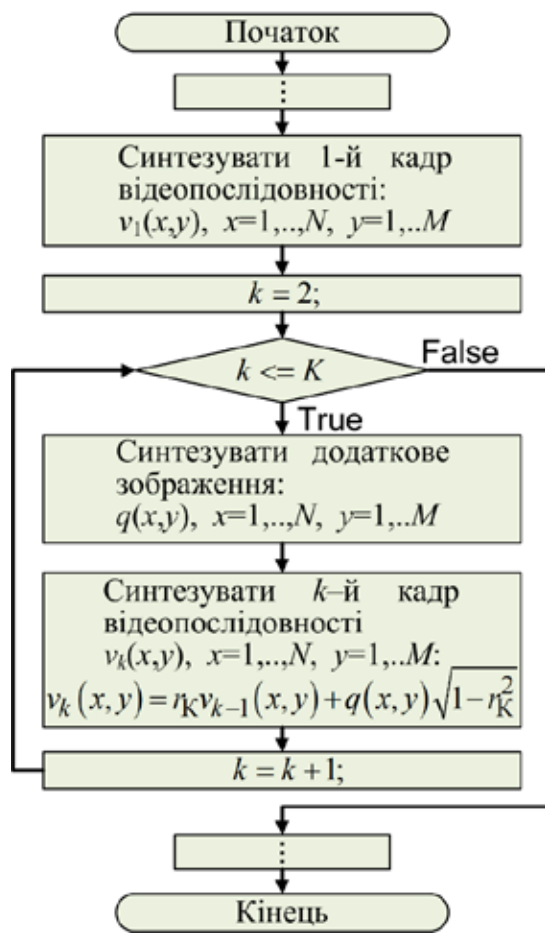


Рис. 4. Алгоритм синтезу кадрів відеопослідовності

масив використовувався для зберігання пікселів зображення, що покладалося поточним кадром (масив A); другий масив використовувався для зберігання зображення, що покладалося попереднім кадром послідовності (масив B); третій масив застосовувався для зберігання додаткового зображення (масив C), яке використовувалося для формування кадрів з 2-го по K -й. При синтезі значень елементів масивів A , B , C було обрано тип даних з плаваючою комою, а сформовані зображення A , B , C загалом представляли собою центровані гаусівські випадкові поля із заданим середньоквадратичним відхиленням σ_v . З метою формування зображень, що характеризуються відмінним від 0 середнім значенням яскравості пікселів, використовувалася константа $mean$, яка відігравала роль середньої яскравості зображення. Значення параметрів σ_v та $mean$ обирались таким чином, щоб після додавання $mean$ до значень елементів масиву A отримані значення потрапляли в діапазон $[0;1]$. Після цього за допомогою функції $im2uint8()$ значення пікселів перетворювалися до цілих значень із діапазону $[0;255]$ і зберігалися в двовимірний

масив D , який по суті представляв собою 8-бітне зображення, що записувалося на жорсткий диск ПК. При цьому значення елементів масиву A не змінювалися, а описана процедура реалізовувалася в MATLAB таким чином:

$$D = im2uint8(A + mean); \quad (6)$$

Процес моделювання в MATLAB відбувався так:

1) для першого кадру ($k=1$) за допомогою обраного алгоритму (Алгоритм № 1 або Алгоритм № 2) синтезувалося перше зображення і зберігалось в масив A . За допомогою виразу (6) формувалася масив D , який представляв собою 8-бітне зображення, що за допомогою функції $imwrite()$ зберігалось на жорсткий диск;

2) для кадрів $k=2, \dots, K$ синтез виконувався в циклі згідно з алгоритмом, який показано на рис. 4. На початку кожної ітерації за допомогою раніше обраного алгоритму (Алгоритм № 1 або Алгоритм № 2) виконувався синтез додаткового зображення C . Крім того на початку кожної ітерації значення елементів масиву A присвоювалися відповідним елементам масиву B (оновлюючи тим самим вміст масиву B , який призначений для зберігання попереднього кадру послідовності). За допомогою виразу аналогічного (5) виконувався розрахунок пікселів зображення поточного кадру (де $v_{k-1}(x,y)$ – це піксель попереднього кадру, тобто відповідний елемент масиву B ; $q(x,y)$ – це піксель додаткового зображення, тобто відповідний елемент масиву C ; $v_k(x,y)$ – піксель поточного кадру, тобто відповідний елемент масиву A). За допомогою виразу (6) формувалася масив D , і отримане зображення записувалося на диск. Після цього відбувався перехід до наступної ітерації.

На рис. 5 показано результати моделювання для різних параметрів моделі, розміри текстурних зображень 100×140 пікселів. На рис. 5.а приведені (зліва направо) автокореляційна функція текстурного зображення, 1-й кадр, 25-й кадр та 100-й кадр відеопослідовності, а також гістограма 1-го кадру; при синтезі послідовності зображень покладалася такі значення параметрів моделі: $\sigma_v = 0,15$, $r_p = 0,98$, $r_c = 0,98$, $r_k = 0,99$. На рис. 5.б показані аналогічні результати для таких параметрів моделі: $\sigma_v = 0,12$, $r_p = 0,9$, $r_c = 0,9$, $r_k = 0,95$. На рис. 5.в приведені результати для випадку: $\sigma_v = 0,1$, $r_p = 0,7$, $r_c = 0,7$, $r_k = 0,9$. Середнє значення пікселів для всіх приведених варіантів покладалося рівним 0,5, що при переведенні значень пікселів за допомогою функції $im2uint8()$ до цілого типу даних, відповідає середньому значенню яскравості 128.

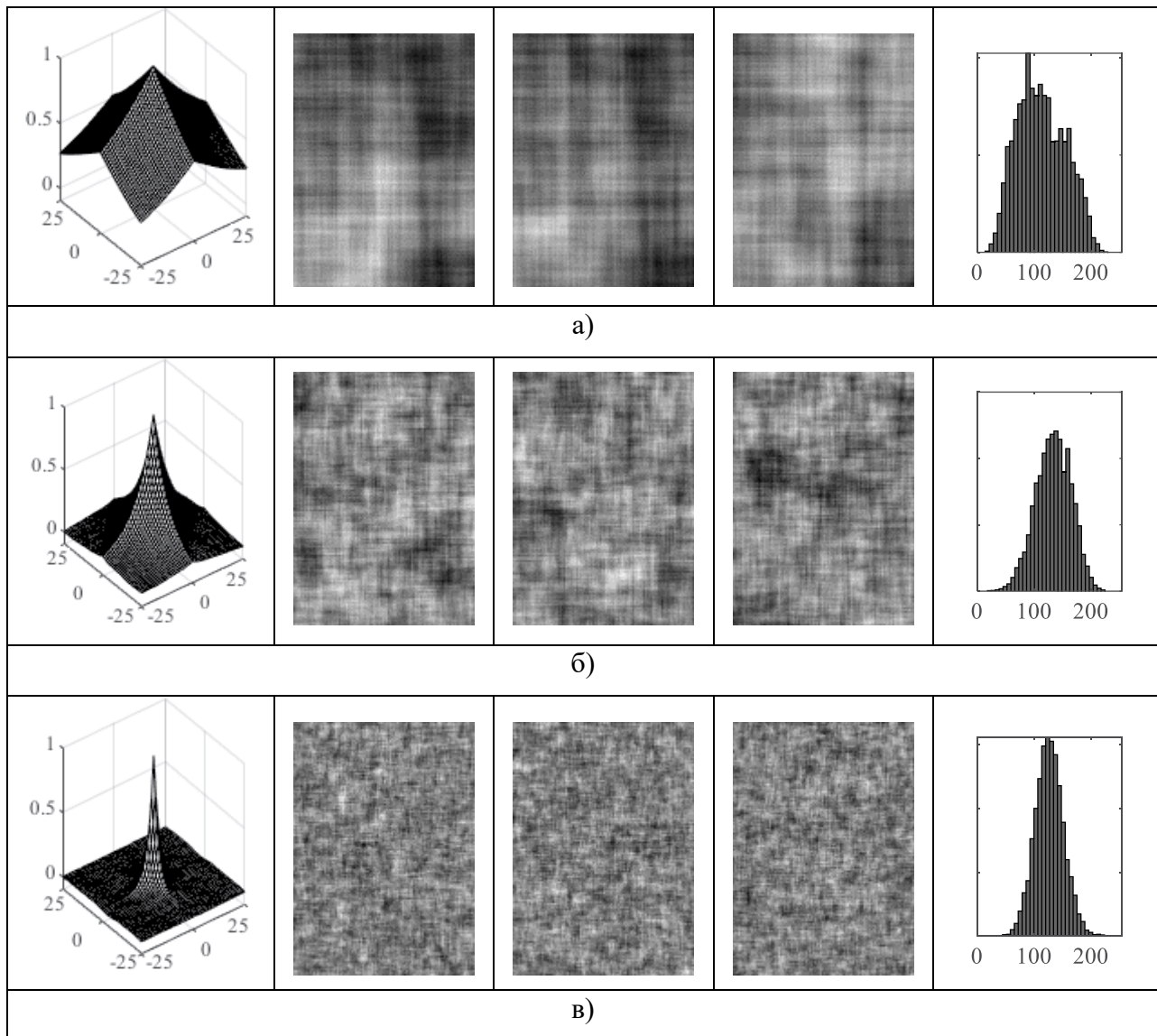


Рис. 5. Результати моделювання

В таблиці 1 приведені часові витрати синтезу тестових відеопослідовностей при використанні *Алгоритму № 1* та *Алгоритму № 2*, які застосовувалися для створення множини однорідних текстурних зображень, що в подальшому використовувалися для формування відеофайлів відповідного формату (розміри відеофайлів приведені в табл. 1). Покладалися такі параметри моделі: $\sigma_v = 0,12$, $mean = 0,5$, $r_p = 0,9$, $r_c = 0,9$, $r_k = 0,95$. Дослідження роботи алгоритмів проводилося в MATLAB на ПК із такими характеристиками: Intel i7-3770K 3.5 ГГц, 32 ГБ RAM, NVIDIA GeForce GTX 1660 Super, Win7. Розміри окремого зображення для кожного окремого експерименту обиралися: 720×576, 1024×768, 1920×1080. Для кожного формату зображення TIFF, JPEG, PNG, BMP формувалося 1500 зображень. Синтезована

послідовність зображень зберігалася на жорсткому диску. За допомогою функції *VideoWriter()* формувався відеофайл тривалістю 1 хвилину, кількість кадрів на секунду 25. При створенні відеофайлів обиралися такі параметри: 1) тип файлу 'MPEG-4', Quality=100, даний формат передбачає використання компресії, що дозволяє суттєво зменшити розмір відеоданих; 2) 'Grayscale AVI' для створення файлу відеопослідовності в градіях сірого, без компресії.

Загалом, результати моделювання демонструють незначні відмінності в швидкодії синтезу послідовності зображень за допомогою *Алгоритму № 1* та *Алгоритму № 2*. Відмінності в часі синтезу множини зображень при виборі різних графічних форматів може бути пояснена необхідністю проведення додаткових процедур комп-

Часові витрати синтезу відеопослідовностей

Розмір кадру: Стовпці × Рядки	Середня тривалість синтезу одного кадру (сек)		Тривалість синтезу всієї послідовності зображень (сек)		Сформований відеофайл. Розмір файлу / тривалість створення файлу (сек)	
	Алг.№ 1	Алг.№ 2	Алг.№ 1	Алг.№ 2	MPEG-4	Grayscale AVI
TIFF						
720×576	0,0351	0,0352	52,6345	52,7615	67,5 МБ	593 МБ
					18,1435	4,5941
1024×768	0,0731	0,0698	109,7112	104,7633	129 МБ	1,09 ГБ
					33,3483	5,6866
1920×1080	0,1838	0,1734	275,6510	260,1296	344 МБ	2,89 ГБ
					87,3245	14,2475
JPEG						
720×576	0,0405	0,0398	60,7552	59,6310	75,8 МБ	593 МБ
					19,7234	8,3854
1024×768	0,0821	0,0781	123,2212	117,1154	145 МБ	1,09 ГБ
					35,6519	14,0599
1920×1080	0,2113	0,2005	316,8998	300,7562	385 МБ	2,89 ГБ
					90,7549	35,4442
PNG						
720×576	0,0557	0,0539	83,5559	80,7756	67,5 МБ	593 МБ
					18,9263	9,6163
1024×768	0,1097	0,1048	164,6081	157,1832	129 МБ	1,09 ГБ
					34,5060	16,0793
1920×1080	0,2823	0,2727	423,5250	409,1239	344 МБ	2,89 ГБ
					86,1868	39,8819
BMP						
720×576	0,0337	0,0331	50,4858	49,6514	67,5 МБ	593 МБ
					19,0803	5,0061
1024×768	0,0704	0,0646	105,5438	96,9165	130 МБ	1,09 ГБ
					33,9088	5,7949
1920×1080	0,1735	0,1672	260,3223	250,7643	344 МБ	2,89 ГБ
					87,0921	13,1189

ресії даних. Формат BMP не передбачає проведення компресії, тому послідовність зображень в даному форматі синтезувалася найшвидше, однак при цьому розмір окремого зображення був найбільшим. Вибір відповідного графічного формату, розміру кадру відеопослідовності, кількості кадрів за секунду може бути продиктовано вимогами до процедури синтезу відеопослідовностей як при виконанні наукових досліджень, наприклад, в ході вивчення методів та алгоритмів обробки відеопослідовностей, так і для реалізації відповідних практичних задач.

Висновки. В роботі проведено дослідження алгоритму синтезу відеопослідовностей, який побудований на основі авторегресійної моделі для генерування гаусівського випадкового поля, що характеризується сепарабельною автокореляційною функцією. Приведено результати моделювання роботи алгоритму в програмному серед-

овищі MATLAB. Надано приклади синтезованих текстур, які отримані для відповідних вхідних параметрів моделі.

Враховуючи обраний порядок моделі, отримані зображення характеризуються анізотропною автокореляційною функцією, що проявляється візуально у вигляді певної сітчастої структури, яка стає більш помітною при наближенні значень коефіцієнтів кореляції r_p та r_c до 1. Таким чином, це обмежує можливості розглянутого алгоритму до синтезу переліку текстур, які можуть зустрічатися на реальних відеопослідовностях. Для подолання зазначеного недоліку може бути ускладнена процедура синтезу значень пікселів кадру шляхом переходу до авторегресійної моделі вищих порядків, однак при цьому також ускладнюється методика розрахунку параметрів моделі при аналізі навчальних даних та збільшуються часові витрати на синтез текстурних відеопослідовностей.

Перевагою представленого алгоритму є невисока складність обчислювальних процедур, які необхідно виконувати в ході синтезу зображень, а також можливість швидкого розрахунку (на основі аналізу тестових текстур) необхідних параметрів (середньоквадратичного відхи-

лення і середнього значення яскравості пікселів зображення, та коефіцієнтів кореляції відліків зображення вздовж кожної із координат), які використовуються в ході моделювання кадрів відеопослідовності, що забезпечує простоту практичної реалізації алгоритму.

Список літератури:

1. Szeliski R. Computer Vision: Algorithms and Applications. Springer, 2022. 925 p.
2. Кульчицька О. Ю., Січко Т.В. Цифрова обробка зображень та відео. *Прикладні інформаційні технології* : тези доп. всеукр.наук.-практ. конф. м. Вінниця, 29 квітня 2020 р. Вінниця, 2020. С. 110–111.
3. Tan L, Jiang J. Digital signal processing: Fundamentals and Applications. Academic Press Elsevier, 2013. 876 p.
4. Deshpande R.G., Ragha L.L. Performance Analysis of Various Video Compression Standards. *International Conference on Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC)* : proceedings. Jalgaon, 22–24 Dec. 2016. Jalgaon, 2016. P. 148–151.
5. Devi N., Singh V., T. Vinay. A Study of Different Video Compression Techniques and Latest Standards. *International Journal of Trend in Research and Development*. 2017. Vol. 4 (3). P. 312–313.
6. Pawlowski P., Piniarski, Dabrowski A. Selection and Tests of Lossless and Lossy Video Codecs for Advanced Driver-assistance Systems. *IEEE 2018 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)* : proceedings. Poznan, 19–21 Sept. 2018. Poznan, 2018. P. 344–349.
7. Lundstom L.-I. Understanding Digital Television. An Introduction to DVB Systems with Satellite, Cable, Broadband, and Terrestrial TV Distribution. Focal Press, 2006. 316 p.
8. Benoit H. Digital Television. Satellite, Cable, Terrestrial, IPTV, Mobile TV in the DVB Framework. Focal Press, 2008. 289 p.
9. Nixon M.S., Aguado A.S. Feature Extraction and Image Processing for Computer Vision. Elsevier Academic Press, 2019. 650 p.
10. Solomon C., Breckon T. Fundamentals of Digital Image Processing. A Practical Approach with Examples in MATLAB. Wiley-Blackwell, 2011. 352 p.
11. Ding D., Ram S., Rodriguez J.J. Image Inpainting Using Nonlocal Texture Matching and Nonlinear Filtering. *IEEE Transactions on Image Processing*. 2019. Vol. 28, No. 4. P. 1705–1719.
12. Lu S., Han C. An Improved Image Inpainting Method Based on Patch Matching and Texture Synthesis. *IEEE Chinese Control And Decision Conference (CCDC)* : proceedings. Nanchang, China, 03–05 Jun. 2019. Nanchang, 2019. P. 5933–5938.
13. Zhang L., Chang M. Image Inpainting for Object Removal Based on Adaptive Two-Round Search Strategy. *IEEE Access*. 2020. Vol. 8. P. 94357–94372.
14. Thakur U.S., Ray B. Image coding using parametric texture synthesis. *IEEE 18th International Workshop on Multimedia Signal Processing (MMSP)* : proceedings. Montreal, 21–23 Sept. 2016. Montreal, 2016.
15. Balle J., Stojanovic A., Ohm J.-R. Models for Static and Dynamic Texture Synthesis in Image and Video Compression. *IEEE Journal of Selected Topics in Signal Processing*. 2011. Vol. 5, No. 7. P. 1353–1365.
16. Yang K., Liu D., Chen Z., Wu F., Li W. Spatiotemporal Generative Adversarial Network-Based Dynamic Texture Synthesis for Surveillance Video Coding. *IEEE Transactions on Circuits and Systems for Video Technology*. 2021. Vol. 32, No. 1. P. 359–373.
17. Li. Guilin, Chen P., Guo S. Texture Blending for Photorealistic Composition on Mobile AR Platform. *15th International Symposium on Medical Information and Communication Technology (ISMICT)* : proceedings. Xiamen, 14–16 Apr. 2021. Xiamen, 2021. P. 92–97.
18. Vishnevyy S.V., Zhuk S.Ya. Two-Stage mutual causal filtration and segmentation of heterogeneous images. *Radioelectronics and Communications Systems*. 2011. Vol. 54, No. 10. P. 554–565.
19. Gonzalez R.C., Woods R.E. Digital Image Processing. Pearson. 2018. 1022 p.
20. Ramya S., Shibu N.V., Malathi P. A Survey on Texture Synthesis and Its Approaches. *International Conference on Computing Technologies and Intelligent Data Engineering (ICCTIDE'16)* : proceedings. Kovilpatti, 07–09 Jan. 2016. Kovilpatti, 2016. P. 1–4.
21. Huang Z., Lin X., Chen C. Fast Texture Synthesis for Discrete Example-Based Elements. *IEEE Access*. 2020. Vol. 8, P. 76683–76691.
22. Luo G., Zhu Y., Guo B. Fast MRF-Based Hole Filling for View Synthesis. *IEEE Signal Processing Letters*. 2018. Vol. 25, No. 1. P. 75–79.
23. Haindl M., Havlicek V. BTF Compound Texture Model with Fast Iterative Non-parametric Control Field Synthesis. *14th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)* : proceedings. Las Palmas de Gran Canaria, 26–29 Nov. 2018. Las Palmas de Gran Canaria, 2018. P. 98–105.

24. Wang L., Chen W., Yang W., Bi F., Yu F.R. A State-of-the-Art Review on Image Synthesis With Generative Adversarial Networks. *IEEE Access*. 2020. Vol. 8. P. 63514–63537.
25. Fadaeddini A., Majidi B., Eshghi M. A Case Study of Generative Adversarial Networks for Procedural Synthesis of Original Textures in Video Games. *IEEE 2nd National and 1st International Digital Games Research Conference: Trends, Technologies, and Applications (DGRC)* : proceedings. Tehran, 29–30 Nov. 2018. Tehran, 2018. P. 118–122.
26. Wang Z., Li M., Xia G. Conditional Generative ConvNets for Exemplar-Based Texture Synthesis. *IEEE Transactions on Image Processing*. 2021. Vol. 30. P. 2461–2475.
27. Zhang L., Liu Y. Image Generation Based on Texture Guided VAE-AGAN for Regions of Interest Detection in Remote Sensing Images. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* : proceedings. Toronto, 06–11 Jun. 2021. Toronto, 2021. P. 2310–2314.
28. Maarouf A., Hachouf F., Kharfouchi S. Exemplar-based Texture Synthesis Using Two Random Coefficients Autoregressive Models. *Image Analysis and Stereology*. 2023. Vol. 42., No. 1. P. 37–49.
29. Vaishali D., Ramesh R., Anita Christaline J. 2 D autoregressive model for texture analysis and synthesis. *International Conference on Communication and Signal Processing* : proceedings. Melmaruvathur, 03–05 Apr. 2014. Melmaruvathur, 2014. P. 1135–1139.
30. Zaporozhets D.A., Vyshnevyi S.V. Filtration of image sequences distorted by noise using two-stage processing technique. *Radioengineering Problems, Signals, Devices and Systems* : proceedings of the XII-th international scientific and technical conference, Kyiv, 13–15 Dec. 2023. Kyiv, 2023. P. 8–10.
31. Khandelia A., Gorecha S., Lall, B., Chaudhury S., Manthur M. Parametric Video Compression Scheme using AR based Texture Synthesis. *Sixth Indian Conference on Computer Vision, Graphics and Image Processing* : proceedings. Bhubaneswar, 16–19 Dec. 2008. Bhubaneswar, 2008. P. 219–225.
32. Racape F., Doshkov D., Koppel M., Ndjiki-Nya P. 2D+T Autoregressive Framework for Video Texture Completion. *IEEE International Conference on Image Processing (ICIP)* : proceedings. Paris, 27–30 Oct. 2014. Paris, 2014. P. 4657–4661.
33. Dolloff J., Doucette P. The Sequential Generation of Gaussian Random Fields for Applications in the Geospatial Sciences. *ISPRS International Journal of Geo-Information*. 2014. Vol. 3(2). P. 817–852.
34. Xiao Te, Zhang L., Li D., Cao Z. Efficient 3-D Random Field Simulation with Separable Correlation Functions. *7th International Symposium on Geotechnical Safety and Risk (ISGSR)* : proceedings. Taipei, 11–13 Dec. 2019. Taipei, 2019. P. 358–363.

Vyshnevyi S.V., Zhurba A.V. SYNTHESIS OF VIDEO SEQUENCES CONTAINING A HOMOGENEOUS TEXTURE WITH A SEPARABLE AUTOCORRELATION FUNCTION

The article addresses the problem of synthesizing video sequences in which all frames contain a homogeneous texture with identical statistical characteristics. A brief review of practical applications of texture image generation methods in the context of video data processing is presented. The study focuses on a video sequence synthesis algorithm based on the use of an autoregressive model of a Gaussian random field with a separable autocorrelation function that accounts for the spatial-temporal correlation of samples. The article provides analytical expressions and block diagrams of algorithms that can be used to implement the proposed video sequence generation approach using an appropriate programming language. The application of the presented texture sequence synthesis algorithm requires prior estimation of the autoregressive model parameters through analysis of training data containing homogeneous texture. This involves estimating the correlation coefficients along each coordinate, as well as calculating the mean brightness and the standard deviation of the pixel values of the texture. The algorithm was modeled in the MATLAB software environment. The article presents samples of synthesized textures, time performance indicators for video file generation across different frame formats, autocorrelation function plots, and histograms of the resulting images. It was found that the frames of the synthesized video sequences are characterized by an anisotropic autocorrelation function, which in some cases may be considered a drawback. To address this issue, it is necessary to increase the complexity of the texture synthesis procedure by applying a higher-order model. The results of this study can be used in scientific research and practical applications where the synthesis of video sequences containing texture is required.

Key words: video sequence, texture, autoregressive model, Gaussian random field, algorithm, digital image.

Дата надходження статті: 29.07.2025

Дата прийняття статті: 28.08.2025

Опубліковано: 27.10.2025

Волк М.О.

Харківський національний університет радіоелектроніки

Бугрій А.М.

Харківський національний університет радіоелектроніки

Ковтун Є.І.

Харківський національний університет радіоелектроніки

Самойлов І.А.

Харківський національний університет радіоелектроніки

Брестовицький Р.М.

Харківський національний університет радіоелектроніки

Соробей Б.В.

Харківський національний університет радіоелектроніки

МЕТОДИ РОЗПОДІЛЕННЯ ОБЧИСЛЮВАЛЬНИХ ЗАДАЧ У ГЕТЕРОГЕННИХ КОМП'ЮТЕРНИХ СИСТЕМАХ

Гетерогенні обчислювальні системи, які поєднують у собі обчислювальні елементи різної архітектури – центральні процесори (CPU), графічні процесори (GPU), матриці FPGA, системи на кристалі (SoC), спеціалізовані інтегральні схеми (ASIC), а також периферійні edge пристрої – стають основою сучасної інфраструктури розподілених обчислень. Така інфраструктура потребує нових підходів до ефективного розподілу обчислювальних задач між обчислювальними елементами з різними обчислювальними характеристиками, енергоспоживанням, архітектурною сумісністю та пропускну здатністю мережевого з'єднання.

У статті здійснено ґрунтовний огляд сучасних методів розподілення обчислювальних задач у гетерогенних комп'ютерних системах, що поєднують ресурси з різними архітектурами, продуктивністю та енергетичними характеристиками. Зважаючи на швидке зростання обсягів даних, різноманітність обчислювальних завдань і необхідність їх паралельного виконання, задача ефективного планування в подібних середовищах набуває критичного значення. Розглянуто класичні алгоритми, евристичні та метаевристичні підходи, включаючи HEFT, Min-Min, генетичні алгоритми (GA), рої частинок (PSO), мурашині алгоритми (ACO), а також сучасні методи глибокого навчання, зокрема методи з підкріпленням (Deep RL), що демонструють значну адаптивність до змін середовища.

Особлива увага приділяється energy-aware методам, які дозволяють досягати балансу між продуктивністю та енергоспоживанням, що є особливо актуальним для IoT-пристроїв, мобільних обчислень і edge-середовищ. У статті запропоновано класифікацію задач за архітектурними (тип зв'язності, залежності) та прикладними (реальний час, HPC, AI) ознаками, що дозволяє встановити доцільність застосування конкретних моделей планування. Наведені у статті порівняльні таблиці методів дозволяють фахівцям обрати найдоцільніший підхід до планування. Отримані результати можуть бути корисними при побудові систем автономного управління завданнями у хмарних, fog- або edge-середовищах.

Ключові слова: гетерогенні обчислювальні системи, розподілення задач, машинне навчання, методи розподілу ресурсів, хмарні системи, енергоефективність, евристичні методи.

Постановка проблеми. З ростом обсягу даних, розвитку штучного інтелекту та переходу до edge-обчислень, з'явилася потреба обробляти величезні обсяги інформації в розподіленому, але енергоефективному середовищі. Гетерогенні комп'ютерні системи дозволяють масштабувати обчислення та оптимізувати продуктивність за рахунок використання спеціалізованих пристроїв. Наприклад: GPU є ефективними для паралельних задач (графіка, матричні обчислення, нейронні мережі); FPGA забезпечують низьку затримку й гнучке налаштування; CPU обробляє послідовні, менш структуровані задачі; ASIC/TPU – вузькоспеціалізовані та високоефективні; edge-пристрої – критичні для low-latency задач у smart city (розумний будинок) та IoT (Інтернет речей).

Проблема полягає у виборі ефективної політики розподілу задач, яка враховує множину факторів: час виконання задачі на конкретному пристрої; архітектурна сумісність і вартість переносу даних; енергоспоживання; QoS-вимоги (затримка, доступність); динамічність навантаження.

Задача розподілу є NP-складною [1], а отже вимагає застосування наближених, адаптивних або навчальних методів. Зокрема, у середовищах, де задачі можуть динамічно надходити, а стан ресурсів – змінюватися в реальному часі, – необхідні онлайн-алгоритми з можливістю самонавчання.

Крім того, у контексті глобального тренду на декарбонізацію IT-сектору, енергоефективність розподілу обчислень також стає критично важливою метою.

Аналіз останніх досліджень і публікацій. У літературі представлено широкий спектр підходів до задачі планування (scheduling) в гетерогенних середовищах. Їх умовно можна поділити на кілька напрямів.

Серед класичних напрямків одним із базових алгоритмів є HEFT (Heterogeneous Earliest Finish Time) [1], який забезпечує ефективне планування DAG-завдань з урахуванням часу передачі та обчислення. HEFT продовжує залишатись популярним через простоту й добру якість рішень у статичних системах.

Гарну рекомендацію серед класичних напрямків мають метаевристичні алгоритми. Genetic Algorithm (GA) дозволяє шукати розподіл, мінімізуючи глобальну цільову функцію (makespan, енергія, QoS). Ant Colony Optimization (ACO) має користь при великій кількості зв'язаних задач. Particle Swarm Optimization (PSO) є методом,

адаптованим для багатовимірних оптимізацій із динамічним середовищем. Сьогодні дослідники використовують комбінації відомих метаевристичних методів для підвищення ефективності. Наприклад, автори статті [2] застосовували гібрид GA + PSO для гетерогенних середовищ із динамічним надходженням (arrival rate) задач. Результати показали покращення на 12–18% щодо стандартного HEFT. Гібридний метод на основі використання GA та методу випадкового лісу (Random forest) запропоновано у роботі [3]. Проведений аналіз показав, що комбінований метод показує зменшення часу виконання на 10–12%.

Останні роки активно розвиваються моделі та методи на базі машинного навчання. До найвідоміших методів відносяться наступні: Reinforcement Learning (RL) – дозволяє приймати рішення на основі стану середовища; Imitation Learning дозволяє агенту, який приймає рішення, навчатись від експертної системи (наприклад, Oracle scheduler); Graph Neural Networks (GNN) – дозволяє представляти DAG та відносини між задачами у вигляді графа. Останні публікації у цьому напрямку також використовують гібридні підходи. Наприклад, для QoS-aware сценаріїв в [4] запропоновано метод HRLHS (Hybrid Reinforcement Learning-based Heuristic Scheduling), заснований на комбінації методів RL та PSO.

Ще один з сучасних напрямків досліджень стосується енергозатратності методів управління ресурсами. Так в [5] запропоновано метод DVFS (Dynamic Voltage and Frequency Scaling) з інтегрованим плануванням, що дозволяє зменшити споживання енергії при незначному падінні продуктивності. Важливо також планувати не лише самі задачі, але й консолідувати ресурси. Проблемам дослідження енергоспоживання в системах розподілених ресурсів присвячено роботи [6, 7].

Стандартні методи багатоцільової оптимізації також застосовуються до розподілу обчислювальних ресурсів. Основними критеріями виступають, як правило, мінімізація часу (makespan), енергоспоживання, забезпечення QoS (latency, availability), балансування навантаження. Прикладами виступають методи NSGA-II або Pareto front [8].

У 2025 році у роботі [9] було представлено найбільш повний систематичний огляд моделей розподілу задач у гетерогенних чипах, класифікуючи їх за критеріями централізації, часового horizon (offline / online), використання машинного навчання, оптимізаційних цілей (QoS, енергія, latency).

Постановка завдання. Метою цієї статті є систематизація та узагальнення існуючих моделей і методів розподілення обчислювальних задач у гетерогенних комп'ютерних системах, з урахуванням їх ефективності, сфери застосування та обмежень. Основна увага приділяється аналізу підходів до планування задач у середовищах із неоднорідними ресурсами, зокрема в контексті задач реального часу, обчислень у хмарі, високопродуктивних обчислень та енергоефективних систем. Стаття також має на меті визначити перспективні напрямки досліджень і сформулювати рекомендації щодо вибору відповідних методів для різних класів обчислювальних задач.

Виклад основного матеріалу. У сфері гетерогенних обчислювальних систем існує широкий спектр підходів до розподілення задач, які розрізняються за своєю складністю, адаптивністю та ефективністю. Одними з найпоширеніших є евристичні алгоритми, зокрема HEFT та Min-Min. Метод HEFT ґрунтується на побудові розкладу задач, орієнтуючись на оцінку найранішого можливого часу завершення завдання на кожному обчислювальному ресурсі. Цей підхід добре працює у випадках, коли задачі мають залежності, представлені у вигляді орієнтованих ациклічних графів (DAG). Алгоритм демонструє високу продуктивність за помірної обчислювальної складності, що робить його зручним для застосування в реальному часі та наукових симуляціях. Метод Min-Min, у свою чергу, фокусується на першочерговому виконанні задач з найменшим прогнозованим часом виконання, що дозволяє досягти високої швидкодії, хоча й часто призводить до менш ефективного завантаження системи.

Значного поширення також набули метаевристичні методи, серед яких провідне місце займають генетичні алгоритми (GA), мурашині алгоритми (ACO) та метод рою частинок (PSO). Генетичні алгоритми базуються на принципах природного добору і використовують операції схрещування, мутації та селекції для пошуку оптимальних або близьких до оптимальних рішень. Їхня гнучкість дозволяє ефективно враховувати різноманітні обмеження задачі, проте вони можуть вимагати значних обчислювальних ресурсів для досягнення збіжності. Мурашині алгоритми моделюють процеси колективного навчання колоній комах і демонструють гарні результати при вирішенні задач з багатьма альтернативними маршрутами, що актуально для планування з урахуванням залежностей. PSO, натхненний поведінкою зграї птахів, швидко сходиться до рішень за умов низь-

кої складності проблеми, проте вимагає ретельного налаштування параметрів і може застрягати в локальних мінімумах.

У сучасних дослідженнях дедалі частіше використовуються методи машинного навчання, особливо навчання з підкріпленням (reinforcement learning, RL), зокрема його глибокі варіанти (deep RL). Однією з відомих реалізацій такого підходу є система Decima, яка дозволяє здійснювати динамічне планування DAG-завдань у розподілених середовищах шляхом навчання агента на основі досвіду з попередніх розкладів. Перевагою цих методів є здатність до самонавчання та адаптації до змін у системі, включаючи змінну доступність ресурсів або флуктуації у навантаженні. Проте застосування глибоких нейронних мереж потребує значної кількості даних для тренування та обчислювальних ресурсів, що обмежує їх використання в деяких практичних сценаріях.

Окремий напрямок досліджень становлять energy-aware методи, які враховують енергоспоживання при розподілі задач. Такі підходи часто реалізуються через застосування технологій динамічного масштабування напруги та частоти (DVFS), що дозволяє знижувати енергоспоживання процесорів без істотної втрати продуктивності. Наприклад, під час виконання задач, не критичних до затримки, система може зменшити частоту процесора, що істотно знижує загальне споживання енергії. Енергетично-ефективне планування є особливо важливим у контексті IoT, edge computing та мобільних пристроїв, де обмеження акумуляторів чи теплові ліміти суттєво впливають на працездатність системи.

У багатьох сучасних розробках спостерігається тенденція до інтеграції гібридних підходів. Зокрема, поєднання RL з евристичними методами, як-от HEFT, дозволяє досягти як високої адаптивності, так і гарантованого початкового рівня продуктивності. Такі гібридні рішення можуть динамічно змінювати свою стратегію в залежності від змін в середовищі виконання або типу задачі. У перспективі саме гібридні моделі, що поєднують інтелектуальне планування, енергетичні обмеження та евристичну стабільність, можуть стати основою для повністю автономних систем розподілу задач.

Проведемо аналіз застосування моделей та методів для різних класів задач. Застосування моделей і методів розподілення задач у гетерогенних системах значною мірою залежить від природи самих задач, архітектури обчислювального середовища та пріоритетів розробників –

зокрема, чи важливіше досягти мінімального часу виконання, забезпечити енергоефективність, гарантувати дотримання дедлайнів чи адаптивно реагувати на змінне навантаження. Залежно від цих параметрів формуються характерні переваги та недоліки різних груп методів, що відображено в типових сценаріях їхнього використання.

Для задач без залежностей між компонентами – так званих незалежних задач – найчастіше застосовуються прості евристичні алгоритми, такі як HEFT чи Min-Min. Їхньою перевагою є швидкість розрахунку та достатньо висока ефективність при статичному розподілі. Проте при збільшенні кількості задач ці методи втрачають гнучкість, оскільки не враховують динамічні зміни у стані системи. У таких випадках доповнення класичних методів навчальними підходами або гібридними схемами може дати кращі результати.

Для задач, що мають структуру DAG (орієнтований ациклічний граф), характерна складніша топологія виконання, тому ефективне планування вимагає врахування як порядку виконання, так і обчислювальних витрат. Саме тут алгоритм HEFT виявляється найбільш популярним, оскільки він добре справляється з такими обмеженнями. Водночас зростає інтерес до застосування навчання з підкріпленням, зокрема у випадках, коли задача виконується в динамічному середовищі або ресурсна доступність змінюється в реальному часі.

Для систем реального часу ключовим є дотримання часових обмежень, тому найкраще зарекомендували себе методи з високою передбачуваністю поведінки. Тут RL демонструє цікаві результати завдяки своїй здатності навчатися оптимальним реакціям на зміну навантаження, проте він потребує попереднього тренування. Евристичні підходи, такі як EDF (earliest deadline first), у поєднанні з HEFT також залишаються конкурентоспроможними за рахунок своєї простоти.

У контексті енергообмежених середовищ, таких як IoT або edge-пристрої, пріоритет надається методам, орієнтованим на енергоспоживання. Energy-aware підходи, включаючи DVFS та планування із врахуванням електроспоживання, значно переважають інші методи. Також певне поширення мають варіанти RL, які включають енергетичні фактори до функції винагороди, однак їх використання вимагає тонкого налаштування та профілювання пристроїв.

Щодо задач машинного навчання та штучного інтелекту, які часто характеризуються складною структурою, високим навантаженням та змінними

параметрами, найбільш ефективними виявилися моделі глибокого навчання, зокрема Deep RL. Вони дозволяють знаходити ефективні стратегії планування навіть у випадках, коли залежності між підзадачами є складними або змінюються в процесі виконання. Однак ці методи також мають високу вартість у плані обчислень і часу тренування.

Нарешті, в задачах високопродуктивних обчислень (HPC), де зазвичай використовуються великі обсяги даних і обчислювальні навантаження, важливо забезпечити максимальне завантаження всіх доступних ресурсів. Тут переважають класичні HEFT-подібні підходи, які демонструють хорошу продуктивність у стабільних середовищах. Проте з ростом складності завдань поступово зростає роль метавристичних алгоритмів і навіть навчальних моделей, які забезпечують адаптивність і масштабованість.

Результати дослідження областей застосування різних підходів відображено у таблиці 1.

Наприклад, у системах автономного водіння inference-моделі розподіляються між on-board GPU та edge сервером: задачі з дедлайном <20 мс залишаються локально, решта – вивантажуються.

У таблиці 2 представлено приблизний розподіл поширеності застосування методів планування у гетерогенних системах для різних класів задач з урахуванням гібридних підходів, тобто коли для даного типу задач використовується комбінація з методів, які відносяться до різних груп.

На основі аналізу публікацій оцінено частоту використання основних підходів до планування задач у гетерогенних системах (рис. 1).

З огляду на стрімкий розвиток гетерогенних обчислювальних систем, зростаючі вимоги до адаптивності, енергоефективності та інтелектуального управління ресурсами, подальші дослідження у цій сфері повинні орієнтуватися не лише на покращення існуючих алгоритмів, але й на формування нових міждисциплінарних рішень. Серед ключових напрямків, що мають потенціал значного впливу, варто передусім виділити поглиблену інтеграцію методів глибокого навчання у процес планування обчислень. Особливо перспективними виглядають моделі навчання з підкріпленням у поєднанні з контекстною інформацією про задачі та системні обмеження, що дозволяє формувати розподіли з врахуванням реальних умов виконання.

Однією з важливих тенденцій майбутнього є створення самонавчальних систем планування, які здатні не лише адаптуватися до змін у струк-

Таблиця 1

Розподіл методів планування за типами задач

Тип задачі	Характеристика	Оптимальні методи
Незалежні задачі (bag-of-tasks)	Немає залежностей між задачами	Min-Min, GA, RL
DAG-завдання (workflow)	Завдання із залежностями	HEFT, ACO, GNN+RL
Реального часу (real-time)	Час виконання обмежений дедлайном	EDF + RL, A3C
Edge/IoT задачі	Обмеження по енергії, латентності, зв'язку	Energy-aware PSO, DVFS, кластеризація задач
ML/AI inference	GPU-heavy, багатопоточні	Static + profile-based mapping, reinforcement scheduling
High-performance computing (HPC)	Великі симуляції, CPU-GPU співпраця	Graph partitioning, Task co-location

Таблиця 2

Розподіл методів планування за типами задач

Тип задачі	HEFT/Min-Min	GA/PSO/ACO	Deep RL	Energy-aware
Незалежні задачі	70%	50%	20%	10%
DAG-завдання	80%	40%	30%	15%
Реального часу	30%	20%	50%	10%
IoT/Енергообмежені	10%	25%	30%	80%
ML/AI	20%	10%	70%	40%
HPC	60%	30%	40%	20%

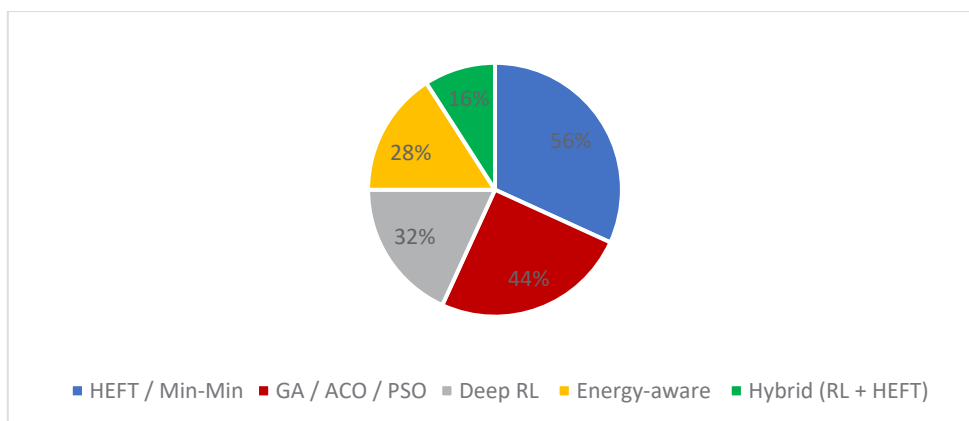


Рис. 1. Використання методів у публікаціях (відсоток від загальної кількості)

турі задач, але й прогнозувати поведінку ресурсів на основі історичних даних. Зокрема, розвиток моделей із короткотерміною пам'яттю (LSTM) або трансформерів для передбачення часу виконання та енергоспоживання задач відкриває нові горизонти для динамічного планування в режимі онлайн. Такі системи потенційно здатні працювати ефективно навіть в умовах відмов або непередбачених змін у мережевій структурі.

Окрему увагу заслуговує розробка гібридних моделей, які поєднують класичні евристичні з сучасними навчальними підходами. Наприклад, використання HEFT чи Min-Min на першому етапі

для грубого попереднього планування може доповнюватися більш тонким коригуванням через RL-агента або енергетичний модулятор. Такий багаторівневий підхід дозволяє досягти балансу між швидкістю та адаптивністю системи.

Перспективним вбачається і впровадження онтологій ресурсів та семантичного опису задач. Ще один напрямок, який варто досліджувати глибше – це федеративне навчання у плануванні, що дозволяє уникнути централізації при збиранні даних. Не менш актуальним є питання надійності й безпеки при розподіленні задач у системах із низьким рівнем довіри до окремих вузлів (напри-

клад, у відкритих edge-мережах або volunteer computing).

Підсумовуючи, можна стверджувати, що подальший розвиток у цій галузі буде все більше орієнтований на адаптивність, контекстну обізнаність, енергоефективність і безпеку. Це вимагає міждисциплінарного підходу, що об'єднує теорію розподілених систем, штучний інтелект, семантичне моделювання та прикладні задачі з реального світу.

Висновки. У статті було проведено системний огляд сучасних моделей та методів розподілення обчислювальних задач у гетерогенних комп'ютерних системах. З'ясовано, що з розвитком гібридних систем, включаючи CPU, GPU, FPGA, ASIC, edge вузли, значно ускладнюється задача ефективного планування обчислень.

Розглянуті підходи демонструють наступне. Класичні евристики (HEFT, Min-Min) залишаються важливими базовими методами, особливо в статичних задачах із відомими графами

виконання. Metaheuristic-алгоритми (GA, PSO, ACO) забезпечують високу адаптивність, особливо в задачах з кількома обмеженнями або метриками. Методи глибокого навчання (RL, GNN-RL) відкривають нові можливості для гнучкого онлайн-розподілу задач в умовах високої варіативності й динаміки. Edge/IoT-сценарії вимагають врахування нових параметрів – енергоспоживання, латентності, надійності зв'язку.

Зазначено, що вибір конкретної моделі повинен враховувати тип задачі (незалежна, реального часу), архітектурну топологію системи, доступні метрики, а також можливість часткового або повного навчання на історичних даних.

До перспективних напрямків досліджень належать: гібридизація RL та евристичних scheduler-ів, врахування семантичного профілю ресурсів, безпечне планування у відкритих гетерогенних середовищах, побудова стандартних benchmark-середовищ для відтворення результатів.

Список літератури:

1. Topcuoglu H., Hariri S., Wu M. Performance-effective and low-complexity task scheduling for heterogeneous computing. *IEEE Transactions on Parallel and Distributed Systems*. 2002. № 13(3). P. 260–274. <https://doi.org/10.1109/71.993206>
2. Wang Y., Li K., Xu Y. Hybrid Genetic and Particle Swarm Algorithm for Energy-aware Task Scheduling. *Future Generation Computer Systems*. 2022. № 108. P. 480–492. <https://doi.org/10.1016/j.future.2020.03.003>
3. Волк М.О., Курочкін В. С., Запороженко А.П., Паронікян П.А. Гібридний метод розподілу ресурсів в хмарних системах. *Системи управління, навігації та зв'язку*. 2024. № 2(76). С. 70–83. doi:10.26906/SUNZ.2024.2.070
4. Shao Q., Zhang Y., Wang H. HRLHS: Hybrid Reinforcement Learning-based Heuristic Scheduling in Edge Computing. *Scientific Reports*. 2024. № 14(1), 8461. URL: <https://www.nature.com/articles/s41598-025-94068-0> (дата звернення: 08.07.2025)
5. Kazemi F., Barzegar B., Motameni H., Yadollahzadeh-Tabari M. An energy-aware scheduling in DVFS-enabled heterogeneous edge computing environments. *The Journal of Supercomputing*. 2025. № 8, 1078. DOI: 10.1007/s11227-025-07432-2
6. Mamchych O., Volk M. A. A unified model and method for forecasting energy consumption in distributed computing systems based on stationary and mobile devices. *Radioelectronic and Computer Systems*. 2024. № 2. P. 120–135. DOI: <https://doi.org/10.32620/reks.2024.2.10>.
7. Mamchych O., Volk M. Smartphone Based Computing Cloud and Energy Efficiency. *12th International Conference on Dependable Systems, Services and Technologies (DESSERT)*. Athens. Greece. 2022. P. 1–5. DOI: 10.1109/DESSERT58054.2022.10018740
8. Zhou M., Li J., Chen Y. Multi-objective scheduling using NSGA-II in heterogeneous IoT systems. *Expert Systems with Applications*. 2024. № 239, 120184. <https://doi.org/10.1016/j.eswa.2024.120184>
9. Miao C., Ren Z., Qian Z. Scheduling in Heterogeneous Computing Systems: A Survey. *Electronics*. 2025. № 14(6), 1191. <https://doi.org/10.3390/electronics14061191>

Volk M.O., Buhrii A.M., Kovtun Ye.I., Samoilov I.A., Brestovytskyi R.M., Sorobey B.V. METHODS OF DISTRIBUTION OF COMPUTATIONAL TASKS IN HETEROGENEOUS COMPUTER SYSTEMS

Heterogeneous computing systems that combine computing elements of different architectures – central processing units (CPUs), graphics processing units (GPUs), FPGAs, systems on a chip (SoCs), application-specific integrated circuits (ASICs), and peripheral edge devices – are becoming the basis of modern distributed computing infrastructure. Such infrastructure requires new approaches to effectively distribute computing

tasks between computing elements with different computing characteristics, power consumption, architectural compatibility, and network connection bandwidth.

The article provides a thorough review of modern models and methods for distributing computing tasks in heterogeneous computer systems that combine resources with different architectures, performance, and power characteristics. Given the rapid growth of data volumes, the variety of computing tasks, and the need for their parallel execution, the problem of effective scheduling in such environments is becoming critical. Classical algorithms, heuristic and metaheuristic approaches, including HEFT, Min-Min, genetic algorithms (GA), particle swarm optimization (PSO), ant algorithms (ACO), as well as modern deep learning methods, in particular methods with reinforcement (Deep RL), which demonstrate significant adaptability to environmental changes, are considered.

Special attention is paid to energy-aware methods that allow achieving a balance between performance and energy consumption, which is especially relevant for IoT devices, mobile computing and edge environments. The article proposes a classification of tasks by architectural (type of connectivity, dependencies) and application (real-time, HPC, AI) features, which allows establishing the feasibility of using specific planning models. The comparative tables of methods presented in the article allow specialists to choose the most appropriate approach to planning. The results obtained can be useful when building autonomous task management systems in cloud, fog or edge environments.

Key words: heterogeneous computing systems, task distribution, machine learning, resource allocation methods, cloud systems, energy efficiency, heuristic methods.

Дата надходження статті: 05.07.2025

Дата прийняття статті: 11.07.2025

Опубліковано: 27.10.2025

Гончаров Д.С.

Чорноморський національний університет імені Петра Могили

ІНТЕЛЕКТУАЛЬНІ МОДЕЛІ ТА МЕТОДИ МОНІТОРИНГУ СТАНУ ЗДОРОВ'Я ЛЮДИНИ НА ОСНОВІ ДАНИХ СЕНСОРНИХ МЕРЕЖ

Стаття присвячена актуальним проблемам традиційних систем моніторингу стану здоров'я, такі як інформаційне перевантаження та втрата від сигналів тривоги медперсоналу, відсутність безперервності та точності даних, а також обмежений доступ до медичної допомоги. Її метою є систематизація та аналіз сучасних інтелектуальних моделей та методів моніторингу стану здоров'я людини з використанням сенсорних мереж. Зокрема, розглянуто основні типи бездротових сенсорних мереж (англ. Wireless Sensor Network, WSN) та носимих пристроїв для збору фізіологічних даних, а також концепція інтернет-медичних систем (англ. Internet of Medical Things, IoMT), що створює взаємопов'язану екосистему для автоматизованого збору та аналізу медичних даних.

У статті визначено етапи попередньої обробки сенсорних даних, які є критично важливими для ефективності інтелектуальних систем. Детально проаналізовано провідні алгоритми машинного навчання (англ. Machine Learning, ML) та глибокого навчання, що застосовуються для обробки, аналізу та інтерпретації медичних даних. З'ясовано, що ці алгоритми дозволяють виявити аномалії, класифікувати активність, прогнозувати загострення хронічних захворювань та створювати точніші системи для ранньої діагностики.

Окрім того, у статті розкрито ключові невирішені проблеми, такі як безпека та конфіденційність даних, інтероперабельність та інтеграція різнорідних даних, адаптивність та персоналізація моделей, обмеження проактивного прогнозування патологій та інтерпретованість моделей штучного інтелекту. З'ясовано, що вирішення цих проблем може полягати у впровадженні підходу роботи з великими даними, засобів штучного інтелекту для проактивного прогнозування та провадження спеціальних тренінгів для медперсоналу, направлених на їхнє ознайомлення з принципами роботи моделей ШІ, які використовуються у системах моніторингу. Зроблено висновок, що доцільність та ефективність таких підходів потребує подальшого дослідження.

Ключові слова: інтелектуальні моделі, моніторинг здоров'я, сенсорні мережі, бездротові сенсорні мережі, інтернет медичних речей, штучний інтелект, машинне навчання.

Постановка проблеми. Традиційні системи моніторингу стану здоров'я людини стикаються з великою кількістю суттєвих обмежень та низькою ефективністю, які перешкоджають наданню своєчасної медичної допомоги. Такі проблеми створюють передумови для пошуку нових інтелектуальних рішень на основі сенсорних мереж.

Одною з головних проблем є інформаційне перевантаження та втрата медперсоналу від сигналів тривоги. Медичні працівники часто працюють на межі своїх можливостей, що супроводжується надмірним обсягом даних та ускладнює своєчасне виявлення найважливішої інформації про пацієнта. Така ситуація призводить до затримок надання необхідної пацієнту допомоги, що сприяє ускладненню поточного стану його здоров'я та до погіршення результатів його лікування [11, с. 698].

Також, значна частина традиційних систем спостереження за станом здоров'я пацієнтів схильні до частих хибних сигналів тривоги. Велика кількість таких спрацьовувань може призвести до звикання медперсоналу до недійсних тривожних сигналів. Таким чином, медпрацівники можуть почати повільніше реагувати на них або перестати звертати на них увагу взагалі. А це, в свою чергу, може призвести до небезпечних наслідків, таких як різке погіршення стану або до смерті пацієнта [11, с. 672].

Ще однією проблемою є відсутність безперервності та точності даних, отриманих з сенсорів. Це пов'язано з тим, що багато існуючих систем моніторингу стану пацієнтів є стаціонарними, що ускладнює безперервний супровід пацієнта такими системами.

Значною проблемою, також, є обмежений доступ до медичної допомоги та нерівність у її

наданні. Доступ до традиційної медичної допомоги суттєво ускладнюється різними чинниками, включаючи нестачу медичних працівників, яка спричинена військовою агресією Росії проти України (багато лікарів мобілізовано), значні відстані, які багато пацієнтів вимушені подолати для отримання медичних послуг, проблеми з транспортом, висока вартість лікування (в тому числі медикаментів), що впливає на рішення пацієнтів вчасно звернутися за медичною допомогою, та відсутність надійного широкосмугового доступу до інтернету, необхідного для телемедичних послуг [12, с. 67].

Ще однією проблемою є нерівність у доступі до медичних технологій, оскільки не кожен пацієнт з огляду на різні чинники може собі дозволити вільний доступ до медичних технологій, що посилює існуючі диспропорції [13, с. 3]. Таким чином, вирішення цих технологічних та доступових чинників за допомогою інтелектуальних систем є не просто питанням підвищення ефективності, а досягнення фундаментальної рівності в охороні здоров'я та соціальної справедливості.

Аналіз останніх досліджень і публікацій. Проблема моніторингу стану здоров'я людини активно досліджується в останні десятиліття, що зумовлено прогресом у галузі біомедичної інженерії, сенсорних технологій, інформатики та штучного інтелекту. Велика кількість наукових робіт присвячена розробці портативних носимих пристроїв для збору основних медичних показників пацієнтів.

Дослідження [1] присвячене використанню бездротових сенсорних мереж для моніторингу стану здоров'я. Авторами розглянуто можливості бездротового та віддаленого моніторингу стану пацієнтів, що покращує спостереження за перебігом захворювання, для впровадження індивідуального підходу в медицині. Особливу увагу приділено застосуванню вебслужб на пристроях (англ. Web Services on Devices, WSD) для моніторингу життєво важливих показників, допомоги особам похилого віку та спостереження за навколишнім середовищем в медичних установах. Проте, подальшого дослідження потребують такі питання, як: безпеки та конфіденційності даних, надійності та сумісності систем, ефективного управління енергоспоживанням та контролю якості зібраних даних.

У роботі [2] представлено систематичний огляд бачення та досвіду роботи медичних працівників з віддаленим моніторингом стану пацієнтів (англ. Remote Patient Monitoring, RPM). Таким чином, автор дійшов висновку, що фахівці вважають RPM здійсненням і важливим у роботі медичних

закладів. Окрім того, були виявлені такі переваги використання RPM у медичних закладах, як безперервний моніторинг за станом пацієнтів, покращення впевненості медперсоналу у контролі за ситуацією. Однак, впровадження RPM передбачає й деякі складнощі, наприклад, збільшення робочого навантаження на медпрацівників, підвищення тривожності пацієнтів, неточність даних, технічні труднощі, фінансові питання та проблеми конфіденційності; деякі з яких можна вирішити шляхом правильної організації роботи медзакладу. Таким чином, згідно з цією роботою актуальним напрямом подальших досліджень є вивчення впливу RPM на робоче навантаження та тривожність медперсоналу та пацієнтів, а також інтеграція штучного інтелекту (ШІ) та блокчейну для підвищення ефективності, з урахуванням проблем сумісності, доступу, безпеки даних та нормативної бази.

В роботі [3] описується, як саме використання медичних сенсорів та ШІ впливають на розвиток клінічних досліджень. Автори підкреслюють, що ці технології роблять клінічні випробування більш точними, ефективними та зручними для пацієнтів, надаючи дані в режимі реального часу та створюючи чіткішу картину стану здоров'я.

Дослідження [4] присвячене інтеграції даних з носимих сенсорів зі штучним інтелектом для моніторингу здоров'я в режимі реального часу та прогнозу діагностики. Основна увага зосереджена на тому, як це поєднання впливає на охорону здоров'я в цілому, шляхом забезпечення безперервного моніторингу та раннього виявлення проблем зі здоров'ям, а також можливість з допомогою використання моніторингових систем безперервного спостереження за перебігом захворювання та побудови персоналізованих планів лікування для пацієнтів. Проте, залишаються відкритими питання конфіденційності та безпеки даних, забезпечення точності та надійності сенсорних даних та прогнозів ШІ.

У роботі [5] від американського агентства Food and Drug administration (FDA) обговорюється роль штучного інтелекту та машинного навчання у медичному програмному забезпеченні. Основною темою є вплив на охорону здоров'я цих технологій та підхід FDA до їхнього регулювання.

У статті [8] описується вплив використання засобів штучного інтелекту на моніторинг здоров'я, причому автор не обмежується одними тільки носимими пристроями. Таким чином, у роботі було досліджено те, як ШІ забезпечує прогностичні та неінвазивні рішення догляду за пацієнтами в режимі реального часу.

Постановка завдання. Метою цього дослідження є систематизація та аналіз сучасних інтелектуальних моделей та методів моніторингу стану здоров'я людини на основі даних сенсорних мереж. Для досягнення мети поставлено такі завдання:

- розглянути основні типи сенсорних мереж та носимих пристроїв, що застосовуються для збору фізіологічних та поведінкових даних у сфері охорони здоров'я;
- проаналізувати провідні інтелектуальні алгоритми машинного та глибокого навчання, що використовуються для обробки, аналізу та інтерпретації сенсорних даних у медичній сфері;
- визначити ключові переваги та поточні обмеження існуючих підходів до інтелектуального моніторингу здоров'я а також окреслити перспективні напрямки подальших досліджень у цій галузі.

Виклад основного матеріалу. Основу інтелектуального моніторингу здоров'я становлять різноманітні архітектури сенсорних мереж. Найбільш поширеною є натільна безпроводна мережа (англ. Body Area Network, BAN), яка складається з невеликих, легких та енергоефективних сенсорів, розміщених на тілі людини або в безпосередній близькості до неї [5, с. 1; 6, с. 5]. Ці сенсори збирають фізіологічні дані, а саме: електрокардіографію (ЕКГ), електроенцефалографію (ЕЕГ), температуру, пульс, рівень глюкози тощо; та передають їх на центральний пристрій, який, у свою чергу, може пересилати дані до хмарного сховища або медичного центру для подальшого аналізу [1, с. 1] (Рис. 1).

Така архітектура має низку переваг, таких як: мобільність, можливість безперервного збору даних у реальному часі та неінвазивність; що робить пристрої з її використанням ідеальними для довгострокового моніторингу стану пацієнтів як у клінічних умовах, так і в домашньому середовищі [12, с. 67].

Наступною, не менш важливою, концепцією є інтернет медичних речей (англ. Internet of Medical Things, IoMT), що є розширенням інтернету речей у сфері охорони здоров'я [4, с. 963]. IoMT поєднує медичні пристрої, сенсори, програмне забезпечення та хмарні платформи для створення взаємопов'язаної екосистеми, яка забезпечує автоматизований збір, передачу та аналіз медичних даних [3, с. 145]. Це дозволяє здійснювати віддалений моніторинг, телемедицину та персоналізований підхід до охорони здоров'я [10, с. 5].

Також, варто згадати про якість та ефективність інтелектуальних моделей, які значною мірою залежать від належного збору та попередньої обробки сенсорних даних.

Значною проблемою є інтероперабельність та інтеграція різнорідних даних. Інтеграція даних з різних типів сенсорів, пристроїв та медичних інформаційних систем залишається складним завданням. Відсутність уніфікованих стандартів та протоколів ускладнює створення комплексних систем моніторингу. Отже, використання даних з сенсорів в реальному часі призводить до накопичення дуже великих обсягів сирих різнотипних даних, які необхідно відповідним чином обробити. Таким чином, для вирішення цієї проблеми доцільно звернутись до такого підходу роботи з даними, як великі дані (англ. Big Data). Використання методів роботи з великими даними дозволяє обробити зростаючий обсяг інформації, що генерується медичними пристроями, електронними медичними записами та різноманітними сенсорами, який створює величезний потік даних, який неможливо обробляти традиційними методами. Великі дані дозволяють: збирати та консолідувати дані з різних джерел, ефективно зберігати величезний обсяг інформації, аналізувати складні взаємозв'язки та виявляти приховані закономірності та прогнозувати розвиток захворювань та оцінювати ризики.

Сирі дані, отримані від сенсорів, часто містять шум, артефакти та пропуски, що може негативно



Рис. 1. Безперервний потік даних від пацієнта до лікаря

вплинути на точність аналізу. Етапи попередньої обробки даних передбачають фільтрацію шумів, нормалізацію та масштабування, сегментацію даних та виділення атрибутів.

Першим етапом обробки даних є фільтрація шумів, що передбачає застосування цифрових фільтрів для видалення небажаних компонентів сигналу, які не несуть релевантної інформації. Наступним кроком є нормалізація та масштабування, що спрямоване на приведення даних до єдиного діапазону або розподілу, аби забезпечити коректну роботу алгоритмів машинного навчання шляхом усунення впливу різномірності даних. Далі відбувається сегментація, що полягає у розбитті безперервних потоків даних на значущі дискретні сегменти або вікна, що є передумовою для їхнього подальшого аналізу.

Останній етап – виділення ознак (атрибутів) – є особливо важливим етапом, оскільки його метою є трансформація сирих даних у більш інформативне представлення, придатне для використання інтелектуальними моделями.

Традиційні підходи до виділення ознак включають розрахунок статистичних параметрів, а також використання частотних ознак, отриманих, наприклад, шляхом перетворення Фур'є, та часово-частотних ознак. Проте, сучасні методології, зокрема, застосування згорткових нейронних мереж (англ. Convolution neural network, CNN), дозволяють автоматично вилучати ієрархічні ознаки без потреби у ручному проєктуванні. Цей підхід суттєво оптимізує процес та підвищує загальну ефективність системи.

Окрім того, ключову роль у первинній обробці даних та класифікації сенсорних даних відіграє машинне навчання. Найбільш поширеними алгоритмами, які використовуються для аналізу медичних даних, є такі методи, як: метод опорних векторів, випадкові ліси, K-найближчих сусідів та логістична регресія.

Методи опорних векторів ефективні для завдань класифікації, таких як розрізнення нормальних та аномальних показників ЕКГ або для виявлення певних станів активності на основі даних акселерометра [6, с. 5729]. Такий метод, як випадкові ліси, добре справляється з великими обсягами даних і може бути використаний для прогнозування ризику розвитку захворювань або класифікації поведінкових патернів [1, с. 2760].

Також, простим, але ефективним, алгоритмом для класифікації та регресії є K-найближчих сусідів, який може бути використаний для ідентифікації певних фізіологічних станів або відсутності

певного стану захворювання на основі зібраних показників.

Що стосується логістичної регресії, то вона часто застосовується для бінарної класифікації, наприклад, визначення наявності або відсутності певного стану захворювання на основі зібраних показників.

Такі ML-алгоритми успішно застосовуються для вирішення таких завдань, як виявлення аномалій у фізіологічних показниках, класифікація активності, прогнозування загострень хронічних захворювань та персоналізована оцінка ризиків [1, с. 2760, 9, с. 302].

Більш того, зі зростанням складності та обсягів сенсорних даних, особливо часових рядів, все більшого значення набувають методи глибокого навчання. Глибокі нейронні мережі здатні автоматично витягувати складні ознаки з сирих даних, що робить їх особливо ефективними для обробки багатовимірних та нелінійних даних, характерних для медичного моніторингу за станом здоров'я. До ключових архітектур глибокого навчання належать: згорткові нейронні мережі, рекурентні нейронні мережі (англ. Recurrent neural networks, RNN) зокрема мережі довгої короточасної пам'яті (англ. Long Short-Term Memory, LSTM) та автокодувальники.

Хоча CNN традиційно використовуються для обробки зображень, вони також дуже ефективні для аналізу часових рядів, де можна виявити локальні закономірності та особливості в даних сенсорів [6, с. 5729].

RNN ідеально підходять для обробки послідовних даних, оскільки вони здатні враховувати залежності між попередніми та поточними значеннями. LSTM широко застосовуються для прогнозування фізіологічних показників, аналізу динаміки стану здоров'я та виявлення складних часових закономірностей у даних [1, с. 2761].

Автокодувальники можуть використовуватися для зменшення розмірності даних, виявлення аномалій або для генерації систематичних даних, що є корисним для навчання моделей у випадках обмеженості реальних медичних даних.

Отже, за допомогою застосування глибокого навчання для ранньої діагностики захворювань можна створити більш точні та надійні системи, безперервного моніторингу критичних станів, персоналізованої оцінки ефективності лікування та прогнозування розвитку ускладнень [1, 9]. Інтеграція таких алгоритмів з даними, отриманими від IoT-пристроїв, сформує основу для інтелектуальних систем підтримки прийняття

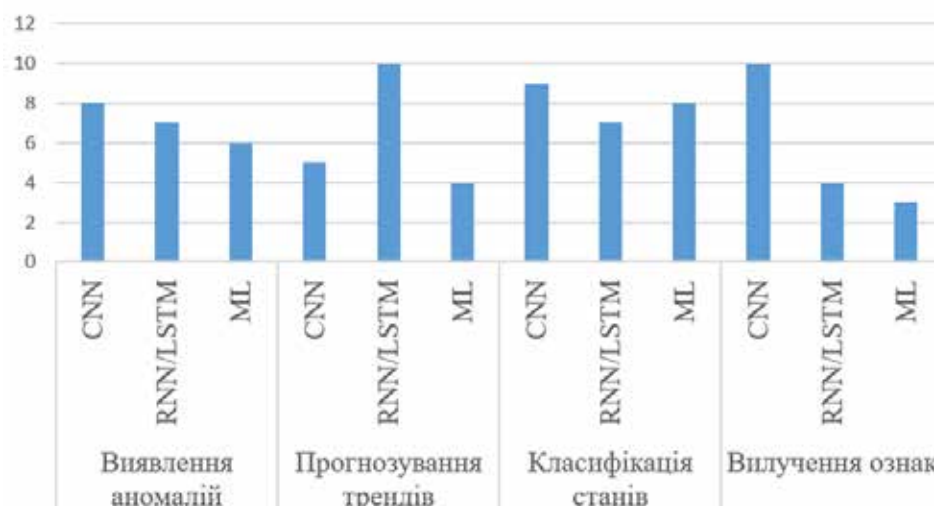


Рис. 2. Застосування алгоритмів за завданням

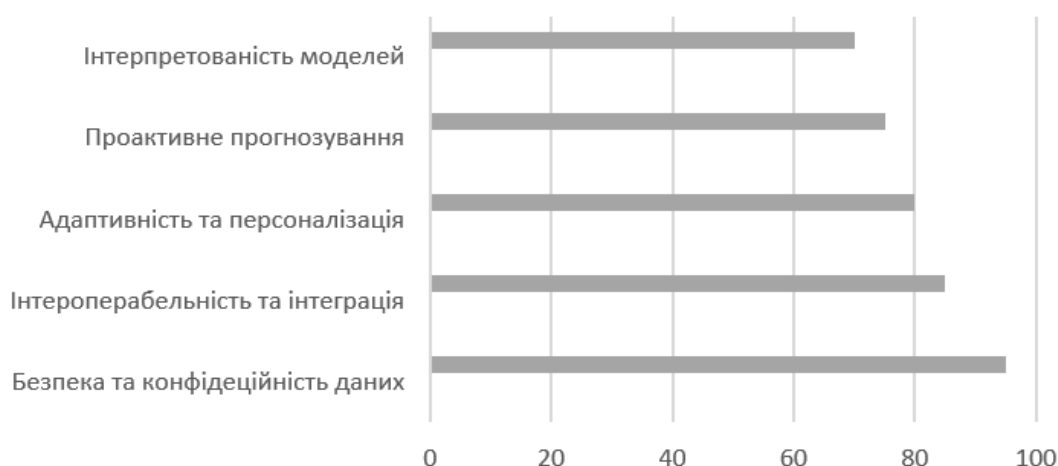


Рис. 3. Ключові перешкоди на шляху до впровадження

рішень у медицині, що дозволить лікарям отримувати більш повну та своєчасну інформацію для оптимального лікування пацієнтів. В наукових дослідженнях з'являються численні платформи та системи, що використовують штучний інтелект та IoT для моніторингу стану здоров'я. Такі пристрої включають пристрої збору даних, шлюзи та комунікаційні модулі, хмарні обчислювані платформи, інтелектуальні аналітичні модулі, інтерфейси користувача.

Більш того, збір та обробка чутливих медичних даних створює серйозні виклики щодо їхньої безпеки та конфіденційності. Хоча перспективні підходи, такі як федеративне навчання та децентралізовані реєстри на основі блокчейну дають надію на їх впровадження. Залишається необхідність проводити подальші дослідження для їхнього

широкого та надійного впровадження в медичну практику.

Наступним важливим аспектом є адаптивність та персоналізація. Велика кількість існуючих моделей сфокусовані на загальних патернах, тоді як адаптація інтелектуальних систем до індивідуальних фізіологічних особливостей кожного пацієнта та забезпечення їхньої надійності в умовах змінних зовнішніх факторів потребує подальших розробок. Вирішення цих проблем може полягати у проведенні тренінгів для медичного персоналу та ознайомлення працівників з принципом роботи ШІ. Головною метою таких тренінгів є подолання недовіри персоналу до новітніх технологій. Також потрібно ознайомити медперсонал з розумним використанням ШІ, яке допоможе у вирішенні складних та суперечних моментів в лікуванні

пацієнтів. Неналежне використання складних систем може призвести до помилок, що негативно вплине на здоров'я пацієнтів. Для цього тренінги повинні охоплювати протоколи дій у надзвичайних ситуаціях, наприклад що робити, якщо система дає збій або видає невірні дані. Медичні технології швидко розвиваються. Тому тренінги не мають бути одноразовими. Загалом доцільність та ефективність таких тренінгів потребує подальшого дослідження.

Окрім того, хоча досягнуто значного прогресу у виявленні аномалій, існуючі системи моніторингу все ще мають обмеження в контексті проактивного прогнозування розвитку патологій на ранніх стадіях, до появи явних симптомів. Для проактивного прогнозування необхідно збирати та аналізувати не лише явні клінічні показники, а й такі дані, які можуть свідчити про початок патологічного процесу задовго до появи симптомів. Впровадження засобів ШІ для проактивного прогнозування є революційним кроком, який дозволить перейти від лікування вже виявлених хвороб до запобігання їх виникненню. Такий крок значно покращить якість життя та знизить навантаження на систему охорони здоров'я. Ефективність такого підходу потребує подальшого дослідження.

Останньою, але не менш важливою проблемою є інтерпретованість моделей. Складність моделей глибокого навчання часто призводить до їхньої закритості, що ускладнює розуміння причин прийняття рішень та довіру до них з боку медичного персоналу. Спеціальні тренінги для персоналу допоможуть подолати цю проблему, а саме здобути базове розуміння роботи засобів ШІ та ліпше зрозуміти обмеженість ШІ. Такі тренінги дадуть розуміння персоналу, що впровадження засобів ШІ призначене не на заміну людини на роботизовані машини, а покликані допомогти у вирішенні складних питань.

Висновки. У цій статті було проведено комплексний аналіз сучасних інтелектуальних моделей і методів моніторингу стану здоров'я людини

з використанням сенсорних мереж, що дозволило висвітлити ключові аспекти та перспективи розвитку цієї галузі.

Систематизовано та проаналізовано основні типи сенсорних мереж та носимих пристроїв, таких як BAN та концепція IoMT. Визначено, що ці технології відіграють ключову роль у безперервному зборі фізіологічних та поведінкових даних, забезпечують мобільність, точність та неінвазивність моніторингу, що є критично важливим для довгострокового спостереження за пацієнтом як у клініці, так і вдома.

Детально розглянуто етапи попередньої обробки сенсорних даних, а саме фільтрацію шуму, нормалізацію, масштабування та сегментацію. Особливу увагу приділено знаходженню ознак, підкреслюючи переваги застосування CNN для автоматичного вилучення ієрархічних ознак, що оптимізує процес та підвищує ефективність системи.

Проаналізовано провідні алгоритми ML та глибокого навчання, що застосовуються для обробки, аналізу та інтерпретації медичних даних. Визначено, що такі ML-алгоритми, як методи опорних векторів, випадкові ліси, K-найближчих сусідів та логістична регресія ефективні для виявлення аномалій, класифікації активності та прогнозування ризиків. Доведено, що глибокі нейронні мережі, зокрема CNN, RNN та LSTM, здатні автоматично витягувати складні ознаки та створювати більш точні та надійні системи для ранньої діагностики та прогнозування розвитку ускладнень.

Виявлено ключові невирішені проблеми та обмеження у сфері інтелектуального моніторингу здоров'я, такі як безпека та конфіденційність даних, інтероперабельність та інтеграція різнорідних даних, адаптивність та персоналізація моделей, обмеження у проактивному прогнозуванні патологій, а також інтерпретованість моделей ШІ. Запропоновано напрями для подальших досліджень, що включають впровадження федеративного навчання для безпеки даних.

Список літератури:

1. An IoT-Based Framework for Personalized Health Assessment and Recommendations Using Machine Learning / S. K. Jagatheesaperumal et al. *Mathematics*. 2023. Vol. 11, no. 12. P. 2758. URL: <https://doi.org/10.3390/math11122758> (date of access: 18.07.2025).
2. Anmulwar S., Gupta A. K., Derawi M. Challenges of IoT in Healthcare. *IoT and ICT for Healthcare Applications*. Cham, 2020. P. 11–20. URL: https://doi.org/10.1007/978-3-030-42934-8_2 (date of access: 18.07.2025).
3. Ehizogie Paul Adeghe, Chioma Anthonia Okolo, Olumuyiwa Tolulope Ojeyinka. A review of wearable technology in healthcare: Monitoring patient health and enhancing outcomes. *Open Access Research Journal of Multidisciplinary Studies*. 2024. Vol. 7, no. 1. P. 142–148. URL: <https://doi.org/10.53022/oarjms.2024.7.1.0019> (date of access: 22.07.2025).

4. Guo X., Li Y. Intelligent Health in the IS Area: A Literature Review and Research Agenda. *Fundamental Research*. 2023. URL: <https://doi.org/10.1016/j.fmre.2023.04.008> (date of access: 20.07.2025).
5. Moon J., Ju B.-K. Wearable Sensors for Healthcare of Industrial Workers: A Scoping Review. *Electronics*. 2024. Vol. 13, no. 19. P. 3849. URL: <https://doi.org/10.3390/electronics13193849> (date of access: 19.07.2025).
6. Muin S., Mosalam K. M. Structural Health Monitoring Using Machine Learning and Cumulative Absolute Velocity Features. *Applied Sciences*. 2021. Vol. 11, no. 12. P. 5727. URL: <https://doi.org/10.3390/app11125727> (date of access: 22.07.2025).
7. Recent advances in smart wearable sensors for continuous human health monitoring / M. B. Kulkarni et al. *Talanta*. 2024. Vol. 272. P. 125817. URL: <https://doi.org/10.1016/j.talanta.2024.125817> (date of access: 22.07.2025).
8. Sensor Networks for Structures Health Monitoring: Placement, Implementations, and Challenges—A Review / S. Mustapha et al. *Vibration*. 2021. Vol. 4, no. 3. P. 551–584. URL: <https://doi.org/10.3390/vibration4030033> (date of access: 22.07.2025).
9. Smart Healthcare System Based on AIoT Emerging Technologies: A Brief Review / C. Prabha et al. *Lecture Notes in Electrical Engineering*. Singapore, 2023. P. 299–311. URL: https://doi.org/10.1007/978-981-99-0601-7_23 (date of access: 22.07.2025).
10. The Impact of AI on the Development of Multimodal Wearable Devices in Musculoskeletal Medicine / G. Olson et al. *HSS Journal®: The Musculoskeletal Journal of Hospital for Special Surgery*. 2025. URL: <https://doi.org/10.1177/15563316251344945> (date of access: 18.07.2025).
11. Wireless sensor network for structural health monitoring: A contemporary review of technologies, challenges, and future direction / M. Abdulkarem et al. *Structural Health Monitoring*. 2019. Vol. 19, no. 3. P. 693–735. URL: <https://doi.org/10.1177/1475921719854528> (date of access: 22.07.2025).
12. Yaghoubi M., Ahmed K., Miao Y. Wireless Body Area Network (WBAN): A Survey on Architecture, Technologies, Energy Consumption, and Security Challenges. *Journal of Sensor and Actuator Networks*. 2022. Vol. 11, no. 4. P. 67. URL: <https://doi.org/10.3390/jsan11040067> (date of access: 22.07.2025).
13. Zakerabasali S., Ayyoubzadeh S. M. Internet of Things and healthcare system: A systematic review of ethical issues. *Health Science Reports*. 2022. Vol. 5, no. 6. URL: <https://doi.org/10.1002/hsr2.863> (date of access: 23.07.2025).

Honcharov D.S. INTELLIGENT MODELS AND METHODS OF MONITORING HUMAN HEALTH BASED ON SENSOR NETWORK DATA

The article is devoted to the current issues in traditional health monitoring systems, such as medical staff information overload and fatigue from alarms, lack of data continuity and accuracy, and limited access to health care. Its aim is to systematize and analyse modern intelligent models and methods for monitoring human health using sensor networks. In particular, the main types of wireless sensor networks (WSNs) and wearable devices for collecting physiological data are considered. Moreover, the concept of Internet medical systems (IoMT), which creates an interconnected ecosystem for automated collection and analysis of medical data is considered too.

The article identifies the stages of pre-processing of sensory data, which are critical for the efficiency of intelligent systems. The leading machine learning (ML) and deep learning algorithms used for processing, analyzing and interpreting medical data are analyzed in detail. It was found that these algorithms allow to detect abnormalities, classify activity, predict the exacerbation of chronic diseases and create more accurate systems for early diagnosis.

In addition, the article reveals key unresolved issues such as data security and privacy, interoperability and integration of disparate data, adaptability and personalization of models, limitations in proactive prediction of pathologies, and interpretability of artificial intelligence models. It was found that the solution to these problems may consist in implementing a big data approach, artificial intelligence tools for proactive forecasting, and conducting special trainings for medical personnel aimed at familiarizing them with the principles of AI models used in monitoring systems. It is concluded that the feasibility and effectiveness of such approaches requires further investigation.

Key words: intelligent models, health monitoring, sensor networks, wireless sensor networks, internet of medical things, artificial intelligence, machine learning.

Дата надходження статті: 26.07.2025

Дата прийняття статті: 04.08.2025

Опубліковано: 27.10.2025

Горбаченко С.А.

Національний університет «Одеська юридична академія»

Овчинников М.Ю.

Національний університет «Одеська юридична академія»

ЗАХИСТ ІНТЕЛЕКТУАЛЬНОЇ ВЛАСНОСТІ У СФЕРІ ЦИФРОВОГО КОНТЕНТУ: ВИКЛИКИ ТА РІШЕННЯ ДЛЯ РЕКЛАМНОГО МЕНЕДЖМЕНТУ

У статті розглянуто проблематику порушення прав інтелектуальної власності при застосуванні мультимедійного контенту в онлайн-рекламі, що набуває особливої актуальності в умовах стрімкого зростання як обсягів самої реклами, так і її мультимедійної складової. У сучасному цифровому середовищі реклама є однією з найбільш динамічних галузей, що використовує різноманітні формати – банери, відеоролики, інтерактивні креативи, нативну рекламу, AR-фільтри, короткі відео для TikTok та Instagram Reels, а також email-розсилки з вбудованими медіа. В усіх цих форматах активно застосовуються зображення, відео, музика, 3D-моделі та голосові озвучення, які часто є об'єктами авторського права. Проаналізовано основні типові порушення, пов'язані з використанням неліцензованого контенту, а також ризики, які вони становлять для рекламодавців і агентств: фінансові санкції, блокування контенту на платформах, репутаційні втрати. Рекламна діяльність постійно ускладнюється стає багаторівневою, що зумовлює необхідність перегляду та уточнення основних законодавчих дефініцій, зокрема в галузі авторських прав. В умовах активного використання мультимедійного контенту в рекламі, важливо не лише правильно визначати права на об'єкти авторського права, а й вдосконалювати правові механізми їхнього розпорядження, щоб забезпечити захист від можливих порушень і забезпечити правову визначеність для рекламодавців. Запропоновано концепцію інформаційної системи, яка дозволяє здійснювати автоматизований моніторинг дотримання авторських прав у рекламних кампаніях. Описано архітектуру системи та принципи її роботи з використанням AI-моніторингу, цифрових водяних знаків і ліцензійних баз. Наведено приклад практичного застосування системи в рамках рекламного менеджменту. Стаття орієнтована на фахівців у галузі цифрового маркетингу, реклами та інформаційних технологій, які прагнуть забезпечити правову безпеку використання мультимедійного контенту.

Ключові слова: аудіовізуальні твори, авторське право, реклама, рекламний менеджмент, комп'ютерні науки, програмування, веброзробка, метадані, стеганографія, цифровий водяний знак, нейронні мережі, захист інформації.

Постановка проблеми. У сучасних умовах цифрової трансформації бізнесу мультимедійний контент – зображення, відео, аудіо, анімація – стає ключовим інструментом у реалізації рекламних кампаній. Використання таких об'єктів підпадає під правову охорону відповідно до Закону України «Про авторське право і суміжні права», який забороняє їхнє використання без дозволу правовласника, зокрема в цифровому середовищі. Згідно із Законом «Про захист від недобросовісної конкуренції», неправомірне використання чужих креативних рішень, зображень, відео чи музики в рекламних матеріалах може розглядатися як

недобросовісна комерційна практика [1, 2, 3]. Це особливо актуально в умовах стрімкого зростання обсягів мультимедійного контенту в онлайн-рекламі. Поширення соціальних мереж, платформ потокового відео та мобільних додатків призвело до стрімкого зростання обсягів створення та споживання мультимедійної інформації. Водночас це породжує численні виклики, пов'язані із захистом прав інтелектуальної власності (ІВ) у сфері рекламного менеджменту. Дуже часто маркетингові агентства та окремі бренди використовують мультимедійні матеріали без належного дотримання авторських прав – як навмисно, так

і через необізнаність або відсутність ефективних інструментів перевірки [4, 5]. Це може призвести до юридичних санкцій, фінансових втрат і репутаційних ризиків. Актуальність проблеми зумовлюється як збільшенням кількості порушень, так і ускладненням виявлення нелегального використання цифрових активів. У зв'язку з цим особливого значення набуває розробка та впровадження інформаційних систем, здатних забезпечити автоматизований контроль за дотриманням авторських прав на мультимедійні об'єкти в межах рекламних кампаній. Такий підхід дозволяє не лише захистити інтереси правовласників, а й формує правову культуру в рекламному середовищі. Проаналізуємо основні загрози порушення прав інтелектуальної власності у цифровому маркетингу та обґрунтуємо різні підходи до створення інформаційної системи, яка б забезпечувала моніторинг і превентивний захист мультимедійного контенту в рамках рекламного менеджменту [6]. З розвитком цифрових технологій та широким впровадженням інтернет-реклами використання мультимедійного контенту стало невід'ємною частиною рекламного менеджменту. Це призвело до того, що виникла потреба захисту прав авторів рекламних продуктів та тих мультимедійних даних, які в цій рекламі використовувалися. Яскраві візуальні образи, динамічні відеоролики та емоційна музика підвищують ефективність комунікації з цільовою аудиторією, але одночасно підвищують ризики порушення прав інтелектуальної власності [7, 8]. Оскільки рекламні кампанії часто реалізуються в стислі строки й за участі великої кількості виконавців (дизайнерів, таргетологів, SMM-фахівців) [9], контроль за дотриманням авторських прав ускладнюється. Відсутність централізованої системи верифікації та відстеження використаного контенту створює «сірі зони», у яких порушення можуть залишитися непоміченими до моменту претензії з боку правовласника. Однак, у будь-якому випадку, це призведе до певних негативних наслідків (табл. 2). Таким чином, виникає потреба в розробці спеціалізованих інформаційних систем, які б дозволяли рекламним менеджерам виявляти потенційні порушення ще на етапі підготовки матеріалів або запуску кампанії.

Аналіз останніх досліджень і публікацій. Проблематика захисту інтелектуальної власності (ІВ) у цифровому середовищі залишається актуальною як у правових, так і в технічних дослідженнях. У фокусі уваги науковців захист мультимедійного контенту, зокрема відео, зображень, музики,

створених або використовуваних у цифровому маркетингу та рекламі. У публікаціях Всесвітньої організації інтелектуальної власності (World Intellectual Property Organization – WIPO) акцентовано увагу на необхідності адаптації правових механізмів до нових реалій цифрової економіки, особливо у сфері онлайн-реклами [10]. Зазначено, що традиційні правові інструменти, такі як авторське право і товарні знаки, потребують підтримки через технологічні рішення – цифрові підписи, верифікацію ліцензій, автоматизований моніторинг [11]. Окрему увагу приділено застосуванню систем DRM (Digital Rights Management) для захисту відео- та аудіоконтенту, а також цифрових водяних знаків, що дозволяють вбудовувати унікальні ідентифікатори у медіафайли [12]. Такі технології вже активно використовуються платформами на кшталт YouTube (через систему Content ID) і Meta (через Rights Manager) для автоматичного виявлення порушень у рекламному контенті [13, 14]. Сучасні дослідження також висвітлюють можливості використання штучного інтелекту для захисту авторських прав. У роботі Лім Дж. Х. (Lim J. H.) описано інтеграцію цифрових водяних знаків безпосередньо в архітектуру генеративних моделей [15]. Це відкриває нові перспективи для контролю над використанням AI-генерованого контенту у сфері реклами. У контексті рекламного менеджменту дослідники зазначають, що порушення авторських прав часто є результатом відсутності внутрішніх процедур контролю у агентствах або нехтування правовими вимогами через комерційний тиск [6, 16]. Низка українських авторів також звертає увагу на необхідність поєднання правових та інформаційних підходів до захисту ІВ, зокрема у проектах із впровадження електронних сервісів та систем контролю авторських прав [17, 18, 19]. Таким чином, існує очевидна потреба у міждисциплінарному підході до проблеми – на перетині інформаційних технологій, права та рекламного менеджменту. Це створює підґрунтя для розробки спеціалізованих інформаційних систем, здатних автоматично виявляти та запобігати порушенням ІВ у межах рекламних кампаній.

Постановка завдання. Метою статті є розробка та обґрунтування концепції інформаційної системи захисту інтелектуальної власності на мультимедійні дані в контексті цифрового маркетингу і реклами, що дозволяє автоматизувати процес виявлення та запобігання порушенням авторських прав у рекламних кампаніях.

Виклад основного матеріалу. Розглянемо сучасні технології захисту інтелектуальної влас-

ності у цифровому середовищі. З метою протидії порушенням прав інтелектуальної власності в онлайн-середовищі активно впроваджуються інноваційні технології, які дозволяють виявляти та запобігати несанкціонованому використанню мультимедійних об'єктів. Ці технології вже застосовуються у сфері медіа, але їх потенціал для рекламного менеджменту наразі використовується недостатньо. Розглянемо типові порушення інтелектуальної власності у сфері реклами та систематизуємо їх. Наведені в таблиці 1 дії можуть бути спричинені не лише злим наміром, а й низкою об'єктивних та суб'єктивних чинників. Насамперед це стосується недостатньої обізнаності працівників рекламної галузі з питаннями авторського права, відсутності спеціалізованих юридичних підрозділів у невеликих агентствах, а також надмірної динаміки ринку, яка вимагає створення великої кількості контенту в стислі терміни. У таких умовах перевірка кожного елемента на предмет ліцензії може розглядатися як другорядна задача.

Водночас деякі виконавці навмисно ігнорують правові вимоги задля оптимізації витрат або швидшого виконання замовлення, що ставить під загрозу не лише ефективність кампанії, а й репутацію замовника.

DRM-системи (Digital Rights Management). Системи цифрового управління правами дозволяють контролювати використання контенту через спеціальні цифрові ліцензії. DRM забезпечує шифрування контенту, що унеможливорює його нелегальне копіювання [20]. Також дозволяє керувати правами доступу, визначаючи, хто, коли й на яких умовах може використовувати матеріал. Цей метод є ефективним для відео- та аудіоконтенту, що розповсюджується через професійні платформи (наприклад, Netflix, Spotify). Однак для відкритих рекламних платформ (YouTube, Instagram) пряме впровадження DRM складне – через масовий відкритий доступ і відсутність централізованих точок контролю.

Цифрові водяні знаки. Цей метод побудований на принципі прихованого вбудовування унікального коду у зображення, відео або аудіофайл. Використання цифрових водяних знаків відкриває низку важливих переваг у сфері захисту мультимедійного контенту. Однією з ключових є можливість відстеження джерела походження файлу навіть у випадках, коли він зазнав копіювання, редагування або повторної публікації на інших платформах. Завдяки цьому водяний знак виступає своєрідним «цифровим відбитком пальця», який дозволяє ідентифікувати першоджерело незалежно від зовнішніх змін у файлі [21]. Крім того, технологія забезпечує надійний доказ авторства в разі виникнення юридичних суперечок. Наявність вбудованого у файл ідентифікатора, що містить дані про правовласника, дату створення та умови використання, може бути використана як електронний доказ у судових процесах. Ще однією суттєвою перевагою є можливість автоматизованого виявлення копій за допомогою спеціалізованого програмного забезпечення або API. Це дозволяє здійснювати моніторинг великих обсягів даних в онлайн-середовищі – на вебсайтах, у соціальних мережах, хмарних сховищах – і оперативно реагувати на факти несанкціонованого використання контенту. Особливо актуальним цей підхід є у сфері реклами, де один і той самий візуальний матеріал часто проходить через ланцюг виконавців – агентства, фрилансери, підрядники, замовники – і може бути змінений або адаптований під різні платформи. У таких умовах цифрове водякування забезпечує контроль за життєвим циклом креативу на всіх етапах його використання.

Технології на базі штучного інтелекту. AI-моделі відіграють ключову роль у масштабному моніторингу цифрового контенту. Серед популярних рішень Content ID (YouTube) – автоматично ідентифікує музику та відео, що порушують авторські права; Facebook Rights Manager – аналогічна сис-

Таблиця 1

Типові порушення інтелектуальної власності у рекламі

Тип порушення	Приклад або контекст використання
Використання неліцензованих зображень або відео з відкритих джерел	Зображення, знайдені через Google Images, Pinterest, соціальні мережі
Застосування музичних фрагментів без ліцензії	Музика у відео для TikTok, Instagram Reels, YouTube Shorts без зазначення авторства або прав
Копіювання оригінального дизайну або брендових елементів	Імітація унікального стилю, логотипів, упаковки або графіки конкурента
Використання чужих рекламних ідей, концептів або слоганів, захищених авторським правом	Запозичення творчих підходів із попередніх рекламних кампаній без дозволу

Таблиця 2

Наслідки порушень прав інтелектуальної власності у сфері реклами

Наслідок порушення	Опис або приклад
Юридичні претензії та судові позови	Подання скарг з боку правовласників, що може призвести до судових розглядів і фінансових санкцій
Блокування акаунтів або контенту на платформах	Автоматичне видалення відео або блокування сторінки рекламодавця за скаргами через Content ID або Rights Manager
Фінансові штрафи за нелегальне використання контенту	Оплата компенсацій за порушення авторських прав відповідно до чинного законодавства
Втрати репутації бренду або рекламного агентства	Негативний резонанс у ЗМІ або соціальних мережах у разі виявлення плагіату чи крадіжки контенту
Неможливість масштабування кампаній	Відмова платформ у запуску реклами, обмеження охоплення або блокування акаунтів

Таблиця 3

Функціональні модулі інформаційної системи захисту ІВ

Модуль	Опис функціоналу
Модуль перевірки контенту	Автоматично аналізує зображення, відео або аудіо на наявність захищених елементів, використовуючи AI-алгоритми та бази авторських робіт (Pixabay, Getty, Adobe Stock, YouTube Content ID API тощо).
Модуль цифрових водяних знаків	Дозволяє вбудовувати водяні знаки у власні матеріали для подальшого моніторингу їх використання в мережі.
Блок управління ліцензіями	Зберігає дані про придбані ліцензії на контент, терміни дії, обмеження по використанню для уникнення ненавмисних порушень.
Аналітичний блок	Надає звіти для маркетологів і керівників проєктів щодо правового статусу використаного контенту.
Інтеграційний API-модуль	Дозволяє підключати систему до зовнішніх платформ (Facebook Ads Manager, Google Ads, YouTube Studio) для попередньої перевірки креативів перед публікацією.

тема для Meta Ads та Instagram; AI-сервіси аналізу зображень (наприклад, Pixsy, SmartFrame) – дозволяють знаходити копії фото по всьому інтернету. Використання таких інструментів дає змогу рекламним менеджерам виявляти потенційно проблемний контент до моменту запуску реклами, а також – автоматично блокувати чи замінювати матеріали з ознаками порушення прав.

Інформаційна система для захисту інтелектуальної власності у рекламному середовищі. Враховуючи потреби рекламних агентств та особливості цифрового контенту, пропонується концепція інформаційної системи, що забезпечує автоматизований контроль за використанням мультимедійних об'єктів у рамках рекламної діяльності. Така система може інтегруватися в існуючі бізнес-процеси, підвищуючи правову безпеку кампаній. Розглянемо архітектуру цієї системи.

Логіка роботи системи полягає у наступному:

1. Рекламний фахівець завантажує медіафайл у систему.
2. Модуль перевірки аналізує файл:
 - порівнює з базами авторських об'єктів;
 - перевіряє метадані;

– шукає вбудовані водяні знаки.

3. Якщо контент проходить перевірку – система фіксує це та дозволяє публікацію.

4. Якщо виявлено ризики – виводиться попередження з пропозицією:

- завантажити ліцензійний файл;
- придбати ліцензію;
- замінити контент.

5. Вся історія перевірок зберігається в аналітичному модулі.

Система складається з кількох функціональних модулів, які представлені у таблиці 3.

У рамках цієї статті було реалізовано обмежену версію інформаційної системи, яка виконує базову перевірку мультимедійного контенту. Поточна реалізація дозволяє завантажити зображення через веб-інтерфейс і автоматично перевірити його на наявність копій серед захищених зображень, що зберігаються в базі даних (у вигляді файлів у папці).

Перевірка відбувається шляхом обчислення хешу SHA-256 для завантаженого зображення та порівняння з гешами зображень, які вже є в системі. Якщо хеші збігаються – виводиться повідомлення про виявлене порушення.

У повнофункціональній версії системи, що розробляється в межах дисертаційного дослідження, передбачається:

- зберігання лише хеш-значень у структурованій базі даних (наприклад, MySQL або PostgreSQL);
- можливість періодичного автоматичного сканування рекламних платформ на наявність збігів;
- інтеграція з системами цифрових водяних знаків та ліцензійного контролю.

Формальний опис алгоритму.

1. Хеш-функція $SHA-256$. Для кожного зображення I обчислюється хеш:

$$H(I) = SHA256(I), \quad (1)$$

де $H(I)$ – хеш-значення зображення, I – бінарне представлення файлу зображення.

2. Алгоритм порівняння. Зображення I_{new} вважається збігом, якщо існує зображення I_j у базі, для якого:

$$H(I_{new}) = H(I_j), \quad (2)$$

де $H(I_{new})$ – хеш завантаженого зображення, $H(I_j) \in H_{1070}$ – множина хешів у базі.

Програмна реалізація мовою PHP.

З метою демонстрації принципу роботи було розроблено прототип системи перевірки мультимедійних даних на предмет можливого пору-

шення авторських прав. Реалізація виконана мовою PHP із використанням вбудованих функцій для роботи з файлами, хеш-алгоритмами та метаданими [22, 23].

У фрагменті коду, представленому нижче (див. Рисунок 1), можна побачити використання наступних функцій:

- `fileHash()` – для обчислення унікального гешу кожного завантаженого зображення;
- `detectFileType()` – для визначення типу файлу (зображення, відео, аудіо тощо);
- `getImageMetadata()` – для отримання метаданих із графічних файлів;
- `move_uploaded_file()` – для безпечного збереження завантажених файлів на сервері;
- `scandir()` – для отримання списку зображень у локальній «базі авторських зразків».

Завдяки цим функціям система автоматично аналізує кожен медіафайл одразу після завантаження, порівнюючи його хеш із тими, що вже зберігаються у базі. Якщо відбувається збіг – виводиться попередження про можливе порушення прав інтелектуальної власності.

Інтерфейс та тестування системи. Користувачський інтерфейс реалізовано у вигляді веб-сторінки з сучасним оформленням: форма завантаження файлу розміщена в центрі екрана

```

61 function detectFileType($file_path) {
62     $finfo = finfo_open(FILEINFO_MIME_TYPE);
63     return finfo_file($finfo, $file_path);
64 }
65
66 function fileHash($file_path) {
67     return hash_file('sha256', $file_path);
68 }
69
70 function getImageMetadata($file_path) {
71     $meta = @exif_read_data($file_path, 0, true);
72     if (!$meta) return "Метадані недоступні або відсутні.";
73     $output = "<ul>";
74     foreach ($meta as $section => $values) {
75         foreach ($values as $key => $val) {
76             $output .= "<li><strong>$section.$key:</strong> " . htmlspecialchars($val) . "
77             </li>";
78         }
79     }
80     $output .= "</ul>";
81     return $output;
82 }
83
84 if ($_SERVER['REQUEST_METHOD'] === 'POST' && isset($_FILES['media'])) {
85     $file = $_FILES['media'];
86     $filename = basename($file['name']);
87     $target_path = $upload_dir . $filename;
88
89     if (move_uploaded_file($file['tmp_name'], $target_path)) {
90         echo "<h3>✔ Файл завантажено: $filename</h3>";
91         $mime = detectFileType($target_path);
92         echo "<p><strong>Тип файлу:</strong> $mime</p>";
93         echo "<img src='uploads/$filename' alt='Зображення' class='preview'>";
94     }
95 }

```

Рис. 1. Програмний код (частина)

(Рисунок 2), а результати перевірки виводяться горизонтально – зображення ліворуч, метадані праворуч. Такий підхід дозволяє зручно робити скріншоти результатів перевірки для подальшого документування чи демонстрації.

Для тестування системи у папку з базою зображень було попередньо завантажено три графічних файли. Далі були виконані наступні кроки:

1. Перший **тест**: завантажено зображення, яке не входить до бази.

- Система успішно прийняла файл;
- Тип контенту був визначений як image/jpeg;
- Хеш не збігся з жодним у базі – система повідомила, що порушень не виявлено;
- Було зроблено скріншот результату (див. Рисунок 3).

2. Другий тест: завантажено зображення, яке є у базі.

– Система також розпізнала файл як зображення;

– Хеш збігся з одним із наявних у папці image_db;

– Виведено попередження про збіг і показано зображення, з яким знайдено збіг;

– Виконано додатковий скріншот інтерфейсу (див. Рисунок 4).

Кожна дія супроводжувалася логуванням у файл check_log.txt із вказанням часу, імені файлу, типу і результату перевірки.

Запропонована система має низку суттєвих переваг, що роблять її цінним інструментом для фахівців у сфері реклами та управління інтелектуальною власністю. По-перше, вона забезпечує превентивний захист – тобто дозволяє виявляти можливі порушення ще на етапі підготовки рекламних матеріалів, до їх публікації. Це дозво-

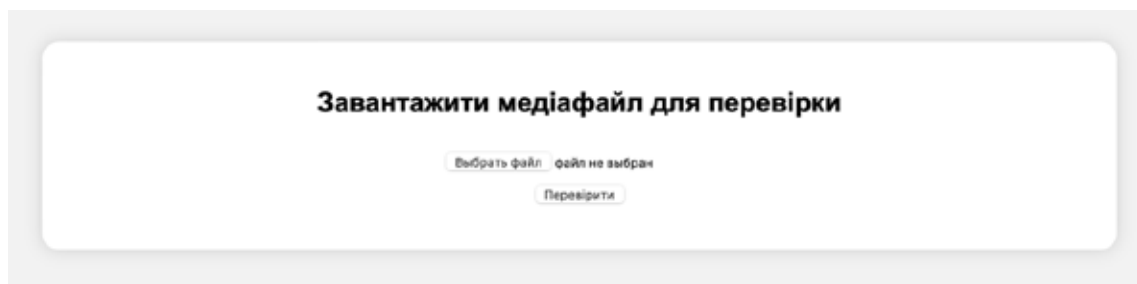


Рис. 2. Інтерфейс системи

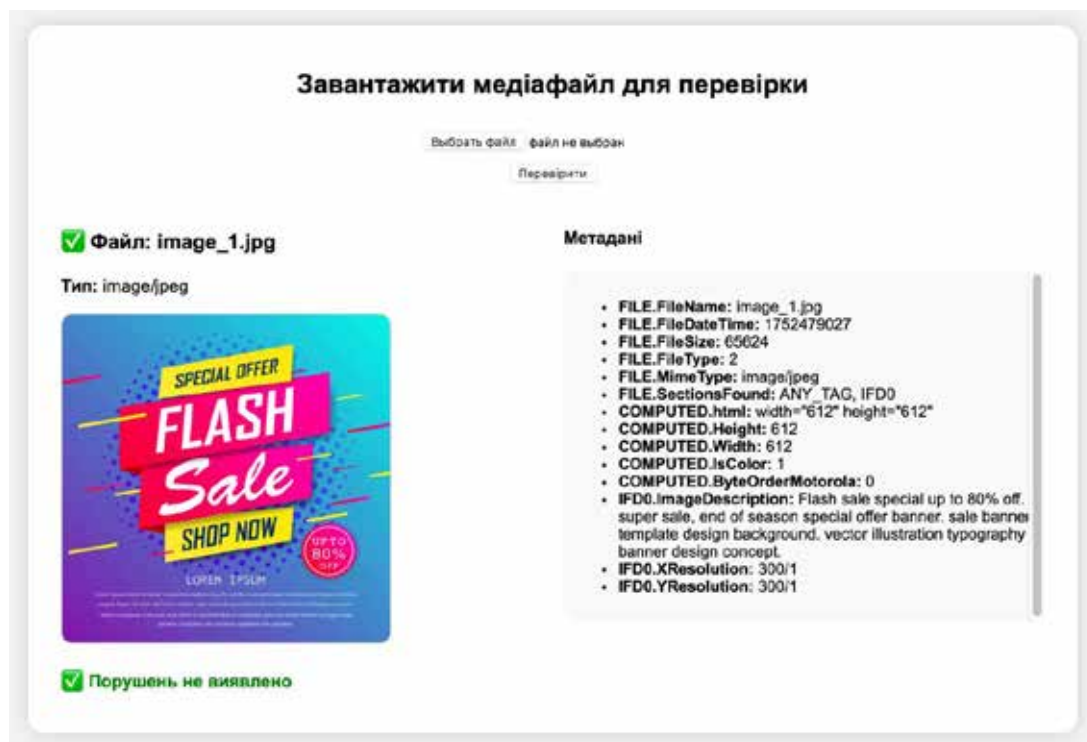


Рис. 3. Результат роботи, коли файл не порушує авторських прав

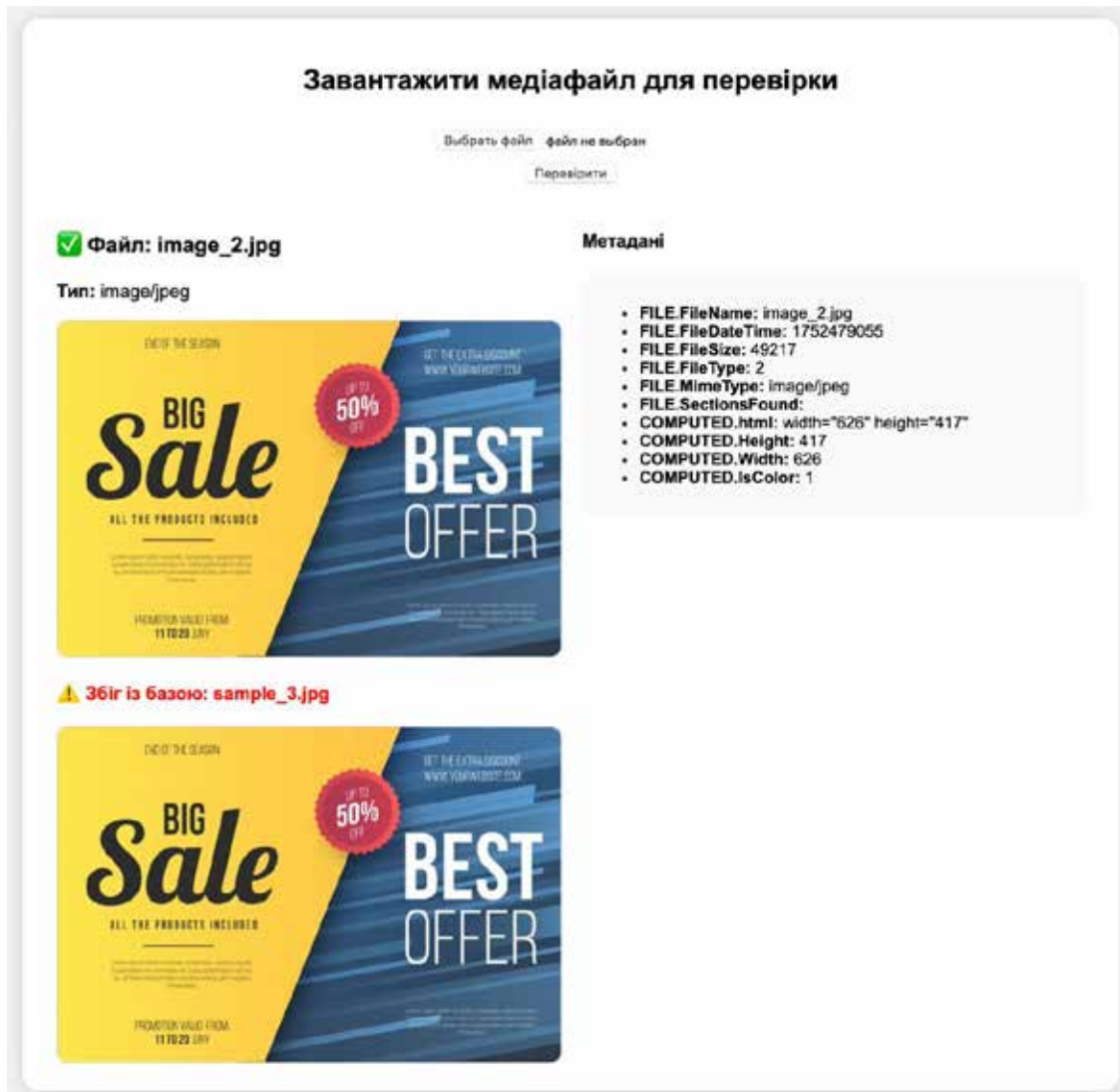


Рис. 4. Результат роботи коли файл порушує авторські права

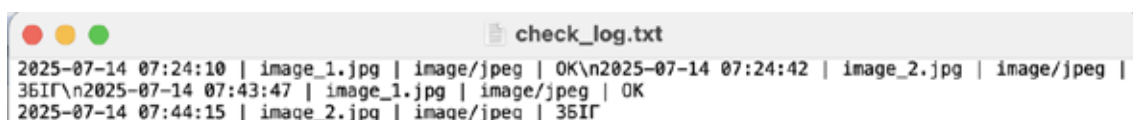


Рис. 5. Результати роботи у лог файлі

ляє уникнути конфліктів з правовласниками, блокування контенту або санкцій з боку рекламних платформ [24, 25].

По-друге, використання такої системи суттєво зменшує як репутаційні, так і фінансові ризики. Компанія, яка демонструє дотримання авторського права, формує довіру у споживачів, партнерів і регуляторів. У свою чергу, уникнення штрафів або судових витрат має прямий економічний ефект.

Третьою важливою перевагою є автоматизація рутинних процесів – система самостійно виконує

перевірки контенту, що значно спрощує роботу маркетологів, дизайнерів і контент-менеджерів. Вони можуть зосередитися на креативній складовій, не витрачаючи час на юридичну перевірку вручну.

І нарешті, впровадження подібної системи сприяє формуванню правової культури у команді. Постійна взаємодія з інструментом, який нагадує про авторські права та ліцензії, поступово підвищує обізнаність працівників і зміцнює етичні стандарти роботи з контентом [26].

Приклад використання системи у рекламному менеджменті

Щоб продемонструвати ефективність запропонованої інформаційної системи, розглянемо умовний кейс запуску рекламної кампанії для нового продукту в соціальних мережах.

Вхідні умови. Рекламне агентство готує відеоролик для Instagram і Facebook Ads, у якому використано: фонову музику з відкритого джерела; стокове зображення для обкладинки; короткий відеофрагмент, знятий фрилансером.

Перед запуском кампанії весь контент завантажувється у внутрішню систему перевірки авторських прав.

Результати перевірки.

Фонова музика – виявлено, що файл є уривком треку, захищеного авторським правом і зареєстрованого у YouTube Content ID. Система пропонує замінити його на безкоштовний трек із ліцензією Creative Commons або придбати офіційну ліцензію.

Зображення обкладинки – виявлено, що воно належить Getty Images і потребує ліцензування. Система видає попередження та посилання на умови використання. Саме цю частину роботи системи продемонстровано у цій статті.

Відеофрагмент від фрилансера – відсутні водяні знаки або метадані про авторство. Система пропонує додати водяний знак і зафіксувати факт створення у базі, щоб уникнути майбутніх суперечок.

4.3. Рішення

Після повідомлення:

- агентство замінює музику на ліцензійну;
- купує зображення із правами на комерційне використання;
- додає цифровий підпис на власне відео та реєструє його в базі внутрішніх креативів.

Після цього кампанія успішно проходить модерацию в Meta Ads і запускається без юридичних ризиків.

Висновки. У статті розглянуто одну з найбільш актуальних проблем сучасного цифрового середовища – порушення прав інтелектуальної

власності під час використання мультимедійного контенту в інтернет-рекламі. Проаналізовано, як необережне або навмисне застосування зображень, відео, музики чи концептуальних елементів без належного ліцензування може призвести до юридичних наслідків: фінансових штрафів, блокування рекламних оголошень, видалення акаунтів і суттєвих репутаційних втрат для бренду.

Запропоновано концепцію інформаційної системи, яка забезпечує превентивний моніторинг і контроль за правомірністю використання мультимедійних об'єктів у рекламних кампаніях. Ключовою перевагою системи є її здатність автоматично перевіряти завантажений контент шляхом обчислення геш-значень та їх порівняння з базою відомих об'єктів інтелектуальної власності. Це дозволяє виявляти потенційні порушення до моменту публікації креативів, що мінімізує ризики й витрати на виправлення ситуацій постфактум.

Розроблений прототип системи демонструє можливість інтеграції таких механізмів, як аналіз метаданих файлів, використання цифрових водяних знаків, логування усіх операцій перевірки, формування звітності для маркетологів та керівників.

Система також має значний освітній ефект: її впровадження сприяє формуванню правової культури в команді, підвищуючи обізнаність про важливість дотримання авторського права.

У перспективі подальших досліджень доцільно зосередитися на:

- розробці повнофункціональних модулів для роботи з відео- та аудіофайлами;
- інтеграції з API популярних рекламних платформ (Meta Ads, Google Ads, TikTok Ads);
- створенні централізованої бази цифрових підписів правовласників;
- використанні штучного інтелекту для семантичного аналізу креативів.

Таким чином, реалізація подібної системи може стати важливим кроком у розвитку відповідального рекламного середовища, де повага до інтелектуальної власності є не лише юридичним обов'язком, а й етичною нормою.

Список літератури:

1. Про авторське право і суміжні права: Закон України від 01.12.2022 р. № 2811-IX. *Верховна Рада України*. Дата оновлення: 15.04.2023 р. URL: <https://zakon.rada.gov.ua/laws/show/2811-20#Text> (дата звернення: 05.07.2025).
2. Про захист від недобросовісної конкуренції: Закон України, чинна редакція від 16.10.2020. URL: <https://zakon.rada.gov.ua/laws/show/236/96-%D0%B2%D1%80#Text> (дата звернення: 05.07.2025).
3. Про рекламу: Закон України, чинна редакція від 27.04.2024 р. URL: <https://zakon.rada.gov.ua/laws/show/270/96-%D0%B2%D1%80#Text> (дата звернення: 05.07.2025).

4. Ульянова Г. О. Захист авторських прав у правовідносинах, які виникають у зв'язку з рекламою : дис. – ступеня канд. юрид. наук: спец. 12.00. 03 «Цивільне право і цивільний процес, 2008.
5. Ульянова Г. О. Захист прав творців реклами від неправомірного використання рекламних матеріалів. *Наукові праці Національного університету Одеська юридична академія*. 2012. №. 12. С. 360-369. URL: <http://naukovipraci.nuoua.od.ua/arhiv/tom12/37.pdf> (дата звернення: 05.07.2025).
6. Краузе О. І., Піняк І. Л., Зяйлик М. Ф. Нормативно-правове регулювання рекламної діяльності: охорона авторського права. *Галицький економічний вісник*. 2024. Т. 88. №. 3. С. 41-49. DOI: https://doi.org/10.33108/galicianvisnyk_tntu2024.03.041
7. Горбаченко С. А. Особливості рекламних кампаній продуктів та послуг у сфері кібербезпеки. Конкурентоспроможна модель інноваційного розвитку економіки України: матеріали VII Міжнар. наук.-практ. конф., м. Кропивницький, 07–08 лист. 2024 р. / М-во освіти і науки України, Центральноукраїн. нац. техн. ун-т. Кропивницький: ЦНТУ, 2024. 410 с. С. 52.
8. Федорова Н. Реклама як об'єкт авторського права. *Теорія і практика інтелектуальної власності*. 2022. № 2. С. 33-38. DOI: <https://doi.org/10.33731/22022.259742> URL: <http://uran.inprojournal.org/article/view/259742> (дата звернення: 05.07.2025).
9. Горбаченко С. А. ІТ-маркетинг : навчальний посібник. Одеса, 2024. 218 с. URL: <https://hdl.handle.net/11300/27929> ; DOI: 10.61563/978-617-8182-04-5
10. Dimita G. WIPO Committee on Development and Intellectual Property. URL: <https://surl.lu/ursfyn>
11. Jorgenson L., Fink C. WIPO's contributions to international cooperation on intellectual property. *Journal of International Economic Law*. 2023. Vol. 26. No 1. P. 30-34. DOI: <https://doi.org/10.1093/jiel/jgac049>
12. Kadian P., Arora S. M., Arora N. Robust digital watermarking techniques for copyright protection of digital data: A survey / *Wireless Personal Communications*. 2021. Vol. 118. No 4. P. 3225-3249.
13. YouTube Help Center, Content ID. URL: <https://support.google.com/youtube/answer/2797370?hl=en> (дата звернення: 01.07.2025).
14. Meta Rights Manager. Meta for creators. URL: <https://creators.facebook.com/tools/rights-manager/> (дата звернення: 01.07.2025).
15. Lim J. H. et al. Protect, show, attend and tell: Empowering image captioning models with ownership protection. *Pattern Recognition*. 2022. Vol. 122. P. 108285. DOI: <https://doi.org/10.1016/j.patcog.2021.108285>
16. Li, J. Impact of Digital Marketing on Consumer Behavior. *SHS Web of Conferences*. 2024. Vol. 207. P. 01026. <https://doi.org/10.1051/shsconf/202420701026>
17. Бутнік-Сіверський О. Б. и др. До питання захисту авторських прав у цифровому середовищі. Наукові дослідження та інновації в галузі суспільно-гуманітарних наук: збірник матеріалів І Всеукраїнської науково-практичної інтернет-конференції (м. Мелітополь, 24 листопада 2021 р.). ТДАТУ: ред. кол. Ломейко ОП, Єременко ОА, Михайлов В. В [та ін.]. Ч. 1. Мелітополь: ТДАТУ, 2021. 581 с. С. 504.
18. Bedratyi I., Shevchshyn O., Astakhov V. Правові аспекти порушення прав інтелектуальної власності в цифровому середовищі: перспективи рекодифікації. *Social Development: Economic and Legal Issues*. 2025. Is. 4. DOI: <https://doi.org/10.70651/3083-6018/2025.4.09>
19. Клименко О. С., Христищенко Н. П. Особливості юридичної відповідальності за порушення прав інтелектуальної власності в цифровому середовищі. *Академічні візії*. 2025. №. 42. 10 с. DOI: <https://doi.org/10.5281/zenodo.15170759> URL: <https://academy-vision.org/index.php/av/article/view/1798>
20. Пуздрановський І. В., Салієва О. В. Аналіз сучасних DRM-систем керування цифровими правами : дис. ВНТУ, 2023.
21. Овчинников М. Ю., Слатвінська В. М. Технічний та нормативно-правовий аспекти захисту об'єктів інтелектуальної власності на ринку аудіо-візуальної продукції. *Вісник Херсонського національного технічного університету*. 2024. №. 3 (90). С. 259-270. DOI: <https://doi.org/10.35546/kntu2078-4481.2024.3.32>
22. Овчинников М. Ю. Додавання та обробка звуку в метаданих файлів за допомогою PHP. *Наука і техніка сьогодні*. 2025. №. 4 (45). С. 1332-1344. DOI: [https://doi.org/10.52058/2786-6025-2025-4\(45\)-1332-1344](https://doi.org/10.52058/2786-6025-2025-4(45)-1332-1344)
23. Маханьов Р. О., Кобилін О. А. Дослідження впливу метаданих для розпізнавання та класифікації зображень. *Computer-integrated technologies of automation of technological processes*. 2024. С. 67.
24. Золотар, А. С. Захист прав на інтелектуальну власність в цифровому середовищі : дис. ... д-ра філософії : 081 Право / Золотар Аліна Сергіївна. – Київ, 2023. – 193 с. URI: <https://ela.kpi.ua/handle/123456789/63213>
25. Булеца С. Б., Чепис О. І. Авторське право в умовах цифрової трансформації. Актуальні проблеми інтелектуальної власності в Україні та Європейському Союзі: матеріали V міжнародної науково-практичної конференції. Будапешт. 2021. С. 20-25.
26. Ковальова В. В. До питання формування правової культури майбутніх менеджерів з адміністративної діяльності. *Наукові записки. Серія «Психолого-педагогічні науки»*. 2021. № 1. С. 125-130. URI: <http://lib.ndu.edu.ua/dspace/bitstream/123456789/2063/1/17.pdf>.

Gorbachenko S.A., Ovchinnikov M.Yu. INTELLECTUAL PROPERTY PROTECTION IN DIGITAL CONTENT: CHALLENGES AND SOLUTIONS FOR ADVERTISING MANAGEMENT

The article examines the issue of intellectual property rights infringement when using multimedia content in online advertising, which is becoming particularly relevant in the context of its rapid growth. In the modern digital environment, advertising is one of the most dynamic industries, using a variety of formats – banners, videos, interactive creatives, native advertising, AR filters, short videos for TikTok and Instagram Reels, as well as email newsletters with embedded media. All of these formats actively use images, videos, music, 3D models and voiceovers, which are often objects of copyright. The main typical violations associated with the use of unlicensed content are analyzed, as well as the risks they pose for advertisers and agencies: financial sanctions, blocking of content on platforms, reputational losses. Advertising activities in the modern digital environment are constantly becoming more complicated and multi-level, which necessitates the revision and clarification of basic legislative definitions, in particular in the field of copyright. In the context of the active use of multimedia content in advertising, it is important not only to correctly determine the rights to copyright objects, but also to improve the legal mechanisms for their disposal in order to provide protection against possible violations and provide legal certainty for advertisers. The concept of an information system is proposed, which allows for automated monitoring of copyright compliance in advertising campaigns. The architecture of the system and the principles of its operation using AI-monitoring, digital watermarks and license bases are described. An example of the practical application of the system within the framework of advertising management is given. The article is aimed at specialists in the field of digital marketing, advertising and information technologies who seek to ensure the legal security of the use of multimedia content.

Key words: audiovisual works, copyright, advertising, advertising management, computer science, programming, web development, metadata, steganography, digital watermark, neural networks, information protection.

Дата надходження статті: 21.07.2025

Дата прийняття статті: 28.07.2025

Опубліковано: 27.10.2025

Грезіна О.М.

Національний університет «Одеська юридична академія»

Гурін Є.П.

Національний університет «Одеська юридична академія»

КІБЕРБЕЗПЕКА МОДУЛІВ JAVASCRIPT-ЗАСТОСУНКІВ

У статті розглянуто ключові проблеми кібербезпеки, притаманні модульним архітектурам JavaScript-застосунків, зокрема CommonJS та ECMAScript Modules (ESM). Аналізується модульність, яка дозволяє масштабувати й прискорювати розробку, водночас створює нові вектори атак: ін'єкційні вразливості, *relative path attacks*, компрометацію залежностей та атаки на ланцюги постачання (*supply chain attacks*). Підкреслено, що сучасна екосистема JavaScript значною мірою залежить від сторонніх бібліотек і плагінів, що ускладнює аудит та підвищує ризики появи шкідливого коду, зокрема у перехресних залежностях. Проведено порівняльний аналіз CommonJS та ESM з погляду стійкості до архітектурних атак: динамічний синтаксис *require* у CommonJS збільшує ризики ін'єкцій, тоді як статичний *import* у ESM спрощує аудит і підвищує безпеку.

Описано інструменти та методи зниження ризиків: статичний і динамічний аналіз, трасування виконання, перевірка хеш-сум і цифрових підписів, а також застосування Zero Trust-підходів. Особливу увагу приділено інтеграції засобів безпеки у CI/CD-процеси та DevOps-середовище: автоматизовані сканери (*npm audit*, *Snyk*, *Trivy*), контроль змін через *git commit signing*, політики доступу та *sandbox-оточення*. Наведено приклади реальних інцидентів компрометації відомих бібліотек (*event-stream*, *ua-parser-js*), що підкреслюють вразливість навіть широко використовуваних модулів.

У висновках зазначено, що безпека модульних систем має стати невід'ємною частиною розробки. Рекомендовано застосовувати багаторівневий підхід до аудиту залежностей, впроваджувати політики «нульової довіри» та системно інтегрувати безпекові практики у життєвий цикл проєкту. Це дозволяє мінімізувати ризики атак на ланцюги постачання та підвищити надійність і стійкість JavaScript-застосунків у складному середовищі сучасної веброзробки.

Ключові слова: кібербезпека, JavaScript-модулі, JavaScript, архітектура програмного забезпечення, програмне середовище, веброзробка, CommonJS, ECMAScript Modules (ESM), аудит безпеки, інструменти безпеки.

Постановка проблеми. В екосистемі JavaScript-застосунків модульність стала стандартом проєктування, що дозволяє масштабувати проєкти, підвищувати повторне використання коду та пришвидшувати розробку. Проте, одночасно із зростанням кількості сторонніх залежностей, що підключається через пакетні менеджери (*npm*, *yarn*), суттєво зростає і поверхня потенційної атаки. Модульні архітектури, зокрема CommonJS та ECMAScript Modules (ESM), мають як технічні переваги, так і структурні вразливості, що уможливають такі атаки, як ін'єкції коду, атаки через відносні шляхи, компрометація залежностей та атаки на ланцюги постачання (*supply chain attacks*). Попри наявність численних інструментів для аналізу безпеки, проблема комплексної кібербезпеки JavaScript-модулів залишається відкри-

тою. Особливо складним є контроль перехресних залежностей, відсутність належного контролю цілісності модулів, складність поєднання старих і нових архітектур (CommonJS та ECM).

Аналіз останніх досліджень і публікацій. Науковий дискурс приділяє значну увагу проблемам кібербезпеки, що виникає у процесі використання JavaScript-модулів у веброзробці. В деяких працях розглянуто принципи модульності та виявлено основні вектори атак, зокрема ін'єкційні вразливості, ризики, пов'язані з неконтрольованим використанням директив *require/import*, а також зростання кількості атак на ланцюги постачання (*supply chain attacks*). Зокрема, у матеріалах OWASP здійснено систематизацію типів атак, пов'язаних із використанням модульних залежностей, таких як атаки по відносному

шляху (Path Traversal) [8]. У звітах компаній Trend Micro та інших дослідницьких організацій наведено приклади компрометації екосистеми *npm*, зокрема випадки впровадження шкідливого коду до широко використовуваних пакетів *event-stream* і *ua-parser-js* [9], [12]. Офіційна документація Node.js, а також аналітичні публікації AppSignal містять порівняння архітектур CommonJS та ECMAScript Modules (ESM), зосереджуючись на технічних особливостях, синтаксисі та механізмах імпорту [5], [6], [7]. Водночас, питання безпеки таких архітектур, зокрема уразливості на рівні реалізації модулів, здебільшого залишаються поза увагою. Таким чином, аналіз актуальних джерел дозволяє виділити кілька невирішених проблем, а саме відсутність цілісної моделі забезпечення безпеки модулів JavaScript у середовищі з великою кількістю зовнішніх залежностей; брак порівняльного аналізу CommonJS і ESM з погляду їхньої стійкості до атак архітектурного рівня; нехтування контекстом інтеграції безпекових практик у DevOps-середовище і CI/CD-процеси.

Постановка завдання. Метою статті є аналіз ключових загроз безпеки, притаманних JavaScript-модулям, порівняння архітектур CommonJS і ESM з погляду кібербезпеки, та огляд інструментів, що дозволяють зменшити ці ризики.

Виклад основного матеріалу. JavaScript є саме тією мовою програмування, якою розробляється значна частина сучасних вебзастосунків. Вона використовується у всіх частинах застосунку – як у клієнтській (frontend), де вона є абсолютним монополістом, адже інші мови програмування не використовуються на стороні клієнта, так і у серверній (backend), конкуруючи з іншими популярними мовами, такими як Java, C#, PHP тощо. Будучи універсальним інструментом для розробки, екосистема JavaScript має численну кількість додаткових плагінів, бібліотек та утиліт, які пришвидшують процес розробки та спрощують масштабізацію проєкту [1].

Модульність є архітектурним підходом, що включає в себе розподілення програмного коду проєкту на логічні, ізольовані блоки коду – модулі, що можуть повторно використовуватись та масштабуватись. Використання сторонніх модулів у сучасних застосунках є невід’ємною складовою будь-якого проєкту [4]. IT-ринок вже відійшов від «ванільних» засобів розробки, надаючи перевагу готовим рішенням, які використовуються у проєкті одразу після імпорту, надаючи готові інструменти для розробника. В інфраструктурі JavaScript є декілька популярних модульних систем, які роз-

вивались відповідно до потреб спільноти розробників та середовища виконання. Найбільш поширеними є CommonJS [5], та ECMAScript Modules (ESM) [5].

CommonJS є стандартом для серверного JavaScript, і є основною системою в середовищі Node.js, залишаючись домінуючою в багатьох серверних архітектурах. Специфіка даної архітектури полягає у синхронному підході до розробки та одноразовому завантаженні модуля безпосередньо в момент його виклику, що є ефективним підходом для серверних застосунків, але не оптимальним для браузеру.

ECMAScript Modules (ESM) – інша модульна система, що є офіційною архітектурою JavaScript та затверджена стандартом ES6. Є універсальною системою, що розроблена з урахуванням вимог як серверної архітектури, так і браузерної. На відміну від CommonJS, орієнтована на асинхронний підхід до розробки та статичний аналіз імпортів перед їх завантаженням безпосередньо у проєкт. Проте така універсальність має й свої недоліки – архітектура налаштовується складніше, ніж у CommonJS, а синтаксис є менш читабельним. Також слід зазначити, що ESM є сучаснішою модульною системою.

У реальному програмному середовищі вибір модульної системи є не лише питанням зручності, але й подальшої стратегії розробки, що визначає архітектуру застосунку, сумісність із зовнішніми залежностями, продуктивність виконання коду, а також можливості аудиту безпеки проєкту.

У Node.js (до 14 версії) домінуючим стандартом залишався CommonJS, що був інтегрований у саму структуру середовища. Однак із появою ESM, як офіційного стандарту модульної архітектури ES6, Node.js почав підтримувати модулі ES6, використовуючи ключову директиву `package.json`, дозволяючи використовувати `import/export` у JS-файлах, проте накладаються й інші обмеження:

- неможливість поєднання синтаксисів: `require(...)` та `import` неможливо використовувати в одному проєкті;
- потреба у розширенні `*.mjs` – модульного файлу `javascript`;
- відсутність автоматичного резолвінгу директорій.

Саме через ці обмеження перехід до ESM-архітектури є ускладненим – проєктам, що створені на базі CommonJS-архітектури, необхідна зворотна сумісність з ESM [7]. Тому перехід є поступовим, і доволі часто потребує проміжних

рішень (використання `webpack`-модулів, трансляція через `Babel` та ін.).

У браузерях ESM-архітектура є очевидним вибором, завдяки нативній підтримці більшістю сучасних браузерів проєкт не вимагає додатково налаштування, а модуль оголошується через тег з відповідним типом файлу:

```
<script type=»module» src=»module.js»></script>
```

Не дивлячись на усі переваги модульної структури, така система розширює потенційну поверхню атаки. У сучасній екосистемі JavaScript-застосунків, яка базується інтенсивному використанні сторонніх пакетів, основними джерелами загроз виступають як вразливості у внутрішньому коді модулів, так і компрометація зовнішніх залежностей. Одним з найпоширеніших видів загроз є ін'єкційні атаки (*injection attacks*), які реалізуються через недбале використання динамічних імпортів, інтерполяцій у шляхи до модулів або відсутності валідації вхідних даних. Вони можуть призвести до виконання довільного коду, підміни модулів або несанкціонованого доступу до середовища виконання. У модульній архітектурі особливо критичними є *relative path attacks* [7] (атаки по відносному шляху), коли шлях до модуля формується з урахуванням введених користувачем даних, що дає змогу спрямувати *import* або *require()* на шкідливий файл.

Іншою загрозою, яка набула вже системного характеру у сучасних вебзастосунках, є *supply chain attacks* (атаки на ланцюги постачання ПЗ). У контексті JavaScript це виражається в інтеграції у проєкт залежностей з модульних менеджерів (наприклад *npm*), де можуть бути опубліковані зловмисні пакети, що імітують безпечні, або ж відбувається компрометація вже існуючих. Відомі випадки, коли автори популярних бібліотек передавали контроль над проєктам третім особам, які згодом впроваджували бекдори, криптомайнери чи алгоритми витоку даних [9].

Окрему категорію становлять вразливості пов'язані з «довіреним» кодом – випадки, коли модуль або залежність розміщена у внутрішньому коді, але реалізовує ризиковану логіку: небезпечні виклики через *eval()*, доступ до глобальних об'єктів (наприклад *global*, *window*), а також спроби обходу політик *CORS* та *CSP*. Такі дії ускладнюють застосування механізмів *sandboxing* та унеможливають гарантування контрольованого середовища виконання.

Крім того, все більшої актуальності набувають перехресні залежності (*transitive dependencies*), ті, що імпортуються не напряму, а через інші модулі.

Саме на цьому рівні часто приховується шкідливий код, оскільки аудит охоплює лише поверхневу частину дерева залежностей. Це створює передумови для атак типу «*dependency confusion*», коли завантажується у відкритий реєстр модуль з ідентичним ім'ям, що використовується в приватному середовищі, змушуючи систему завантажити саме фальшивий модуль.

Вразливості можуть виникати через відсутність контролю над життєвим циклом модулів, зокрема у випадку, коли не розробляється механізм верифікації цілісності, або коли дозволено імпорт модулів з небезпечного джерела (наприклад через *Content Delivery Network (CDN)* без *Transfer Layer Security (TLS)*). У сукупності ці фактори вимагають від розробників не лише технічної обізнаності, а й системного підходу до менеджменту залежностей та інтеграції інструментів аудиту безпосередньо у процес розробки. Забезпечення безпеки у JavaScript-модулях не може обмежуватись лише вибором модульної системи. Воно повинне охоплювати багаторівневу перевірку залежностей, застосування захисних шаблонів проєктування, жорстке управління середовищем виконання та регулярний аналіз коду.

Враховуючи складність сучасних JavaScript-застосунків та їх залежність від зовнішніх модулів, процес розробки не може починатись без аналізу безпеки компонентів проєкту. Визначення потенційних вразливостей, бекдорів чи шкідливої поведінки у модулях вимагає системного підходу, який охоплює як автоматизовані інструменти, так і ручні техніки перевірки. Серед методів аналізу вирізняють статичний та динамічний аналіз.

Статичний аналіз передбачає дослідження вихідного коду модулів без їх безпосереднього виконання. Застосування такого методу дає змогу виявити вразливості на ранніх етапах, зокрема небезпечні конструкції (*eval*, *Function*, динамічні імпорти), порушення встановлених політик безпеки або порушення *best practices*. Інструменти типу *ESLint*, *SonarQube*, *Semgrep* забезпечують високий рівень точності при виявленні синтаксичних аномалій, відсутності перевірок на вхідні дані та потенційного витоку інформації. Однак недоліком статичного методу є неможливість виявлення поведінкових загроз, що активуються лише в час виконання, або вразливостей, прихованих у перехресних залежностях. Саме тому статичний аналіз доцільно поєднувати з іншими техніками.

Динамічний аналіз включає запуск модулів у контрольованому середовищі (наприклад *sandbox* або віртуальна машина) з метою моні-

торингу їхньої фактичної поведінки. Аналізуються виклики до операційної системи, мережеві з'єднання, файловий ввід/вивід, звернення до глобального контексту тощо. У результаті аналізу можливо виявити спроби з'єднання з C&C-сервером, приховане завантаження зовнішніх скриптів, несанкціоновані модифікації DOM, або внутрішніх елементів застосунку.

Аналіз залежностей полягає у створенні відповідного графа імпорту з подальшою перевіркою кожного вузла на наявність вразливостей, підозрілої активності, або сумнівного походження. Такі інструменти як *npm audit*, *Snyk*, *Retire.js*, *OSSIndex*, *Dependabot* автоматизують цей процес, здійснюючи порівняння з публічними базами вразливостей. Попри доведену ефективність, навіть такі технології не можуть повною мірою гарантувати відсутність небезпечного коду, оскільки не здійснюють повного семантичного аналізу, а покладаються на сигнатурні бази – такий підхід пов'язаний з потребою у швидкодії цих інструментів, проводячи повний семантичний аналіз, кожна перевірка займала б значну кількість часу.

Трасування виконання (*runtime tracing*) є однією з технік аналізу застосунку, функціонал якої полягає у відображенні роботи програми у детальних логах виконання коду, включаючи інформацію про завантажені модулі, порядок викликів, передані параметри та обробку помилок. Цей підхід забезпечує детальний інсайт у внутрішню логіку модуля, дозволяючи ідентифікувати потенційні приховані залежності, які можуть мати шкідливий характер. У поєднанні з моніторингом системних подій, трасування дозволяє також виявляти спроби виконання шкідливого коду, що активується лише у специфічному середовищі або на конкретній платформі. У цьому аспекті особливу роль відіграють системи профілювання, такі як *Chrome DevTools* або *Node.js Inspector*, які забезпечують самоконтроль застосунку в процесі виконання програми [10].

Важливим методом захисту є перевірка цілісності модулів шляхом зіставлення хеш-сум [10] або цифрових підписів. Це дозволяє визначити, чи було модифіковано код після завантаження, а також унеможливити використання підроблених або змінених версій коду. У сучасних інфраструктурах CI/CD цей підхід реалізовується у через механізми перевірки lock-файлів (*package-lock.json* для *npm*, *yarn.lock* для *yarn*), або впровадження механізмів *checksum enforcement*, що блокує встановлення модулів, які не відповідають очікуваному хешу.

Комбіноване застосування методів, що згадані вище, забезпечує багатоаспектний аналіз безпеки модулів, що особливо важливо в контексті зростаючої кількості атак на ланцюги постачання. Побудова багаторівневої моделі верифікації – від статичної перевірки до динамічного моніторингу, дозволяє мінімізувати ризики компрометації застосунку через використання вразливих або залежностей, що передбачають у своєму функціоналі зловмисні дії.

Відкритість найпопулярнішої екосистеми для керування модулями *npm* робить застосунки більш вразливими до атак, що полягає у скомпрометованих та/або шкідливих залежностях. Однак основною загрозою залишається недбалість та некомпетентність розробника. Більшість таких атак є успішною саме через безвідповідальність при розробці застосунку.

Безпека модульних систем є ключовим аспектом у цілісності, конфіденційності та надійності програмного забезпечення. Кібербезпека залежностей є особливо актуально в умовах зростаючої складності клієнт-серверних архітектур і широкого використання сторонніх залежностей. Більш традиційна модульна система *CommonJS* характеризується низкою вразливостей, що зумовлені її динамічною природою. Зокрема використання функції *require()* для імпорту модулів створює різноманітні вектори можливих атак, пов'язані з ін'єкцією шкідливого коду через маніпуляції з шляхами до файлів або залежностей. Одним із факторів ризику є кешування модулів на глобальному рівні, притаманне *CommonJS* – інфраструктурі, що унеможливорює ізоляцію компонентів та створює додаткові умови для атак потворного використання об'єктів, але вже з модифікованим станом.

На противагу дещо застарілому підходу *CommonJS*, система *ESM* [2], виступаючи сучасним стандартом модульності, запроваджує статичний синтаксис *import/export*, що дозволяє інструментам аналізу коду ефективно визначати залежності ще до виконання програми, тим самим мінімізуючи ризики ін'єкції коду через динамічні компоненти. Також забезпечується краща ізоляція модулів і підтримку політик безпеки на рівні завантаження даних, зокрема завдяки *CSP* (*Content Security Policy*) та *sandbox*-механізмів.

Особливу небезпеку для обох архітектурних систем становлять для обох систем становлять атаки на ланцюги постачання (*supply chain attacks*), що реалізуються через компрометацію сторонніх бібліотек або публікацію та подальший

експорт у проєкти шкідливих бібліотек із назвами, що схожі на популярні бібліотеки. У контексті таких загроз принципово важливим є впровадження методів превентивного захисту, таких як:

- локальне «дзеркалювання» репозиторіїв
- використання інструменту аудиту залежностей (npm audit, snyk)
- фіксація версій через package-lock.json
- обмеження динамічного імпорту
- перевірка джерел.

Сучасна практика розробки має базуватись на парадигмі «безпечної за замовчуванням» модульності, з акцентом на декларативність, контрольованість і перевірюваність залежностей.

Одним з найяскравіших прикладів компрометації програмних модулів є інцидент з пакетом event-stream у 2018 році. Після того, як автор передавав права на підтримку пакету іншому розробнику, до коду було додано залежність flatmap-stream, яка містила у собі шкідливий код для викрадення криптовалют з гаманців, пов'язаних з проєктом Сорау. Компрометація залежностей залишалась непоміченою протягом кількох тижнів, встигнувши розповсюдитись open-source середовищем. Інший випадок стосувався бібліотеки ua-parser-js у 2021 році, яка була скомпрометована, після чого було впроваджено майнери криптовалют в офіційні релізи проєктів. Пакет мав мільйони завантажень щотижня, а отже, атака потенційно торкнулася широкого кола користувачів. Такі інциденти доводять, що навіть перевірені, широко використовувані модулі можуть стати джерелом критичних загроз, особливо в умовах відсутності механізмів багаторівневої перевірки довіри. Подібні атаки часто проходять через оновлення патчів чи міnorних версій, які автоматично встановлюються при слабкій конфігурації залежностей [14].

Переходячи до DevOps середовища, критично важливо інтегрувати безпеку в усі етапи життєвого циклу розробки – особливо актуальним це є для CI/CD пайплайнів. Для запобігання впровадженню шкідливих або вразливих модулів необхідним є застосування автоматизованих інструментів аналізу безпеки на етапі побудови та тестування. Такі засоби, як Snyk та інші згадані інструменти не лише надають можливість автоматично виявляти відомі уразливості, але й блокувати подальше розгортання застосунку у випадку порушення політик безпеки проєкту. Окрему роль відіграє також сканування контейнерів перед розгортанням: такі інструменти, як Trivy, Gype або Anchore дозво-

ляють виявити шкідливі залежності всередині Docker-контейнерів. Крім того, рекомендується використовувати сигнатуру перевірку вихідного коду (git commit signing) та механізми контролю змін (code owners, pull request rules) для унеможливлення несанкціонованих інтеграцій у проєкт. Виконуючи попередні рекомендації, CI/CD трансформується з інструмента автоматизації у потужний бар'єр, що стоїть між розробкою та експлуатацією проєкту і здатен стримувати загрози до моменту їх інтеграції в проєкт [12].

Переходячи до так званих політик розробки, у серйозних проєктах однозначно слід використовувати Zero Trust Policy (політика «нульової» довіри) [12]. У контексті JavaScript-розробки вони означають категоричну відмову від «довіри за-замовчуванням», навіть до коду з того самого проєкту, чи офіційних бібліотек. У рамках цієї парадигми кожен компонент – включно з npm-пакетами, CI/CD-інтеграціями та навіть скриптами в браузері – усе розглядається як потенційно небезпечний код. Технічно, Zero Trust [1] реалізується через обмеження прав виконання, суворої ізоляції модулів верифікацію джерел та мікро-сегментацію. Наприклад, розгортання в sandbox-оточеннях дозволяє обмежити доступ до файлової системи та мережі на рівні модуля. Ще одним принципом є «верифікація за кожним запитом» – перед використанням нової залежності має бути перевірена її цілісність, історія змін, кількість мейнтейнерів і наявність автоматизованих тестів. Застосування Zero Trust суттєво знижує ймовірність поширення атаки в разі успішної компрометації одного з модулів і створює додатковий рівень захисту при експлуатації вразливостей типу supply chain.

Висновки. Вибір архітектурного підходу для проєкту є одним з найважливіших рішень перед початком проєкту. Модульність, як архітектурний підхід, є одним з найпопулярніших у сучасній веброзробці. Такий підхід дозволяє розробити швидко та масштабовану систему, однак відкриває широкий спектр атак на ланцюги постачання програмного забезпечення проєкту. Залежність від тисяч сторонніх бібліотек, слабкий контроль змін у відкритих репозиторіях і висока швидкість релізів формують складне середовище, у якому навіть незначна уразливість може спричинити катастрофічні наслідки. У цьому контексті безпека JavaScript-модулів має розглядатися як невід'ємна частина розробницького процесу – з інтегрованим аудитом, скануванням уразливостей, політиками довіри та захистом на рівні CI/

CD. Особливої актуальності набувають підходи Zero Trust, які ставлять під сумнів будь-які припущення про безпеку за замовчуванням. Подальше вдосконалення інструментів аналізу, посилення

прозорості в open-source та впровадження нових стандартів формують основу для більш стійкої, контрольованої та безпечної модульної екосистеми у JavaScript.

Список літератури:

1. Williams A. T., Chen S. AI-Powered Policy Formulation and Enforcement in Zero Trust Frameworks. 2025 International Conference on Cyber Security and Data Science (CSDS). 2025. URL: <https://surl.li/pvrygq> DOI: <https://doi.org/10.1109/CSDS59213.2025.123456>
2. Zhong Y., Li T., Ding Z., Wang Y., Zhang X.-Y. An empirical study of bloated dependencies in CommonJS packages. arXiv preprint arXiv:2405.17939. 2024. DOI: <https://doi.org/10.48550/arXiv.2405.17939>
3. The Universal Language: JavaScript's Journey from Client-Side to Everywhere. URL: <https://surl.li/axhzdh>
4. JavaScript модулі. URL: <https://uk.javascript.info/modules>
5. CommonJS. URL: <https://nodejs.org/docs/latest/api/modules.html>
6. ECMAScript Modules. URL: <https://nodejs.org/docs/latest/api/esm.html>
7. Deep dive into CommonJS and ESM. URL: <https://surl.lu/gmdxko>
8. Path traversal. URL: <https://surl.li/mibsmnt>
9. Package hacking to steal bitcoin wallets. URL: <https://surl.lu/guvczn>
10. Node.js debugging. URL: <https://nodejs.org/en/learn/getting-started/debugging>
11. Secure Hash Algorithms. URL: https://en.wikipedia.org/wiki/Secure_Hash_Algorithms
12. CI/CD security cheat sheet. URL: <https://surl.lt/xeejyi>
13. Zero trust policy. URL: <https://surl.li/xrfrpf>
14. Compromised npm Package: event-stream. URL: <https://surl.li/ktmsop>

Hrezina O.M., Hurin Ye.P. CYBERSECURITY OF JAVASCRIPT APPLICATION MODULES

The article discusses key cybersecurity issues inherent in modular JavaScript application architectures, including CommonJS and ECMAScript Modules (ESM). It analyzes modularity, which allows for scalability and acceleration of development, while at the same time creating new attack vectors: injection vulnerabilities, relative path attacks, dependency compromise, and supply chain attacks. It emphasizes that the modern JavaScript ecosystem is largely dependent on third-party libraries and plugins, which complicates auditing and increases the risks of malicious code, particularly in cross-dependencies. A comparative analysis of CommonJS and ESM is conducted in terms of resistance to architectural attacks: the dynamic require syntax in CommonJS increases the risks of injections, while the static import in ESM simplifies auditing and increases security.

Tools and methods for reducing risks are described: static and dynamic analysis, execution tracing, verification of hash sums and digital signatures, as well as the use of Zero Trust approaches. Particular attention is paid to the integration of security tools into CI/CD processes and DevOps environments: automated scanners (npm audit, Snyk, Trivy), change control via git commit signing, access policies and sandbox environments. Examples of real incidents of compromise of well-known libraries (event-stream, ua-parser-js) are given, which emphasize the vulnerability of even widely used modules.

The conclusions indicate that the security of modular systems should become an integral part of development. It is recommended to apply a multi-level approach to dependency auditing, implement «zero trust» policies and systematically integrate security practices into the project life cycle. This allows you to minimize the risks of attacks on supply chains and increase the reliability and stability of JavaScript applications in the complex environment of modern web development.

Key words: cybersecurity, JavaScript modules, JavaScript, software architecture, programming environment, web development, CommonJS, ECMAScript Modules (ESM), security audit, security tools.

Дата надходження статті: 30.06.2025

Дата прийняття статті: 11.07.2025

Опубліковано: 27.10.2025

Дадиверін В.В.

Вінницький національний технічний університет

Бісікало О.В.

Вінницький національний технічний університет

АНАЛІЗ ТА ІМПЛЕМЕНТАЦІЯ ІДЕЇ НАВЧАННЯ ВЕЛИКИХ МОВНИХ МОДЕЛЕЙ ЗА АНАЛОГІЄЮ З ДИТЯЧИМ КОГНІТИВНИМ РОЗВИТКОМ

Дослідження присвячено проблематиці недостатньої ефективності великих мовних моделей у питаннях розв'язання математичних завдань, послідовного багатокрокового аналізу та міркування щодо можливих варіантів вирішення. У статті розкрито недосконалість сучасних LLM у питаннях послідовного мислення, попри їхню здатність до генерації природного та осмисленого тексту, вони досі залишаються суттєво обмеженими при вирішенні задач, які потребують логічного міркування або обґрунтування. З'ясовано, що такі обмеження нерідко є наслідком структури навчального процесу мовної моделі, бо той не має нічого спільного з процесами, які відбуваються під час формування когнітивних та ментальних моделей людини.

У цій роботі розглянуто підхід навчання за подібністю до когнітивного розвитку людини, який базується на поєднанні наступних концепцій: *Curriculum learning* – навчання з поетапним збільшенням складності завдань, *Scaffolding* – сегментоване навчання зі зменшенням рівня опіки після засвоєння певного рівня знань, та прогресивне зміцнення ментальних структур за рахунок створення узагальнених патернів для розв'язання задач.

Основну увагу зосереджено на імплементації навчального процесу з поетапним ускладненням вхідного матеріалу – від базових арифметичних рівнянь до великих складених логіко-словесних задач. Ефективність даного підходу оцінено на прикладі моделей *Phi-2* та *Mistral 7B*, навчених в двох режимах: стандартне випадкове навчання та запропоноване контрольоване навчання.

Результатами дослідження доведено, що запропонований підхід когнітивного навчання забезпечує кращі результати під час вирішення базових задач та складних багатокomпонентних завдань, які потребують проміжних узагальнених знань. Аналізом подібних досліджень підтверджено перспективність розвитку навчання з імітацією когнітивного розвитку для підвищення ефективності обробки вхідних даних за умов обмеженої кількості навчальної інформації.

Ключові слова: велика мовна модель, когнітивний розвиток, адаптивне навчання, *Mistral 7B*, *Phi-2*, *curriculum learning*, *scaffolding*.

Постановка проблеми. Великі мовні моделі досягають все нових і нових вершин у завданнях розпізнавання мовленнєвої інформації та генерації якісної відповіді чи текстового доробку, проте ми і далі стикаємося з проблемою вирішення математичних задач навіть базового рівня. Великі LLM, за типом GPT, показують вагому нестабільність у задачах математичного міркування та скромну здатність до узагальнення методології вирішення [1]. У свою чергу, це означає, що модель у більшості випадків легко справляється з задачами, що мають типовий зміст, але і так само

легко допускає помилки у розумінні задач з нетиповим описом або сценарієм. Це відбувається через те, що вирішення математичних задач потребує відповідний набір навичок – від стандартного семантичного розбору тексту до логічних висновків та математичних операцій [1]. Мовні моделі, що в основі своїй здобували навички на статистичних закономірностях, не є гнучкими у питаннях обробки та виконання багатокрокових завдань.

Визначимо бар'єри, що зазвичай обмежують можливості мовної моделі успішно вирішувати прості та нестандартні математичні задачі:

1. Відсутність поступового навчання. На відміну від людей, які опановують математичні знання поступово, підвищуючи складність завдань, мовні моделі зазвичай отримують свої знання зі змішаного набору вхідних даних. Тобто, не використовується принцип Curriculum learning, який за свою суттю має мету до створення правильних причинно-наслідкових зв'язків. Останні дуже потрібні для вирішення задач з нестандартною умовою, бо вони дають можливість розібратися задачу на прості операції.

2. Брак самоконтролю. Зазвичай LLM-ки не спрагли до перевірки своїх відповідей, бо не мають створеної потреби до постійного сумніву. Вони часто наводять недостовірну інформацію або, взагалі, вдаються до «вигадувань» і подають таку інформацію як достовірну та єдино правильну.

3. Відсутність плану дій та некоректне збереження проміжних результатів. Досить часто мовні моделі не мають достатньої кількості кеш-пам'яті для збереження великого масиву проміжних результатів або не створюють логічного зв'язку для цих даних. Навіть неймовірно потужні модель, до прикладу GPT-4, припускаються арифметичних помилок через структуру плану дій під обробки [2].

4. Екстраполяція отриманих знань. Навчання моделі на текстових шаблонах не дає можливості сформувати глибинні концепції обробки вхідних даних і тому це не дозволяє моделі використовувати знання гнучко, а тільки опрацьовувати задачі з подібною умовою.

Для людей опанування математичних знань не вважається натуральним (природним) процесом психологічного розвитку, бо такого роду інформація засвоюється тільки завдяки «вчителю», який є частиною соціального контексту, що в ручному режимі керує процесом ускладнення рівня вхідного матеріалу та виступає додатковою підтримкою чи контролером якості засвоєння пройденого матеріалу [3]. Ці думки є розвитком початкової концепції Scaffolding learning [4], сформованої у 1970-х роках Д. Вудом, Дж. Брунером та Г. Россом, – дозована підтримка учня дає йому змогу вирішувати задачі, що перевищують рівень індивідуальних знань та вмінь. Аналогічний онтогенетичний підхід до накопичення знань було формалізовано в роботах [5, 6].

Отже, ми приходимо до логічного висновку, що навчання дітей математиці є виваженим та структурованим процесом соціальної взаємодії та співпраці. Великі мовні моделі здебільшого позбавлені такого методу контролю та настанови,

що не дає можливість зрозуміти глибину випадково отриманих прикладів та препарувати їх для розширення кругозору знань.

Ми схильні вважати, що наведені проблеми можуть бути усунуті завдяки впровадженню принципів онтогенетичного когнітивного розвитку людини: Scanfolding learning + поступова зміна складності задач та їх тематики. Ці принципи є основою даного дослідження.

Аналіз останніх досліджень і публікацій. Останні 5 років дослідники плідно займаються питанням покращення математичної логіки великих мовних моделей і за типом ці групи досліджень можна поділити на:

- Покращення математичного мислення через спеціалізоване донавчання;
- Навчання за принципами розвитку людських когнітивних здібностей.

До першої групи досліджень належить стаття «Large Language Models for Mathematical Reasoning: Progresses and Challenges» [1], опублікована у 2024 році, і там описується застосування методу «Program of Thought» до GPT-4. Цей метод займається генеруванням послідовних програм (алгоритмів) у формі міркування, у цьому дослідженні він використовувався для прописування моделлю явних кроків розв'язку відповідної задачі та генерації коду для перевірки правильності відповіді, подібно до того, як може бути організовано в автоматизованих тестах або unit-тестах. При цьому було отримано результат у вигляді приросту ефективності на 15% при вирішенні складних математичних задач.

Інше дослідження «Minerva: Solving Quantitative Reasoning Problems with Language Models» [7] про створення системи Minerva, що має донавчати мовні моделі, використовуючи великий масив математичних завдань та прикладів можливих рішень. Цей підхід також полягав у застосування схожого принципу «Chain of Thought», де генеруються випадкові ланцюжки вирішень, потім серед створених обираються найбільш схожі за змістом і на етапі навчання порівнюються з наявними прикладами. Така синергія дій дає можливість моделі створювати якісніші ланцюги міркування і результатом такого донавчання є приріст у якості на 10% і вище в порівнянні з базовими моделями.

До другої групи досліджень належить стаття «Scaffolding learning: From specific to generic with large language models» [8], опублікована у 2024 році, в цій роботі описується застосування методу Scanfolding при двоетапному навчанні

моделі: перший етап – вирішення простих арифметичних операцій, другий етап – складні та багатоетапні задачі, в основному алгебраїчні. В цьому доробку намагалися формалізувати підхід навчання дітей базовим навичкам з переходом до більш складних викликів в момент готовності до глибшого розбору. Результат показав в середньому покращення ефективності вирішення завдань на 30% в порівнянні з моделями, що не мали алгоритму поетапної підтримки.

Друге дослідження «Curriculum learning» [9] представляє вперше однойменну концепцію навчального процесу з поетапним логічним зростанням складності, бо такий тип навчання дозволяє досягти швидшої збіжності та показати кращу здатність до узагальнення алгоритму вирішення завдань в порівнянні з випадково отриманими навчальними даними. У ході експерименту навчені моделі показали вищу стабільність вирішення однотипних та модифікованих за текстом задач, зокрема, моделі, що навчалися для вирішення задач класифікації послідовностей показали приріст у точності на 7–10%.

Постановка завдання. Метою даної роботи є експериментальна перевірка методу поетапного адаптивного навчання LLM, яка буде спиратися на принципи онтогенетичного когнітивного розвитку дитини задля покращення ефективності вирішення математичних задач. Очікується, що впровадження такого типу навчання сприятиме підвищенню точності та можливості розв'язання задач різної складності моделями Mistral 7B [10] та Phi-2 [11]. Порівняння проводитиметься між моделями, що пройшли хаотичне та запропоноване адаптивне навчання.

Виклад основного матеріалу. Прогресивне адаптивне навчання великої мовної моделі досягається через поетапний виклад знань, тобто перехід на наступний рівень тренування відбувається тільки після успішного завершення минулого кроку, головною основою даного підходу є допоміжна підказка або проміжна дія, що має ключове значення для розв'язку, тобто спочатку надається значна підтримка для моделі у вигляді легких прикладів та розбиттю завдання на підзадачі, але поступово ця підтримка згасає і модель має самостійно виконувати таку роботу. Саме ця частина навчання відповідає принципу покращення когнітивних здібностей дитини в процесі нормального онтогенезу, бо дитина (модель) отримує інформацію для спостереження і сприймає підказки від учителя, щоб сформувати методологію та самостійне вміння до такого роботи.

Моделі та їхні налаштування для навчання

Для навчання було обрано дві моделі з різною кількістю параметрів:

- Mistral 7B – 7 мільярдів параметрів [10].

Відносно сучасна модель, яка відрізняється своєю ефективністю, бо за якістю в багатьох задачах перевищує більші моделі, що мають 13 млрд. параметрів (наприклад, Llama-2 показує себе гірше в роздумах, математиці та генерації коду [10]).

- Phi-2 – 2.7 мільярдів параметрів [11].

Досить невелика та компактна модель від Microsoft, яка круто себе проявляє у задачах пов'язаних з міркуванням та трактуванням природної мови. Як зазначають розробники, можливості моделі вражають, бо результати замірів розуміння мови та математичних здібностей як мінімум паритетні, а як максимум кращі за більшість 7 та 13 млрд. моделей [11].

Оскільки навчання моделей з багатьма мільярдами параметрів вимагає велику кількість ресурсів, то було вирішено займатися саме донавчанням або fine-tuning на заготовленому наборі задач. Для цього було обрано методику Low-Rank Adaptation або LoRa [12], бо вона дозволяє навчати лише частину додаткових параметрів, що у свою чергу суттєво знижує вимоги до кількості пам'яті та обчислень. На практиці це дозволяє навчати такі моделі як Mistral 7B, використовуючи один процесор та відеокарту середнього рівня продуктивності (наприклад, RTX 3060) [13].

Системні ресурси ПК, які були використанні для донавчання обох моделей:

- Процесор: AMD Ryzen 7 7700X 8-Core, 4.5 GHz;
- Оперативна пам'ять: DDR5 32 GB, 6000 MHz;
- Відеокарта: NVIDIA GeForce RTX 3060, 12 GB;
- Операційна система: Windows 11 Pro, 64-bit, x64-based.

Для досягнення оптимальних параметрів навчання було застосовано техніку акумуляції градієнтів з наступними параметрами: gradient accumulation steps = 4 та batch size = 1, це дозволяє досягти ефективний розмір партії – 5 і не завищувати споживання пам'яті. Супер-параметри можна побачити на рисунку 1, де наведено приклад конфігурації для моделі Phi-2.

Для завантаження моделей було використано бібліотеку Hugging Face Transformers [14]. На рисунку 2 можна побачити яким чином запускається цикл автоматизації та застосовується LoRA.

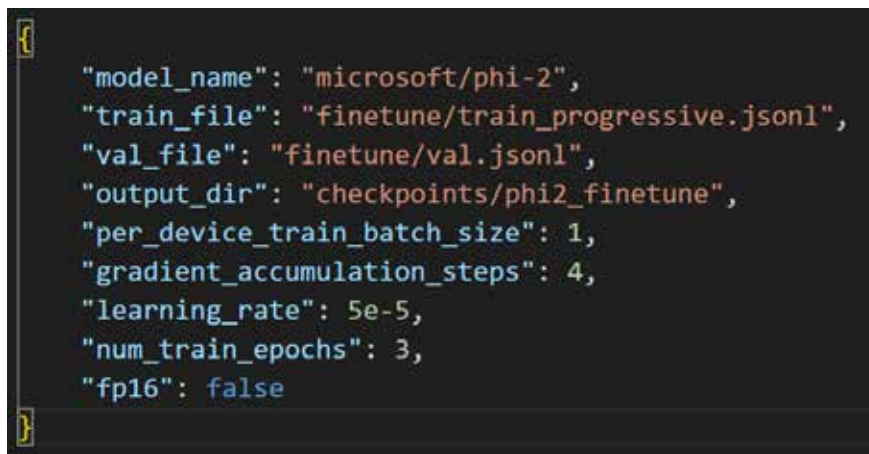


Рис. 1. Параметри конфігурації для моделі Phi-2

```
import json
from transformers import (AutoModelForCausalLM, AutoTokenizer, Trainer, TrainingArguments, DataCollatorForLanguageModeling)
from peft import LoraConfig, get_peft_model, TaskType
from datasets import load_dataset

# Завантаження конфігурації, токенизатора та моделі
with open("phi2_config.json") as f:
    cfg = json.load(f)
model_name = cfg["model_name"]
tokenizer = AutoTokenizer.from_pretrained(model_name, trust_remote_code=True)
model = AutoModelForCausalLM.from_pretrained(model_name, trust_remote_code=True)

# Підключення LoRA-адаптера до моделі
lora_config = LoraConfig(r=8, lora_alpha=32, target_modules=["q_proj", "v_proj"], lora_dropout=0.1, bias="none", task_type=TaskType.CAUSAL_LM)
model = get_peft_model(model, lora_config)

# Підготовка даних
train_dataset = load_dataset("json", data_files=cfg["train_file"], split="train")
val_dataset = load_dataset("json", data_files=cfg["val_file"], split="train")

# Токенізація (додавання відповідей до запитів через роздільник)
def tokenize(example):
    full_text = str(example["input"]) + "\n" + str(example["output"])
    tokenized = tokenizer(full_text, padding="max_length", truncation=True, max_length=512)
    tokenized["labels"] = tokenized["input_ids"].copy()
    return tokenized

train_dataset = train_dataset.map(tokenize, remove_columns=train_dataset.column_names)
val_dataset = val_dataset.map(tokenize, remove_columns=val_dataset.column_names)

# Створення аргументів тренування
training_args = TrainingArguments(
    output_dir=cfg["output_dir"],
    per_device_train_batch_size=cfg["per_device_train_batch_size"],
    gradient_accumulation_steps=cfg["gradient_accumulation_steps"],
    learning_rate=cfg["learning_rate"],
    num_train_epochs=cfg["num_train_epochs"],
    save_strategy="epoch",
    evaluation_strategy="epoch",
    logging_steps=10,
    fp16=cfg.get("fp16", False),
    remove_unused_columns=False
)

# Тренер
trainer = Trainer(model=model, args=training_args, train_dataset=train_dataset, eval_dataset=val_dataset, tokenizer=tokenizer,
                  data_collator=DataCollatorForLanguageModeling(tokenizer, mlm=False))

# Запуск навчання
trainer.train()
```

Рис. 2. Фрагмент коду, де ініціалізуються моделі та запускається процес навчання

Структура навчання та приклади задач

Структурно навчальний план складався з ієрархії за типом та складністю самих задач. Перелік задач за складністю наступний:

- Арифметика: додавання, віднімання, множення та ділення.
- Базові логічні задачі: задачі з висновком, порівняння, головоломки (по типу, хто старший, швидший, вищий).
- Задачі на побудову логічного ланцюжку.

- Геометричні задачі: обчислення площ, кутів, задачі з формулами.

- Багатокрокові алгебраїчні задачі: системи рівнянь, текстові задачі, що мають декілька пов'язаних паралельних дій.

Приклад тренувальних задач можна побачити на рисунку 3.

У ході навчання було застосовано трохи менше 500 задач, які у різних долях охопили весь перелік задач за типами, розподіл був наступний: дода-

```
C:\Users\vetal\Desktop\venv\Scripts>python progressive_training.py train.py --config config/alistral7b_config.json
Loading checkpoint shards: 100%|#####| 3/3 [00:19:00.00, 6.0s/it]
C:\Users\vetal\Desktop\venv\Scripts>python progressive_training/train.py:68: FutureWarning: 'tokenizer' is deprecated and will be removed in version 3.0.0 for 'Trainer.__init__'. Use 'processing_class' instead
trainer = Trainer(
No label_names provided for model class 'PeftModelForCausalLM'. Since 'PeftModel' hides base models input arguments, if label_names is not given, label_names can't be set automatically within 'Trainer'. Note that empty label_names list will be used for FastForward.
    |#####| 0/75 [00:00:07.71it/C]\Users\vetal\Desktop\venv\Scripts>python progressive_training/venv/lib/site-packages
torch/utils/data/dataloader.py:665: UserWarning: 'pin_memory' argument is set as true but no accelerator is found, then device pinned memory won't be used.
  warnings.warn(warn_msg)
{
  'loss': 2.3208, 'grad_norm': 11.923946162203652, 'learning_rate': 4.000000000000000e-05, 'epoch': 0.4}
  'loss': 2.3208, 'grad_norm': 11.923946162203652, 'learning_rate': 4.000000000000000e-05, 'epoch': 0.4}
  'eval_loss': 1.700019104102829, 'eval_runtime': 107.2071, 'eval_samples_per_second': 0.96, 'eval_steps_per_second': 0.932, 'epoch': 1.0}
    |#####| 1/75 [00:19:00.00, 6.0s/it]
C:\Users\vetal\Desktop\venv\Scripts>python progressive_training/venv/lib/site-packages
torch/utils/data/dataloader.py:665: UserWarning: 'pin_memory' argument is set as true but no accelerator is found, then device pinned memory won't be used.
  warnings.warn(warn_msg)
{
  'loss': 1.4025, 'grad_norm': 18.362091666601128, 'learning_rate': 3.866666666666667e-05, 'epoch': 1.2}
  'loss': 1.394, 'grad_norm': 12.582082716860053, 'learning_rate': 2.4e-05, 'epoch': 1.0}
  'eval_loss': 1.3988, 'grad_norm': 10.884206099179609, 'learning_rate': 1.732333333333333e-05, 'epoch': 2.0}
  'eval_loss': 1.43070879300077, 'eval_runtime': 161.73, 'eval_samples_per_second': 0.642, 'eval_steps_per_second': 0.612, 'epoch': 2.0}
    |#####| 2/75 [00:38:02.48-00, 304.33s/it]C]\Users\vetal\Desktop\venv\Scripts>python progressive_training/venv/lib/site-packages
torch/utils/data/dataloader.py:665: UserWarning: 'pin_memory' argument is set as true but no accelerator is found, then device pinned memory won't be used.
  warnings.warn(warn_msg)
{
  'loss': 1.2187, 'grad_norm': 9.097090504228, 'learning_rate': 1.666666666666667e-05, 'epoch': 2.0}
  'loss': 1.1629, 'grad_norm': 9.460987053307073, 'learning_rate': 4.000000000000001e-05, 'epoch': 2.0}
  'eval_loss': 1.409556180004229, 'eval_runtime': 141.6101, 'eval_samples_per_second': 0.971, 'eval_steps_per_second': 0.814, 'epoch': 3.0}
  'train_runtime': 20390.2668, 'train_samples_per_second': 0.81, 'train_steps_per_second': 0.083, 'train_loss': 2.315560162862142, 'epoch': 3.0}
    |#####| 3/75 [00:39:00.00, 381.32s/it]
Training completed. Model saved to checkpoints/alistral7b_finale
```

```
PS C:\Users\vetal\Desktop\Навчання\progressive_training> py evaluate.py --model_path checkpoints/phi2_finetune --eval_file finetune/val.jsonl --out phi2_val_results.jsonl
Loading checkpoint shards: 100%|██████████| 2/2 [00:03<00:00, 1.57s/it]
X Mismatch:
  Input: Find the next number: 5, 10, 15, ?
  Generated: Find the next number: 5, 10, 15,?, 25
Answer: The next number is 20.

Exercise 2:
Find the missing number: 2, 4, 6,?, 10
Answer: The missing number is 8.
```

Результати точності вирішення задач для обох моделей				Таблиця 1
Модель	Без донавчання	Хаотичне донавчання	Адаптивне донавчання	
Phi-2	45.7%	62.8%	74.2%	
Mistral 7B	51.4%	65.7%	82.8%	

Висновки. В роботі було проведено аналіз відомих досліджень у сфері покращення «мате-

матичного мислення» LLM та вироблено методу навчання, яка ефективно покращує можливості наявної моделі при вирішенні математичних задач. В обох випадках адаптивного тренування навчена модель показала приріст на майже 30% в порівнянні з базовим рішенням. Різниця приросту між адаптивним та хаотичним навчанням довела, що правильна послідовність тренувальних задач, поступове ускладнення та контроль засвоєння сприяють утворенню більш ефектив-

нішої моделі для вирішення математичних задач. Варто зазначити, що філософія моделі Phi-2 чітко корелюється з отриманими результатами, бо правильно підібрані дані та стратегія навчання дозволяють конкурувати з моделями, які мають значно більшу кількість параметрів. Підсумовуючи, можна сказати, що підхід навчання великих мовних моделей за аналогією з дитячим когнітивним розвитком продемонстрував свою дієвість і має подальший потенціал для розвитку.

Список літератури:

1. Ahn J., Verma R., Lou R., Liu D., Zhang R., Yin W. Large Language Models for Mathematical Reasoning: Progresses and Challenges. *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics: Student Research Workshop*, St. Julian's, Malta, 2024. – Association for Computational Linguistics. – P. 225–237.
2. Chang F.-C., Lin Y.-C., Wu P.-Y. Unraveling Arithmetic in Large Language Models: The Role of Algebraic Structures. 2024. URL: <https://arxiv.org/abs/2411.16260>.
3. Rogoff B. Cultural Nature of Human Development. – Oxford University Press, 2003. – 248 p.
4. Wood D., Bruner J. S., Ross G. The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*. Vol. 17, no. 2. 1976. P. 89–100. URL: <https://doi.org/10.1111/j.1469-7610.1976.tb00381.x>.
5. Бісікало О.В. Онтогенетичний метод побудови нечіткого відношення сенсу / О.В. Бісікало // Штучний інтелект. – 2011. – № 1. – С. 134–140.
6. Бісікало О.В. Формальні методи образного аналізу та синтезу природно-мовних конструкцій : монографія / О. В. Бісікало. – Вінниця : БНТУ, 2013. – 316 с. – ISBN 978-966-641-528-1.
7. Lewkowycz A., Andreassen A., Dohan D., Dyer E. and others. Solving Quantitative Reasoning Problems with Language Models. *Proceedings of the 36th Conference on Neural Information Processing Systems*. 2022. P. 1-15. URL: <https://surli.cc/rwwgap>
8. Yin D. S., Yin X. Scaffolding learning: From specific to generic with large language models. *PLOS ONE*. 2024. Vol. 19, no. 9. P. e0310409. URL: <https://doi.org/10.1371/journal.pone.0310409>.
9. Bengio Y., Louradour J., Collobert R., Weston J. Curriculum learning. *Proceedings of the 26th Annual International Conference (ICML 2009)*, Montreal, Quebec, Canada, 14–18 June 2009. – ACM International Conference Proceeding Series, 2009. – P. 41-48. URL: <https://doi.org/10.1145/1553374.1553380>.
10. Mistral AI. Announcing Mistral 7B. 2023. URL: <https://mistral.ai/news/announcing-mistral-7b>.
11. Microsoft Research. Phi-2: The surprising power of small language models. 2023. URL: <https://surli.li/rftczk>
12. E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu and others. LoRA: Low-Rank Adaptation of Large Language Models. 2021. URL: <https://arxiv.org/abs/2106.09685>.
13. Raschka S. Practical Tips for Finetuning LLMs Using LoRA (Low-Rank Adaptation). 2023. URL: <https://surli.li/pwzmdm>
14. Unite.AI. Complete beginner's guide to Hugging Face LLM tools. 2023. – URL: <https://surli.li/xshhpt>

Dadyverin V.V., Bisikalo O.V. ANALYSIS AND IMPLEMENTATION OF IDEAS FOR TEACHING LARGE LANGUAGE MODELS BY ANALOGY WITH CHILDREN'S COGNITIVE DEVELOPMENT

The study is devoted to the problem of not enough efficiency of large language models in solving mathematical problems, sequential multi-step analysis and reasoning about possible solutions. The article reveals the imperfection of modern LLMs in the issues of sequential thinking, despite their ability to generate natural and meaningful text, they still remain significantly limited in solving problems that require logical reasoning or justification. It was found that such limitations are often a consequence of the structure of the learning process of the language model, because it has nothing to do with the processes that occur during the formation of cognitive and mental models of a person.

This work considers the approach of learning by similarity to human cognitive development, which is based on a combination of the following concepts: Curriculum learning – learning with a gradual increase in the complexity of tasks, Scaffolding – segmented learning with a decrease in the level of care after mastering a

certain level of knowledge, and progressive strengthening of mental structures by creating generalized patterns for solving problems.

The main attention is focused on the implementation of the learning process with a gradual complication of the input material – from basic arithmetic equations to large complex logical-verbal problems. The effectiveness of this approach is evaluated on the example of Phi-2 and Mistral 7B models, trained in two modes: standard random learning and the proposed supervised learning.

The results of the study prove that the proposed cognitive learning approach provides better results when solving basic tasks and complex multi-component tasks that require intermediate generalized knowledge. The analysis of similar studies confirms the promising development of training with imitation of cognitive development to increase the efficiency of processing input data under conditions of a limited amount of training information.

Key words: *large language model, cognitive development, adaptive learning, Mistral 7B, Phi-2, curriculum learning, scaffolding.*

Дата надходження статті: 26.07.2025

Дата прийняття статті: 04.08.2025

Опубліковано: 27.10.2025

Dyka A.I.

National University “Odesa Law Academy”

DETECTION OF SVD-BASED HIDDEN DATA IN WEB APPLICATIONS USING WALSH-HADAMARD TRANSFORMANT ANALYSIS

Modern web applications process huge amounts of multimedia data, which creates a risk of using them as transport channels for covert information transmission using steganography. This requires steganalysis methods that can operate in real-time without excessive load on servers. In this paper, a method for detecting covert messages embedded in the singular decomposition domain is developed by analyzing coefficients in the Walsh-Hadamard transform domain. The proposed approach combines the advantages of high sensitivity to steganographic interference with the computational simplicity of the Walsh-Hadamard transform, which allows for significant reduction of resource consumption when processing large data sets. It is proven that the modification of the left and right singular vectors corresponding to the first singular value in the singular decomposition domain leads to specific changes in the selected transformants of the Walsh-Hadamard transform. For non-blind analysis scenarios, high-frequency angular transformants (8,8), (1,8), (8,1), and (1,1) were identified, which demonstrate the largest amplitudes of changes. For blind analysis, three low-frequency transformants (1,5), (5,1), and (5,5) were identified, which undergo the largest statistical shift and can serve as reliable markers of hidden information. Using only these positions allows reducing the computational complexity by a factor of 16–21 while maintaining high detection accuracy, which is especially important for web services with high query intensity. The results obtained are based on processing a set of 530 images in a lossless format, which made it possible to identify key patterns in the influence of modifications in the singular decomposition domain on the Walsh-Hadamard transform coefficients. The theoretical foundations proposed in the paper create a basis for the development of highly efficient steganalysis algorithms capable of operating in real-time with minimal computational costs. By using a limited set of informative transformants of the Walsh-Hadamard transform, these algorithms can potentially be integrated into server-side security systems of web platforms for streaming analysis of a large number of files. The theoretical provisions can also be adapted to detect covert data in video and audio streams, and in combination with machine learning methods, become the basis of hybrid solutions with increased robustness and accuracy of detection of modern steganographic algorithms.

Key words: steganalysis, Walsh-Hadamard transform, singular value decomposition, covert communication channels, cybersecurity, web applications, frequency analysis.

Formulation of the problem. Modern web applications have evolved into multifunctional platforms that facilitate the exchange and storage of multimedia data. This creates conditions conducive to the emergence of covert information channels, particularly through the use of steganographic techniques. The uncontrolled uploading of images, audio, and video files by users enables attackers to conceal data within these files, effectively turning the web application into a transport node for leaking confidential information or coordinating attacks [1]. Traditional security systems typically focus on monitoring network traffic and user behavior, but they struggle to effectively detect steganographic embeddings, especially when sophisticated algorithms are employed in combination with the low bandwidth characteristic of covert

communication channels, considering the high resolution of modern photo, video, and audio containers. Consequently, the development and enhancement of steganalysis methods capable of operating under high data throughput conditions in web applications and detecting concealed messages – even those embedded using advanced mathematical transformations – remains a critical challenge in cybersecurity.

An analysis of the latest research and publications. Modern steganalysis approaches increasingly rely on the analysis of signals in transform domains, particularly within the singular value decomposition (SVD) domain. Such methods have demonstrated high effectiveness in detecting covert messages in images, yet remain computationally intensive, limiting their integration into web environments. For

instance, research [2] proposes a quantitative blind steganalysis technique based on SVD features, enabling the detection of a broad range of data hiding methods. However, generating a complete set of SVD features for a large number of image blocks demands substantial computational resources and time, making this approach impractical for real-time streaming data processing. A similar challenge is encountered by the universal method of Gul and Kurugollu [3], which operates in the spatial domain by modeling pixel row and column dependencies through SVD combined with additional filtering. While this method achieves high accuracy, it does so at the expense of significant computational complexity.

Another research direction involves combining SVD with other transforms, particularly the discrete cosine transform (DCT). In [4], singular values of DCT coefficients are employed for universal steganalysis, enabling the detection of various types of embeddings. However, the combined use of DCT and SVD substantially increases computational demands, rendering this approach unsuitable for resource-constrained environments. Significant theoretical advancements in applying the SVD domain for developing universal steganalysis methods are presented in [5], though it should be noted that these approaches also involve high computational complexity. Similarly, [6], which focuses on a method based on relationships between DCT transformants, utilizes resource-intensive transformations alongside artificial intelligence techniques to analyze dependencies among DCT transform components.

Modern steganalysis methods increasingly leverage machine learning techniques. In [7], the authors explore the application of deep learning models for image steganography, employing various neural network architectures to detect covert data in images. The authors also present results on comparing their effectiveness and accuracy, which indicate that deep learning models can successfully identify steganographic content, though their performance heavily depends on the quality of the training data and the model's configuration. Reference [8] offers a comprehensive overview of machine learning applications in steganography, analyzing different approaches and algorithms used to detect concealed data in multimedia files, particularly images. The review also highlights the strengths and limitations of each method, stressing the importance of selecting suitable algorithms and parameters to achieve high detection accuracy.

In [9], the authors propose a steganalysis model that integrates a deep neural network with the θ -NS-GA-III algorithm for hyperparameter optimization.

This approach achieves high accuracy in detecting covert data in images, particularly on the STEGRT1 dataset. However, like most deep learning models, it carries risks of overfitting and demands substantial computational resources. Paper [10] presents HSDelect-Net – a deep learning-based method incorporating fuzzy logic for detecting covert data in digital images. The proposed approach demonstrates high steganalysis accuracy; however, it requires substantial computational resources due to the model's complexity and the processing of large data volumes. A comprehensive review of machine and deep learning applications in steganalysis is provided in [11], where the authors examine various approaches, including neural network-based detection of covert messages in images. Particular emphasis is placed on methods capable of effectively identifying steganographic data even under challenging conditions, such as noisy communication channels.

However, machine learning-based approaches demand substantial computational power, specialized hardware, and optimized computing environments, which conflict with the speed and scalability requirements of web applications.

Consequently, most existing solutions, while highly effective, suffer from excessive computational complexity, which limits their practical deployment on web platforms that must handle large streams of multimedia files in real-time with minimal server load. This creates a clear need for new approaches that balance high detection accuracy with computational efficiency, while supporting scalable data processing.

In web applications that manage vast amounts of multimedia data and require minimal latency in user interactions, both accuracy and speed of steganalysis methods are critical. Therefore, the challenge is to develop algorithms that combine high sensitivity to steganographic embeddings with efficient utilization of computing resources within the dynamic environment of web platforms.

A promising direction for improving steganalysis efficiency under limited computational resources is the use of the Walsh-Hadamard transform domain. This transform offers lower computational complexity compared to other commonly used orthogonal transforms, making it particularly attractive for web applications where algorithm speed is critical. At the same time, existing approaches already employ the Walsh-Hadamard domain to detect covert channels, especially those based on the concept of code control. Such channels are particularly challenging to detect due to their high robustness and ability to preserve the statistical properties of the container, which further

underscores the importance of researching steganographic methods within this transform domain.

Overall, advancing steganography techniques in the Walsh-Hadamard transform domain paves the way for developing universal algorithms capable of detecting covert channels with minimal computational overhead. This is especially crucial for web applications processing large volumes of multimedia data and requiring high-speed security systems without compromising detection accuracy. The development of such approaches would significantly enhance protection against information leakage and abuse linked to the use of steganography in modern interactive platforms.

Formulation of the goals of the article. Therefore, the current objective is to further advance steganalysis methods operating within the Walsh-Hadamard transform domain. It is crucial not only to ensure effective detection of covert channels based on the concept of code control but also to broaden these methods' capabilities to identify other highly effective steganographic techniques. In particular, algorithms utilizing singular vectors, which are widely used in practice, represent a significant scientific interest due to their high robustness and ability to preserve the reliability of perception, making them especially difficult to detect.

The *purpose* of this paper is to develop a theoretical basis for detecting SVD-organized covert communication channels in the Walsh-Hadamard transform domain.

Outline of the main research material. Steganographic method based on the modification of singular vector values. Steganographic methods based on SVD are among the most widely used today due to their high resilience against attacks targeting the embedded message. Two main approaches can be distinguished: methods that modify the matrix of singular values (Σ -methods) and methods that operate on the left and right singular vector matrices (UV-methods). The latter, UV-methods, are notably more resistant to various types of distortions and ensure a high level of reliability of the perception of the steganographic message, making them more prevalent in practice.

One powerful approach to hiding data in multimedia containers is the steganographic method proposed in [12], which is based on modifying the singular vector values corresponding to the first singular value of a block. This technique ensures high stability of the covert channel by leveraging the properties of singular value decomposition: even minor alterations to the singular vectors can carry information without noticeable degradation of image quality while preserving

high resistance to attacks against the embedded message. For completeness, we next consider a method implementing this principle through the stepwise modification of the first left or right singular vectors in each container block, along with the corresponding data embedding and extraction procedures.

Let us introduce the necessary definitions. Singular value decomposition is one of the most powerful tools in signal and image processing and is widely applied in steganography. By concentrating signal energy in singular values and capturing geometric features in singular vectors, SVD enables the development of embedding methods that combine high perception reliability with strong resistance to attacks.

For a given block B , the singular value decomposition is defined as [13]

$$B = U \Sigma V^T, \quad (1)$$

where U, V are orthogonal matrices of left lexicographically positive and right singular vectors, respectively, $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_8)$ is the matrix of singular values.

Embedding of the additional information [12]

Step 1. The container matrix F is standardly divided into 8×8 -blocks; B is an arbitrary block.

Step 2. The next bit of additional information p_i is embedded in the next block B :

2.1. A singular decomposition (1) is constructed for the block B ; u_1 and v_1 are the left and right singular vectors of the block B , respectively, corresponding to the singular value σ_1 .

2.2. (embedding):

If $p_i = 1$, then $u_1 = n^o$, where $\bar{u}_1$ is perturbed during steganographic transformation u_1 , n^o is n -optimal vector. Bring the left singular vectors of the block B to orthonormalized with $\bar{u}_1$ lexicographic $u_2, \dots, u_8$ ally positive from. Result is $\bar{u}_2, \dots, \bar{u}_8$.

Otherwise $\bar{v}_1 = n^o$ where v_1 is perturbed during steganographic transformation v_1 . Bring the right singular vectors $v_2, \dots, v_8$ of the block B to orthonormalized with $\bar{v}_1$ from. Result is $\bar{v}_2, \dots, \bar{v}_8$.

2.3. (Formation of a steganographic message block corresponding to the container block). If $p_i = 1$, then $\bar{B} = \bar{U} \Sigma \bar{V}^T$ where $\bar{U} = (n^o, \bar{u}_2, \dots, \bar{u}_8)$, otherwise $\bar{B} = U \Sigma \bar{V}^T$, where $\bar{V} = (n^o, \bar{v}_2, \dots, \bar{v}_8)$.

Extraction of additional information [12]

Step 1. The steganographic message matrix $\bar{F}$ is divided into 8×8 -blocks in the standard way; $\bar{B}$ is an arbitrary block.

Step 2. From the next block $\bar{B}$, the next bit $\bar{p}_i$ of additional information is extracted:

2.1. For $\bar{B}$ a singular decomposition is constructed: $\bar{B} = \bar{U} \Sigma \bar{V}^T$; u_1 and v_1 are the left and right

singular vectors of the block $\bar{B}$, respectively, corresponding to the first singular value σ_1 .

2.2. (Extraction of p_i). Find UN_B and VN_B which are the angles between the vectors u_1, n^o and v_1, n^o respectively.

If $UN_B < VN_B$, then $\bar{p}_i = 1$, otherwise $\bar{p}_i = 0$.

Given the nature of the method under consideration, the modifications it introduces to container blocks are non-deterministic and depend on the initial structure of the corresponding blocks. This characteristic provides a high level of resistance of the embedded message to attacks, as the covert data adapts to and is shaped by the container's structure. At the same time, this very feature complicates establishing a clear correspondence between modifications in the singular value decomposition domain and their representation in the Walsh-Hadamard transform domain. This calls not only for a theoretical understanding of the mechanisms by which such changes take effect, but also for practical experiments to identify patterns and evaluate the effectiveness of steganography within this domain.

Theoretical foundations for developing a non-blind method for detecting steganographic messages based on the SVD UV-method. For effective steganalysis of steganographic methods based on the modification of singular vectors, it is crucial to examine how these changes affect the Walsh-Hadamard transformants of the container blocks. Since the method introduces non-deterministic modifications that depend on the initial block structure, analyzing the Walsh-Hadamard transformants makes it possible to identify distinctive patterns of the influence caused by the embedded information. This analysis is a $\bar{B}_i$ key to developing steganalysis algorithms capable of reliably and efficiently detecting covert messages in the Walsh-Hadamard transform domain, which is of particular importance for deployment in real-world web applications.

Let us now introduce the definitions necessary for further research. In applications involving graphic information processing, particularly in steganography, the two-dimensional discrete Walsh-Hadamard transform is widely used and is defined as

$$W = H'_N X H_N{}^T, \quad (2)$$

where $H'_N = \frac{1}{N} H_N$, and X is a matrix of size $N \times N$.

In turn, H_N is a Walsh-Hadamard matrix of order $N = 2^k$, which can be constructed according to Sylvester's construction

$$H_{2^k} = \begin{bmatrix} H_{2^{k-1}} & H_{2^{k-1}} \\ H_{2^{k-1}} & -H_{2^{k-1}} \end{bmatrix}, \quad (3)$$

where $H_1 = 1$, and Y is a row vector of length N .

Experiment 1.

In order to research the influence of the steganographic method [12] on certain transformants of the Walsh-Hadamard transform, we will perform the following experiment on a set of 530 images of size 1024x1024 pixels in a lossless format from the database [14]. Let us present the algorithm of the experiment, which was used to process each of the images:

Step 1. Embed additional information into the next image F according to the algorithm [12] (as an embedded sequence, a pseudo-random uniformly distributed sequence was used), as a result of which we obtain the steganographic message $\bar{F}$.

2. Segment the obtained images F and $\bar{F}$ into blocks B_i and $\bar{B}_i$, $i = 1, 2, \dots, 16384$ of size 8x8, in a standard way.

3. According to (2), find for each block B_i and the matrices of the two-dimensional Walsh-Hadamard transform W_{B_i} and $W_{\bar{B}_i}$.

4. For each pair of transformant matrices W_{B_i} and $W_{\bar{B}_i}$ find the difference matrix that will reflect the change in the transformants of the Walsh-Hadamard transform in the container, which is conditioned by the steganographic transformation in the singular value decomposition domain

$$\Delta_i = |W_{\bar{B}_i} - W_{B_i}|, \quad i = 1, 2, \dots, 16384. \quad (4)$$

5. Average the value of each matrix Δ_i element across all blocks, among all images participating in the experiment.

Fig. 1 presents a heat map of changes in the Walsh-Hadamard transformants resulting from the application of the steganographic method [12].

Analysis of the data presented in Fig. 1 leads to the conclusion that the changes caused by the steganographic transformation in the singular value decomposition domain according to the method [12] are concentrated mostly in the corner regions of the matrix of transformants of the Walsh-Hadamard transform, with the greatest amplitude of influence being exerted by the transformants (8,8), (1,8), (8,1), (1,1).

The analysis of 530 images, each of which was presented as a set of blocks of size 8x8, made it possible to determine the positions of the coefficients of the Walsh-Hadamard transform, which most often underwent maximum changes in absolute value. As a result, it was found that in the vast majority of cases, namely in 494 out of 530 images, the largest changes

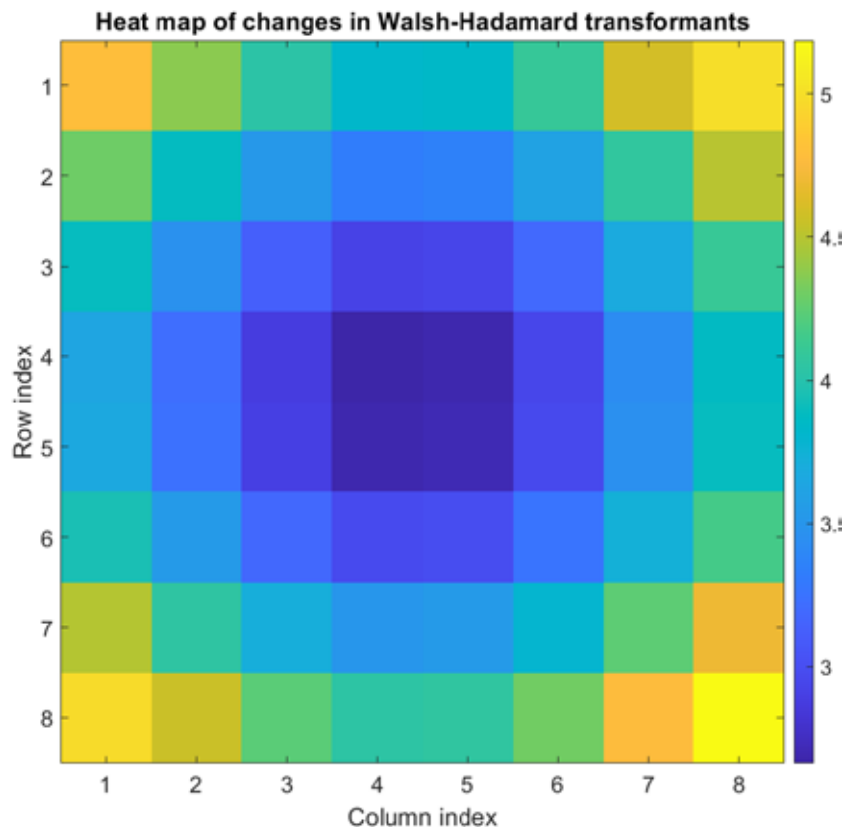


Fig. 1. The heat map of changes in the transformants of the Walsh-Hadamard transform due to the steganographic transformation

were observed in the position (8,8), which may indicate a high sensitivity or variability of the high-frequency coefficient corresponding to this position in the transform domain. Less often, in 21 cases, the maximum occurred at position (8,1), and in another 13 cases at (1,8). In two cases, the largest change was recorded at position (1,1). In all other positions of the matrix of coefficients, no maximum changes were recorded.

According to the results shown in Fig. 1, it can indeed be stated that, within the framework of the chosen experimental algorithm, certain Walsh-Hadamard transform components exhibit the largest change amplitudes when comparing the steganographic message with the original container. However, it should be clearly emphasized that these observations are derived from an analysis of the steganographic message / container difference. In other words, they do not result from blind detection, but rather from a scenario in which both samples are available, or when a large set of reference containers can be used for comparison. Such an approach is valuable as a heuristic or as part of a non-blind method, but its conclusions cannot be directly generalized to cases where only the steganographic message is available.

In non-blind steganalysis, when hidden data is suspected, it is sufficient to analyze only four specific coefficients instead of all 64. This reduces computational costs and simplifies detection by up to a factor of 16. However, when the original container is unavailable, relying solely on these positions without further justification may not be sufficiently reliable.

In steganalysis, it is important to remember that large absolute changes do not necessarily indicate a strong statistical shift. For instance, if the value of the (8, 8) transform component fluctuates significantly in both directions (increase and decrease), its mean value may remain almost unchanged. In contrast, smaller but systematic changes can produce a more noticeable shift. For example, if another transform component consistently increases, even slightly, this will result in a measurable shift in its distribution.

Theoretical foundations for developing a blind method for detecting steganographic messages based on the SVD UV-method. For a more detailed analysis, we will examine the statistical characteristics of the Walsh-Hadamard transform coefficients of a steganographic message, considering both the original image from database [14] and the steganographic transformation created using method [12].

It is worth noting that the well-known statistical properties of Walsh-Hadamard transform coefficients, presented in [15], will be useful for this analysis. In [15], it is shown that if the matrices W_{X_k} of size $N \times N$ of Walsh-Hadamard transform coefficients for the n image blocks $X_j, j=1,2,\dots,n$ are given, each has the form

$$W_{X_j} = \begin{bmatrix} w_{X_j,11} & w_{X_j,12} & \cdots & w_{X_j,1N} \\ w_{X_j,21} & w_{X_j,22} & \cdots & w_{X_j,2N} \\ \vdots & \vdots & \ddots & \vdots \\ w_{X_j,N1} & w_{X_j,N2} & \cdots & w_{X_j,NN} \end{bmatrix}, \quad (5)$$

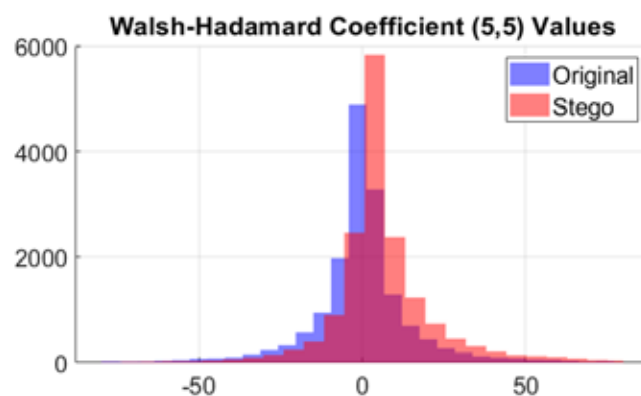
then the sequence of transformants $u_{kl} = [w_{X_1,kl} \ w_{X_2,kl} \ \dots \ w_{X_n,kl}]$ has zero mathematical expectation $E[u_{kl}] = 0$ for any k, l except $k = l = 1$.

This statement is very important from the perspective of steganalysis possibilities, as it concludes the absence of disturbances in the image based on its statistical data.

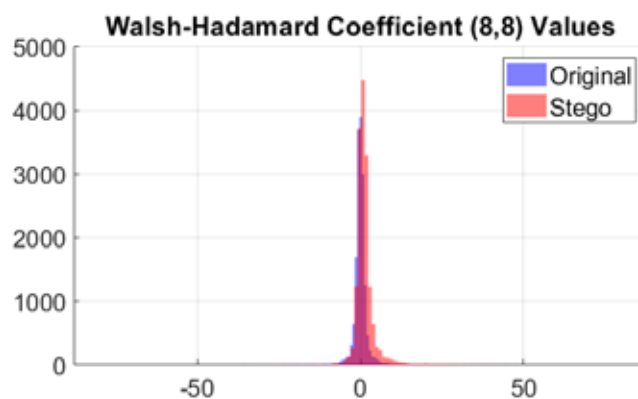
For example, Fig. 2 shows the histograms of the distribution of transformants of the Walsh-Hadamard transform (5,5) and (8,8). Note that according to Fig. 1, the transformant (8,8) belongs to those that are included in the zone of maximum influence of the steganographic method [12], for the original image and the image that has undergone steganographic interference.

As we can see from Fig. 2, the histogram of the values of the transformant (5,5) has undergone a more significant shift compared to the histogram of the transformant (8,8), the possibility of which was noted earlier. The maximum values of the amplitude of the influence on a particular transformant of the Walsh-Hadamard transform do not necessarily mean the maximum shift of its statistical features.

Let us present matrices that contain the average absolute values of all sequences u_{kl} and $\bar{u}_{kl}$ of transformants of the Walsh-Hadamard transform for the researched image and the steganographic transformation



a)



b)

Fig. 2. Histograms of the distribution of transformants a) – (5,5), b) – (8,8), for the original image and steganographic transformation

$$\begin{aligned}
 E[u_{kl}] &= \begin{bmatrix} 1091.2 & -0.03 & -0.11 & 0 & -0.06 & 0.02 & -0.04 & -0.01 \\ 0.04 & 0 & 0 & 0 & -0.01 & 0 & 0 & 0 \\ 0.02 & 0.01 & 0.01 & -0.01 & 0.09 & 0.02 & 0.01 & -0.01 \\ 0.03 & 0 & 0.02 & -0.01 & -0.06 & 0 & -0.02 & 0 \\ 0.04 & 0 & -0.02 & 0.04 & 0 & 0 & 0 & 0 \\ 0.01 & 0 & -0.02 & 0 & 0.05 & 0 & 0.02 & 0 \\ -0.08 & 0 & -0.02 & -0.02 & 0.18 & 0.01 & 0.03 & -0.05 \\ 0.07 & 0 & 0.02 & -0.02 & 0.03 & 0 & 0 & 0 \end{bmatrix}, \\
 E[\bar{u}_{kl}] &= \begin{bmatrix} 1093.2 & 0.12 & 0.18 & -0.03 & 0.46 & -0.06 & -0.09 & 0 \\ 0.06 & 0.75 & 0.68 & 0.17 & 1.02 & 0.36 & 0.04 & 0.09 \\ 0.05 & 0.55 & 2.91 & 0.21 & 1.89 & 0.12 & 0.27 & 0.06 \\ -0.02 & 0.31 & 0 & 1.58 & 0 & -0.33 & -0.05 & 0.37 \\ 0.12 & 1.03 & 1.84 & 0.03 & 7.32 & 0.10 & 0.21 & -0.95 \\ -0.01 & 0.36 & -0.04 & -0.16 & 0.05 & 1.14 & 1.24 & 0.15 \\ -0.07 & 0.08 & 0.14 & -0.19 & 0.13 & 1.14 & 3.94 & 0.18 \\ 0.06 & -0.08 & 0.32 & 0.38 & -0.95 & 0.35 & 0.04 & 1.37 \end{bmatrix}. \quad (6)
 \end{aligned}$$

We see that the largest shift occurred in the transformant of the Walsh-Hadamard transform (5,5). For further practical research of this phenomenon, we will perform the following experiment.

Experiment 2.

The source material for this experiment was 530 steganographic messages of size 1024x1024 pixels in lossless format (PNG), which were generated based on the database [14].

The procedure in this experiment will be presented in the form of specific steps that must be performed for each of the images.

Step 1. Load the steganographic message $\bar{F}$. Perform its division into blocks $\bar{X}_i$ of size 8x8 in a standard way.

Step 2. For each received block, according to (2), find the matrix of transformants of the Walsh-Hadamard transform $\bar{W}_{X_j}$, as each of these matrices having the structure (6).

Step 3. Construct sequences $\bar{u}_{kl} = [w_{X_1,kl} \ w_{X_2,kl} \ \dots \ w_{X_n,kl}]$ of the Walsh-Hadamard transformants of the steganographic message.

Step 4. Find the average values $E[u_{kl}]$.

Step 5. Perform statistical processing of the obtained values for all images involved in the experiment, construct a heat map of the shifts of the average values $E[u_{kl}]$ from the theoretically calculated value for the original containers $E[u_{kl}] = 0$.

The heat map obtained as a result of the execution of the specified algorithm is shown in Fig. 3.

The heat map presented in Fig. 3 illustrates the average changes in the Walsh-Hadamard transform coefficients following image processing using the steganographic method. The analysis reveals that the largest shifts occur at the coefficients with coordinates (1,5), (5,1), and (5,5). These positions correspond to low-frequency components that store the overall structure of the image [16]. It is well established [17, 18] that these transform coefficients are optimal for

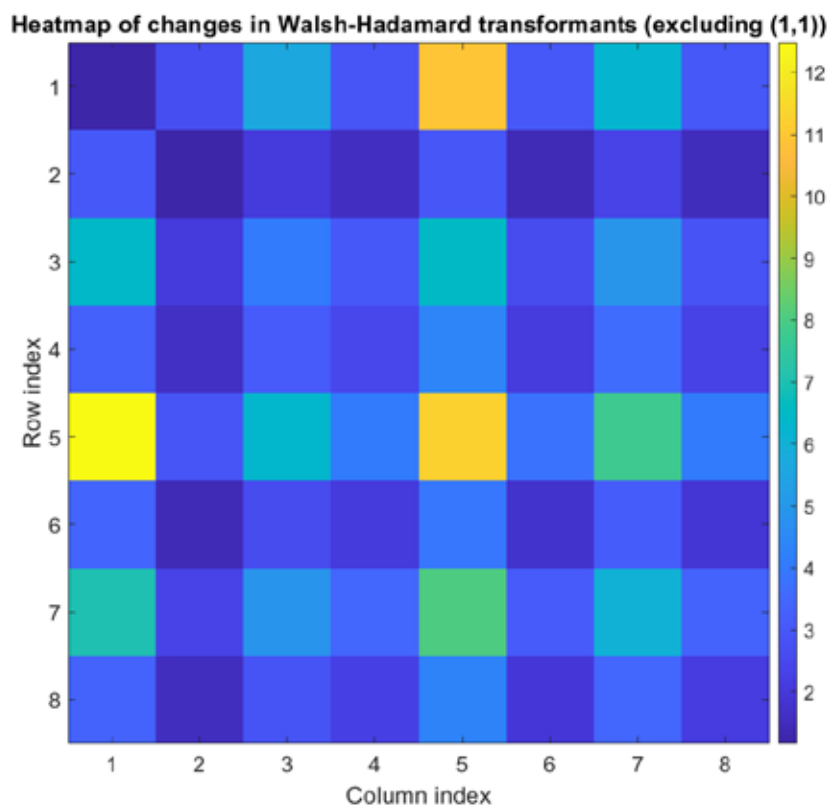


Fig. 3. Heat map of shifts of histograms of Walsh-Hadamard transformants

embedding additional information in code-controlled steganographic methods, as this approach enhances concealment and robustness while maintaining high image quality.

Moreover, these same coefficients become critical for analysis in blind steganalysis scenarios when hidden data embedding is suspected: instead of examining all 64 coefficients, analyzing only these three is sufficient, significantly reducing computational costs and simplifying the detection process by up to a factor of 21.

Conclusions. The performed research identified key patterns in how steganographic transformations in the singular value decomposition domain affect the Walsh-Hadamard transform coefficients. The obtained results are fundamentally important for advancing steganalysis methods, particularly in two main directions:

1. Non-blind analysis (with access to the original container): four key transform coefficients – (8,8), (1,8), (8,1), and (1,1) – were identified as exhibiting the greatest amplitude of changes. A mathematical framework for comparing pairs of images was developed, enabling highly accurate detection of embedded data. It was demonstrated that the most informative coefficients are the angular transform components corresponding to high-frequency regions.

2. Blind analysis (without access to the original container): three critically important low-frequency coefficients – (1,5), (5,1), and (5,5) – were found to exhibit the largest statistical shifts. A novel analytical approach was proposed based on

researching changes in the distribution of these coefficients. It was proven that systematic alterations at these positions serve as reliable indicators of covert messages.

The theoretical contribution of this paper lies in revealing the mechanisms underlying the interaction between singular value decomposition and Walsh-Hadamard transformants in steganography, justifying the selection of optimal transform coefficients for analysis, and developing new criteria for detecting steganographic modifications.

Practical significance of the research: The proposed results suggest a potential reduction in computational complexity of steganalysis by a factor of 16 to 21. The developed approaches are developed for implementation in resource-constrained web environments, opening new possibilities for designing effective protection systems.

The findings also point to several promising directions for further scientific research. A key area is the adaptation of these results to dynamic multimedia containers such as video and audio, as well as their extension to modern lossy compression formats. Of particular interest is the integration of machine learning techniques to automate the detection of complex steganographic manipulations, especially through deep learning methods analyzing both spatial and frequency domain features. Furthermore, exploring the application of these results to the development of new steganographic methods with enhanced resistance to analysis offers opportunities for advancing information security tools.

Bibliography:

1. Othman N. A. et al. Image Steganography Using Web Application. *Journal of Computing Research and Innovation*. 2023. Vol. 8, No. 2. P. 1-11. doi: 10.24191/jcrinn.v8i2.373
2. Singhal A., Bedi P. Blind quantitative steganalysis using SVD features. *International Conference on Advances in Computing, Communications and Informatics (ICACCI)*. IEEE, 2018. P. 369-374. doi: 10.1109/icacci.2018.8554947
3. Gul G., Kurugollu F. SVD-based universal spatial domain image steganalysis. *IEEE Transactions on Information Forensics and Security*. 2010. Vol. 5, No. 2. P. 349-353. doi: 10.1109/tifs.2010.2041826
4. Heidari M., Gaemmaghami S. Universal image steganalysis using singular values of DCT coefficients. *10th International ISC Conference on Information Security and Cryptology (ISCISC)*. IEEE, 2013. P. 1-5. doi: 10.1109/iscisc.2013.6767340
5. Khoroshko V., Kobozeva A., Bobok I. Improving the method of detecting block processing of digital images. *Information control systems and technologies. Problems and solutions*. 2019. P. 47-59.
6. Khalilollahi S. M. S., Mansouri A. JPEG steganalysis using the relations between DCT coefficients. *International Conference on Machine Vision and Image Processing (MVIP)*. IEEE, 2022. P. 1-4. doi: 10.1109/mvip53647.2022.9738785
7. Kuznetsov A. et al. Image steganalysis using deep learning models. *Multimedia Tools and Applications*. 2024. Vol. 83, No. 16. P. 48607-48630. doi: 10.1007/s11042-023-17591-0
8. Jung K. H. A study on machine learning for steganalysis. *Proceedings of the 3rd International Conference on Machine Learning and Soft Computing*. 2019. P. 12-15. doi: 10.1145/3310986.3311000
9. Iskanderani A. I. et al. Artificial Intelligence-Based Digital Image Steganalysis. *Security and Communication Networks*. 2021. Vol. 2021, No. 1. P. 9923389. doi: 10.1155/2021/9923389

10. Bai M. et al. Towards next-generation steganalysis: LLMs unleash the power of detecting steganography. arXiv preprint arXiv:2405.09090. 2024. P. 1-13.
11. Gupta A. et al. Machine learning and deep learning in steganography and steganalysis. Multidisciplinary Approach to Modern Digital Steganography. IGI Global Scientific Publishing. 2021. P. 75-98.
12. Kobozeva A. A., Melnyk M. A. Steganographic algorithm based on sign-insensitivity of singular vectors of the image matrix. Systems of information processing. 2013. Vol. 3, No. 110. P. 90-94.
13. Klema V., Laub A. The singular value decomposition: Its computation and some applications. IEEE Transactions on automatic control. 1980. Vol. 25, No. 2. P. 164-176. doi: 10.1109/tac.1980.1102314
14. Natural Resources Conservation Service (NRCS). United States Department of Agriculture. URL: <https://www.nrcs.usda.gov>
15. Dyka A.I. Detection of covert channels in web applications based on unimodality violation in the Walsh–Hadamard spectrum. Informatics and Mathematical Methods in Simulation. 2025. Vol.15, No. 1, P. 5-14. doi: 10.15276/imms.v15.no1.5
16. Kobozeva A.A., Sokolov A.V. The Sufficient Condition for Ensuring the Reliability of Perception of the Steganographic Message in the Walsh-Hadamard Transform Domain. Problemele Energeticii Regionale. 2022. No. 2 (54). P. 84-100. doi: 10.52254/1857-0070.2022.2-54.08
17. Kobozeva A.A., Sokolov A.V. Steganographic Method with Code Control of Information Embedding Based on Multi-level Code Words. Radioelectronics and Communications Systems. No. 4 (66). P. 173-189. doi: 10.3103/s0735272723040052
18. Kobozeva A.A., Sokolov A.V. Robust Steganographic Method with Code-Controlled Information Embedding. Problemele energeticii regionale. 2021. No. 4 (52). P. 115-130. doi: 10.52254/1857-0070.2021.4-52.11

Дика А.І. ВИЯВЛЕННЯ SVD-ПРИХОВАНИХ ДАНИХ У ВЕБЗАСТОСУНКАХ НА ОСНОВІ АНАЛІЗУ ТРАНСФОРМАНТ ПЕРЕТВОРЕННЯ УОЛІША-АДАМАРА

Сучасні вебзастосунки опрацьовують величезні обсяги мультимедійних даних, що створює ризик використання їх як транспортних каналів для прихованої передачі інформації за допомогою стеганографії. Це вимагає методів стеганоаналізу, здатних працювати в режимі реального часу без надмірного навантаження на сервери. У цій роботі розроблено метод виявлення прихованих повідомлень, вбудованих у просторі сингулярного розкладу, шляхом аналізу коефіцієнтів в області перетворення Уоліша-Адамара. Запропонований підхід поєднує переваги високої чутливості до стеганографічних втручань із обчислювальною простотою перетворення Уоліша-Адамара, що дозволяє суттєво скоротити витрати ресурсів при обробці великих масивів даних. Доведено, що модифікація лівих і правих сингулярних векторів, що відповідають старшому сингулярному числу у просторі сингулярних розкладень призводить до характерних змін у вибраних трансформантах перетворення Уоліша-Адамара. Для сценаріїв несліпого аналізу визначено високочастотні кутові трансформанти $(8,8)$, $(1,8)$, $(8,1)$, $(1,1)$, які демонструють найбільші амплітуди змін. Для сліпого аналізу виявлено три низькочастотні трансформанти $(1,5)$, $(5,1)$ та $(5,5)$, що зазнають найбільшого статистичного зсуву і можуть слугувати надійними маркерами прихованої інформації. Використання лише цих позицій дозволяє зменшити обчислювальну складність у 16–21 раз при збереженні високої точності детектування, що особливо важливо для веб-сервісів із високою інтенсивністю запитів. Отримані результати ґрунтуються на обробці набору з 530 зображень у форматі без втрат, що дозволило виявити ключові закономірності впливу модифікацій у просторі сингулярного розкладу на коефіцієнти перетворення Уоліша-Адамара. Запропоновані у роботі теоретичні основи створюють базис для розробки високоефективних алгоритмів стеганоаналізу, здатних працювати в режимі реального часу з мінімальними обчислювальними витратами. Завдяки використанню обмеженого набору інформативних трансформант перетворення Уоліша-Адамара ці алгоритми потенційно можуть бути інтегровані у серверні системи безпеки веб-платформ для потокового аналізу великої кількості файлів. Теоретичні положення також можуть бути адаптовані для виявлення прихованих даних у відео- та аудіопотоках, а у поєднанні з методами машинного навчання – стати основою гібридних рішень з підвищеною стійкістю та точністю детектування сучасних стеганографічних алгоритмів.

Ключові слова: стеганоаналіз, перетворення Уоліша-Адамара, сингулярний розклад, приховані канали зв'язку, кібербезпека, вебзастосунки, частотний аналіз.

Дата надходження статті: 11.08.2025

Дата прийняття статті: 04.09.2025

Опубліковано: 27.10.2025

Добровольський Ю.Г.

Чернівецький національний університет імені Юрія Федьковича

Дячук Р.Л.

Чернівецький національний університет імені Юрія Федьковича

ДОСЛІДЖЕННЯ ГЕНЕРАЦІЇ ВИПАДКОВОЇ ПОСЛІДОВНОСТІ ЧИСЕЛ НА ОСНОВІ ТЕМНОВОГО СТРУМ ФОТОДІОДА ПРИ РІЗНИХ ТЕМПЕРАТУРАХ

У статті розглянуто актуальне питання сучасної програмної інженерії, яке полягає у розробці та дослідженні зовні непередбачуваних псевдовипадкових числових послідовностей, які можуть бути покладені в основу розробки протоколів кібербезпеки. В нашому випадку в якості джерела випадкових чисел виступає темновий струм фотодіода. Відомі дослідження розглядали певні напівпровідникові вироби в якості як джерела псевдовипадкової послідовності. Але фотодіоди не були розглянуті.

Метою дослідження є дослідження генерації випадкової послідовності чисел на основі темного струму фотодіода, які отриманих при різних температурах. Для досягнення мети дослідження було розроблено установку для вимірювання темного струму фотодіода ФД-288 при температурі 10, 30 та 50 °C з частотою відбору 20000 значень темного струму в секунду. Також розроблений алгоритм генерації псевдовипадкової послідовності чисел, основних на значеннях темного струму фотодіода. Досліджено генерацію випадкової послідовності чисел на основі вимірних значень зворотного (темного) струму фотодіода, отриманих при різних температурах. А саме при 10, 30 та 50 °C.

В результаті проведених досліджень встановлено, що найкращі результати досягнуті при генерації випадкової послідовності чисел на основі темного струму фотодіода, отриманої при 50 °C. А саме – мінімальне значення складає 0,0 при вимозі 0.39; максимальне значення складає 0.68 при вимозі 0.39; середнє значення 0.39 при вимозі 0.39; середньоквадратичне відхилення 0.18 при вимозі 0.04; процент відсутніх елементів 0 при вимозі 0. Показано, що фотодіод, при температурі 50 °C, може бути джерелом для генерації послідовності псевдовипадкових чисел. При цьому розподіл чисел є рівномірно-хаотичним, а кожне число рівномірно представлене не частіше 0.18%, що менше за ідеальні умови, але достатньо для практичного застосування.

Подальші дослідження передбачають вивчення можливостей створеного обладнання з точки зору застосування в якості джерела ентропії інших фотодіодів та алгоритму для генерації ПВП на основі темного струму фотодіода та підвищення швидкодії методу.

Ключові слова: програмна інженерія, кібербезпека, фотодіод, темновий струм, випадкові числа.

Постановка проблеми. Інформаційні технології забезпечують практично весь цифровий інформаційний потік у сучасному світі. Однак, цей потік має бути надійно захищений з огляду на інтереси як окремої людини користувача, так і держави у цілому. Це завдання загалом вирішується, як відомо, засобами кібербезпеки, методи якої ґрунтуються, у великій мірі, на технологіях програмної інженерії. Зокрема, вельми актуальним з цієї точки зору, є випадкові числові послідовності, які покликані убезпечити потоки вхідних і вихідних даних. На разі застосовуються багато методів генерації випадкових числових

послідовностей. Наприклад, серед найвідоміших методів є використання обсягу стекової/купної пам'яті комп'ютера, або точне вимірювання поточного значення часу з точністю до наносекунд (час Unix), які дають набір випадкових значень, які можна використати для генерації відповідної випадкової послідовності. Зовнішня периферія комп'ютера, така як USB, клавіатура, або миша також можуть слугувати джерелами випадкових чисел. Загалом, усі ці джерела називають джерелами ентропії, або, інакше, хаосу, оскільки вважається, що їх, ці випадкові послідовності, не можна передбачити. Але, значення

випадкових послідовностей, отримані вище наведеними методами, не можна вважати абсолютно випадковими. Вони генеруються у певних діапазонах, підкоряються певним коливанням. Усі ці особливості можливо передбачити. Щоби такого не трапилося такі послідовності чисел можна перетворити на справді випадкові числа, як що їх піддати криптографічним перетворенням. Наприклад за допомогою клітинних автоматів [1]. Таки перетворення створюють рівномірно розподілені випадкові значення кожного числа в послідовності. Також вони характеризуються нерівномірним розподілом значення джерела ентропії. Отримані таким чином числові послідовності прийнято називають псевдовипадковими, оскільки вони створюються детерміновано із ентропії, і не є дійсно випадковими.

Отже, джерело ентропії (хаосу), будь якого походження, можна використовувати для створення випадкових числових послідовностей. Це дозволяє отримати непередбачуваність цифрових значень, які містяться у такій послідовності. Вони використовуються в протоколах безпеки мереж, для крипто безпечних з'єднань, створення криптографічних ключів, відслідковування цілісності системи і в багатьох інших цілях [2].

Підсумовуючи вище наведене, можна напевне стверджувати, що розробка надійних, крипто стійких протоколів, які в своїй основі мають генерацію непередбачуваних, псевдовипадкових числових послідовностей є дуже актуальним завданням для сучасної програмної інженерії.

Аналіз останніх досліджень і публікацій. Сучасні генератори псевдовипадкових числових послідовностей, які базуються лише на певному програмному забезпеченні, не можуть здолати певної передбачуваності [3]. Окрім того, існуючи технології створення псевдовипадкових послідовностей, є загальнодоступні. Як що взяти генерацію ентропії, отриману за допомогою мови Java [4], то такі послідовності можуть бути, певним чином, теоретично, доступні для зламу алгоритму шифрування. Застосування квантових обчислень в кібератаках [5, 6], збільшує ефективність таких атак практично до успішного рівня.

Відомі, загалом, два підходи, за допомогою яких створюються послідовності випадкових чисел. Перший підхід використовує спеціалізовані прилади, основані на фізичних джерелах шуму [7]. Їх недолік полягає у необхідності додаткового приладового обладнання (адаптерів) та спеціалізованого програмного забезпечення для комунікації із комп'ютером.

Інший підхід базується на використанні фізичних подій, у комп'ютерних пристроях. З точки зору надійності, такий підхід є більш виправданим. Наприклад, використання в якості генератора псевдовипадкової послідовності лічильника тактової частоти процесора. Його недолік полягає у тому, що він є чутливим до зовнішніх факторів, і, таким чином, до зовнішнього впливу на процедуру генерування випадкової послідовності чисел [8]. Оптичний маніпулятор миші також може бути запропонований для генерації випадкової числової послідовності [9]. Оптичний маніпулятор миші створює досить неоднорідний розподіл випадкових чисел, про те, має малу швидкість генерації (біля 1 кбіт/с), що не дозволяє його використання у швидкісних системах шифрування.

У [10] запропонований інший метод генерації випадкової числової послідовності. Зокрема, автори використовують веб-камеру, а саме піксели, з яких веб-камера створює зображення, в якості джерела зовнішньої непередбачуваності. Як показано авторами у [10] динаміка інтенсивності пікселів, які містяться у матриці, побудованій на основі приладу з зарядовим зв'язком (ПЗЗ)-матриця веб-камери, є абсолютно непередбачуваною. Її не видно звичайним оком, разом з тим, вона чітко фіксуються як вимірювальною апаратурою, так і програмними методами.

Спираючись на вище наведене, було розглянуто питання щодо використання в якості джерела псевдовипадкової послідовності чисел не веб-камери (пікселів, які генерують зображення), а фізичні властивості ПЗЗ матриці, на основі якої побудована веб-камера. Такі матриці створюються за звичай на основі кремнію, як що веб-камера працює у видимому діапазоні спектру. Постало питання, а чи доцільно використання в якості джерела псевдовипадкової числової послідовності не інтенсивність пікселів веб-камери, значення темного струму кремнієвого фотодіоду. Ця величина надзвичайно динамічна і її конкретна величина у кожен момент часу змінюється. Також темновий струм фотодіоду надзвичайно чутливий до температури та інших зовнішніх впливів і є дійсно непередбачуваною величиною [11]. Раніше, у [12], розглядалися певні напівпровідникові вироби в якості як джерела псевдовипадкової послідовності. Але не фотодіоди.

Постановка завдання. Метою роботи є: дослідження генерації випадкової послідовності чисел на основі зворотного струм фотодіода, отриманих при різних температурах.

Виклад основного матеріалу. Раніше ми досліджували можливість та ефективність генерації псевдовипадкової послідовності чисел, заснованому на фіксації темного струму кремнієвого фотодіода ФД-288. Ці дослідження показали, що ряд випадкових чисел у діапазоні $[0, 9]$, які відповідають значенням темного струму фотодіода, і цифровані за допомогою USB-адаптера на базі Arduino (112593), мають швидкість генерації, інакше – продуктивність (Р) 980 байт/сек. Що є дуже мало. Вимоги криптографії вказують, щоби випадкові числа хаотичної послідовності були у діапазоні $[0, 256]$. Це відповідає типу даних byte, які притаманні мови програмування Java. Підсилення сигналу темного струму на 5 Вольт, переводить характеристику розподілу по значенню псевдовипадкових чисел у бік «рівномірно-хаотичної». А саме, кожне отримане число у послідовності представлено рівномірно. Його повторення зустрічається у діапазоні 0.3–0.7 %, при ідеальному варіанті: $1/256 = 0.004$, або 0.4%.

З урахуванням вище наведеного, було запропоновано підвищити температуру фотодіода для збільшення розкиду значень темного струму. Було проведено вимірювання темного струму фотодіода ФД-288 при температурі 10, 30 та 50 °C, що відповідає умовам його експлуатації за допомогою установки, блок-схема якої наведена на рисунку 1.

Температуру на фотодіоді знижували до 10 °C і підвищували до 30 та 50 °C за допомогою термоелектричного охолоджувача. Зменшення чи збільшення температури на фотодіоді здійснювали змінюючи полярність живлення на охолоджувачі. Температура фотодіода контролювалася в цокольній області його корпусу, за допомогою термо-

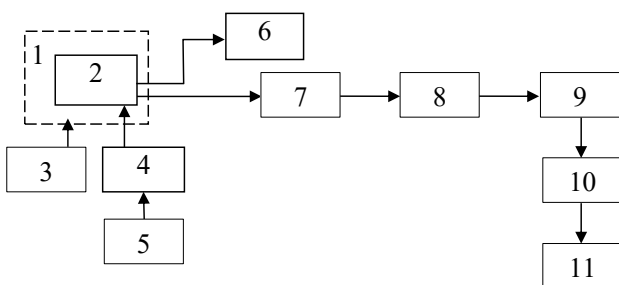


Рис. 1. Блок-схема установки для вимірювання темного струму фотодіода ФД-288 при температурі 10, 30 та 50 °C

1 – світло不透кий екран, 2 – фотодіод, 3 – блок живлення фотодіода, 4 – термоелектричний охолоджувач, 5 – блок живлення термоелектричного охолоджувача, 6 – термометр цифровий Тензор-42, 7 – ППЧН, 8 – контролер Arduino (112593), 9 – USB-адаптер, 10 – USB порт комп'ютера, 11 – процесор комп'ютера

метра цифрового Тензор-42, який забезпечує вимірювання температури у діапазоні 0–200 °C із дискретністю показів 0,01 °C із межею основної відносної похибки $\pm 0,05$ °C. Темновий струм вимірювався при зміщенні на р-п переході фотодіоду 1 В.

Далі, виміняні значення темного струму у режимі on-line за допомогою прецизійного перетворювача струм-напруга (ППЧН) перетворюються у напругу (U_T) і, використовуючи USB-адаптер на основі Arduino (112593), передавалися до комп'ютера, як це показано на рис. 1. В комп'ютері значення U_T відображалися у вигляді гістограм.

Далі, за допомогою засобів програмного забезпечення (ПЗ) мови програмування Java (версія 17), із використанням бібліотеки javax-usb, розроблений алгоритм генерації псевдовипадкової послідовності чисел (ПВЧ), основних на значеннях темного струму фотодіода (рис. 2).

У подальшому цей список випадкових чисел досліджується на відповідність статистичним характеристикам за допомогою бібліотеки DescriptiveStatistics. Для графічного відображення статистичних характеристик використовувалась бібліотека jfree.chart та jfree.data.

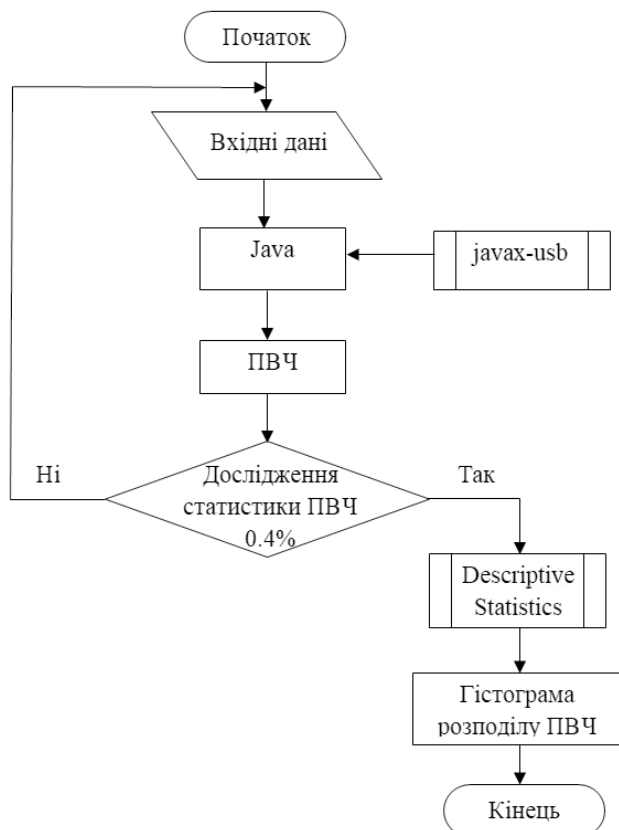


Рис. 2. Блок-схема алгоритму генерації псевдовипадкової послідовності чисел

При кожній температурі було виміряно по 20000 значень темного струму в секунду. Динаміка темного струму при трьох температурах – 10,

30 та 50 °С наведена на рисунках. 3, 4 та 5 відповідно.

З рисунків 3–5 видно, що при температурі 10 °С діапазон виміряних значень темного струму складає від 10^{-11} до $8 \cdot 10^{-9}$ А, при температурі 30 °С діапазон виміряних значень темного струму складає від 10^{-11} до $8 \cdot 10^{-7}$ А, а при температурі 50 °С діапазон виміряних значень темного струму складає від 10^{-11} до $7 \cdot 10^{-7}$ А. При цьому

найбільший розкид значень спостерігається при 10 °С, а найбільша густина розподілу значень темного струму спостерігається при 50 °С.

З отриманого набору даних (значень темного струмку, в подальшому перетворених у напругу і переданого до комп'ютера) були згенеровані ПВП згідно алгоритму, наведеному на рис. 2.

На рис. 6. наведена гістограма розподілення випадкових чисел по значенню для температури 10 °С.

З рисунку видно, що не всі елементи випадкової послідовності (байти) представлені одноковою часткою присутності. Байти у діапазоні [5 ... 25]

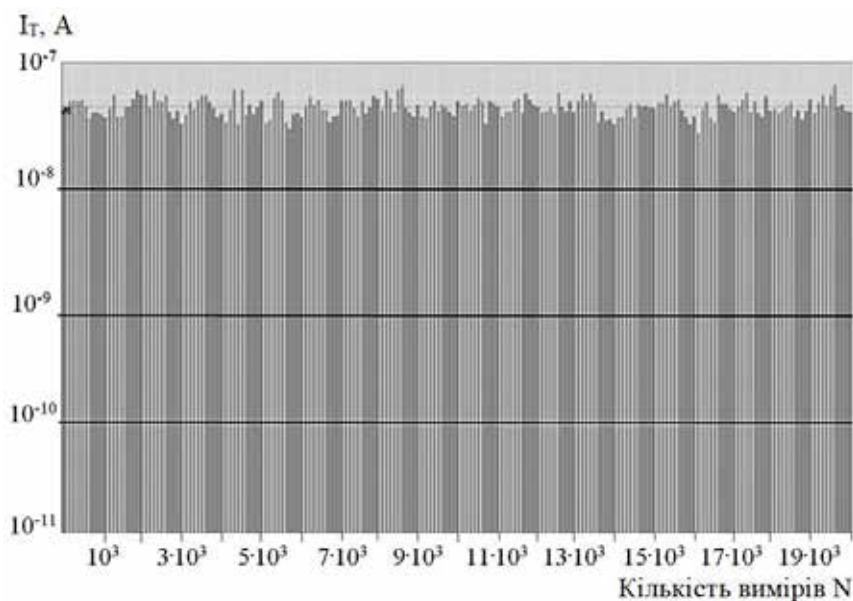


Рис. 3. Динаміка темного струму при температурі 50 °С

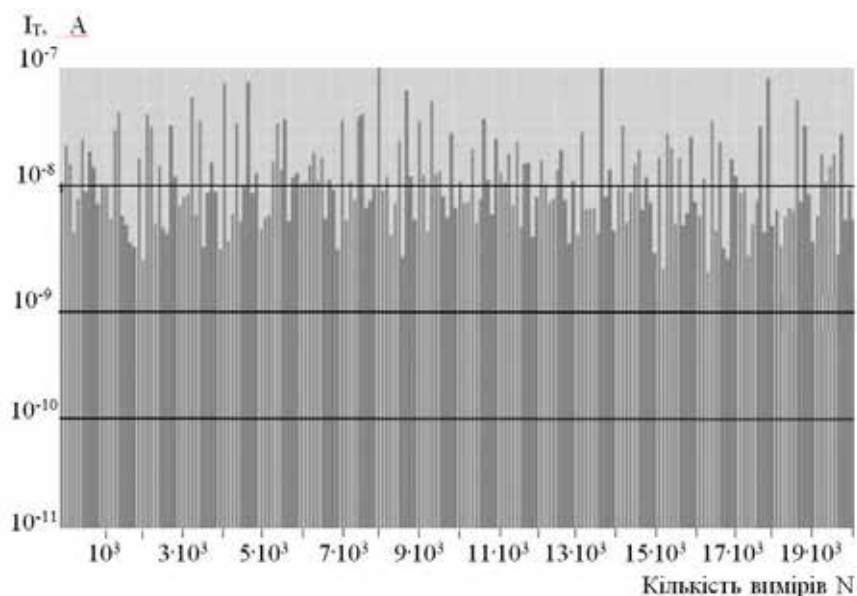


Рис. 4. Динаміка темного струму при температурі 30 °С

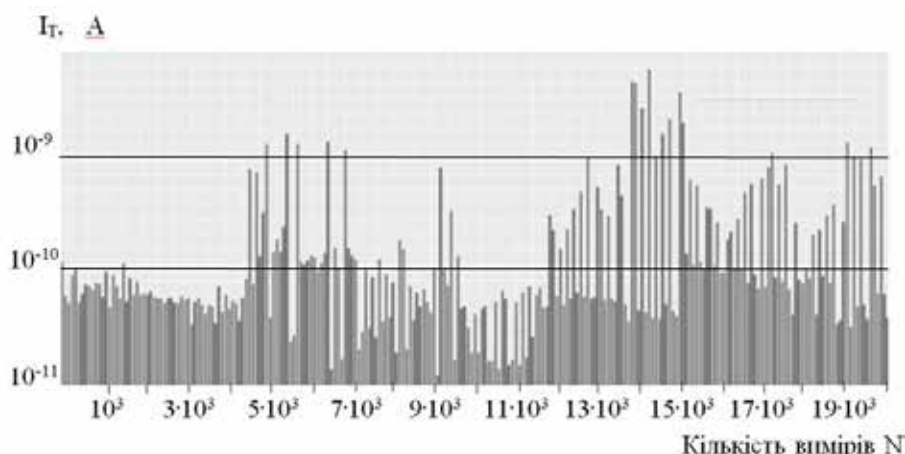


Рис. 5. Динаміка темного струму при температурі 10 °C

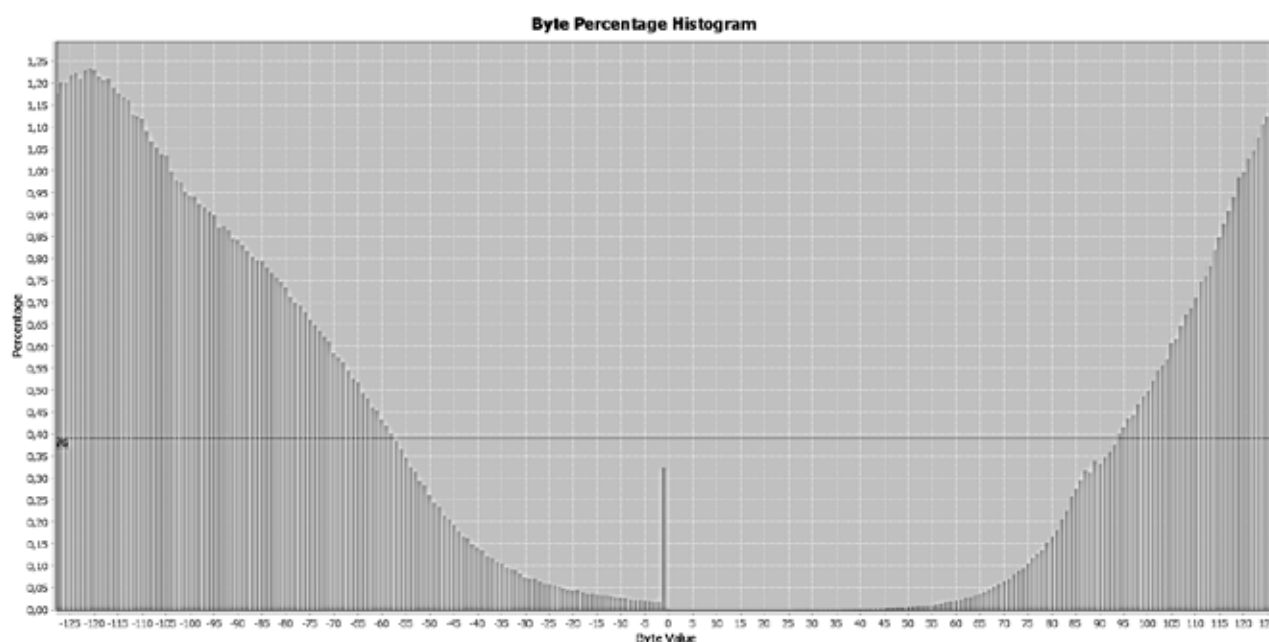


Рис. 6. Гістограма розподілення випадкових чисел, що були згенеровані, по значенню величини для температури 10 °C

взагалі відсутні у послідовності, а байти у діапазоні [120 ... 125] займають приблизно 1,2 % частки від усіх елементів послідовності. Хоча всі числа носять абсолютно випадковий характер, такий розподіл не є задовільним з точки зору вимог програмної інженерії до криптозахисту.

На рис. 7 наведена гістограма розподілення випадкових чисел по значенню для температури 30 °C.

Для температури 30 °C видно з гістограми, що розподілення дещо вирівнялась – немає відсутніх елементів. Проте круті нерівності гістограми не дають нам змогу характеризувати цей розподіл, як рівномірний. Скоріше всього це хаотичний розподіл.

На рис. 8 наведена гістограма розподілення випадкових чисел по значенню для температури 50 °C.

Для температури 50 °C видно з гістограми, що розподілення дещо стала плавною. Це дає нам змогу характеризувати цей розподіл, як рівномірно-хаотичний. І для деяких випадків (особливо у стеганографії) такий розподіл є цілком задовільний.

Проте візуальний огляд гістограми не дає нам кількісних показників якості згенерованих ПВЧ. За допомогою мови Джава, зокрема статистичної бібліотеки DescriptiveStatistics, вдалось обчислити параметри, що характеризують якість згенерованої ПВЧ для кожної температури.

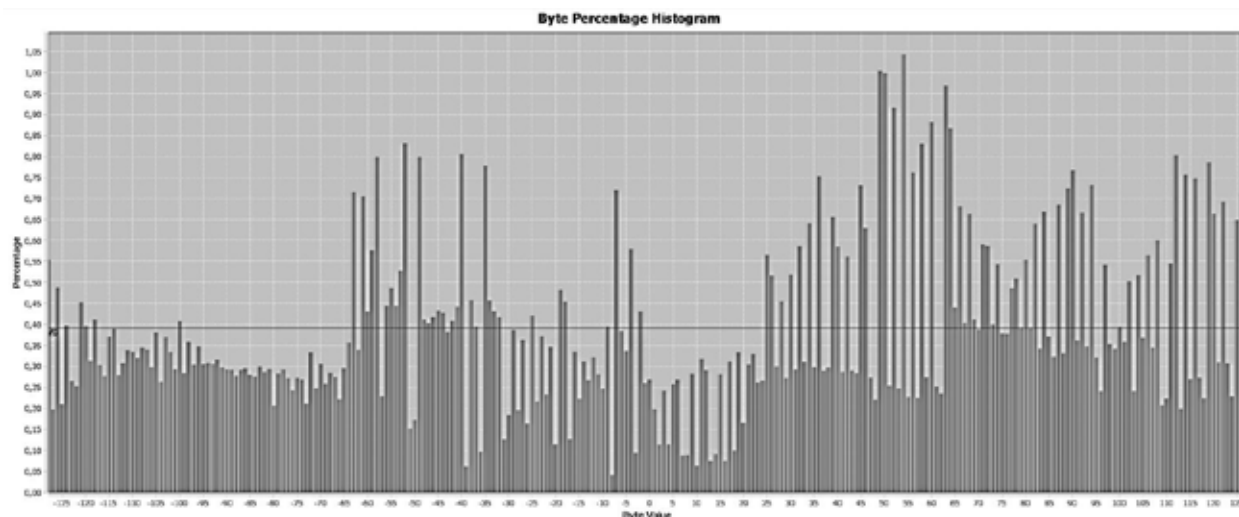


Рис. 7. Гістограма розподілення випадкових чисел, що були згенеровані, по значенню величини для температури 30 °C

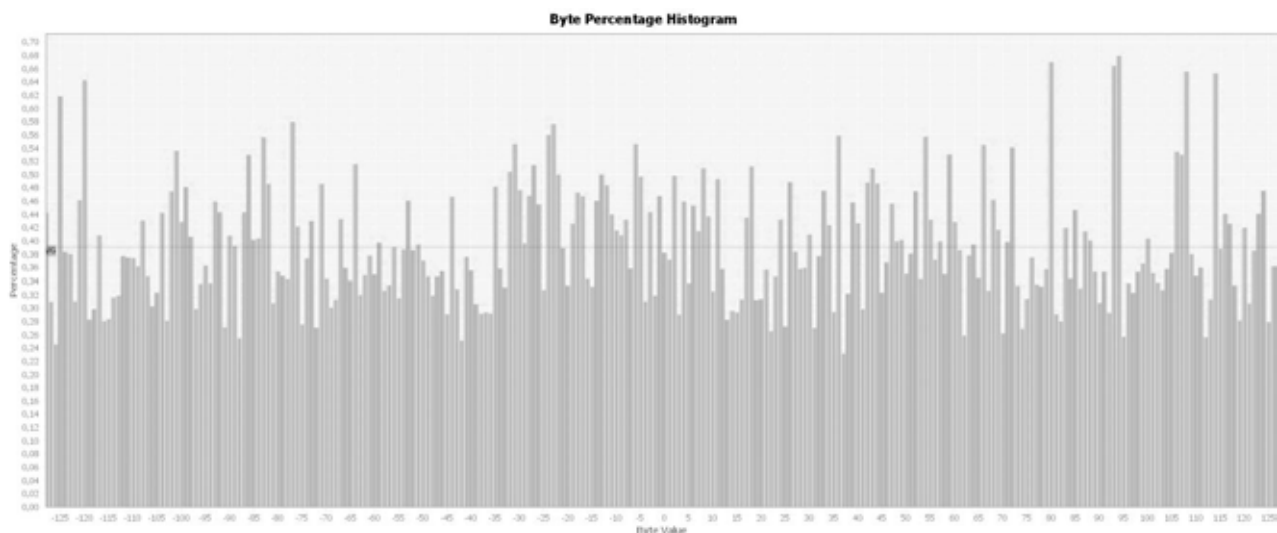


Рис. 8. Гістограма розподілення випадкових чисел, що були згенеровані, по значенню величини для температури 50 °C

Порівняльні статистичні характеристики згенерованих послідовностей, отриманих з значень темного струму при різних температурах, наведені у таблиці 1.

Нагадаємо, що для ідеального рівномірного розподілу [13]:

- мінімальне значення $\min = 0.39$;
- максимальне значення $\max = 0.39$;
- середнє значення $\text{mean} = 0.39$;
- середньоквадратичне відхилення $\text{st dev} = 0.04$;
- процент відсутніх елементів $\text{absent} = 0\%$.

Самою показовою характеристикою є середньо квадратичне відхилення. Чим менше відхилення, тим менше розкид значень, тим більш рівномірним є розподіл.

Як видно з таблиці 1, середньо квадратичне відхилення зменшується із зростанням температури. Зокрема встановлено, що найкращі результати досягнуті при генерації випадкової послідовності чисел на основі темного струму фотодіода, отриманої при 50 °C. А саме – мінімальне значення складає 0,0 при вимозі 0.39; максимальне значення складає 0.68 при вимозі 0.39; середнє значення 0.39 при вимозі 0.39; середньоквадратичне відхилення 0.18 при вимозі 0.04; процент відсутніх елементів 0 при вимозі 0.

Показано, що фотодіод, при температурі 50 °C, може бути джерелом для генерації послідовності псевдовипадкових чисел. При цьому розподіл чисел є рівномірно-хаотичним, а кожне число

Статистичні характеристики згенерованих послідовностей

Температура	10°C	30°C	50°C	Вимоги ідеального розподілу ПВП
Статистичні параметри розподілу, %	min: 0.0 max: 1.27 mean: 0.39 std dev: 0.85 absent: 8%	min: 0.0 max: 1.05 mean: 0.39 std dev: 0.33 absent: 0	min: 0.0 max: 0.68 mean: 0.39 std dev: 0.18 absent: 0	min: 0.39 max: 0.39 mean: 0.39 std dev: 0.04 absent: 0

рівномірно представлене не частіше 0.18%, що менше за ідеальні умови, але достатньо для практичного застосування.

Подальші дослідження передбачають вивчення можливостей створеного обладнання з точки зору застосування в якості джерела ентропії інших фотодіодів та алгоритму для генерації ПВП на основі темного струму фотодіода та підвищення швидкодії методу.

Висновки. У даному дослідженні розроблено установку для вимірювання темного струму фотодіода ФД-288 при температурі 10, 30 та 50 °C з частотою відбору 20000 значень темного струму в секунду. Також розроблений алгоритм генерації псевдовипадкової послідовності чисел, основних на значеннях темного струму фотодіода. Досліджено генерацію випадкової послідов-

ності чисел на основі вимірних значень зворотного (темного) струму фотодіода, отриманих при різних температурах. А саме при 10, 30 та 50 °C. Встановлено, що найкращі результати досягнуті при генерації випадкової послідовності чисел на основі темного струму фотодіода, отриманої при 50 °C. А саме – мінімальне значення складає 0,0 при вимозі 0.39; максимальне значення складає 0.68 при вимозі 0.39; середнє значення 0.39 при вимозі 0.39; середньоквадратичне відхилення 0.18 при вимозі 0.04; процент відсутніх елементів 0 при вимозі 0. Показано, що фотодіод, при температурі 50 °C, може бути джерелом для генерації послідовності псевдовипадкових чисел. При цьому розподіл чисел є рівномірно-хаотичним, а кожне число рівномірно представлене не частіше 0.18%.

Список літератури:

1. Cicuttin A. De Micco L. Crespo M. L. et al. Looking for suitable rules for true random number generation with asynchronous cellular automata. *Nonlinear Dynamics*. 2022. Vol. 111, P. 2711-2722. DOI: 10.1007/s11071-022-07957-8.
2. Aljahdal A. O. Random Number Generators Survey. *International Journal of Computer Science and Information Security*. 2020. Vol. 18, No. 10. P. 14-22. DOI: 10.5281/zenodo.4249406.
3. Martinez F. Attacks on Pseudo Random Number Generators Hiding a Linear Structure. *Cryptographers' Track at the RSA Conference 2022*. (Virtual Event, United States., Mar 2022). P. 145-168.
4. Class SecureRandom. All Implemented Interfaces. URL: <https://docs.oracle.com/javase/8/docs/api/java/security/SecureRandom.html> (дата звернення: 21.06.2025).
5. Остапов С.Е., Добровольський Ю.Г. Квантова інформатика та квантові обчислення: навч. посібник. Чернівці: ЧНУ ім. Ю. Федьковича, 2021. – 99 с.
6. Jacak J.E., Jacak W.A., Donderowicz W.A. Wojciech A. Jacak L. Quantum random number generators with entanglement for public randomness testing. *Scientific Reports*. 2020. Vol. 10, № 164. P. 121-130. DOI: 10.1038/s41598-019-56706-2.
7. Бараннік В. В., Сідченко С. О., Бараннік Н. В., Хіменко А. М. Метод маскування уцілювання службових даних в системах компресії відеозображень. *Радіоелектронні і комп'ютерні системи*. 2021. № 2 (98). С. 51-63 DOI: 10.32620/reks.2021.2.05.
8. Бараннік В. В., Бараннік Н. В., Ігнат'єв О. О., Хіменко А. М. Метод непрямого приховування інформації в процесі стиснення відеозображень. *Радіоелектронні і комп'ютерні системи*. 2021. № 4. С. 119-131 DOI: 10.32620/reks.2021.4.10.
9. Ostapov, S., Diakonenko, B., Fylypiuk, M., Hazdiuk, K., Shumylyak, L., Tarnovetska, O. Symmetrical Cryptosystems based on Cellular Automata. *International Journal of Computing*. 2023. № 22. P. 15-20. DOI: 10.47839/ijc.22.1.2874.
10. Diachuk R., Dobrovolsky Y., Hanzhelo D., Prokhorov H., Trembach D. Research the Level of Chaotic and Reliability in Webcam-generated Random Number Sequences. *Security of Infocommunication Systems and Internet of Things*. 2024. Vol. 2, № 1. P. 1-6. DOI: 10.31861/sisiot2024.1.01004.

11. Rogalski A. Infrared and Terahertz Detectors. Third Edition. London, CRC Press, 2022, 1066 P.
12. Soorat R., Ashok M. K. Vudayagiri Hardware Random number Generator for cryptography. arXiv. 2015. DOI: 10.48550/arXiv.1510.01234.
13. Веригіна І. В., Островська О. В. Теорія ймовірностей та математична статистика: Частина 2. Випадкові величини: Лекції і практикум. Київ, КПІ ім. Ігоря Сікорського, 2021. 77 с. URL: <https://surl.li/kuhkuk> (дата звернення: 21.06.2025).

Dobrovolsky Yu.G., Diachuk R.L. STUDY OF RANDOM NUMBER SEQUENCE GENERATION BASED ON DARK CURRENT OF A PHOTODIODE AT DIFFERENT TEMPERATURES

The article considers a topical issue of modern software engineering, which consists in the development and study of externally unpredictable pseudo-random number sequences that can be used as a basis for the development of cybersecurity protocols. In our case, the dark current of a photodiode acts as a source of random numbers. Known studies have considered certain semiconductor products as sources of a pseudo-random sequence. But photodiodes were not considered.

The purpose of the study is to study the generation of a random sequence of numbers based on the dark current of a photodiode, which are obtained at different temperatures. To achieve the goal of the study, a device was developed for measuring the dark current of a photodiode HD-288 at temperatures of 10, 30 and 50 °C with a sampling frequency of 20000 dark current values per second. An algorithm for generating a pseudo-random sequence of numbers based on the dark current values of the photodiode has also been developed. The generation of a random sequence of numbers based on the measured values of the reverse (dark) current of a photodiode obtained at different temperatures was investigated. Namely at 10, 30 and 50 °C.

As a result of the conducted research, it was found that the best results were achieved when generating a random sequence of numbers based on the dark current of a photodiode obtained at 50 °C. Namely – the minimum value is 0.0 at the requirement of 0.39; the maximum value is 0.68 at the requirement of 0.39; the average value is 0.39 at the requirement of 0.39; the standard deviation is 0.18 at the requirement of 0.04; the percentage of missing elements is 0 at the requirement of 0. It is shown that a photodiode, at a temperature of 50 °C, can be a source for generating a sequence of pseudo-random numbers.

In this case, the distribution of numbers is uniformly chaotic, and each number is uniformly represented no more than 0.18%, which is less than ideal conditions, but sufficient for practical application. Further research involves studying the capabilities of the created equipment in terms of using other photodiodes as a source of entropy and an algorithm for generating PVP based on the dark current of the photodiode and increasing the speed of the method.

Key words: software engineering, cybersecurity, photodiode, dark current, random numbers.

Дата надходження статті: 20.06.2025

Дата прийняття статті: 30.06.2025

Опубліковано: 27.10.2025

Дубко В.О.

Таврійський національний університет імені В.І. Вернадського

ЗАСТОСУВАННЯ ІНДИКАТОРНИХ ПРОЦЕСІВ ПРИ МОДЕЛЮВАННІ ТА ДОСЛІДЖЕННІ ПОКАЗНИКІВ СТОХАСТИЧНИХ СИСТЕМ ІЗ ЗМІННОЮ СТРУКТУРОЮ

Ця робота є розширенням та продовженням наших досліджень, пов'язаних з використанням випадкових індикаторних процесів $\chi_l(t), l=1, (n-1), \chi_l(t) \in \{0,1\}$ при моделюванні динамічних систем з стрибкоподібною змінною структурою. В основу покладено можливість розкладу довільного n -компонентного процесу на «ортогональні» компоненти з використанням несумісних процесів, утворених за певним алгоритмом комбінацій $\chi_l(t)$. Було продемонстровано ефективність застосування індикаторних процесів при моделюванні явищ, які можна трактувати як їх еволюцію у випадковому середовищі, характеристики якого змінюються стрибками у випадкові моменти часу. Закон цих змін узгоджувався з законом Пуассона. Для моделі дворівневого стану системи, встановлено зв'язок з моделлю Каца, що використовувалась для опису одновимірного процесу блукання частинки з постійною за величиною швидкістю, але з зміною напрямку її руху у випадкові моменти часу. Досліджені моделі побудовані на основі стохастичних рівнянь Іто зі стрибкоподібними змінами коефіцієнтів, побудовані рівняння для характеристичних функцій та знайдено їх розв'язок. Спираючись на основні результати цих робіт, в даній статті детальніше розглянуто варіанти представлення індикаторних випадкових процесів, зосереджено увагу на процесах, пов'язаних з випадковими інтервалами часу перебування системи у певних можливих станах, розподілі показників системи, залежних від цих інтервалів. Як приклад такого показника, вибрано надійність багатоелементної системи, яка може перебувати у певному k -му стані, для якого відомий аналітичний вираз зміни її надійності в залежності від часу функціонування при фіксованих показниках k -го стану «зовнішнього середовища». У такій трактовці надійність системи вже є випадковою функцією від часу. Для дворівневого випадку встановлені явні вирази щільності розподілу випадкового часу та надійності, їх зв'язок з класичними розв'язками телеграфного рівняння. Вказано на особливості опису багаторівневих, багатоелементних систем, шляхи пошуку розподілів характеристик аналогічних тим, що були розглянуті в роботі для системи з двома можливими станами.

Ключові слова: індикаторні процеси, зовнішнє середовище, випадковий розподіл часу, надійність системи, процеси Пуассона, телеграфне рівняння.

Постановка проблеми. При побудові моделей реальних систем використовуються різні методи моделювання. Чим більший набір цих методів, тим більш широкий клас систем стає доступним для відображення та дослідження математичними методами. Так застосування методів і висновків класичної теорії стохастичних диференціальних рівнянь (СДР) [3] з невідповідними коефіцієнтами, наприклад, не завжди дозволяє моделювати системи у випадковому середовищі (ВС). Вплив зміни стану ВС на реальну систему потребує розгляду і моделей з стрибкоподібною змінною коефіцієнтів у СДР (системи з змінною структурою). Змістовні результати вдається одержати і при запровадженні нових понять, додаткових математичних структур [2].

Аналіз останніх досліджень і публікацій.

Існує безліч робіт присвячених функціонуванню систем у випадкових середовищах. Широкий огляд, перелік, класифікація та застосування таких моделей, виконано у [6]. Особливістю побудованих та розглянутих нами моделей, є використання несумісних випадкових індикаторних процесів при моделюванні стохастичних характеристик системи із стрибкоподібною змінною станів. У даній роботі, окрім того, що ми спираємось на наші попередні результати [1], [2], використовуємо і певні аналітичні результати та висновки з інших робіт [4], [5], на які посиляємось по мірі необхідності, при викладі основного матеріалу.

Постановка завдання. Метою цієї роботи є розвиток напрямку моделювання систем з змінною структурою, заснованого на використанні несумісних процесів, сформованих на основі набору індикаторних випадкових процесів [2], у тому числі, при знаходженні аналітичних розв'язків для розглянутих моделей. Ефективність такого напрямку була продемонстрована на прикладі стохастичних рівнянь Іто, в наших роботах [1-2]. Ця стаття є продовження досліджень у цьому напрямку, та пов'язана з демонстрацією запропонованого алгоритму в [2] при дослідженні стохастичних моделей показників системи, у котрих відбуваються випадкові зміни показників «середовища функціонування».

Виклад основного матеріалу. Основні положення, моделі та їх дослідження. Для визначеності, покладемо, що зміни стану системи, яку моделюємо, пов'язані з випадковими k станами «зовнішнього середовища» (ЗС), зміна яких відбувається у випадкові моменти часу за законом Пуассона. Будемо розглядати такі динамічні характеристики як час перебування системи та зміна її надійності для k -го стану ЗС. Систему розглядаємо як набір з k підсистем. Реакцію системи на зміну стану ЗС трактуємо як перехід до k -ї підсистеми, яка забезпечує роботу системи у цьому стані. Причому вважаємо, що зміна динамічних показників k -ї підсистеми відбувається лише в період її роботи.

Випадкові індикаторні процеси.

Наведемо деякі необхідні означення та властивості випадкових індикаторних процесів (ВІП) [2, с. 81].

Випадкові процеси $\chi_l(t), l = \overline{1, (n-1)}$ утворюють набір ВІП [2, с. 81], якщо кожен із них може приймати лише значення 1, з ймовірністю $p(\chi_l(t) = 1) = p_l(t)$, та 0, з ймовірністю $(1 - p_l(t))$.

Із визначення $\chi_l(t)$ слідує, що

$$(1 - \chi_l(t))\chi_l(t) = 0. \quad (1)$$

ВІП можна надати аналітичне представлення, наприклад таке: $\chi_l(t) = 2^{-1}[1 + \cos(\pi g_l(t))]$ – де $g_k(t)$ різні $\forall k$, цілочисельні, загалом, випадкові, процеси, які приймають обов'язково як парні, так і не парні значення.

Спираючись на $\chi_l(t), l = \overline{1, (n-1)}$, побудуємо послідовність процесів [2]:

$$y_1(t) = \chi_1(t), \dots, y_k(t) = \chi_k(t) \prod_{l=1}^{k-1} (1 - \chi_l(t)), \dots, \\ y_n(t) = \prod_{l=1}^{n-1} (1 - \chi_l(t)); \quad (2)$$

$y_k(t)$ приймають лише значення $\{0, 1\}$. Тобто, $y_k(t)$ теж є випадковими індикаторними процесами, та, враховуючи властивість (1), утворюють набір несумісних, ортогональних процесів:

$$y_l(t)y_k(t) = 0, \forall l \neq k.$$

Крім того, як доведено в [2, с. 81, 82] (Лема 1, 2)

$$y_l^\nu(t) = y_l(t), \forall l, k = \overline{1, n}, \nu > 0 \forall a_l \chi_k(t),$$

$$\sum_{k=1}^n y_k(t) = 1. \quad (3)$$

Остання рівність в (3) може розглядатись як наслідок властивостей $y_k(t)$, або алгебраїчної рівності: $a_1 + \sum_{k=1}^{n-1} a_k \prod_{l=1}^{k-1} (1 - a_l) + \prod_{l=1}^{n-1} (1 - a_l) = 1$, набору чисел.

Змінюючи порядок вибору при формуванні послідовностей типу (2), можемо побудувати $(n-1)!$ різних наборів ортогональних процесів. В подальшому, будемо розглядати послідовність в позначеннях (2).

Враховуючи (3), встановлюємо несумісності і таких процесів:

$$z_k(t) = f_k(t)y_k = f_k(t)\chi_k(t) \prod_{l=1}^{k-1} (1 - \chi_l(t)), z_n(t) = \\ = f_n(t)y_n = \prod_{l=1}^{n-1} (1 - \chi_l(t)) \\ z_k(t)z_{k'}(t) = \delta_{k,k'}, \forall t,$$

де $f_k(t), k = \overline{1, n}$ послідовності довільних процесів. (4) можна трактувати як $p(\chi_l(t) = 1) = p_l(t)$ розклад довільного процесу $F(t) = \{f_1(t), \dots, f_n(t)\}$ по базису ортогональних процесів (2). Як відмічено у [2, с. 86], при заданому наборі $p(z_k(t))$ існує можливість визначити ймовірності, які фігурують в (1).

Розподіл часу знаходження системи у фіксованому стані.

Перейдемо до розгляду випадкових процесів

$$x_k(t) = \int_0^t y_k(\tau) d\tau, \quad (5)$$

$$\sum_{l=1}^n x_k(t) = \int_0^t d\tau \sum_{l=1}^n y_k(\tau) = t, \quad 0 \leq x_k(t) \leq t.$$

Як бачимо, $x_k(t)$ відповідає випадковому часу знаходження системи у стані $z_k(t)$ до поточного моменту спостереження t .

Далі будемо працювати з таким набором незалежних індикаторних процесів

$$\chi_k(t) = 2^{-1}[1 + \cos(N_k(t)\pi)], \quad (6)$$

де $N_k(t)$ – незалежні процеси Пуассона інтенсивності λ_k , $k = \overline{1, (n-1)}$.

Відповідні елементи в (2) набувають вигляду:

$$y_k(t) = \chi_k(t) \prod_{l=1}^{k-1} (1 - \chi_l(t)) = 2^{-k} [1 + \cos(N_k(t)\pi)] \prod_{l=1}^{k-1} [1 - \cos(N_l(t)\pi)]. \quad (7)$$

Властивості випадкових індикаторних функцій приводять до рівностей $\chi_k^v = \chi_k, \delta(\cdot)$ $y_k^v(t) = y_k(t), \forall v > 0$. Тому, за необхідності, достатньо знаходити лише рівняння для χ_k , та $y_k(t)$.

Приклад 1. Розглянемо дворівневу систему: $y_1(t) = \chi(t)$, $y_2(t) = 1 - \chi(t)$. Почнемо з встановлення щільності розподілу випадкового часу знаходження системи у стані $y_1(t)$.

З урахуванням (5), (6) та (7),

$$x_1(t) = \int_0^t y_1(\tau) d\tau = 2^{-1} \int_0^t [1 + \cos(N(\tau)\pi)] d\tau = 2^{-1} t + 2^{-1} \int_0^t \cos(N(\tau)\pi) d\tau. \quad (8)$$

Перейдемо до нових змінних:

$$x(t) = x_1(t) - 2^{-1} t = 2^{-1} \int_0^t [\cos(N(\tau)\pi)] d\tau. \quad (9)$$

Процес $x(t)$ часто ототожнюється у літературі з координатою частинки, що рухається з постійною за величиною швидкістю на лінії, і змінює напрямок свого руху в випадкові моменти часу [4], [5]. У нашому випадку «швидкість» $c = 2^{-1}$, а зміна напрямку визначається функцією $(-1)^{N(t)} = \cos(N(t)\pi)$. $N(t)$ – Пуассонівський процес інтенсивності λ .

Характеристична функція $\bar{I}_1(\beta, t)$ для $x_1(t)$, з урахуванням того, що t не є випадковою величиною, допускає представлення:

$$\int_{-\infty}^{\infty} e^{i\beta x_1} \rho_1(x_1, t) dx_1 = \bar{I}_1(\beta, t) = e^{i\beta 2^{-1} t} \int_{-\infty}^{\infty} e^{i\beta x} \rho(x, t) dx \quad (10)$$

де $\rho(x, t)$ – щільність розподілу випадкового процесу (9), а $\rho_1(x_1, t)$ відповідає, випадковому процесу (8).

Виконаємо обернене перетворення для (10):

$$\begin{aligned} \rho_1(x_1, t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\beta x_1} I_1(\beta, t) d\beta = \int_{-\infty}^{\infty} \rho(x, t) dx \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\beta x} e^{i\beta 2^{-1} t} e^{-i\beta x_1} d\beta = \\ &= \int_{-\infty}^{\infty} \rho(x, t) dx \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\{-i\beta[x_1 - x - 2^{-1} t]\} d\beta = \\ &= \int_{-\infty}^{\infty} \rho(x, t) \delta([x - (x_1 - 2^{-1} t)]) dx = \rho(x_1 - 2^{-1} t, t) \end{aligned} \quad (11)$$

$\rho(x, t)$ – фундаментальний розв'язок телеграфного рівняння

$$\begin{cases} \frac{\partial^2 \rho(x, t)}{\partial t^2} + 2\lambda \frac{\partial \rho(x, t)}{\partial t} - 2^{-2} \frac{\partial^2 \rho(x, t)}{\partial x^2} = 0 \\ \rho(x, 0) = \delta(x), \quad \frac{\partial}{\partial t} \rho(x, t)|_{t=0} = 0. \end{cases} \quad (12)$$

Розв'язок (12) наведений у багатьох роботах. Використаємо аналітичний вигляд розв'язку (12), наприклад, із роботи [5, стор.142], адаптуючи його до наших позначень:

$$\begin{aligned} \rho(x, t) &= e^{-\lambda t} 2^{-1} [\delta(2^{-1} t + x) + \delta(2^{-1} t - x) + \\ &+ e^{-\lambda t} \left[\lambda I_0(2\lambda \sqrt{2^{-2} t^2 - x^2}) + \frac{\partial}{\partial t} I_0(2\lambda \sqrt{2^{-2} t^2 - x^2}) \right] \theta(2^{-1} t - |x|)]; \quad (13) \\ x &\in (-\infty, \infty), t > 0. \end{aligned}$$

де використанні позначення: $I_0(\cdot)$ – модифікована функція Бесселя нульового порядку, $\delta(\cdot)$ – дельта функція Дірака, а $\theta(\cdot)$ – одинична функція Хевісайда.

Повертаючись до (11) та врахувавши (5), отримуємо:

$$\begin{aligned} \rho_1(x_1, t) &= \rho(x_1 - 2^{-1} t, t) = \\ &= \frac{e^{-\lambda t}}{2} [\delta(x_1) + \delta(t - x_1) + e^{-\lambda t} \left[\lambda I_0(2\lambda \sqrt{x_1(t - x_1)}) + \right. \\ &\left. + \frac{\partial}{\partial t} I_0(2\lambda \sqrt{x_1(t - x_1)}) \right] \times \theta(t - x)]; \quad x_1 \in (0, \infty), t > 0. \end{aligned} \quad (14)$$

Позначимо через $\rho_2(x_2, t)$ щільність розподілу для випадкового часу знаходження системи у другому стані:

$$\begin{aligned} x_2(t) &= \int_0^t (1 - \chi(\tau)) d\tau = 2^{-1} \int_0^t [1 - \cos(N(\tau)\pi)] d\tau = \\ &= 2^{-1} t - 2^{-1} \int_0^t [\cos(N(\tau)\pi)] d\tau, \end{aligned}$$

Перейдемо до нової змінної:

$$x(t) = 2^{-1} t - x_2(t) = 2^{-1} \int_0^t [\cos(N(\tau)\pi)] d\tau,$$

Бачимо, що знову можемо скористатись розв'язком рівняння (12), за тим же алгоритмом, як і при знаходженні $\rho_1(x_1, t)$. Повертаючись до початкових змінних, прийдемо до такого представлення $\rho_2(x_2, t)$:

$$\begin{aligned} \rho_2(x_2, t) &= e^{-\lambda t} 2^{-1} [\delta(t - x_2) + \delta(x_2) + \\ &+ e^{-\lambda t} \left[\lambda I_0(2\lambda \sqrt{(t - x_2)x_2}) + \frac{\partial}{\partial t} I_0(2\lambda \sqrt{(t - x_2)x_2}) \right] \theta(t - x_2)]; \quad (15) \\ x_2 &\in (0, \infty), t > 0. \end{aligned}$$

Розподіл надійності.

Нехай

$$P(t) = \sum_{k=1}^n z_k(t) = \sum_{k=1}^n y_k(t) P_k \left(\int_0^t y_k(\tau) d\tau \right), \quad (16)$$

де $P_k(t)$ відомі функції.Характеристична функції процесу $P(t)$ (16), з урахуванням (4), допускає представлення:

$$\begin{aligned} M[\exp\{i\beta P(t)\}] &= M[\exp\{i\beta \sum_{k=1}^n y_k(t) P_k \left(\int_0^t y_k(\tau) d\tau \right)\}] = \\ &= \sum_{k=1}^n M[y_k(t) \exp\{i\beta P_k \left(\int_0^t y_k(\tau) d\tau \right)\}]. \end{aligned} \quad (17)$$

У цьому можна впевнитись, скориставшись ортогональністю (4) процесів $z_k(t)$ та властивістю (3):

$$\begin{aligned} M[\exp\{i\beta \sum_{k=1}^n y_k(t) P_k \left(\int_0^t y_k(\tau) d\tau \right)\}] &= \\ &= \prod_{k=1}^n M[\exp\{i\beta y_k(t) P_k \left(\int_0^t y_k(\tau) d\tau \right)\}] = \\ &= M \left[\prod_{k=1}^n \left\{ 1 + y_k(t) \sum_{j=1}^{\infty} \frac{(i\beta P_k \left(\int_0^t y_k(\tau) d\tau \right))^j}{j!} \right\} \right] = \\ &= M \left[1 - \sum_{k=1}^n y_k(t) + y_k(t) \sum_{k=1}^n \sum_{j=0}^{\infty} \frac{(i\beta P_k \left(\int_0^t y_k(\tau) d\tau \right))^j}{j!} \right]. \end{aligned}$$

Тобто, є можливість розглядати (17) як суму доданків

$$I_k(t; \beta) = M[y_k(t) \exp\{i\beta P_k \left(\int_0^t y_k(\tau) d\tau \right)\}], \quad (18)$$

та знаходити рівняння для кожної компоненти (18), взаємозалежності між ними. Позначимо через $g_k(P, t)$ обернені перетворення (18). Тоді рівність (17) замінюється на таку:

$$\rho(P, t) = \sum_{k=1}^n g_k(P, t). \quad (19)$$

Прикладом $P_k(t)$ може бути ймовірність функціонування системи – надійність. В (17), (18) $P_k \left(\int_0^t y_k(\tau) d\tau \right)$ – вже випадкові процеси. Множник $y_k(t)$ можемо трактувати як модель зміни стану «зовнішнього середовища», яке впливає на вибір (включення) робочої підсистеми з характеристикою $P_k(\cdot)$.

Аналітичне представлення зміни надійності системи часто зустрічається у вигляді $P_k(t) = \exp(-\gamma_k t)$, $\gamma_k > 0$. При такому виборі, (18) набуває вигляду:

$$I_k(t; \beta) = M[y_k(t) \exp\{i\beta \exp\{-\gamma_k \int_0^t y_k(\tau) d\tau\}\}]. \quad (20)$$

Приклад 2. Обмежимося представленням (20). Для дворівневого стану системи ($y_1(t) = \chi(t)$, $y_2(t) = 1 - \chi(t)$), у силу рівності (19), щільність розподіл надійності системи матиме вигляд:

$$\rho(P, t) = g_1(P, t) + g_2(P, t), \quad (21)$$

де $g_1(P, t)$, $g_2(P, t)$ визначаються через відповідні обернені перетворення (20).Розглянемо $I_1(\beta, t)$. Дещо змінимо його представлення:

$$\begin{aligned} I_1(\beta, t) &= M[\chi(t) \exp\{i\beta \exp\{-\gamma_1 \int_0^t \chi(\tau) d\tau\}\}] = \\ &= M[\{\chi(t) - \frac{\gamma_1^{-1}}{(i\beta)} \sum_{j=0}^{\infty} \frac{(i\beta)^{j+1}}{(j+1)!} \frac{\partial}{\partial t} (\exp\{-\gamma_1 \int_0^t \chi(\tau) d\tau\})^{j+1} - \\ &\quad - \frac{\partial}{\partial t} (\exp\{-\gamma_1 \int_0^t \chi(\tau) d\tau\})\}] \end{aligned}$$

Враховуючи що $M[\chi(t)] = p_\chi(t)$, та виконавши згортку суми, отримуємо:

$$\begin{aligned} (i\beta)\gamma_1 I_1(\beta, t) &= (i\beta)\gamma_1 p_\chi(t) - \frac{\partial}{\partial t} M[\exp\{i\beta \exp\{-\gamma_1 \int_0^t \chi(\tau) d\tau\}\}] + \\ &+ \frac{\partial}{\partial t} M[\exp\{-\gamma_1 \int_0^t \chi(\tau) d\tau\}] \end{aligned} \quad (22)$$

Обернене перетворення (22), приводить до рівняння:

$$\begin{aligned} -\gamma_1 \frac{\partial}{\partial P} g_1(P, t) &= \gamma_1 p_\chi(t) \frac{\partial}{\partial P} \delta(P) - \\ &- \frac{\partial}{\partial t} \int_0^\infty \delta(P - \exp\{-\gamma_1 x_1\}) \bar{\rho}_1(x_1, t) dx_1 + \\ &+ \delta(P) \frac{\partial}{\partial t} \int_0^\infty \exp\{-\gamma_1 x\} \bar{\rho}_1(x_1, t) dx_1, \quad P \in (0, 1]. \end{aligned} \quad (23)$$

$\bar{\rho}_1(x_1, t)$ вже було встановлено та визначається (14). Тому, відповідні інтеграли в (23) обчислюються. Подальше інтегрування по змінній P , з урахуванням обмеження $P \in (0, 1]$, яке є наслідком представлення $P_k(t) = \exp(-\gamma_k t)$, $\gamma_k > 0$, і веде до аналітичного представлення $g_1(P_1, t)$.

Алгоритм пошуку $g_2(P, t)$ такий же та приводить до рівняння вигляду (22), з відповідними зміною у позначеннях, та заміною $\bar{\rho}_1(x, t)$ на функцію $\tilde{\rho}_2(x, t)$, яка відповідає (16), а $p_\chi(t)$ на $(1 - p_\chi(t))$. Після знаходження $g_1(P, t)$ та $g_2(P, t)$, встановлюється і аналітичне представлення $\rho(P, t)$ в (21).

Послідовність знаходження функцій $g_k(P, t)$, $\forall k$, за умови використання представлення (20), провадиться за алгоритмом використаного

у Прикладі 2. Якщо вдалось встановити щільність $\bar{\rho}_k(x_k, t)$ розподілу процесу $x_k(t)$, то приходимо до рівняння типу (23), яке тепер набуває вигляду:

$$-\gamma_k \frac{\partial}{\partial P} g_k(P, t) = \gamma_k p_{y_k}(t) \frac{\partial}{\partial P} \delta(P) - \\ - \frac{\partial}{\partial t} \int_0^\infty \delta(P - \exp\{-\gamma_k x_k\}) \bar{\rho}_k(x_k, t) dx_k \\ + \delta(P) \frac{\partial}{\partial t} \int_0^\infty \exp\{-\gamma_k x_k\} \bar{\rho}_k(x_k, t) dx_k, P \in (0, 1],$$

де

$$p_{y_k}(t) = p_{x_k}(t) \prod_{l=1}^{k-1} (1 - p_{x_l}), \quad p_{y_n}(t) = \prod_{l=1}^{n-1} (1 - p_{x_l}).$$

Після знаходження усіх $g_k(P, t)$, встановлюється аналітичний вигляд $\rho(P, t)$ за формулою (19).

Зауважимо, що пошук $\bar{\rho}_k(x_k, t)$ для $k > 2$ пов'язаний з громіздкістю рівнянь для його визначення, певними труднощами, які не вдалось нам подолати.

Висновки. Розглянуті приклади використання індикаторних випадкових процесів вказують на можливість застосування, у подальшому, подібних моделей для оптимізації характеристик технічних систем, що функціонують у в умовах випадкових змін середовища. Відмітимо, що розглянуті приклади не обмежують прикладного їх застосування. Так випадковий розподіл часу знаходження системи у певному стані, наприклад, може трактуватись, як час затримки у порах, при русі частинки у пористому середовищі [1].

Список літератури:

1. Дубко В.О. Моделювання розповсюдження домішок у середовищі з центрами затримки за допомогою індикаторних функцій. *Аграрна наука та освіта в XXI столітті: проблеми, перспективи та інновації*: зб. Наукових праць № 9. Ніжин, 2018. С. 313–317.
2. Дубко В.А. Моделювання систем з випадковою структурою за допомогою індикаторних випадкових процесів. *Вчені записки ТНУ імені В.І. Вернадського. Серія: технічні науки*. Том 31 (70) № 4 2020. С. 81–86. DOI: <https://doi.org/10.32838/2663-5941/2020.4/11>
3. Gikhman I.I., and Skorokhod A.V. Introduction to the Theory of Random Processes, *Philadelphia, W. B. Saunders*, 1969. 354 P.
4. Кас, М. Probability and Related Topics in Physical Sciences, *Interscience Publ.: London, UK*, 1959. 266 p.
5. Orsingher E., Ratanov N. Random motions in inhomogeneous media *Theor. Probability and Math. Statist.* Vip. 76, 2007 No. 76, 2008, Pp.141–153. S 0094-9000(08)00738-2.
6. Özekici, S., Soyer, R. Reliability Modeling and Analysis in Random Environments. In: *An Expository Perspective. International Series in Operations Research & Management Science*, Springer, Boston, vol 67. 2004 Pp. 249–273. https://doi.org/10.1007/978-1-4419-9021-1_13

Doobko V.O. APPLICATION OF INDICATOR PROCESSES IN MODELING AND RESEARCH OF INDICATORS OF STOCHASTIC SYSTEMS WITH VARIABLE STRUCTURE

This work is an extension and continuation of our research related to the use of random indicator processes $\chi_l(t), l = \overline{1, (n-1)}$, $\chi_l(t) \in \{0, 1\}$ in modeling dynamic systems with a jump-like structure. The basis is the possibility of decomposing an arbitrary component process into “orthogonal” components using incompatible processes formed according to a certain combination algorithm. The effectiveness of using indicator processes in modeling phenomena that can be interpreted as their evolution in a random environment, the characteristics of which change by jumps at random moments of time, was demonstrated. The law of these changes was consistent with Poisson’s law. For the model of the two-level state of the system, a connection was established with the Katz model, which was used to describe the one-dimensional process of particle wandering with a constant velocity in magnitude, but with a change in the direction of its motion at random moments of time. The studied models are built on the basis of stochastic Ito equations with jump-like changes in coefficients, equations for characteristic functions were constructed and their solution was found. Based on the main results of these works, this article considers in more detail the options for representing indicator random processes, focusing on processes associated with random time intervals of the system being in certain possible states, the distribution of system indicators depending on these intervals. As an example of such an indicator, the reliability of a multi-element system is chosen, which can be in a certain k -th state, for which an analytical expression of the change in its reliability depending on the time of operation is known at fixed indicators of the k -th state of the “external environment”. In such an interpretation, the reliability of the system is already a random function of time. For the two-level case, explicit expressions for the density distribution of random time and reliability are established, their

connection with classical solutions of the telegraph equation. The features of the description of multi-level, multi-element systems are indicated, and ways of finding distributions of characteristics similar to those considered in the work for a system with two possible states are indicated.

Key words: *indicator processes, external environment, random time distribution, system reliability, Poisson processes, telegraph equation.*

Дата надходження статті: 04.09.2025

Дата прийняття статті: 12.09.2025

Опубліковано: 27.10.2025

Здольник М.О.Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»**Степанець О.В.**Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»

МУЛЬТИАГЕНТНИЙ ПІДХІД ДО АРХІТЕКТУРИ ЛОКАЛЬНОЇ СИСТЕМИ АВТОМАТИЧНОГО КЕРУВАННЯ ОБ'ЄКТАМИ ТЕПЛОПОСТАЧАННЯ ПРОМИСЛОВОГО ТА ЖИТЛОВОГО ПРИЗНАЧЕННЯ

У статті досліджено можливості впровадження мультиагентної архітектури у локальні системи автоматичного керування об'єктами теплопостачання, зокрема індивідуальними тепловими пунктами (ІТП) житлового та промислового призначення. Обґрунтовано, що традиційно використовувана централізована структура локальних систем керування ІТП має обмежену здатність до адаптації, масштабування та координації між функціональними модулями, особливо в умовах змінних зовнішніх факторів. З огляду на потреби цифрової трансформації теплової інфраструктури запропоновано концепцію переходу до мультиагентної моделі, в якій кожен функціональний блок ІТП виконує роль автономного програмного агента з чітко визначеною функцією, ціллю керування та інтерфейсами взаємодії. Методологія дослідження базується на системному аналізі контурів керування та принципів функціональної декомпозиції. Для обґрунтування архітектурного підходу проаналізовано структуру двох типових об'єктів: водяного ІТП багатоквартирного житлового будинку та пароводяного ІТП промислового підприємства. Для кожного з них виконано декомпозицію на агентні компоненти, визначено функції, цілі та сценарії взаємодії. Побудовано графи інформаційного обміну між агентами та розроблено приклади поведінки системи в різних режимах – від штатного запуску до аварійних подій або зміни попиту. Концептуально запропоновано технічне рішення реалізації мультиагентної моделі на базі сучасних засобів промислової автоматизації. Результати дослідження демонструють, що перехід до мультиагентної архітектури дозволяє забезпечити адаптивність і гнучкість систем керування, зберігаючи контури регулювання як базові механізми керування. Визначено переваги такої архітектури, зокрема її масштабованість та можливість інтеграції з розподіленими інформаційними платформами. Перспективами подальших досліджень визначено створення прототипів агентних систем для реальних ІТП, застосування машинного навчання для підвищення ефективності рішень та розширення моделі на рівень теплових мереж у концепції Smart Thermal Grid.

Ключові слова: автоматизація технологічних процесів, мультиагентні системи керування, теплопункт, децентралізоване керування, цифрова трансформація.

Постановка проблеми. Попри технічну зрілість технологічних комплексів теплопостачання, зокрема індивідуальних теплових пунктів (ІТП) як об'єктів автоматизації, їхні системи керування здебільшого реалізовані у вигляді централізованої програмної логіки з жорстко визначеними правилами функціонування для кожного контуру. Така структура ефективна за сталих умов, однак позбавлена здатності до локальної оптимізації, координації між модулями чи адаптації до зовніш-

ніх факторів – таких як варіації температурного графіка або навантаження від споживачів. У контексті розвитку інтелектуальних теплових мереж і цифрової трансформації інфраструктури теплопостачання зростає потреба у підходах до автоматизації, які забезпечують гнучкість, автономність і готовність до горизонтальної інтеграції. Мультиагентна архітектура керування відкриває можливість побудови систем, де кожен функціональний модуль ІТП виступає як автономний агент із влас-

ною логікою поведінки та здатністю до взаємодії, що суттєво підвищує загальну адаптивність теплотехнічного об'єкта.

Аналіз останніх досліджень і публікацій. Упродовж останніх років спостерігається зростаючий інтерес до впровадження мультиагентних систем (МАС) у сфері автоматизації об'єктів теплоспоживання [1, с. 1], зокрема систем ОВіК та локальних систем теплопостачання. Дослідження у цій галузі демонструють потенціал МАС щодо підвищення енергоефективності, адаптивності до змінних умов навколишнього середовища та здатності до децентралізованого прийняття рішень. У [2, с. 1] запропоновано агентну архітектуру для децентралізованого керування кондиціонерами. У цьому підході агенти самостійно регулюють положення регулюючих органів, реагуючи на зміни температури в окремих зонах будівлі, що дозволяє ефективно балансувати навантаження та зберігати комфортні умови. Інше дослідження [3, с. 1] присвячено розробці подіє-орієнтованої системи керування HVAC у розумних будівлях з використанням мультиагентної взаємодії. Запропонований підхід дозволяє мінімізувати обмін даними між сенсорами та обчислювальними вузлами, активуючи обчислення лише у відповідь на релевантні події. Аспекти реалізації МАС в умовах обмежених обчислювальних ресурсів також є предметом досліджень [4, с. 1]. Тут описано децентралізовану архітектуру керування ОВіК, орієнтовану на використання у пристроях з малими обчислювальними можливостями, що особливо актуально для IoT-середовищ. Водночас піднімається питання сумісності з промисловими стандартами [5, с. 1], де описано двошарову агентну архітектуру, що відповідає специфікації IEC 61499. Така відповідність дозволяє моделювати поведінку як фізичного, так і кібернетичного шару системи керування у єдиному середовищі, забезпечуючи надійну інтеграцію з промисловою автоматикою.

Значного поширення набувають дослідження, що застосовують у мультиагентних системах методи навчання з підкріпленням. У роботі [6, с. 1] описано систему MARCO, яка реалізує глибоке навчання агентів для оптимізації роботи кондиціонерів та чілерів у великих будівлях. Такий підхід дозволив досягти кращих показників продуктивності та енергоспоживання порівняно з традиційними регуляторами. Комплексний огляд застосування МАС у системах енергоменеджменту будівель наведено в оглядовому дослідженні [1, с. 1], де підкреслено важливість

координації між агентами, що керують різними підсистемами – такими як вентиляція, гаряче водопостачання, відновлювані джерела енергії та накопичувачі – задля досягнення загальної оптимізації енергоспоживання. У роботі [7, с. 1] продемонстровано можливість врахування поведінки користувача в системах керування опаленням за допомогою агентів з глибоким навчанням, що дозволяє гнучко адаптувати режими теплопостачання до реальних умов експлуатації.

Здійснений аналіз сучасних теоретичних та прикладних досліджень свідчить про потенціал мультиагентних систем як архітектурної основи для локальних систем керування. Разом з тим, брак інформації про працюючі на наведених принципах системи керування тепловими пунктами підтверджує доцільність подальших досліджень щодо впровадження мультиагентного підходу у сфері теплопостачання.

Постановка завдання. Метою дослідження є обґрунтування доцільності впровадження мультиагентної архітектури у локальні системи автоматичного керування об'єктами теплопостачання житлового та промислового призначення, зокрема індивідуальними тепловими пунктами, а також демонстрація функціональної декомпозиції технологічного комплексу для відповідності принципам мультиагентного керування.

Виклад основного матеріалу. Методологія дослідження ґрунтується на системному підході до аналізу складних автоматизованих об'єктів, а також на застосуванні концепції моделювання на основі агентів (agent-based modeling, АВМ), яка все частіше використовується у сфері автоматизації систем теплопостачання [1, с. 1; 4, с. 1]. Такий підхід передбачає уявлення складної технічної системи як множини взаємодіючих автономних функціональних елементів (агентів), кожен з яких має власні цілі, правила поведінки, сенсорні входи, засоби керування та механізми комунікації. Застосування принципів функціональної декомпозиції [8, с. 1] дозволяє ідентифікувати ці елементи на основі технологічної структури об'єкта, що забезпечує логічну основу для побудови мультиагентної архітектури.

Для перевірки гіпотези про ефективність мультиагентного підходу проведено структурно-функціональний аналіз двох типових об'єктів – водяного ІТП багатоквартирного житлового будинку та паро-водяного ІТП промислового підприємства. На основі двох відмінних сценаріїв теплопостачання буде показано, як виглядають об'єкти автоматизації у класичному уявленні, готуючи

ґрунт для подальшого переходу до мультиагентної архітектури. У межах аналізу визначено основні контури регулювання, характерні режими роботи та технологічні залежності між окремими функціональними модулями. На основі цього виконано агентну декомпозицію: кожен контур або технологічний модуль розглядається як окремий агент, що виконує певну функцію і взаємодіє з іншими агентами через стандартизовані інтерфейси даних. Визначено базові сценарії взаємодії, що охоплюють як штатні режими роботи, так і нештатні ситуації. Передбачається, що в основі механізмів координації лежать подіє-орієнтовані протоколи, які забезпечують локальне узгодження рішень між агентами та збереження глобальної узгодженості.

Першим ілюстративним об'єктом є тепловий пункт багатоквартирного будинку (рис. 1), який має незалежне підключення до зовнішньої тепломережі через теплообмінники [9]. Уся теплова енергія для опалення та гарячого водопостачання (ГВП) передається від мережі через два окремі теплообмінники – для опалення і для ГВП. Кількість теплоносія первинного контуру для нагріву

води у вторинних контурах регулюється механізованими клапанами, які стоять перед теплообмінниками. У кожному внутрішньому контурі присутня циркуляційна насосна група з двох насосів, які працюють по схемі «основний-резервний». Циркуляційні насоси контуру опалення змінюють свою продуктивність за допомогою частотного перетворювача. Контур опалення має функцію автоматичного підживлювання водою.

Основними контурами керування є: регулювання температури подачі води в систему опалення з корекцією по температурі повітря в навколишньому середовищі, стабілізація перепаду тиску між прямою та зворотною водою системи опалення за допомогою змінної продуктивності циркуляційних насосів, керування насосом і клапаном-соленоїдом підживлювання системи опалення, регулювання температури води на подачі в систему гарячого водопостачання, керування станом циркуляційних насосів системи гарячого водопостачання. Характер навантаження – регулярний, прогнозований. Головна мета автоматизації – забезпечити комфорт мешканців із максимальною енергоефективністю.

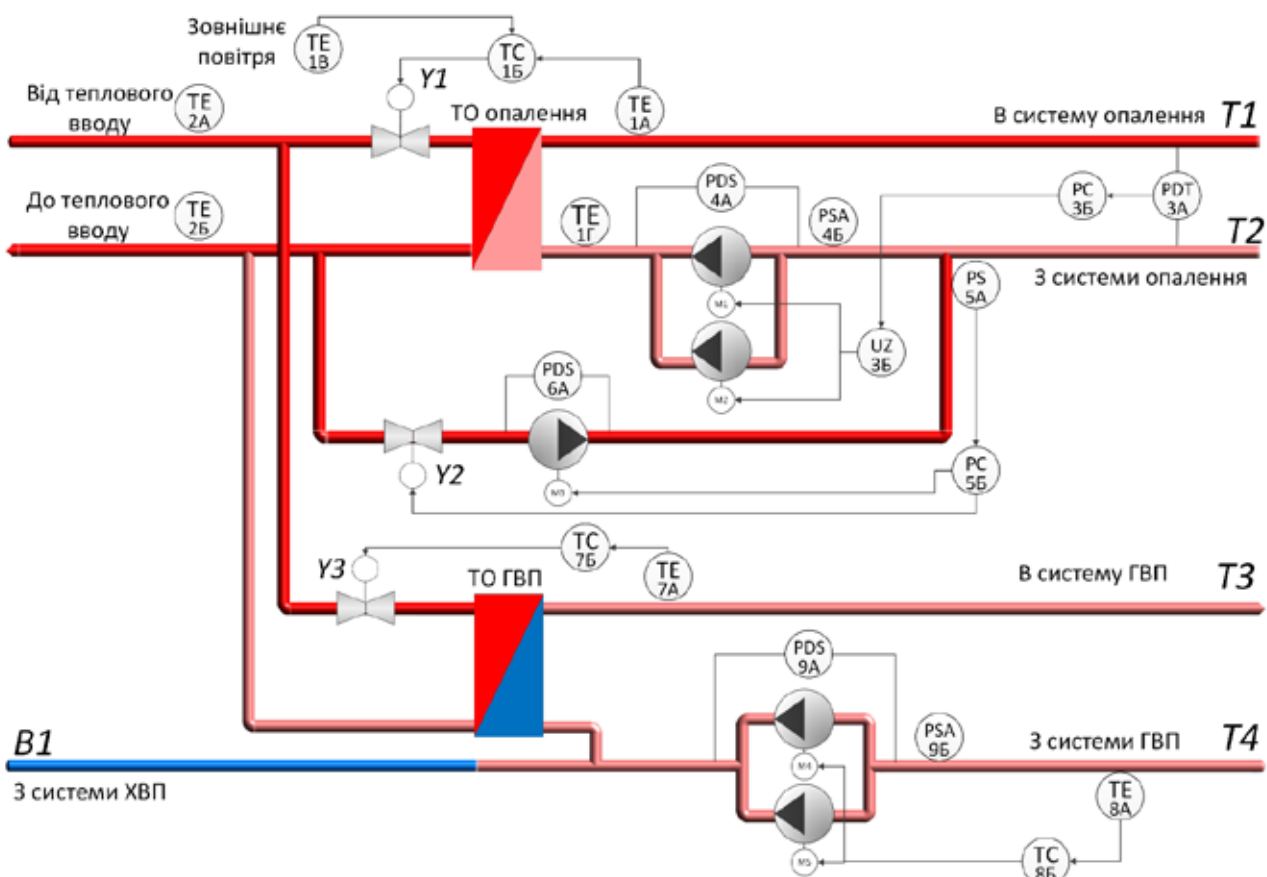


Рис. 1. Схема автоматизації індивідуального теплового пункту житлової будівлі

Пароводяний тепловий пункт [10, с. 54] є частиною технологічної інфраструктури промислового підприємства, яке отримує водяну пару з власної котельні. Тепло передається до опалення та системи ГВП через теплообмінники, а гаряча вода додатково використовується у технологічних процесах, що ускладнює керування. Структурно об'єкт подібний до попереднього, однак у цьому тепловому пункті розміщено бак-акумулятор гарячої води, з якого відбирається гаряча вода для споживачів та на потреби технології. За командою від технології вмикається додатковий циркуляційний насос великої потужності.

Серед контурів керування тут виділено (рис. 2): погодозалежне регулювання температури подачі води в систему опалення, регулювання перепаду тиску між прямою та зворотною водою системи опалення за допомогою змінної продуктивності циркуляційних насосів, керування лінією підживлення системи опалення, регулювання температури води на подачі в бак-акумулятор системи

гарячого водопостачання, регулювання рівня в бак-акумуляторі ГВП, регулювання продуктивності циркуляційних насосів системи гарячого водопостачання та керування додатковим насосом для потреб технологічного устаткування. Характер навантаження – динамічний, зі значним впливом режимів роботи виробничої технології. Ціль автоматизації – забезпечити безперебійне функціонування технології з максимальною енергоефективністю.

Типовим рішенням побудови системи автоматизації є реалізація логіки керування на програмованому логічному контролері. Система має чітко розмежовані контури регулювання. Виконавчі механізми працюють в автоматичному або ручному режимах. Задачі оптимізації (економія енергії, зміна режимів) виконуються через заздалегідь запрограмовані таймери та графіки, які можна сконфігурувати залежно від особливостей конкретного об'єкта.

Аналіз структури об'єктів дослідження дозволив представити кожен з ІТП як множину функціональних модулів з локальною метою керування, обмеженим набором вхідних і вихідних сигналів

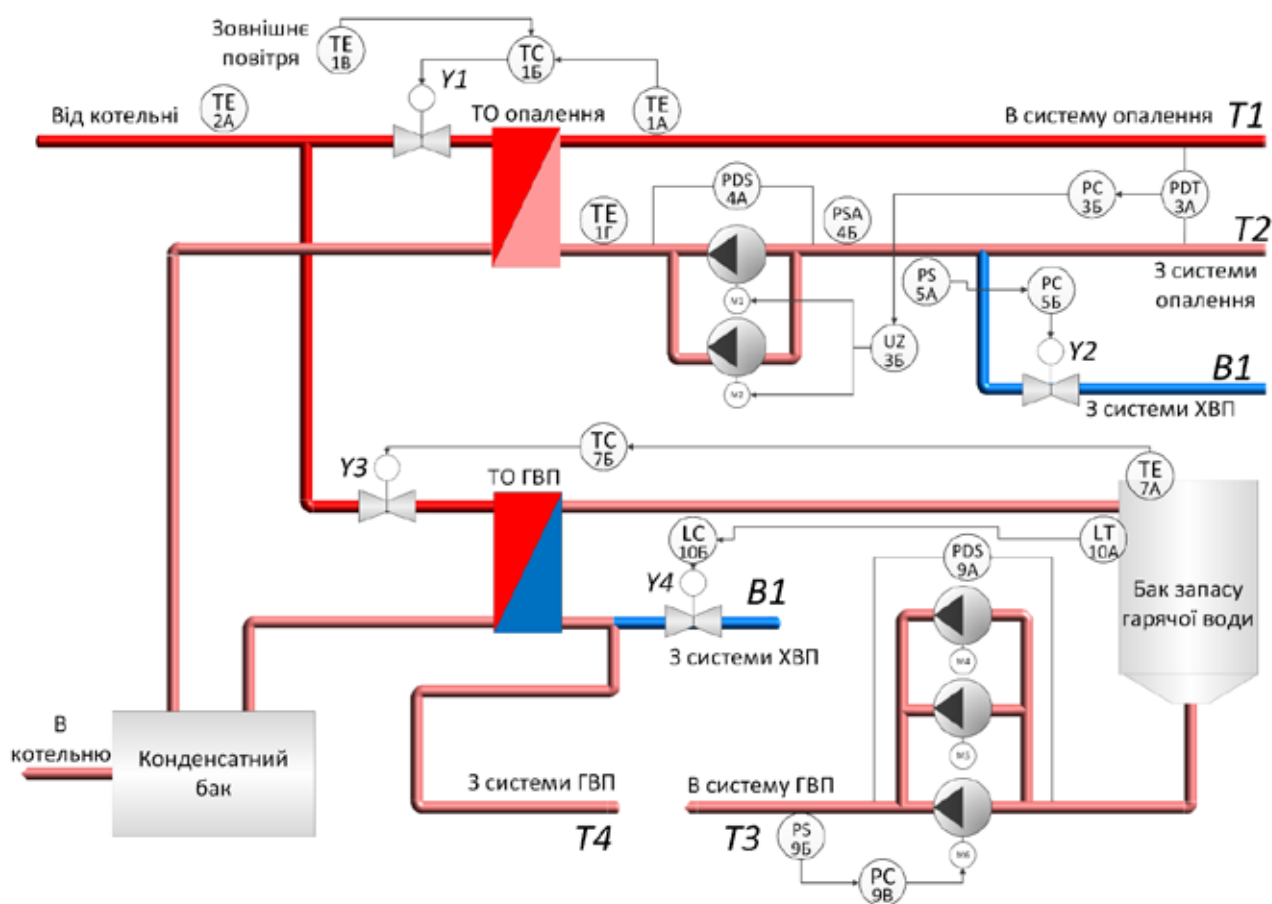


Рис. 2. Схема автоматизації промислового індивідуального теплового пункту

та визначеним інформаційним контекстом взаємодії (таблиці 1 та 2). Така архітектура є придатною до представлення у вигляді мультиагентної системи, в якій кожен контур або підсистема автоматизації функціонує як автономний програмний агент, що приймає рішення на основі локальних даних, але з урахуванням глобального режиму роботи.

На рис. 3 зображено приклад структури взаємодії наведених агентів теплового пункту житлового будинку. Варто відмітити, що направлені зв'язки символізують не стільки ієрархічне підпорядкування, як головний напрям інформаційних потоків у взаємодії агентів між собою. Така архітектура системи потенційно може мати ряд переваг, зокрема: гнучкість – відносно легко додати нові агенти, наприклад, агент розрахунку тарифу; стійкість до збоїв – при виході з ладу одного з агентів, наприклад, оптимізації енерговитрат, інша функціональність не порушується або перебудовується відповідно до обставин; адаптивність – агенти можуть змінювати свою поведінку залежно від сезону, тарифів, історії.

Пароводяний індивідуальний тепловий пункт виробничого підприємства із системою керування, побудованою на основі мультиагентної архітектури, виконує більш складні завдання, тому потребує додаткових агентів (таблиця 2).

Взаємодія агентів зображена у вигляді схеми (рис. 4), в якій окремі програмно-логічні модулі координують спільну діяльність для досягнення локальних показників ефективності та загальної цілі автоматизації.

Для демонстрації працездатності запропонованої мультиагентної архітектури проведено моделювання типових сценаріїв функціонування індивідуальних теплових пунктів у штатних і нештатних умовах. У центрі уваги – механізми взаємодії між агентами, обмін повідомленнями та координація дій на основі локальної інформації. Для кожного сценарію визначено послідовність подій, агенти-ініціатори та агенти-виконавці, а також наслідки для системи в цілому.

Сценарій 1 – «Початковий запуск системи опалення»: Після надходження команди на запуск опалювального сезону, Агент погодозалежного керування активує формування температурної уставки з огляду на поточну температуру навколишнього середовища з урахуванням рекомендацій агентів споживача та оптимізації енерговитрат. Це значення передається агенту температури опалення, який формує команду на відкриття регулюючого клапана в первинному контурі. Одночасно агент насосів ініціює запуск основного насоса з частотним регулюванням і контролює його стан у реальному часі. Агент

Таблиця 1

Агенти теплового пункту багатоквартирного будинку

Агент	Функція	Ціль
Агент температури опалення	Підтримує задану температуру теплоносія у контурі опалення	Мінімізувати відхилення від уставки
Агент погодозалежного керування	Формує уставку температури подачі на основі температури зовнішнього повітря	Забезпечити баланс між комфортом і енергоефективністю
Агент гідравлічного режиму системи опалення	Регулює продуктивність насосної групи для стабілізації перепаду тиску	Забезпечити гідравлічну стійкість, оптимальне енергоспоживання
Агент насосів опалення	Контролює наявність активного насоса, перемикає на резерв у разі аварії або за лічильником мотогодин, періодично вмикає для прокручування при тривалому простоюванні	Забезпечити безперервність роботи, рівномірність зношування
Агент температури гарячого водопостачання	Підтримує температуру гарячої води на подачі в систему ГВП	Підтримувати комфортну температуру ГВП
Агент насосів ГВП	Контролює наявність активного насоса, перемикає на резерв у разі аварії або за лічильником мотогодин, періодично вмикає для прокручування при тривалому простоюванні	Забезпечити безперервність роботи, рівномірність зношування
Агент споживача (будинку)	Інформує систему про стан температури у приміщеннях, рівень комфорту, скарги	Максимізація комфорту мешканців, подання запитів на режим
Агент оптимізації енерговитрат	Виконує аналіз історичних даних і пропонує сценарії енергозбереження	Мінімізувати загальні витрати енергії

Агенти пароводяного індивідуального теплового пункту виробничого підприємства

Агент	Функція	Ціль
Агент постачання пари (джерело)	Контролює витрату пари на об'єкті	Підтримання стабільності тиску пари в котельні, забезпечення пріоритетів
Агент температури опалення	Підтримує задану температуру теплоносія у контурі опалення	Мінімізувати відхилення від уставки
Агент погодозалежного керування	Формує уставку температури подачі на основі температури зовнішнього повітря	Зниження витрат енергії при дотриманні цілей
Агент гідравлічного режиму системи опалення	Регулює продуктивність насосної групи для стабілізації перепаду тиску	Забезпечити циркуляцію та гідравлічну стабільність з урахуванням енергоефективності
Агент насосів опалення	Контролює наявність активного насоса, перемикає на резерв у разі аварії або за лічильником мотогодин, періодично вмикає для прокручування при тривалому простоюванні	Забезпечити безперервність роботи, рівномірність зношування
Агент гарячого водопостачання (бак-акумулятор)	Контролює нагрів води в баку для подачі в систему ГВП	Забезпечення готовності до технологічного споживання
Агент рівня в баку-акумуляторі	Підтримує об'єм води в межах цільового діапазону	Недопущення "сухого ходу" та переповнення
Агент насосів ГВП	Контролює наявність активного насоса, виконує перемикавання за закладеними алгоритмами	Забезпечення своєчасного подання та стабільної подачі
Агент технологічного споживання	Реагує на запити виробництва, керує піковим насосом для потреб технології	Гарантоване постачання ГВП у технологічні цикли
Агент прогнозу навантаження	Формує прогноз потреб у воді та теплі	Підготовка системи до пікових навантажень
Агент оптимізації енерговитрат	Виконує аналіз історичних даних і пропонує сценарії енергозбереження	Мінімізувати загальні витрати енергії

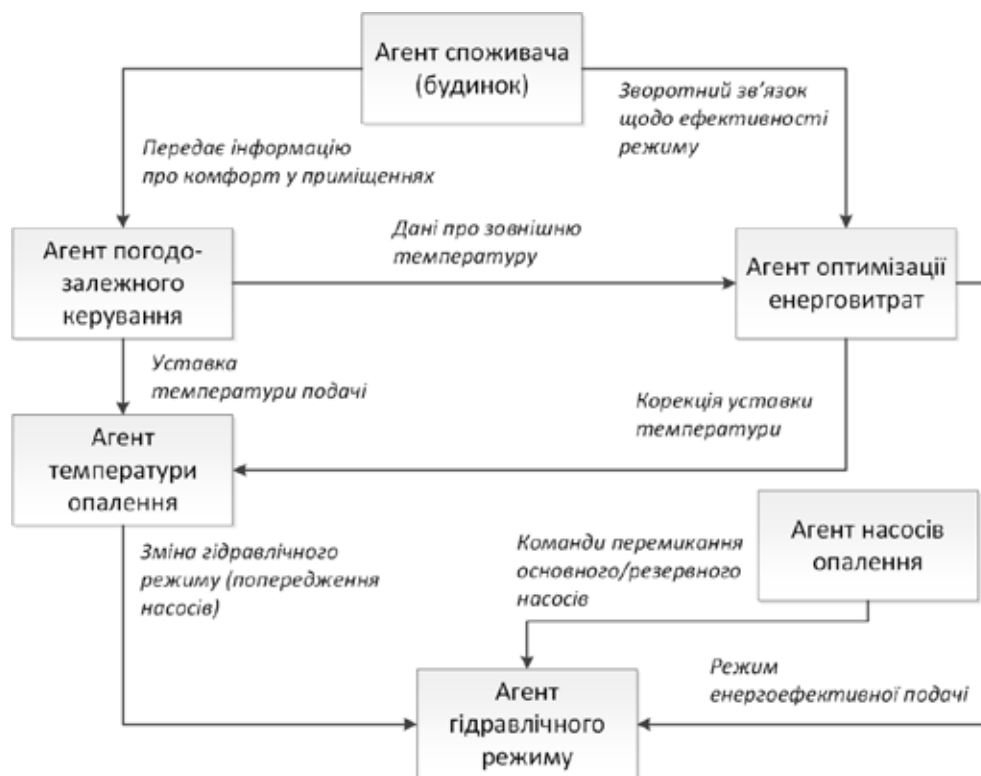


Рис. 3. Структура взаємодії агентів (фрагмент контуру опалення)



Рис. 4. Схема взаємодії агентів промислового ІТП

гідравлічного режиму забезпечує необхідний перепад тиску між подаючим та зворотнім трубопроводами системи опалення для циркуляції теплоносія, впливаючи на продуктивність насоса за допомогою частотного перетворювача.

Сценарій 2 – «Перехід у знижене теплове навантаження вночі», актуальне для споживачів (цехів), в яких люди перебувають не цілодобово. Агент споживача будівлі передає сигнал про зменшення попиту на тепло. Агент формування уставки адаптує температурний графік, після чого нова уставка передається регулятору подачі. Зменшення навантаження фіксується також агентом джерела тепла, який надсилає підтвердження про зниження відбору тепла.

Сценарій 3 – «Аварія основного насоса». Агент насосів отримує аварійний сигнал про зупинку основного насоса та ініціює переключення на резервний агрегат. Тиск в системі стабілізується агентом гідравлічного режиму, а регулятор температури враховує тимчасове зниження теплопередачі. Аварійна подія реєструється та може бути передана на верхній рівень або використана для місцевої діагностики.

Сценарій 4 – «Зниження тиску пари на виході з котельні». Агент постачання пари повідомляє

про зниження тиску пари надмірне сукупне використання пари всіма споживачами. Агент постачання пари ініціює коригування уставок у підпорядкованих агентах (опалення, ГВП), пропонуючи режим енергозбереження. Агент ГВП погоджується на тимчасове незначне зниження температури, узгоджене з агентом технологічного споживання, а агент опалення – на зменшення теплової потужності у нежитлових приміщеннях. Спільна координація дозволяє уникнути аварійного відключення.

Сценарій 5 – «Використання гарячої води з акумулювальної ємності на виробництві». У промисловому ІТП агент технологічного споживання реагує на запити виробництва на подачу гарячої води, беручи до уваги дані агента прогнозу навантаження та передаючи інформацію до агента гарячого водопостачання (бак-акумулятор) та агентів рівня в баку-акумуляторі. Ці агенти забезпечують необхідну кількість гарячої води в потрібний момент часу. Агент насосів активує пікову насосну групу, попередньо перевіривши рівень у ємності через агента рівня. Якщо рівень допустимий – вода подається, а агент ГВП отримує сигнал про зниження температури і ініціює додатковий тепловий вплив через агента прогнозу

навантаження. Агент постачання пари з котельні підтверджує готовність до подачі тепла і встановлює нові обмеження для інших споживачів.

Інші сценарії також можна привести до процедур взаємодій агентів.

Реалізація мультиагентної архітектури локальних систем автоматичного керування ІТП на практиці можлива за рахунок використання сучасних вільнопрограмованих логічних контролерів (ПЛК), протоколів обміну даними та вбудованих можливостей логічної модульності. Особливо доцільною така реалізація є для промислових ПЛК з середовищами розробки, які відповідають принципам стандарту промислового програмування IEC 61131-3, де можна формувати агентоподібні структури завдяки елементам об'єктно-орієнтованого підходу до програмування.

Для реалізації взаємодії між агентами доцільно використовувати легковагі протоколи обміну повідомленнями типу MQTT, які працюють за моделлю взаємодії publish-subscribe. Кожен програмний агент функціонує як окрема логічна одиниця, яка публікує повідомлення про зміни свого стану або результати обробки, а також підписується на канали інших агентів. Це дозволяє формувати подієво-орієнтовану інформаційну шину без централізованого координатора. У випадку реалізації на фізично розподіленій платформі (наприклад, поетапна модернізація окремих частин ІТП), така архітектура дозволяє досягти гнучкості у масштабуванні без потреби реконфігурації всієї системи.

З точки зору практичної реалізації, окремі агенти можуть бути представлені функціональними блоками або модулями з чітко визначеними інтерфейсами обміну даними. Наприклад, кожен агент реалізується як FB-модуль із вхідними параметрами (дані вимірювань, зовнішні команди), внутрішнім станом (логіка прийняття рішень) та вихідними сигналами (керуючі впливи, повідомлення для інших агентів). Обмін між модулями здійснюється через брокер MQTT, реалізований або як окремий компонент на диспетчерському рівні або безпосередньо у контролері за допомогою вбудованих бібліотек.

SCADA-системи, які зберігають свою функцію візуалізації, логування та супервізорного втру-

чання, тут вже не виступають центром прийняття рішень. Вони взаємодіють з агентами через ті ж протоколи обміну (MQTT, OPC UA), отримуючи агреговану інформацію про стан системи та можливість втручання у разі потреби.

Висновки. У роботі було запропоновано концепцію мультиагентної архітектури для локальних систем автоматичного керування об'єктами тепlopостачання на прикладі житлового та промислового індивідуальних теплових пунктів. Продемонстровано, що традиційно централізована логіка ПЛК може бути трансформована у децентралізовану модель із функціонально автономними агентами, які взаємодіють через подієво-орієнтовану інформаційну шину.

Аналіз двох ІТП різного призначення дозволив здійснити функціональну декомпозицію на підсистеми, кожна з яких була представлена як агент з власною ціллю, інтерфейсами взаємодії та механізмом координації з іншими агентами. Формалізація цілей агентів забезпечила основу для побудови логічної моделі їхньої поведінки у штатних та нештатних ситуаціях. Було проілюстровано типові сценарії роботи системи, зокрема адаптацію до змін температури зовнішнього середовища, зміни теплового навантаження, аварійні ситуації та обмеження з боку джерел енергії.

Окрему увагу приділено технічній реалізованості запропонованого підходу. Показано, що мультиагентна архітектура потенційно може бути реалізована на базі сучасних вільнопрограмованих ПЛК із використанням протоколів обміну даними MQTT або OPC UA. Це дозволяє впроваджувати інтелектуальні функції координації та адаптації без необхідності кардинального переоснащення об'єктів.

Перспективи подальших досліджень полягають у розробці прототипів агентних систем для реальних ІТП, верифікації логіки взаємодії на симуляційних моделях, а також у застосуванні методів машинного навчання для вдосконалення стратегій прийняття рішень на рівні окремих агентів. Крім того, подальше узагальнення моделі на більші системи – мережі ІТП або територіально розподілені об'єкти – відкриває можливості інтеграції в концепцію Smart Thermal Grid.

Список літератури:

1. Michailidis P., Michailidis I., Kosmatopoulos E. Review and Evaluation of Multi-Agent Control Applications for Energy Management in Buildings. *Energies*. Вип. 17, № 19. С. 4835. DOI:10.3390/en17194835.
2. Dai M., Li H., Wang S. Multi-agent based distributed cooperative control of air-conditioning systems for building fast demand response. *Journal of Building Engineering*. Вип. 77, 10.2023. С. 107463. DOI:10.1016/j.jobbe.2023.107463

3. Li W., Li H., Wang S. An event-driven multi-agent based distributed optimal control strategy for HVAC systems in IoT-enabled smart buildings. *Automation in Construction*. 2021. Vol. 132. P. 103919. URL: <https://doi.org/10.1016/j.autcon.2021.103919>
4. Su B., Wang S. An agent-based distributed real-time optimal control strategy for building HVAC systems for applications in the context of future IoT-based smart sensor networks. *Applied Energy*. 2020. Вип. 274. С. 115322. DOI:10.1016/j.apenergy.2020.115322
5. Lyu G., Brennan R. W. Multi-agent modelling of cyber-physical systems for IEC 61499-based distributed intelligent automation. *International Journal of Computer Integrated Manufacturing*. 2023. Вип. 38, № 5. С. 596–622. DOI:10.1080/0951192X.2023.2294442
6. Nagarathinam S., Menon V., Vasan A. та ін. MARCO – Multi-Agent Reinforcement learning based Control of building HVAC systems. *Proceedings of the Eleventh ACM International Conference on Future Energy Systems-Energy '20: The Eleventh ACM International Conference on Future Energy Systems*. (Virtual Event Australia, 12.06.2020). Virtual Event Australia : ACM, 2020. DOI:10.1145/3396851.3397694. С. 57–67.
7. McKee E., Du Y., Li F. et al. Deep Reinforcement Learning for Residential HVAC Control with Consideration of Human Occupancy. *2020 IEEE Power & Energy Society General Meeting (PESGM)2020 IEEE Power & Energy Society General Meeting (PESGM)*. (Montreal, QC, Canada, 02.08.2020). Montreal, QC, Canada : IEEE, 2020. URL: <https://doi.org/10.1109/PESGM41954.2020.9281893>. P. 1–5.
8. Van Breemen A. J. N., de Vries T. J. A. Design and implementation of a room thermostat using an agent-based approach. *Control Engineering Practice*. 2001. Vol. 9, no. 3. P. 233–248. URL: [https://doi.org/10.1016/S0967-0661\(00\)00111-8](https://doi.org/10.1016/S0967-0661(00)00111-8)
9. Типові рішення по термомодернізації житлових будинків: альбом технічних рішень / за заг. ред. Головатюк-Унгуряну Ю.В.; уклад. Фаренюк Є.Г. Київ : Фонд енергоефективності, 2020. Том 3. 43 с.
10. Степанець О. В., Ущатовський А. П. Synthesis of automatic control system of industrial steam-water heating unit. *Eastern-European Journal of Enterprise Technologies*. 2015. Vol. 3, no. 2(75). P. 54. URL: <https://doi.org/10.15587/1729-4061.2015.42141>

Zdolnyk M.O., Stepanets O.V. MULTI-AGENT APPROACH TO THE ARCHITECTURE OF LOCAL AUTOMATIC CONTROL SYSTEMS FOR INDUSTRIAL AND RESIDENTIAL HEAT SUPPLY FACILITIES

This article explores the potential for implementing a multi-agent architecture in local automatic control systems for heat supply facilities, particularly individual heating substations (IHS) for residential and industrial applications. It is substantiated that the traditionally used centralized structure of local IHS control systems has limited capacity for adaptation, scalability, and coordination among functional modules, especially under changing external conditions. In view of the needs for digital transformation of thermal infrastructure, a transition to a multi-agent model is proposed, in which each functional block of an IHS operates as an autonomous software agent with a clearly defined function, control objective, and interaction interfaces. The research methodology is based on system analysis of control loops and the principles of functional decomposition. To justify the architectural approach, the structures of two typical facilities are analyzed: a water-based IHS for a multi-apartment residential building and a steam-water IHS for an industrial enterprise. For each facility, agent-based decomposition is performed, with functions, objectives, and interaction scenarios defined for each agent. Information exchange graphs between agents are constructed, and examples of system behavior in various modes—ranging from normal startup to emergency events or changes in demand—are developed. A conceptual technical solution for implementing the multi-agent model using modern industrial automation tools is proposed. The results demonstrate that adopting a multi-agent architecture ensures adaptability and flexibility of control systems while preserving regulatory loops as fundamental control mechanisms. The advantages of such an architecture, particularly its scalability and potential for integration with distributed information platforms, are identified. Future research prospects include the development of agent system prototypes for real IHS, the application of machine learning to enhance decision-making efficiency, and the expansion of the model to the level of thermal networks within the Smart Thermal Grid concept.

Key words: process automation, multi-agent control systems, heating substation, decentralized control, digital transformation.

Зіноватна С.Л.

Національний університет «Одеська політехніка»

Білий Д.О.

Національний університет «Одеська політехніка»

Онщенко Т.В.

Національний університет «Одеська політехніка»

ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ВЗАЇМОДІЇ ПЛАНУВАННЯ МЕНЮ ТА ВИКОРИСТАННЯ СКЛАДСЬКИХ ЗАПАСІВ

У статті розглянуто проблему зменшення розміру відходів під час виготовлення блюд для забезпечення харчування в будь-якій організації (ресторан, кафе, школа тощо). Ефективне використання харчових продуктів є найважливішою задачею сьогодення, оскільки виробництво продуктів вимагає багато ресурсів від землі, зношення невикористаних продуктів погіршує екологічний стан навколишнього середовища, є частиною соціальної відповідальності. Завдання зменшення відходів в закладах, які забезпечують громадське харчування, є актуальною задачею. Запропонована інформаційна модель, яка формалізує представлення даних для планування меню з виготовлення блюд в організації громадського харчування та контролю запасів продуктів. Розроблено схему послідовності етапів взаємодії блоку формування меню та блоку обліку запасів на складі із зворотнім зв'язком. Розроблено інформаційну технологію коректованого формування меню з урахуванням наявних продуктів та можливої заміни продуктів в технологічних картах блюд з метою зниження кількості відходів. Визначено користувачів програмної системи, в основу якої покладено розроблену інформаційну технологію, та визначені доступні для кожного користувача функції програми. Функціональні вимоги до програми описані у вигляді варіантів використання. Запропонована архітектура програмного забезпечення, за допомогою якого виконується контроль та керування взаємодією блоку планування меню та керування запасами. Розроблена технологія дозволяє автоматизувати підбір продуктів під час формування меню та підвищити рівень використання наявних на складі продуктів. Використання запропонованих алгоритмів може підвищити прибуток закладу за рахунок максимального використання складських запасів, зниження витрат за рахунок зменшення кількості закупок продуктів. Також в цілому зменшується кількість відходів за рахунок ефективного використання наявних позицій.

Ключові слова: програмна система, керування запасами, мінімізація відходів, моделювання меню, варіанти використання.

Постановка проблеми. Харчування людини є найважливішою проблемою з точки зору різних аспектів. З одного боку, правильне харчування має забезпечити організм необхідними речовинами для його функціонування, при цьому запобігати потраплянню в організм шкідливих продуктів. Харчування має бути збалансованим для зниження ризику виникнення захворювань, які можуть бути викликані як надлишковим вживанням деяких продуктів (ожиріння), так і їх нестачею (анемія). З другого боку, під час виготовлення їжі має бути підібраний набір продуктів, який забезпечить потрібну якість отриманих блюд та буде відповідати припустимим витратам. На сьогодні проблема еконо-

мії продуктів є однією з найважливіших. Виробництво продуктів вимагає великої кількості ресурсів. Планування харчування, правильне збереження продуктів, використання залишків допомагають зекономити бюджет. Невикористані продукти гниють та збільшують обсяг парникових газів, що має негативний вплив на екологію. Для багатьох людей у світі проблемою є нестача їжі, тому раціональне використання продуктів має соціальну відповідальність. Таким чином, актуальним є використання різних засобів для зниження відходів продуктів під час приготування їжі.

Аналіз останніх досліджень і публікацій. Уживання здорової їжі важливо для кожної

людини. В [1] наведені дані, що відповідно статистики Всесвітньої організації охорони здоров'я, на лікування захворювань, викликаних продуктами харчування, щорічно витрачається близько 50-80 мільярдів доларів США. Згідно [2] ризики для безпеки харчових продуктів виникають на всіх етапах виробництва, транспортування, зберігання і їхньої переробки, при цьому одним з найбільш уразливих етапів є зберігання до й після обробки.

Невірні умови зберігання продуктів приводять до збільшення кількості продуктів, які не можуть бути використані у їжу і мають бути утилізовані.

Втрати продовольства й велика кількість харчових відходів є джерелом економічних, екологічних і соціальних проблем, і в цілому символом неефективності, несправедливості й нестійкості продовольчих систем. Втрати й харчові відходи можуть виникати на різних етапах ланцюжка поставок продовольства: сільськогосподарське виробництво (до збору врожаю/виробництво, збір урожаю); післязбиральна обробка й зберігання; переробка; розподіл; роздрібна торгівля; споживання [3].

Зокрема, в Україні втрати продовольства відбуваються на всіх етапах ланцюжка споживання: від сільськогосподарського виробництва до кінцевого споживання домогосподарствами. З одного боку, продукти викидаються, навіть якщо вони ще придатні до вживання, з іншого боку, більша частина продовольства губиться на етапах виробництва, зберігання й переробки, через недосконалі технології виробництва й нерозвинену складську інфраструктуру для зберігання продуктів [4].

Контроль запасів на складах організацій громадського харчування й магазинів гарантує доступність продукції в потрібний момент, мінімізує відходи й максимізує прибуток [5], можливі різні методи оптимізації запасів продуктів харчування: принцип «першим прийшов першим пішов» (FIFO) – у першу чергу використовуються самі старі запаси; прогнозування попиту на продукцію з використанням історичних даних і аналізу ринку; відстеження строків придатності продуктів; систематична перевірка рівня запасів продуктів, їхніх витрат і закупок.

Для зменшення витрат продуктів має бути забезпечено їх якісне зберігання. Управління складом має багато аспектів: наявність необхідного обладнання, зонування зберігання, доступність, постійний контроль, дотримання санітарних норм з боку співробітників тощо [2].

При цьому мають бути враховані типи продуктів, наприклад, в [3] розглянуті засоби зменшення

витрат щодо молочної продукції, а в [6] щодо сухих харчових матеріалів, які використовуються для виготовлення їжі, в [7] – для швидкопсувних продуктів, в [8] – для м'ясних продуктів.

Для місць громадського харчування витрати продуктів відповідають визначеним меню закладу. Під час формування меню має бути врахована велика кількість факторів. Це збільшення прибутку закладу (попит, ціноутворення, дизайн), збалансованість блюда (поєднання продуктів в самому блюді, з іншими пропонованими блюдами, відповідність вимогам здорового харчування), використання продуктів (вартість, доступність, транспортування, умови збереження, сезонність, локальність) [9].

Під час закупок та витрат продуктів має бути також враховано розмір пакування [10], оскільки це впливає на залишки продуктів у розрізі вартості та кількості (велика упаковка може зекономити фінанси, але привести до їх втрати через неможливість своєчасного використання залишків продукту або їх псування). На кількість відходів також впливає відношення співробітників до обробки продуктів, наприклад, дослідження [11] показали, що працівники харчової промисловості ігнорують важливість меню як інструмента комунікації при плануванні, роботі й мінімізації відходів. Над проблемою скорочення відходів працюють багато країн, для цього залучають новітні технології, зокрема Інтернет речей [12]: «технологічні стратегії містять у собі інтелектуальну кухонну техніку, керування запасами в режимі реального часу й рішення по переробці відходів, такі як біореактори й компостування».

Створення меню з урахуванням можливих витрат є одним з етапів зниження відходів. Планування меню є складною задачею, потребує врахування багатьох критеріїв. Для допомоги у рішенні такого завдання можуть бути використані інформаційні технології та відповідне програмне забезпечення. В [13] описаний досвід використання розробленої системи KNOBIE для підтримки стійких методів планування меню й рецептів. Задіяні в експерименті шеф-кухарі оцінили користь від використання KNOBI у закладах з великими обсягами їжі, де потрібно систематичне обслуговування, але досвід показав, що система може бути неідеальною для невеликих ресторанів. Пропонується подальша адаптація системи, яка зможе збирати дані про наявність інгредієнтів і змінювати або просувати певні позиції, щоб запобігти псуванню й відходам.

Постановка завдання. Метою статті є зниження витрат організації громадського харчування за рахунок автоматизації взаємодії планування меню як початкового етапу виробництва блюд та використання продуктів на продовольчому складі.

Виклад основного матеріалу. Для формалізації представлення даних розроблено інформаційну модель, яка описує основні компоненти предметної області у вигляді множин та кортежів.

Меню представляється множиною кортежів, кожен кортеж містить набір блюд та відповідність до часу: $M = \{ \langle dt, D \rangle \}$, де dt – дата, на яку складається меню, D – множина блюд. Елементами множини D є наступні кортежі: $D = \{ \langle R, q, tm \rangle \}$, де R – рецепт блюда, q – кількість блюд, які потрібно приготувати, tm – час, на який блюда мають бути виготовлені. Опис рецепту складається з наступних компонентів: $R = \langle P, tp, wd, tc \rangle$, де P – множина продуктів, які необхідні для приготування блюда, tp – тип блюда, wd – вага блюда, tc – технологічна карта. Елементами множини P є наступні кортежі: $P = \{ \langle p, w \rangle \}$, де p – опис продукту, w – частка продукту в блюді. Опис продукту має наступний вигляд: $p = \langle pr, qin, pw, pc \rangle$, де pr – назва продукту, qin – вхідна кількість продукту, pw – очікуваний відсоток відходів, pc – опис обробки продукту.

Також для продуктів блюда створюється опис можливих заміन: $PRC = \{ \langle pr, prc, wc \rangle \}$, де prc – продукт, на який може бути замінений pr ; wc – потрібна кількість для заміненого продукту.

Для розрахунку кількості продуктів на складі використовуються множини PH (закупка), IS (видача), WO (списання). Елементи всіх множин мають однаковий склад компонентів: $e = \langle dp, pr, qp, tr, ai \rangle$, де dp – дата дії, qp – кількість продукту, tr – термін дії продукту, ai – додаткова інформація.

На основі наданої моделі сформовано множини сутностей предметної області, зв'язки між ними та побудовано структуру бази даних.

Інформаційна технологія взаємодії планування меню та використання складських запасів включає множини наступних етапів (рис. 1).

1. Кухар формує меню M на основі множини рецептів R , при необхідності створює нові рецепти.

2. На основі множини M формується список необхідних продуктів $LP = \{ \langle pr, qpr \rangle \}$, де qpr – розрахована загальна кількість продуктів необхідна для виготовлення всіх блюд меню.

3. Визначається наявна кількість продуктів на складі $LPR = \{ \langle pr, qprr \rangle \}$. Якщо для всіх продуктів $qprr \geq qpr$, то перехід до п. 6.

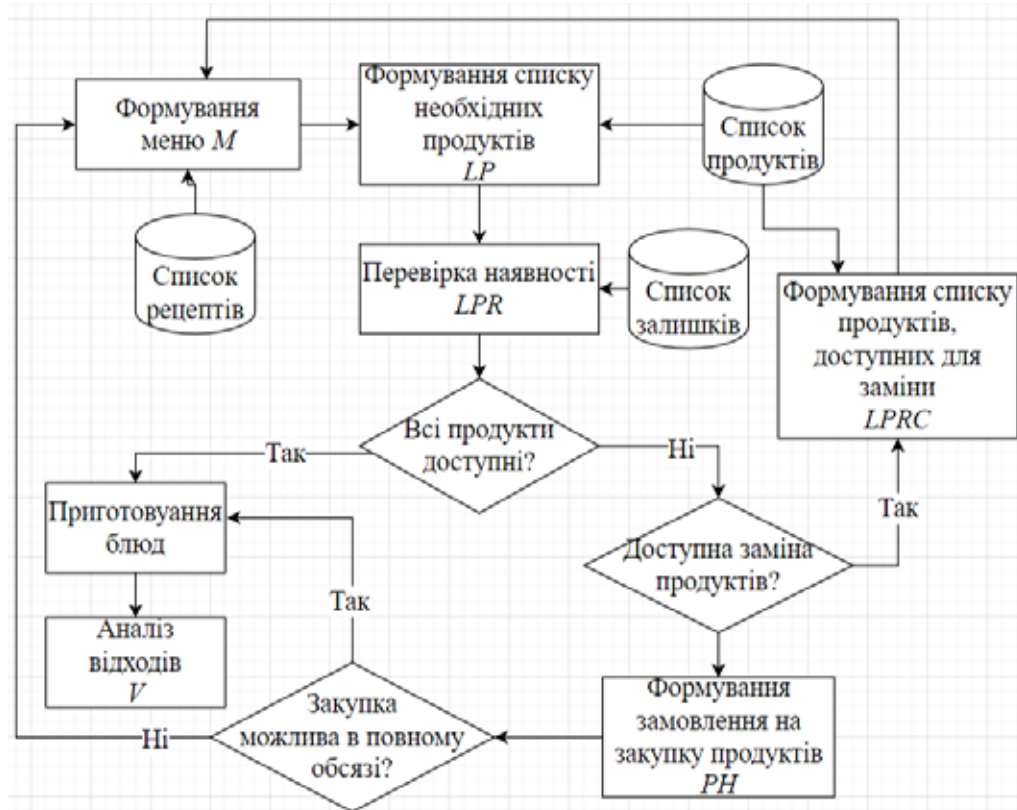


Рис. 1. Послідовність етапів взаємодії планування меню та використання складських запасів

4. Формування списку продуктів з урахуванням можливих замін $LPRC = \{<prc, qprc>\}$ Якщо для всіх продуктів до $qprc \geq qpr$, то перехід до п. 6.

5. Формування замовлення на закупівлю продукту PH . Якщо закупка можлива в повному обсязі та в потрібний термін, перехід до п. 6. Інакше перехід в п. 1.

6. Приготування блюд відповідно M .

7. Аналіз відходів для подальшого корегування рецептів та контролю залишків на складі.

На основі запропонованої інформаційної технології створено програмну систему, функціональні вимоги до якої представлені на діаграмі варіантів використання (рис. 2).

Діаграма містить акторів системи (Кухар та Комірник), які за допомогою програмних засобів створюють вхідні дані та отримують необхідну інформацію, розраховану на основі наявних даних. Надана інформація використовується для прийняття рішень з боку користувачів.

Архітектура програмної системи відповідає шаблону MVC, який є популярним для використання у веброзробках та мобільних застосунках [14].

Відповідно шаблону, програма складається з трьох компонентів (рис. 3). Компонент View відповідає за представлення інформації користувачу, в загальному вигляді реалізує автоматизовані робочі місця кухара та комірника, надає можливість переглядати потрібну інформацію та запитувати зміну даних, а також вводити дані для їх створення та корегування. Компонент Controller відповідає за взаємодію з даними, формує запити на отримання та модифікацію даних, повертає отриманні дані в View. Компонент Model відповідає за роботу з базою даних, містить набір класів, які забезпечують роботу з сутностями предметної області та виконання запитів.

Використання запропонованих у дослідженні алгоритмів дозволяє підвищити економічну ефек-

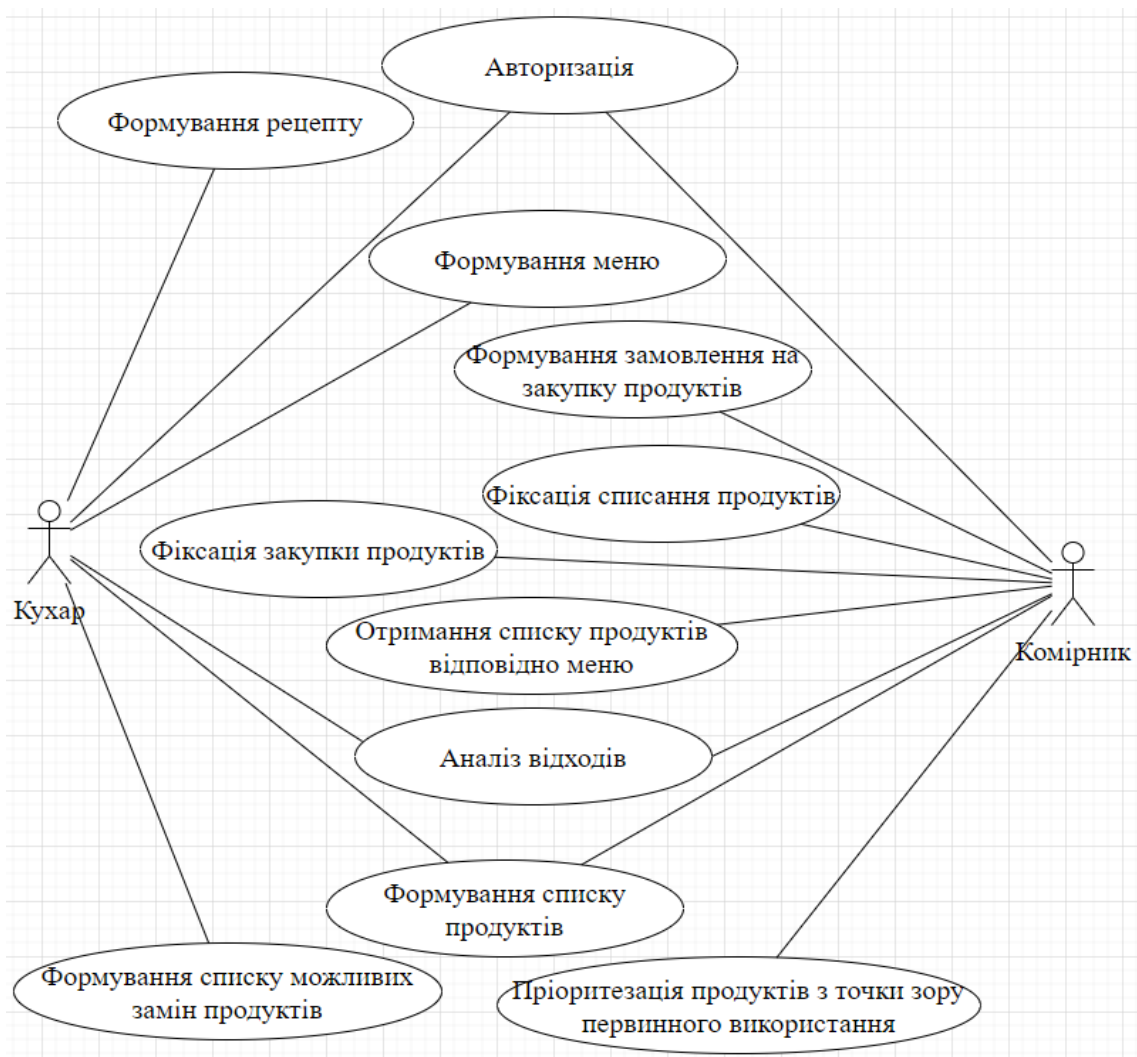


Рис. 2. Діаграма варіантів використання

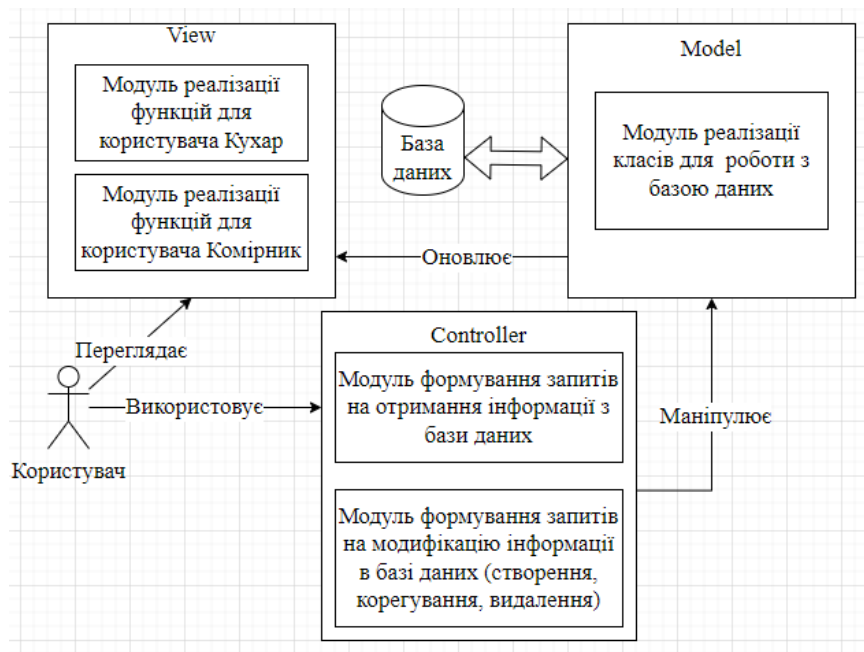


Рис. 3. Архітектура програмної системи

тивність роботи закладів громадського харчування. Це досягається завдяки максимальному залученню до процесу виробництва наявних на складі продуктів, що знижує рівень невикористаних запасів. Рациональне управління залишками на складі сприяє скороченню обсягів закупок нових продуктів, що, у свою чергу, дозволяє суттєво зменшити операційні витрати. Таким чином, оптимізація використання ресурсів не лише знижує рівень харчових відходів, а й безпосередньо впливає на збільшення прибутковості закладу.

Висновки. Дане дослідження пропонує інформаційну технологію, яка складається з послідовності етапів із зворотнім зв'язком та вирішує

проблему максимального використання наявних запасів продуктів на складі закладу, який має забезпечувати громадське харчування. Зберігання та отримання інформації передбачається з бази даних, структура якої побудована на основі моделі взаємодії планування меню та використання складських запасів продуктів. Виконаний аналіз вимог до програмної системи, в основу якої покладена розроблена інформаційна технологія. Спроектовано архітектуру програмної системи, використання якої дасть можливість зменшити кількість відходів продуктів в закладі та за рахунок максимального використання запасів знизити фінансові витрати.

Список літератури:

1. Khan R., Tyagi N., Chauhan N. Safety of food and food warehouse using VIBHISHAN. *Journal of Food Quality*. 2021. DOI: 10.1155/2021/1328332.
2. Guidelines for food storage warehouses. 2016. URL: <https://www.chinaafricasps.org/zh/resources/file/2024/10/17/602216062603333.pdf> (дата звернення 03.06.2025).
3. Nikolicic S., Kilibarda M., Maslarić M., Mirčetić D., Bojic S. Reducing food waste in the retail supply chains by improving efficiency of logistics operations. *Sustainability*. 2021. Vol. 13. DOI: 10.3390/su13126511.
4. Filippov V.Yu., Sagitova R., Voloshchuk L.O., Jack L. Analysis of the current situation and measures to prevent food loss and food waste in Ukraine. *Economic journal Odessa polytechnic university*. 2023. № 3(25). P. 71–87. DOI: 10.15276/EJ.03.2023.9.
5. Selwitz A. Optimizing Food Inventory: 4 Smart Techniques to Minimize Waste & Maximize Fulfillment. URL: <https://surl.li/sfdbfl> (дата звернення 03.06.2025).
6. Sari R., Firiza S., Syahputri K., Rizkya I., Siregar I. Minimizing inventory cost of dried food materials availability: An analyzing in teaching hospital. *IOP Conference Series: Materials Science and Engineering*. 2021. Vol. 1122. DOI: 012116. 10.1088/1757-899X/1122/1/012116.
7. Adedeji A. E. Fresh Food Products Inventory Control Management: the challenges in avoiding perishability at the retailers' shelf. *Master Thesis. Linnaeus University Sweden*. 2020. URL: <https://www.diva-portal.org/smash/get/diva2:1440421/FULLTEXT01.pdf> (дата звернення 03.06.2025).

8. Marques P.A., Mendes de Carvalho A. Reducing shrinkage and stockouts in a food retail store: a Portuguese case study. *Jurnal Teknik Industri*. 2021. Vol. 23, No. 2. P. 93–100. DOI: 10.9744/jti.23.2.93-100.
9. Nebioğlu O. Review of menu management process model with a case study. *Advances in Hospitality and Tourism Research*. 2020. Vol. 8. P. 203–234. DOI: 10.30519/ahtr.668220.
10. van Rooijen M.A., Gerdessen J.C., Claassen G.D.H., de Leeuw S.L.J.M. Optimizing household food waste: The impact of meal planning, package sizes, and performance indicators. *Resources, Conservation and Recycling*. 2024. Vol. 205. P. 107559. DOI: 10.1016/j.resconrec.2024.107559.
11. Xaba P., Marumo-Ngwenya K., Kearney J.E.. The application of menu planning principles in food waste prevention by food handlers in supermarket kitchens in Soweto, South Africa. *African Journal of Hospitality, Tourism and Leisure*. 2024. Vol. 13(2). P. 325–333. DOI: 10.46222/ajhtl.19770720.513.
12. Singh V., Singh A., Tyagi P.K. Utilizing technology for food waste management in the hospitality industry hotels and restaurants technology for food waste management in hotels, restaurants. In book: *Sustainable Disposable Methods of Food Waste in Hospitality Operations* Publisher: IGI Global Publication USA. 2024. DOI: 10.4018/979-8-3693-2181-2.ch019.
13. Coskun A., Genç H.U., Coşkun A. How sustainable is your menu? Designing and assessing an interactive artefact to support chefs' sustainable recipe-planning practices. *Proceedings of the ACM SIGCAS/SIGCHI Conference on Computing and Sustainable Societies*. 2023. P. 90–98. DOI: 10.1145/3588001.3609366.
14. Necula S. Exploring the Model-View-Controller (MVC) architecture: a broad analysis of market and technological applications. *Preprints*. 2024. DOI: 10.20944/preprints202404.1860.v1.

Zinovatna S.L., Bilyi D.O., Onishchenko T.V. INFORMATION TECHNOLOGY INTERACTION OF MENU PLANNING AND USING OF WAREHOUSE INVENTORIES

The article addresses the issue of reducing food waste during the preparation of meals in catering establishments such as restaurants, cafes, schools, and other organizations. Efficient use of food resources is a critical challenge today, as food production consumes significant natural resources, and the disposal of unused products negatively impacts the environment. Moreover, minimizing waste is an important aspect of social responsibility. This paper proposes an information model that formalizes data representation for menu planning and inventory control in public catering organizations. A diagram has been developed to illustrate the interaction between the menu formation module and the inventory management module, including feedback mechanisms. An information technology has been designed for dynamic menu planning that accounts for the availability of products in stock and supports the substitution of ingredients in recipes, with the goal of reducing waste. The users of the proposed software system are identified, along with the functions available to each type of user. Functional requirements for the software are presented as use cases. A software architecture is proposed to manage and coordinate the interaction between menu planning and inventory management modules. The developed technology enables automation of product selection during menu creation, thereby increasing the utilization of available stock. By applying the proposed algorithms, institutions can boost their profits by maximizing the use of warehouse inventory and reducing costs associated with purchasing new products. Additionally, overall waste can be significantly reduced through more efficient use of available ingredients.

Key words: software system, inventory management, waste minimization, menu modeling, use cases.

Дата надходження статті: 26.07.2025

Дата прийняття статті: 07.08.2025

Опубліковано: 27.10.2025

Зінченко Д.Д.

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»

Маріяш Ю.І.

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»

КООРДИНОВАНЕ КЕРУВАННЯ ЧАСТОТОЮ CPU ТА GPU ДЛЯ ПІДВИЩЕННЯ ПРОДУКТИВНОСТІ ТА КОМФОРТУ КОРИСТУВАЧА МОБІЛЬНОГО ПРИСТРОЮ

У даній роботі розглядається одна з ключових проблем сучасної мобільної електроніки – перегрів корпусу пристроїв, що є прямим наслідком зростаючої обчислювальної потужності систем-на-кристалі та обмежених можливостей пасивного тепловідведення. Ця проблема виходить за рамки технічних недоліків, оскільки підвищення температури поверхні вище 45 °C викликає значний фізичний дискомфорт у користувача та суттєво погіршує загальний досвід експлуатації. Метою дослідження є розробка та експериментальна оцінка нового координованого підходу до динамічного регулювання температури, який би ефективно вирішував фундаментальний конфлікт між прагненням до максимальної продуктивності та необхідністю забезпечення комфорту. На відміну від поширених промислових рішень, що базуються на примітивних позиційних регуляторах або ігнорують внесок графічного процесора, запропоновано комплексну систему керування на основі цифрового Пропорційно-Інтегрально-Диференціального регулятора. Наукова новизна підходу полягає у впровадженні механізму координованого розподілу керуючої дії між центральним та графічним процесорами. Цей розподіл здійснюється динамічно, пропорційно до їхнього реального внеску в поточне навантаження, який оцінюється за допомогою більш точної метрики фактичного навантаження (*effective load*), а не спрощеної утилізації. Ефективність розробленої системи було перевірено на комерційному смартфоні в умовах різних сценаріїв навантаження (CPU-орієнтований бенчмарк *Whetstone* та графічно-інтенсивний *3DMark Wild Life*) та порівняно зі штатною системою регулювання. Результати експериментів продемонстрували значні переваги запропонованого методу: було досягнуто одночасного приросту середньої продуктивності на 3,5–8,2% та зменшення частки часу, протягом якої корпус пристрою був перегрітим, на 55–65%. Крім того, забезпечено значно вищу стабільність продуктивності, що є критичним для вимогливих додатків. Таким чином, доведено, що розроблений підхід дозволяє досягти кращого балансу між обчислювальною потужністю та тепловою безпекою, що має високу практичну цінність для підвищення якості користувацького досвіду сучасних мобільних пристроїв.

Ключові слова: динамічне керування тепловим режимом, ПІД-регулятор, регулювання температури корпусу, система-на-кристалі, мобільний пристрій, тротлінг продуктивності.

Постановка проблеми. Сучасні мобільні пристрої, що функціонують на базі складних систем-на-кристалі (англ. *system on a chip*, SoC), виконують дедалі більш ресурсомісткі обчислювальні завдання [1]. Поширення вимогливих додатків, таких як ігри з високою якістю графіки та застосунок доповненої реальності, стимулює постійне масштабування продуктивності мобільних процесорів [2]. Цей процес, що часто реалізується шляхом збільшення кількості обчислювальних ядер та підвищення їх тактових частот, неминуче призво-

дить до зростання концентрації потужності та, як наслідок, до значних теплових викликів [3].

Ключова проблема полягає в тому, що компактні корпуси мобільних пристроїв мають обмежені можливості для розсіювання тепла, що генерується високопродуктивними SoC. Ця неефективність тепловідведення є прямою причиною підвищення температури зовнішньої оболонки пристрою [4]. Нагрівання корпусу не є суто технічним недоліком; воно має прямий і вимірюваний вплив на користувача. Дослідження пока-

зують, що температура поверхні, яка перевищує 45 °C, викликає значний дискомфорт у понад 75% користувачів [5]. Більше того, перегрів призводить до негативних наслідків, включаючи фізичний біль та погіршення емоційного стану, що суттєво знижує загальну задоволеність від використання пристрою [6]. Відчуття дискомфорту може зберігатися до однієї години навіть після завершення роботи з пристроєм, що підкреслює довготривалий негативний вплив цієї проблеми [7]. Таким чином, виникає фундаментальний конфлікт між прагненням до максимальної обчислювальної продуктивності та необхідністю забезпечення фізичного комфорту та позитивного користувацького досвіду, що робить керування температурою корпусу критично важливим завданням.

Аналіз останніх досліджень і публікацій. Для вирішення проблеми перегріву в мобільних системах активно розробляються та впроваджуються методи динамічного керування температурним режимом (Dynamic Thermal Management, DTM), огляд яких представлено в [8]. Історично, в промислових та вбудованих системах домінуючими підходами є використання позиційних (релейних) регуляторів та пропорційно-інтегрально-диференціальних (ПІД) контролерів [9]. Однак їх застосування в мобільних SoC стикається з низкою обмежень.

Багато існуючих стратегій регулювання температури корпусу зосереджені виключно на керуванні частотою центрального процесора (CPU), ігноруючи внесок графічного процесора (GPU) у загальне тепловиділення [10]. Такий підхід є неповним, оскільки сучасні робочі навантаження, особливо ігрові, інтенсивно використовують обидва обчислювальні модулі. Дослідження [11] переконливо демонструє, що незалежне та неузгоджене «дроселювання» (throttling) CPU та GPU є недоречним для сучасних мобільних завдань. Подібна неузгодженість призводить до суттєвого погіршення продуктивності основного додатка, що прямо шкодить користувацькому досвіду [12].

Хоча необхідність у скоординованій стратегії керування CPU та GPU визнається, залишається відкритою критична проблема: як саме динамічно та інтелектуально розподіляти керуючий вплив між цими двома основними джерелами тепла. Сучасні підходи не мають механізму адаптації стратегії дроселювання залежно від специфіки активного робочого навантаження, де відносний внесок CPU та GPU може кардинально змінюватися.

В академічному середовищі існують більш досконалі методи керування, такі як керування з прогнозуючими моделями (Model Predictive Control) або концепція цифрових двійників, проте ці підходи ще не знайшли широкого застосування в комерційних мобільних пристроях через свою обчислювальну складність та вимоги до ресурсів. Таким чином, існує чітко окреслена науково-технічна прогалина у створенні ефективного, адаптивного та практичного для реалізації механізму спільного керування температурою ЦП та ГП.

Постановка завдання. Метою статті є розробка та оцінка нового підходу до регулювання температури корпусу смартфона, який спільно керує частотами центрального та графічного процесорів, зважуючи і розподіляючи керуючу дію для CPU і GPU.

Виклад основного матеріалу.

1. Архітектура системи керування

Контур керування побудовано на принципі від'ємного зворотного зв'язку (Рисунок 1). У ролі контрольованої змінної (PV), виступає температура корпусу пристрою T_{shell} , значення якої зчитується з вбудованого датчика температури. Цей параметр зіставляється з уставкою (SP), що відповідає пороговому рівню температури, за якого користувач починає відчувати дискомфорт (тобто, 45 °C). Результатом порівняння є сигнал відхилення $e(t) = SP(t) - PV(t)$, що передається на вхід ПІД-регулятора.

У якості регулятора використано класичний ПІД-контролер (1), реалізований відповідно до стандарту ISA [13]. До його складу входить фільтр похідної складової, що зменшує вплив високочастотних коливань, типових для температурних сигналів, на ефективність системи. Налаштування регулятора включають коефіцієнт підсилення K_p , сталі часу інтегрування T_i та диференціювання T_d , а також параметр фільтру N , що представляє собою співвідношення між T_d і сталою часу додаткового фільтруючого кореня. У цій неперервній передатній функції s є комплексною змінною перетворення Лапласа, що використовується для представлення динаміки системи в частотній області.

$$W_{PID} = K_p \left(1 + \frac{1}{T_i s} + \frac{T_d s}{1 + (T_d s) / N} \right) \quad (1)$$

Дану стратегію керування було дискретизовано за допомогою методом білінійного перетворення (перетворення Тастіна) (2), у результаті чого отримано дискретну форму (3), придатну для використання в умовах реального середовища мобільного пристрою.

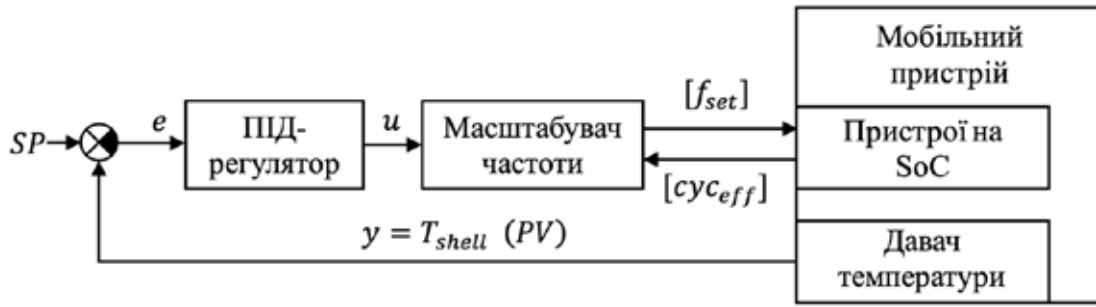


Рис. 1. Структура системи керування

$$s = \frac{2}{T_s} * \frac{1 - z^{-1}}{1 + z^{-1}} \quad (2)$$

$$W(z) = K_p \left(1 + \frac{T_s(z+1)}{2T_i(z-1)} + \frac{2T_d N(z-1)}{z(2T_d + NT_s) + (NT_s - 2T_d)} \right), \quad (3)$$

де T_s позначає період дискретизації. Застосоване перетворення здійснює відображення неперервної площини s у дискретну площину z . В отриманій дискретній моделі змінна z виступає у ролі комплексної змінної Z -перетворення, а вираз z^{-1} розглядається як оператор одиничного часового зсуву (затримки на один такт).

Для розрахунку керуючої дії, прикладеної до кожного з пристроїв на SoC, у контур керування введено масштабувач частоти. Реалізація асиметричного координованого керування полягає в масштабуванні керуючої дії відносно поточного теплового внеску окремого пристрою, який можна опосередковано оцінити за його навантаженням. Зазвичай, поточне навантаження вимірюється метрикою утилізації (utilization), що розраховується як частка часу, коли елемент виконував роботу – був не у стані простою:

$$util = 1 - \frac{t_{idle}}{t_{total}}, \quad (4)$$

де t_{idle} – відрізок часу, протягом якого процесор знаходився у стані простою, t_{total} – відрізок часу всього виміру.

Але, утилізація не показує реального навантаження на систему: навіть якщо процесор очікує відповідь від периферійного пристрою, таке очікування вважається активним і не класифікується як простій [14]. Більше того, ця метрика є залежною від встановленої частоти конкретного пристрою: при збільшенні частоти, утилізація зменшиться через зменшення часу виконання процесів, що може призвести до нестабільностей під час керування.

Для більш точної оцінки поточного навантаження системи було застосовано метрику effective

load – фактичного навантаження на елемент (5), що визначається як пропорція корисних тактових циклів cyc_{eff} – циклів які були витрачені на проведення розрахунків, до загального числа тактових циклів пристрою cyc_{total} за деякий період часу.

$$EL = \frac{cyc_{eff}}{cyc_{total}} \quad (5)$$

Для координованого керування в умовах асиметрії SoC, кожному елементу системи надається зважений коефіцієнт впливу, що визначає силу керуючої дії для цього пристрою. Зваження здійснюється пропорційно до нормованої метрики навантаження:

$$w_i = \frac{EL_i}{\sum_{j=0}^n EL_j}, \quad (6)$$

де w_i – вага керуючої дії для i -го пристрою, EL_i – фактичне навантаження цього пристрою, n – кількість керованих пристроїв.

Оскільки кожен із керованих обчислювальних блоків підтримує власний набір частот, що не є неперервним, модуль масштабування частоти також виконує перетворення зваженої керуючої дії на частоту до встановлення на елементи системи. Для цього, зважена керуюча дія $u \cdot w_i$ лінійно масштабується в допустимому діапазоні частот $[f_{min}, f_{max}]$, для отримання цільової частоти f_{target} (7). Після цього з наявного дискретного набору підтримуваних частот $f_{supported}$, обирається найближче значення f_{set} , яке й подається на пристрій (8).

$$f_{target,i} = f_{min,i} + u \cdot w_i \cdot (f_{max,i} - f_{min,i}) \quad (7)$$

$$f_{set,i} = \operatorname{argmin}_{f \in f_{supported,i}} |f_{target,i} - f| \quad (8)$$

Таким чином, пристрої з вищим реальним внеском у поточне навантаження та потенційно більшим тепловиділенням отримують пропорційно сильнішу корекцію керуючої дії, що дозволяє більш ефективно розподіляти теплові обмеження.

Для коректного налаштування регулятора було проведено ідентифікацію об'єкта керування. Його динаміку було апроксимовано моделлю другого порядку з запізненням і нулем:

$$W_m = K \frac{T_2 s + 1}{T_1^2 s^2 + 2\xi T_1 s + 1} e^{-\tau s}, \quad (9)$$

де K – коефіцієнт передачі, T_1 та T_2 – сталі часу, ξ – коефіцієнт демпфування та τ – час запізнення. В результаті ідентифікації ці параметри набули значень $K = 0.805 \text{ } ^\circ\text{C}/\%$, $T_1 = 338.82 \text{ с}$, $T_2 = 291.31 \text{ с}$, $\xi = 1.69$ і $\tau = 5 \text{ с}$. На основі отриманої моделі параметри ПД-регулятора було розраховано за експрес-методом, запропонованим Chien et al. (2003) (15) для даного типу систем, що дало змогу визначити такі налаштування: $K_p = 9.587 \text{ } \%/^\circ\text{C}$, $T_i = 1136 \text{ с}$, $T_d = 82.75 \text{ с}$, $N = 10$.

2. Експериментальна платформа

Для експериментальної валідації було обрано мобільний пристрій серійного виробництва ZTE Nubia RedMagic 7s Pro, оснащений системою на кристалі Qualcomm SM8450 Snapdragon 8 Gen 1. Ця SoC широко представлена в сучасних комерційних смартфонах, що дозволяє екстраполювати результати дослідження на досить широкий спектр пристроїв. Вона вміщає в собі центральний процесор, графічний процесор, та набір різноманітних апаратних модулів.

Центральний процесор цієї системи має три-кластерну архітектуру, що включає енергоефективний кластер «Little», продуктивний кластер «Big» та високопродуктивне ядро «Prime». У рамках реалізації системи керування, активне регулювання охоплювало лише кластери «Big» та «Prime», оскільки тепловий вплив «Little» кластеру на загальну температуру корпусу був визнаний неістотним за результатами попереднього аналізу. Крім центрального процесора, керуючий вплив також поширювався на графічний процесор, який є значним джерелом тепловиділення під час виконання графічно-інтенсивних обчислень.

Для реалізації зчитування даних та накладення керуючих дій було використано стандартні механізми, доступні в середовищі Android. Зокрема, для моніторингу та зміни тактових частот центрального і графічного процесорів застосовувалися інтерфейси файлової системи «sysfs», які забезпечують доступ до керування енергетичними параметрами апаратних компонентів. Зчитування температури корпусу здійснювалося через відповідний віртуальний сенсор, також доступний через «sysfs». Для отримання інформації про кількість корисних (активних) тактових циклів

використовувалася системна утиліта «perf», яка забезпечує доступ до апаратних лічильників продуктивності системи на кристалі.

Оцінку ефективності запропонованої системи керування було здійснено шляхом її порівняння з штатною системою регулювання температури корпусу у смартфоні. Ця система реалізує позиційний регулятор, який накладає обмеження на тактові частоти CPU та GPU при досягненні температурою корпусу порогу $46 \text{ } ^\circ\text{C}$ та скасовує їх при зниженні температури до $44 \text{ } ^\circ\text{C}$.

Усі експериментальні дослідження проводилися в стандартизованих контрольованих умовах, що забезпечували відтворюваність результатів та коректність порівняння. Температура навколишнього середовища підтримувалася на постійному рівні $25 \text{ } ^\circ\text{C}$. Перед кожним запуском тестів пристрій охолоджувався до температури корпусу $35 \text{ } ^\circ\text{C}$, яка визначалася як початкова точка вимірювання.

3. Методика експерименту

З метою об'єктивного аналізу впливу різних алгоритмів керування на продуктивність та теплову стабільність пристрою було проведено тестування у двох сценаріях з різним комплексом навантажень на компоненти SoC.

Перший сценарій передбачав вимірювання обчислювальної потужності центрального процесора. Для цього застосовувався бенчмарк Whetstone, який реалізує стандартизований набір обчислень з використанням як цілочисельної арифметики, так і операцій з плаваючою комою. Цей тест дозволяє оцінити пікову продуктивність CPU у вигляді кількості виконаних інструкцій Whetstone за одиницю часу (WIPS). З урахуванням гетерогенної архітектури процесора, програмне забезпечення було адаптовано таким чином, щоб забезпечити збалансований розподіл операційного навантаження між потоками, тим самим мінімізуючи вплив внутрішньої конфігурації на результати.

Другий сценарій був орієнтований на моделювання реального користувацького навантаження в умовах високої графічної інтенсивності. Для цього використовувався Wild Life Extreme Stress Test із пакету 3DMark, який є галузевим стандартом для стрес-тестування мобільних графічних підсистем. Він виконує тривалий рендеринг складних 3D-сцен із високою частотою оновлення кадрів, залучаючи як CPU, так і GPU. Основною метрикою, що фіксувалася під час проходження тесту, була частота кадрів (FPS, к/с), яка безпосередньо характеризує здатність системи забезпечу-

вати стабільний рівень продуктивності в умовах тривалого теплового навантаження.

4. Оцінка ефективності та результати

На рисунку 2 представлено температуру корпусу T_{shell} , обмеження тактових частот графічного процесора ($f_{lim, GPU}$), а також кластерів «Big» ($f_{lim, Big}$) та «Prime» ($f_{lim, Prime}$) центрального процесора під час виконання запланованих сценаріїв з штатною та запропонованою системами керування. З цих рисунків помітно, що запропонована система значно зменшила перегрів корпусу пристрою та забезпечила ефективний розподіл керуючої дії між керованими компонентами. Це розподілення дозволило запобігти агресивному зменшенню частоти компонентів, які не виконують велику кількість роботи, аби зберегти їх у режимі підвищеної готовності для обслуговування фонових або майбутніх процесів.

Як видно з даних, наведених у таблиці 1, використання запропонованої системи керування під час виконання бенчмарку Whetstone забезпечило

приріст продуктивності на 3,5% порівняно зі штатним підходом. Окрім цього, було досягнуто суттєвого покращення теплових характеристик: частка часу, протягом якого температура корпусу перевищувала комфортний рівень у 45 °C, зменшилася на 64,6%, а максимальне температурне значення також знизилася.

У випадку бенчмарку 3DMark спостерігалось ще виразніше поліпшення. Продуктивність зросла на 8,2%, тоді як частка часу перегріву скоротилася на 55,8%, що свідчить про більший потенціал покращення теплових характеристик пристрою за умов інтенсивнішого навантаження. Пікова температура корпусу при цьому також виявилася нижчою, ніж у випадку застосування штатної системи регулювання.

Крім того, дані на рисунку 3 свідчать, що застосування плавного регулювання дозволило досягти значно вищого рівня стабільності продуктивності в часі, що є результатом усунення агресивних спадів частоти, викликаних логікою позиційного керування.

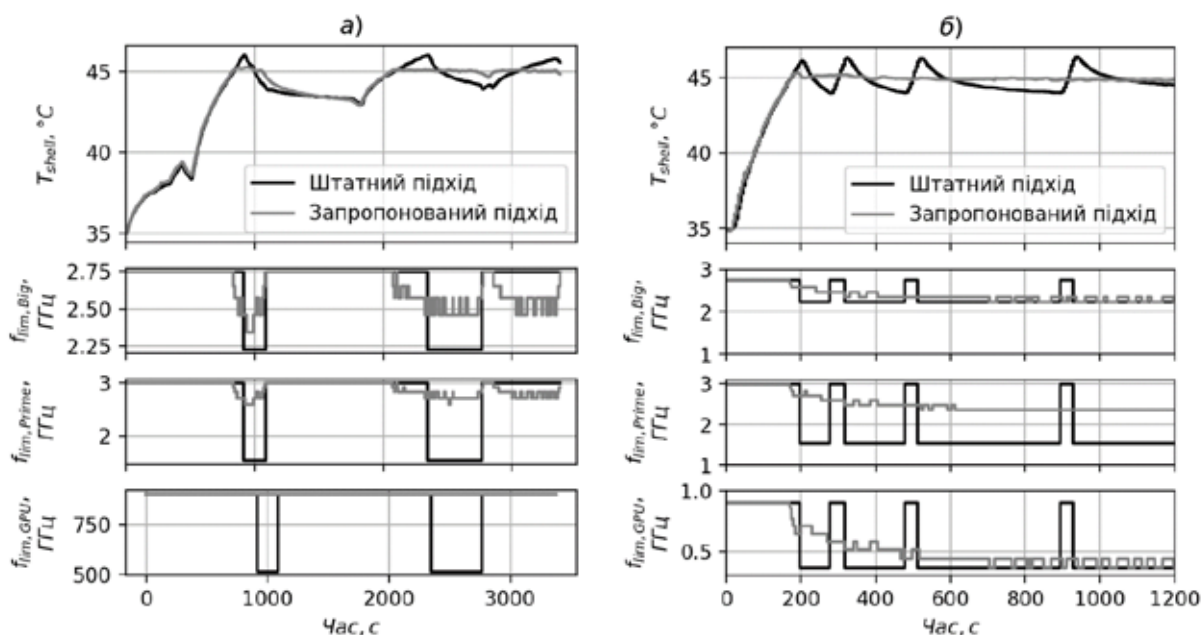


Рис. 2. Порівняння експлуатаційних показників смартфона під час оцінки стратегій керування на бенчмарках Whetstone (а) та 3DMark (б)

Таблиця 1
Порівняння продуктивності та теплових характеристик при виконанні бенчмарків Whetstone та 3DMark для штатної та запропонованої систем керування

Сценарій	Підхід	Результат бенчмарку	Частка часу перегріву, %	Пікова температура, °C
Whetstone	Штатний	4340.2 MWIPS	23.05	46.0
	Запропонований	4491.8 MWIPS	8.16	45.2
3DMark	Штатний	38.93 к/с	23.16	46.3
	Запропонований	42.13 к/с	10.22	45.2



Рис. 3. Порівняння середньої частоти кадрів за раундами бенчмарку 3DMark

Отримані результати підтверджують ефективність запропонованого підходу як з погляду підвищення обчислювальної продуктивності, так і з огляду на покращення теплових характеристик пристрою. Таким чином, впровадження системи інтелектуального керування дає змогу досягти кращого балансу між продуктивністю та тепловою безпекою під час виконання ресурсоємних задач на мобільних пристроях.

Висновки. У даній роботі було вирішено актуальну науково-технічну проблему керування температурою корпусу мобільних пристроїв, що виникає через конфлікт між зростаючою продуктивністю систем-на-кристалі та обмеженими можливостями тепловідведення. Метою дослідження була розробка та оцінка нового підходу до спільного керування частотами центрального та графічного процесорів, який би адаптувався до поточного навантаження.

Запропонований підхід базується на ПІД-регуляторі, ключовою особливістю якого є координованій розподіл керуючої дії між CPU та GPU. На відміну від стандартних методів, які використовують неточну метрику утилізації, розроблена система застосовує показник фактичного навантаження (effective load). Це дозволяє динамічно зва-

жувати вплив на кожен процесор пропорційно до його реального внеску в поточне обчислювальне завдання та, відповідно, у загальне тепловиділення.

Експериментальна валідація на комерційному смартфоні продемонструвала значні переваги розробленої системи порівняно зі штатною логікою позиційного регулювання.

- У сценарії з інтенсивним навантаженням на CPU (бенчмарк Whetstone) було досягнуто приросту продуктивності на 3,5% при одночасному зменшенні часу перегріву корпусу вище 45 °C на 64,6%.

- У змішаному сценарії, що імітує сучасні ігри (бенчмарк 3DMark), переваги були ще більш вираженими: продуктивність зросла на 8,2%, а частка часу перегріву скоротилася на 55,8%.

- Запропонований плавний метод регулювання забезпечив значно вищу стабільність продуктивності, усунувши різкі просідання частоти кадрів, характерні для штатних систем.

Таким чином, розроблена система керування ефективно вирішує поставлене завдання, забезпечуючи кращий баланс між обчислювальною потужністю та комфортом користувача. Вона дозволяє мобільному пристрою підтримувати вищий рівень продуктивності протягом тривалого часу, не допускаючи дискомфортного нагрівання корпусу.

Перспективними напрямками подальших досліджень є розширення системи для охоплення всіх значущих джерел тепла на SoC, включно з нейронним процесором та контролером пам'яті. Ключовим етапом стане перехід від реактивного до проактивного керування шляхом розробки предиктивних алгоритмів, які дозволять прогнозувати та нейтралізувати теплові навантаження ще до їхнього виникнення, що виведе ефективність терморегуляції на якісно новий рівень.

Список літератури:

1. Review of task-scheduling methods for heterogeneous chips / Z. Miao et al. Electronics. 2025. Vol. 14, no. 6. URL: <https://doi.org/10.3390/electronics14061191>.
2. Li X., Li G. An adaptive CPU-GPU governing framework for mobile games on big.little architectures. IEEE transactions on computers. 2021. Vol. 70, no. 9. P. 1472–1483. URL: <https://doi.org/10.1109/TC.2020.3012987>.
3. Panchal V. Thermal and power management challenges in high-performance mobile processors. International journal of innovative research in science, engineering and technology. 2024. Vol. 13. URL: <https://doi.org/10.15680/IJIRSET.2024.1311014>.
4. An overview of thermal challenges and opportunities for monolithic 3D ics / P. Shukla et al. Proceedings of the 2019 great lakes symposium on VLSI. New York, NY, USA, 2019. P. 439–444. URL: <https://doi.org/10.1145/3299874.3319485>.
5. Fire in your hands: understanding thermal behavior of smartphones / S. Kang et al. The 25th annual international conference on mobile computing and networking. New York, NY, USA, 2019. URL: <https://doi.org/10.1145/3300061.3300128>.

6. Does heat matter in phone usage? Antecedents and consequences of mobile thermal satisfaction / A. Zhang et al. Cross-Cultural design. culture and society / ed. by P.-L. P. Rau. Cham, 2019. P. 421–431.
7. The influence of smartphones' operation modes on the superficial temperature distribution in the human auricle region / J. Bauer et al. Journal of thermal analysis and calorimetry. 2018. Vol. 133, no. 1. P. 559–569. URL: <https://doi.org/10.1007/s10973-018-7047-8>.
8. A survey on energy management for mobile and iot devices / S. Pasricha et al. IEEE design & test. 2020. PP. P. 1–1. URL: <https://doi.org/10.1109/MDAT.2020.2976669>.
9. Stepanets O., Mariiash Y. Analysis of influence of technical features of a PID-controller implementation on the dynamics of automated control system. Eastern-European journal of enterprise technologies. 2018. Vol. 3. P. 60–69. URL: <https://doi.org/10.15587/1729-4061.2018.132229>.
10. Chetoui S., Reda S. CasCon: cascaded thermal and electrical current throttling for mobile devices. IEEE embedded systems letters. 2021. PP. P. 1–1. URL: <https://doi.org/10.1109/LES.2021.3079707>.
11. Integrated CPU-GPU power management for 3D mobile games / A. Pathania et al. 2014. URL: <https://doi.org/10.1145/2593069.2593151>.
12. Adaptive thermal property control technique for holistic thermal management of mobile devices / K. Lee et al. IEICE electronics express. 2018. Vol. 15. URL: <https://doi.org/10.1587/elex.15.20180187>.
13. ISA TR5.9-2023. Proportional integral derivative (PID) algorithms and performance / ISA Subcommittee 5.9. International society of automation, 2023.
14. Cockcroft A. Utilization is virtually useless as a metric!. 2006. P. 557–562.
15. O'dwyer A. Handbook of PI and PID controller tuning rules (3rd edition). World Scientific Publishing Company, 2009.

Zinchenko D.D., Mariiash Yu.I. COORDINATED CONTROL OF CPU AND GPU FREQUENCIES TO ENHANCE PERFORMANCE AND USER COMFORT IN MOBILE DEVICES

This paper addresses a key challenge in modern mobile electronics: device shell overheating, a direct consequence of the escalating computational power of Systems-on-Chip and constrained passive thermal dissipation. This issue extends beyond a mere technical drawback, as a surface temperature exceeding 45 °C causes significant physical discomfort for the user and substantially degrades the overall user experience. The objective of this study is to develop and experimentally evaluate a novel coordinated approach to dynamic thermal management that effectively resolves the fundamental trade-off between maximizing performance and ensuring user comfort. In contrast to prevalent industry solutions, which often rely on rudimentary threshold-based controllers or neglect the contribution of the graphics processing unit, we propose a comprehensive control system based on a digital Proportional-Integral-Derivative controller. The scientific novelty of this approach lies in the implementation of a mechanism for intelligent allocation of the control action between the central and graphics processing units. This allocation is performed dynamically, in proportion to their respective contributions to the current workload, which are assessed using a more precise “effective load” metric instead of conventional utilization. The efficacy of the proposed system was evaluated on a commercial smartphone under various workload scenarios (the CPU-intensive Whetstone benchmark and the graphically intensive 3DMark Wild Life) and compared against the stock thermal management system. The experimental results demonstrated the significant advantages of the proposed method: a simultaneous increase of 3.5–8.2% in average performance and a 55–65% reduction in the duration of overheating events were achieved. Furthermore, the system provided significantly higher performance stability, which is critical for demanding applications. Thus, the proposed approach is demonstrated to achieve a superior balance between computational power and thermal safety, holding significant practical value for improving the user experience on contemporary mobile devices.

Key words: dynamic thermal management (DTM), PID controller, skin temperature control, System-on-Chip (SoC), mobile device, performance throttling.

Дата надходження статті: 22.07.2025

Дата прийняття статті: 31.07.2025

Опубліковано: 27.10.2025

Ilin M.O.National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute”**Oleshchenko L.M.**National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute”

SOFTWARE-DRIVEN DATA ANALYSIS OF PROPAGANDA NARRATIVES AND SOCIAL MEDIA ENGAGEMENT

The article introduces a software framework for analyzing war propaganda distributed through Telegram during the full-scale invasion of Ukraine. The dataset comprises 5.3 million messages, 21.5 million text entities (hashtags, mentions, URLs), and 22.8 million user reactions (e.g., emojis, likes) from 116 public channels, the research approaches the problem from a structural and behavioral angle rather than through full-scale classification. The dataset for this research was extracted from a custom PostgreSQL database designed to handle high-volume, complex Telegram data related to Russian propaganda during the first two years of the full-scale invasion of Ukraine. The primary focus has been placed on building a processing pipeline capable of handling large and structurally heterogeneous data streams, most of which have been multilingual, emotionally charged, and inconsistently formatted. Exploratory analysis methods have been combined with basic feature engineering, which has made it possible to identify recurring patterns in hashtags, message structures, and audience reactions. For example, emoji usage has been found to be frequent but unevenly distributed, indicating the use of emotional framing strategies that vary across message types and channels.

Rather than presenting complete modeling results, the work has established a reproducible and modular baseline for future experimentation. Interpretable features, such as text length, hashtag count, emoji diversity, and posting times, have been extracted and organized for straightforward downstream analysis, including narrative classification, engagement prediction, or network modeling. Co-occurrence analyses of text entities and emoji-based response profiling have revealed how emotionally framed narratives have been shaped and reinforced. A few working hypotheses have been proposed, including the idea that emotionally framed content tends to receive more concentrated and homogeneous reactions.

The research has demonstrated that with a carefully designed data pipeline and relatively modest resources, it is possible to extract meaningful structure from adversarial, large-scale communication environments. The goal has not been to solve the propaganda problem in full, but rather to clarify what groundwork has been necessary for studying Telegram-based influence operations at scale. As disinformation ecosystems continue to evolve, a flexible software toolkit and foundational dataset have been offered to support future technical and analytical research on hostile media environments.

Key words: software methods of data analysis, messaging and social media platforms, narratives spreading, Python, narrative and engagement analysis, Big Data, processing heterogeneous streaming data.

Formulation of the problem. The use of propaganda in wartime is not new, but the way it spreads has transformed significantly in the digital age. Messaging platforms like Telegram have become vital tools in modern information warfare. What sets Telegram apart is its hybrid nature: it functions both as a broadcasting platform and a private messaging app. This structure, with public channels that can reach hundreds of thousands of subscribers, is particularly

useful for spreading consistent narratives with minimal external interference.

Since the beginning of Russia’s full-scale invasion of Ukraine in 2022, Telegram has played a central role in the online propaganda ecosystem. Public channels linked to Russian-aligned messaging have pushed out content at high volumes, targeting both domestic and international audiences. These messages range from traditional patriotic appeals to more

subtle or misleading information that aims to distort facts and undermine trust.

While there has been public awareness of these campaigns, the academic and technical communities still face challenges in scaling analysis to the volume and variety of content distributed across platforms like Telegram. Much of the research has focused on more mainstream social media, leaving this relatively unregulated space underexamined [1, 2]. The multilingual, emotional, and highly contextual nature of the content complicates efforts to systematically analyze it.

This research addresses that gap. By focusing on Telegram-based propaganda channels, we present a framework capable of handling large-scale, heterogeneous datasets. Our aim is not just to describe what these channels are doing, but to build software tools and methods that allow others to analyze, monitor, and potentially intervene in digital information conflicts more effectively.

Analysis of recent research and publications.

Over the past few years, researchers have intensely studied how propaganda and disinformation spread on social media. The issue gained urgency after events like Brexit and the 2016 U.S. elections, and especially during the COVID-19 “infodemic,” which highlighted the scale of online misinformation. In response, numerous works from 2020–2025 have focused on detecting deceptive or influential content in online platforms. Propaganda is typically defined as strategically crafted messaging aimed at influencing public opinion for a specific agenda. It often exploits rhetorical and emotional techniques – for example, using loaded language, patriotic appeals, or false dilemmas – to sway an audience. Notably, propaganda is related to but distinct from disinformation, which is deliberately false information spread with the intent to cause harm. Propaganda can be based on truths or falsehoods and may not always be overtly malicious, whereas disinformation by definition involves falsehood and harmful intent [3].

To clarify such concepts, scholars have proposed various misinformation taxonomies. Modern studies have expanded these categorizations to dozens of specific tactics. Likewise, researchers distinguish misinformation (unintentionally incorrect claims) from disinformation (intentional falsehoods) [4].

A variety of computational techniques have been applied to identify propaganda and influence campaigns in social media. One prominent approach is stance detection – determining whether a text agrees, disagrees, or is neutral toward a given claim or narrative. Stance detection plays a key role in fact-check-

ing and rumor verification systems. Beyond stance, scholars are increasingly interested in narrative analysis of disinformation. Rather than treating each message in isolation, narrative-focused studies examine the storylines and themes that link many pieces of content. Another crucial method in propaganda detection is sentiment and emotion analysis. Propagandists frequently use emotional appeals – outrage, fear, national pride, anger – to make their message resonate. Several studies have integrated sentiment detection into misinformation classifiers.

Researchers now leverage stance detection, narrative analysis, sentiment analysis, and taxonomies to categorize different forms of propaganda and disinformation. This rich toolkit, developed in studies of Twitter, Facebook and other platforms, sets the stage for examining newer outlets. In particular, these methods inform our analysis of Russian propaganda on Telegram [5].

Telegram is a hybrid messaging and social media platform whose unique affordances have made it a hotbed for propaganda dissemination in recent years. Telegram enables users to create public channels that can broadcast messages to unlimited subscribers, all while allowing a high degree of anonymity. This anonymity, combined with Telegram’s lax moderation, provides a fertile ground for disinformation actors. Telegram’s minimal content moderation stands in stark contrast to the policies of Facebook, Twitter, or YouTube. In response to bans on mainstream platforms, many Russian state-affiliated actors migrated to Telegram. Telegram’s channel-based architecture leads to decentralized information flows that often form echo chambers. A recent study analyzed over 1,700 Telegram channels and identified “bridge nodes” that link disparate groups by sharing messages across ideological lines [6].

The content of these propaganda posts often revolved around key narratives: justifications for the invasion, denials of atrocities, and conspiracy theories.

Compared to Twitter or Facebook, Telegram also presents hurdles for researchers and policymakers trying to monitor propaganda. Telegram has no official public API for channel messages. Researchers must manually discover relevant Telegram channels and scrape messages, which can lead to sampling biases. Telegram’s openness allowed both genuine citizen reporting and false propaganda to thrive side by side. Analysts have noted that during 2022–2023, the Kremlin’s information operatives treated Telegram as a digital frontline.

Recent advances in NLP, particularly the development of large pre-trained language models, have

enhanced our ability to process and interpret text at scale. A cornerstone of modern text analysis is the BERT family of models. For Russian-language text, researchers commonly use RuBERT or multilingual models like XLM-R. In propaganda detection tasks, fine-tuned BERT models have achieved state-of-the-art results. Alongside transformers, more lightweight text representations still play a role. Word embeddings such as those produced by fastText can efficiently encode words into vectors and have been used in many multilingual text applications.

To extract themes and narratives from text corpora, topic modeling techniques such as LDA and BERTopic are widely used. BERTopic combines transformer embeddings with clustering to generate more coherent topics.

Named Entity Recognition (NER) tools can automatically tag mentions of people, organizations, and places. NER also enables the construction of networks of co-occurring entities. In multilingual settings, NER models or dictionaries need to be adapted for each language [7].

Despite the arsenal of NLP tools available, analysts face several challenges when mining large-scale text data for propaganda. Firstly, there is a scarcity of labeled training data for supervised learning. Creating high-quality annotations for what counts as propaganda vs. normal content is time-consuming. Irony and sarcasm can invert the literal meaning of text, confounding simple sentiment or stance detectors. The evolving nature of language and tactics means any static model can become outdated. Propagandists continually adapt their vocabulary to evade filters. Our methodology will address these challenges by applying tools that balance performance with adaptability [8].

Task statement. This research is structured around a set of guiding questions that emerged during the process of working with large-scale Telegram data in the context of Russian wartime propaganda. Rather than focusing on predictive classification, we concentrated on questions that could be addressed through structural analysis, feature extraction, and descriptive modeling.

The core research questions are as follows.

1. What kinds of structural and content-based patterns can be observed in Telegram propaganda messages, and how are they distributed across channels and over time? This includes recurring hashtags, variations in message length and complexity, and patterns in emoji usage as a proxy for emotional tone and engagement.

2. How can a scalable, relational data model support the systematic analysis of such content across millions of records? Here, the aim is to define a reproducible schema that allows for both flexibility and performance when working with Telegram data in relational form.

3. Which features – textual, temporal, and behavioral can be extracted efficiently and interpreted meaningfully to support future analytical tasks such as narrative detection or engagement modeling?

To address these questions, we developed a modular pipeline for ingesting, exploring, and engineering features from Telegram propaganda content. The goal was to produce a reusable analytical framework that can serve both as a descriptive lens and as a foundation for future modeling efforts.

Outline of the main material of the research.

The dataset used in this research was extracted from a custom-built PostgreSQL database, designed to accommodate the scale and complexity of data collected from Telegram channels disseminating Russian propaganda during the first two years of the full-scale invasion of Ukraine. The dataset comprises information from 116 public Telegram channels, identified as part of a coordinated pro-Kremlin media ecosystem. These channels were selected based on prior research, open-source intelligence, and manual vetting of content history.

In total, the database contains:

- 5,309,851 messages;
- 21,511,942 text entities (e.g., hashtags, mentions, URLs);
- 22,794,701 reactions (emojis, likes, other engagement markers).

The database is structured across four interlinked tables:

- *channels*: contains metadata about each Telegram channel, including its Telegram ID, username, and classification by type (e.g., news, commentary, satire);
- *messages*: stores individual posts along with their metadata—timestamps, author identifiers, message types, edited versions, media attachments (photo, video), and raw text content;
- *text_entities*: extracted elements from each message, such as hashtags, hyperlinks, mentions, and formatted text (bold, italic, etc.), stored with type labels and positional order;
- *reactions*: records emoji-based reactions tied to specific messages, including the emoji itself and the count of users who reacted with it.

This schema allows us to analyze messages not only as standalone units but also in terms of how

they are embedded within broader discursive and engagement contexts. Messages are linked to both their originating channels and associated interactions (entities, reactions), forming a relational structure suitable for large-scale querying, statistical analysis, and modeling.

The temporal attributes stored in both human-readable (timestamp) and Unix (unixtime) formats support granular chronological analysis, enabling us to capture patterns in propaganda bursts, narrative pivots, and audience response cycles over time. The data used in this research was collected through a combination of automated and supervised methods. Public Telegram channels were identified based on prior open-source investigations, civil society monitoring projects, and manual verification of content consistency with known patterns of Russian-aligned propaganda. The collection process relied on Telegram's open API endpoints and automated scraping tools, ensuring that all acquired data remained within the bounds of publicly accessible content. No private groups or encrypted messages were accessed. Only public channels were included, and only data visible without user authentication was collected. This guarantees that the data acquisition process adhered to Telegram's terms of service and respected the platform's intended public-private content boundaries.

The resulting database was designed to support both relational analysis and scalable processing. Data was stored in normalized tables, preserving structural integrity and making it possible to trace any record back to its channel, timestamp, or message ID. Fields such as user identifiers were retained only when they referred to public or institutional accounts (e.g., media outlets or known figures); any personal user data was either excluded or anonymized. Given the sensitive nature of the material – propaganda produced during an active war – the research followed ethical guidelines for handling war-related and potentially harmful content. The research does not reproduce or promote any of the propagandist material. Instead, all analysis is aimed at deconstructing techniques and patterns of influence for academic and counter-disinformation purposes. All results are presented in aggregate, and no attempts are made to deanonymize individuals or trace personal interactions. Furthermore, wherever relevant, disclaimers are applied to visualizations and quoted examples to contextualize their nature as subjects of critical analysis rather than information endorsement.

To process the Telegram data at scale, we developed a modular analytical pipeline consisting of four main stages: data ingestion and profiling, exploratory

statistics, systematic cleansing, and feature engineering. The design emphasizes reproducibility and computational efficiency, recognizing the challenges of working with large volumes of data collected from propaganda-linked Telegram channels.

The dataset is stored in a PostgreSQL database and structured across four normalized tables: channels, messages, text_entities, and reactions. Initial profiling queries confirmed structural integrity, absence of missing values, and alignment between foreign key references (e.g., message_id links across tables). All core fields (text_content, channel_id, message_date) were present across records, enabling immediate downstream analysis.

Sampling and summary queries revealed early insights into message types and engagement patterns. A random sample of messages showed content ranging from daily greetings and calls to action to battlefield reports and emotional appeals.

Many posts featured strong emotive or rhetorical devices, often paired with emoji reactions – most frequently:

- 👍 (over 4 billion reactions);
- 😊, 🙏, ❤️ and 🇷🇺 – each with several hundred million reactions.

These reactions offer a behavioral proxy for audience engagement and emotion-triggering content. Textual entities, particularly hashtags, revealed patterns of narrative segmentation and branding. The most frequent hashtags included campaign-specific tags alongside broader geopolitical markers. This structure enables both time-series tracking and topical clustering in later analysis.

Despite the dataset's scale, quality checks showed no missing values in critical fields. Messages were filtered for completeness and deduplicated where necessary. Repeated content, truncated posts, or bot-generated noise were flagged using length and lexical diversity heuristics.

To support downstream modeling, features were engineered at multiple levels:

- *textual*: message length, punctuation density, stopword ratios, named entity counts, and TF-IDF vectors;
- *behavioral*: reaction counts by emoji, average engagement per channel, emoji entropy;
- *semantic*: hashtags, sentiment polarity (to be derived), rhetorical signal tokens.

The architecture supports stepwise iteration and selective caching of derived variables to ensure scalability across millions of records.

Fig. 1 presents proposed multi-stage pipeline for processing raw Telegram data (channels, messages,

text entities, and reactions) to enable downstream analytical tasks. The process begins with Data Ingestion & Profiling, where the structure is validated, samples are summarized, and foreign key integrity is checked. This is followed by Exploratory Analysis, which focuses on frequency statistics, emoji patterns, and hashtag surfacing.

Data Cleansing is performed to remove duplicates, normalize formats, and identify noisy or incomplete data. These cleaned and profiled data then feed into Feature Engineering, where NLP features, reaction metrics, and temporal features are extracted. The final outputs are used for various Downstream Tasks.

The initial exploratory phase aimed to surface broad patterns across message volume, content structure, and user reactions. This step provided the empirical grounding for refining data cleansing procedures and for selecting relevant features for later modeling. The initial exploratory phase utilized Python tools such as Pandas for data manipulation, Matplotlib and Seaborn for visual analysis, and NumPy for efficient numerical operations. These libraries enabled the identification of broad patterns in message volume, content structure, and user reactions, providing an empirical foundation for refining data cleansing procedures and guiding feature selection for subsequent modeling.

Across the 5.3 million messages, activity remained sustained throughout the two-year window, with noticeable peaks aligning with battlefield events or significant political developments. Due to hardware constraints, full time-series plotting was deferred; timestamp sampling showed regular posting, with specific spikes around symbolic dates and coordinated media pushes.

Sampled messages revealed considerable variability in content structure—from brief emoji-laced posts to lengthy reports or transcripts. Posts often began with rhetorical intensifiers (“!!”, “⚡”) and included hashtags or campaign-specific branding. Some posts lacked any meaningful text, while others spanned several paragraphs.

User reactions, recorded as over 22 million individual counts, show a clear tendency toward emotionally charged content. The top emojis – 👍, 😬, 🙏, ❤️, and 🤔 – are indicative of binary approval (👍), outrage, or emotional resonance (🙏, ❤️).

Particularly noteworthy is the dominance of 👍, which accounted for over 4 billion cumulative reactions, suggesting either:

- the use of 👍 as a default button on many channels,
- or bot-driven amplification, a known tactic in propaganda ecosystems.

Analysis of 21.5 million text entities shows extensive use of hashtags as narrative anchors. To under-

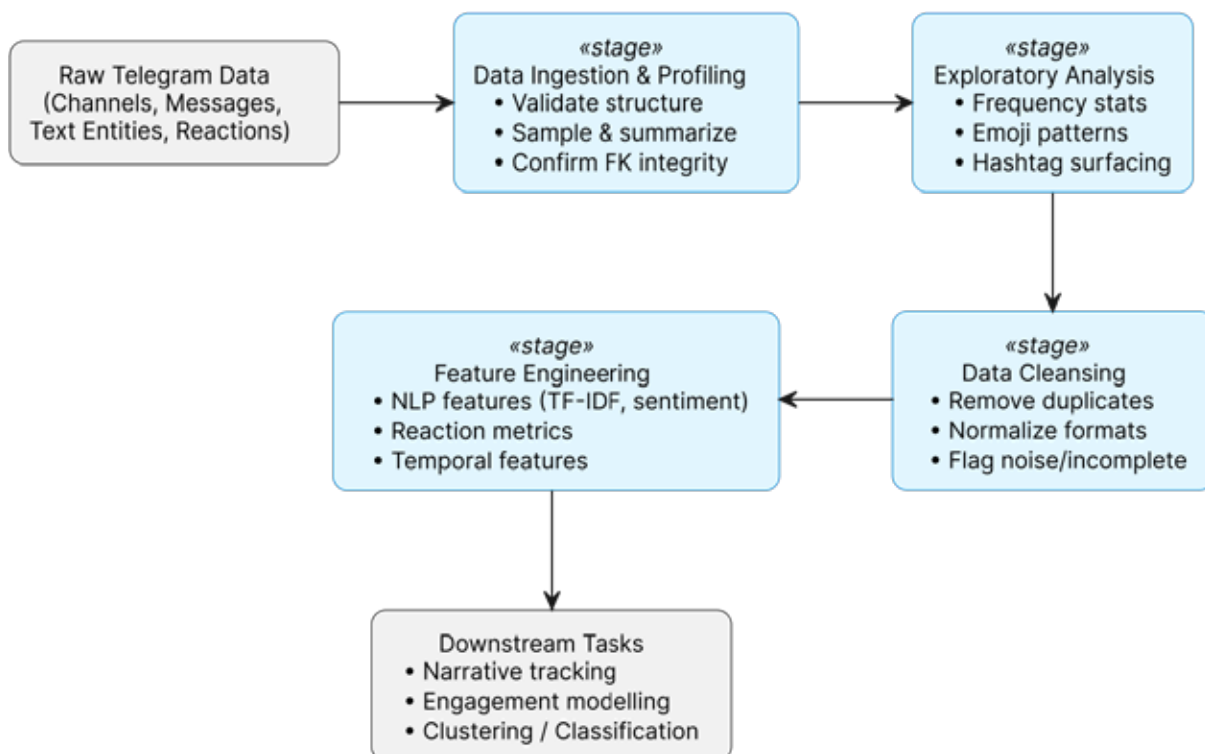


Fig. 1. Proposed overall methodological pipeline

stand how propaganda dissemination varies across actors, we conducted a comparative analysis of activity levels and engagement metrics for individual Telegram channels. Using aggregated data from the relational schema, we computed per-channel counts for total messages, word volumes, and emoji-based reactions. This revealed distinct behavioral signatures and operational scales among the channels. Emoji reactions offer insight into audience resonance. While absolute counts correlate with message volume, certain channels consistently outperform others in reaction density.

Textual entities – particularly hashtags – play a central role in shaping and signaling propaganda narratives. They function not only as indexing devices but also as rhetorical cues that frame a message’s tone, urgency, or political alignment. In our dataset, over 21.5 million such entities were extracted, with hashtags representing a major subset.

To visualize broader relationships, we constructed a network graph where nodes represent hashtags and edges represent co-appearance frequency. Thicker edges denote stronger co-occurrence ties.

This analysis demonstrates that hashtags are not randomly or independently deployed. Rather, they form deliberate semantic bundles that support narrative framing, audience targeting, and automated repetition. To enable downstream analytical tasks such as clustering, classification, or narrative tracking, we engineered a set of interpretable features from the raw Telegram messages. Given the high volume of data and the heterogeneity of message formats, we prior-

itized lightweight, expressive features that capture structural, semantic, and behavioral characteristics of each message.

Building on this infrastructure, we outline several hypothesis for our research:

H1: messages containing emotionally charged hashtags receive significantly higher reaction counts than neutral ones;

H2: channels using more repetitive, short-format messaging (high frequency, low word count) achieve higher cumulative engagement per unit time;

H3: reaction entropy is inversely correlated with message polarity – highly polarized messages attract fewer distinct emoji types.

In follow-up work, we intend to:

- train supervised classifiers to predict reaction volume or narrative type based on engineered features and embeddings;
- integrate transformer-based sentence representations (e.g., RuBERT, XLM-R) for improved text modeling;
- apply graph-based methods to track cross-channel influence and narrative flow;
- expand the dataset to include multilingual propaganda networks, especially channels targeting foreign audiences in English or German.

To support downstream modeling tasks and exploratory narrative analysis, features were engineered at multiple levels – textual, behavioral, temporal, and semantic (Table 1). These features prioritize interpretability and computational efficiency across millions of records.

Table 1

Engineered features

Feature Name	Description	Type
text_length	Character count of the message text	Numeric
word_count	Number of words in the message	Numeric
punctuation_density	Ratio of punctuation marks to total characters	Numeric
stopword_ratio	Fraction of stopwords among all tokens (language-specific)	Numeric
named_entity_count	Count of named entities per message (if processed)	Numeric
tfidf_vector	Sparse vector representation of text content	Vector
hashtag_count	Number of hashtags in message	Numeric
url_count	Number of URLs present	Numeric
reaction_total	Sum of all reactions to the message	Numeric
reaction_entropy	Diversity of emoji reactions (Shannon entropy)	Numeric
emoji_count	Count of emojis in message body	Numeric
avg_channel_engagement	Channel-level average reactions per post	Numeric
sentiment_polarity	To be derived via sentiment analysis model	Numeric
rhetorical_tokens	Binary indicators of known propaganda keywords or formats	Categorical
hour_of_day	Posting hour (0–23)	Categorical
day_of_week	Weekday of message (0=Monday, ..., 6=Sunday)	Categorical
channel_id	Foreign key linking message to its source channel	Categorical

Conclusions. This research presented a structured approach to analyzing large-scale propaganda content disseminated through Telegram during the first two years of Russia's war against Ukraine. Rather than attempting to solve classification or detection tasks directly, we focused on building a reproducible and efficient pipeline for processing, exploring, and organizing Telegram data at scale.

By designing a relational schema and extracting interpretable features from over five million messages, we were able to highlight patterns in messaging structure, hashtag use, and audience reaction. Co-occurrence analyses of text entities and emoji-based response profiling provided early insights into how narratives are framed, repeated, and emotionally charged across different channels. The exploratory phase used Python libraries like Pandas, NumPy, Matplotlib, and Seaborn to identify key patterns in message volume, structure, and reactions. These

insights informed data cleaning and feature selection for later analysis.

Our primary contribution lies in creating a foundation for future work – both technical and analytical. The cleaned, structured dataset and modular feature engineering workflow are intended to support downstream tasks such as narrative classification, engagement prediction, or social graph modeling. The hypotheses outlined in this research may serve as starting points for researchers interested in tracing influence strategies or testing propaganda effectiveness over time.

The research demonstrates that with relatively modest resources, it is possible to extract meaningful structure from complex, adversarial communication environments – provided the data pipeline is thoughtfully designed. As disinformation ecosystems continue to evolve, we hope this work offers a flexible and transparent basis for analyzing them with greater clarity and scale.

Bibliography:

1. Starbird K. Examining the alternative media ecosystem through the production of alternative narratives of mass shooting events on Twitter. *Proceedings of the International AAAI Conference on Web and Social Media*. 2017. Vol. 11(1). P. 230–239. URL: <https://ojs.aaai.org/index.php/ICWSM/article/view/14878> (date of access: 20.04.2025).
2. Zannettou S., Caulfield T., Blackburn J., Stringhini G. Disinformation warfare: Understanding state-sponsored trolls on Twitter and their influence on the Web. *Companion Proceedings of the Web Conference 2020*. P. 218–226. DOI: <https://doi.org/10.1145/3308560.3316495>.
3. Cinelli M., Quattrocioni W., Galeazzi A., et al. The COVID-19 social media infodemic. *Scientific Reports*. 2020. Vol. 10. Article 16598. DOI: <https://doi.org/10.1038/s41598-020-73510-5>.
4. Hardalov M., Nakov P., Koychev I. A survey on stance detection for mis- and disinformation identification. *Information Processing & Management*. 2022. Vol. 59(3). Article 102899. DOI: 10.18653/v1/2022.findings-naacl.94.
5. Tschirky M., Makhortykh M. #Azovsteel: comparing qualitative and quantitative approaches for studying framing of the siege of Mariupol on Twitter. *Media, War & Conflict*. 2023. Vol. 17 (2). P. 163–178. DOI: 10.1177/17506352231184163.
6. Li Q., Liu Q., Liu S., Di X., Chen S., Zhang H. Influence of social bots in information warfare: a case study on @UAWeapons Twitter account in the context of the Russia–Ukraine conflict. *Communication and the Public*. 2023. Vol. 8(2). P. 54–80. DOI: 10.1177/20570473231166157.
7. Lin Z., Xie J., Li Q. Multi-modal news event detection with external knowledge. *Information Processing & Management*. 2024. Vol. 61(3), Article 103697. DOI: 10.1016/j.ipm.2024.103697.
8. Javed D., Jhanjhi N.Z., Khan N.A., Ray S.K., Al Mazroa A.A., Ashfaq F., Das S.R. Towards the future of bot detection: a comprehensive taxonomical review and challenges on Twitter/X. *Computer Networks*. 2024. Vol. 254. Article 110808. DOI: 10.1016/j.comnet.2024.110808.

Льїн М.О., Олещенко Л.М. ПРОГРАМНО-КЕРОВАНІЙ АНАЛІЗ ДАНИХ НАРАТИВІВ ПРОПАГАНДИ ТА ЗАЛУЧЕННЯ В СОЦІАЛЬНИХ МЕРЕЖАХ

У статті представлено програмне рішення для аналізу російської пропаганди, що поширюється через Telegram під час повномасштабного вторгнення в Україну. Набір даних для дослідження містить 5,3 мільйони повідомлень, 21,5 мільйони текстових сутностей (хештеги, згадки, URL) та 22,8 мільйони реакцій користувачів (емодзі, вподобання тощо) зі 116 публічних каналів. Дослідження здійснено з урахуванням структурного та поведінкового аналізу, а не повної класифікації. Було використано дані з бази PostgreSQL, спеціально розробленої для обробки великих обсягів складної інформації з Telegram, пов'язаної з російською пропагандою протягом перших двох років повномасштабного вторгнення. Основну увагу приділено побудові обробного конвеєра, здатного ефективно працювати з великим і неоднорідним потоком даних, більшість яких є багатомовними, емоційно забарвленими та нестан-

дартно структурованими. Було поєднано методи розвідкового аналізу з базовою інженерією ознак, що дозволило виявити повторювані шаблони в хештегах, структурі повідомлень і реакціях аудиторії. Зокрема, було виявлено, що використання емодзі є частим, але розподіляється нерівномірно, що вказує на застосування стратегій емоційного фреймінгу, які варіюються залежно від типу повідомлень і каналів. Замість представлення повних результатів моделювання у роботі було закладено відтворювану та модульну основу для подальших експериментів. Було виокремлено інтерпретовані ознаки, такі як довжина тексту, кількість хештегів, різноманітність емодзі та час публікацій і систематизовано їх для подальшого аналізу, зокрема, для класифікації наративів, прогнозування залучення чи моделювання мереж. Був проведений аналіз співзвучності текстових сутностей та профілювання реакцій на основі емодзі, що дозволило виявити, як емоційно забарвлені наративи формуються та підкріплюються. Було сформульовано кілька робочих гіпотез, зокрема, припущення, що емоційно оформлений контент, як правило, отримує більш концентровані та однорідні реакції.

Було продемонстровано, що за умови використання ретельно спроектованого конвеєра обробки даних і відносно помірних ресурсів можливо виокремити значущі структурні закономірності з великих обсягів ворожого комунікаційного контенту. Метою дослідження було формування основ, необхідних для масштабного дослідження впливових операцій у Telegram. В умовах постійної еволюції дезінформаційних екосистем було запропоновано гнучкий програмний інструментарій і базовий набір даних, покликаний підтримати майбутні технічні та аналітичні дослідження у сфері ворожих інформаційних середовищ.

Ключові слова: програмні методи аналізу даних, платформи обміну повідомленнями, соціальні мережі, поширення наративів, Python, аналіз наративів і залучення, великі дані, обробка неоднорідних потокових даних.

Дата надходження статті: 07.07.2025

Дата прийняття статті: 18.07.2025

Опубліковано: 27.10.2025

Козуб Г.О.

ДЗ «Луганський національний університет імені Тараса Шевченка»

Артеменко О.І.

ПВНЗ «Буковинський університет»

Осадчук С.І.

ПВНЗ «Буковинський університет»

ПАРАЛЕЛЬНЕ ПРОГРАМУВАННЯ В МОБІЛЬНИХ ІГРАХ: МЕТОДИКИ ОПТИМІЗАЦІЇ ГРАФІКИ ТА ІГРОВОЇ ЛОГІКИ

У роботі зроблено спробу визначити ефективні способи використання багатоядерних архітектур та графічних прискорювачів для підвищення продуктивності ігор, зменшення енергоспоживання та покращення якості мобільних ігор для користувачів.

Метою статті є дослідження використання сучасних методів паралельного програмування для оптимізації графіки та ігрової логіки в мобільних іграх, наскільки вони ефективні та практичні в умовах обмежень поширених мобільних пристроїв.

Наукова новизна. Стаття пропонує систематизований аналіз методів паралельного програмування, включаючи GPU-шейдери, CPU-багатопоточність та розподілені обчислення, з акцентом на їх застосування в мобільних іграх. Новизна полягає в порівняльному аналізі цих методів з урахуванням їх впливу на енергоспоживання та продуктивність на пристроях з обмеженими ресурсами. Запропонована архітектурна схема паралельного виконання демонструє ефективний розподіл обчислень між графічними, логічними та фізичними потоками, що сприяє оптимальному використанню ресурсів. Практична реалізація на прикладі паралельної обробки матричних операцій для 3D-графіки підкреслює потенціал прискорення обчислень у реальних ігрових сценаріях.

Результати. Дослідження показало, що паралельне програмування значно підвищує продуктивність мобільних ігор завдяки ефективному використанню GPU для рендерингу та CPU для ігрової логіки. GPU-шейдери забезпечують високу пропускну здатність графічних обчислень, тоді як багатопоточність на CPU оптимізує обробку ігрової логіки, зокрема штучного інтелекту та фізики. Розподілені обчислення зменшують навантаження на пристрій, що особливо важливо для багатокористувацьких ігор. Порівняльний аналіз методів (таблиця 1) виявив їхні переваги та обмеження: GPU-шейдери ускладнюють синхронізацію, CPU-багатопоточність має високі накладні витрати, а розподілені обчислення залежать від мережі. Практична реалізація паралельної обробки 3D-об'єктів продемонструвала прискорення обчислень до 2–3 разів порівняно з послідовним виконанням.

Висновки. Паралельне програмування є ключовим інструментом для створення високопродуктивних мобільних ігор, дозволяючи ефективно використовувати багатоядерні архітектури та графічні прискорювачі. Оптимізація синхронізації потоків, управління пам'яттю та енергоспоживанням залишаються основними викликами. Розподілені обчислення відкривають перспективи для стабільної роботи багатокористувацьких ігор. Майбутні дослідження мають зосередитися на автоматизації паралелізації та інтеграції штучного інтелекту для адаптивної оптимізації продуктивності залежно від характеристик пристрою та навантаження.

Ключові слова: програмування, оптимізація продуктивності, ігрова логіка, мобільна графіка, синхронізація потоків.

Постановка проблеми. Сучасні мобільні ігри характеризуються високими вимогами до графічної якості та складності ігрової логіки, що створює значні виклики для розробників у контексті

обмежених ресурсів мобільних пристроїв. Зростання популярності віртуальної комунікації та інтерактивних ігрових систем вимагає від розробників пошуку ефективних методів оптимізації

продуктивності мобільних додатків. Традиційні послідовні алгоритми не здатні повною мірою використовувати можливості сучасних багатоядерних процесорів та графічних прискорювачів мобільних пристроїв, що призводить до зниження частоти кадрів, збільшення енергоспоживання та погіршення користувацького досвіду. Особливо гостро постає питання балансу між візуальною якістю ігор та їх продуктивністю на пристроях з обмеженими обчислювальними ресурсами та ємністю батареї.

Аналіз останніх досліджень і публікацій. Дослідження в галузі паралельного програмування для мобільних ігор активно розвиваються в останні роки. Бойко О. Ю. [1, с. 5] проаналізував феномен віртуальної комунікації, що є фундаментальною основою для розуміння взаємодії користувачів з мобільними ігровими платформами. Декларативний підхід при створенні мультиплатформних додатків, досліджений Козуб Г. О. та Козуб Ю. Г. [2, с. 12], надає цінні інсайти щодо архітектурних рішень для мобільних ігор, особливо в контексті паралельної обробки різних компонентів системи. Czarnul P. та колеги [3, с. 45] провели комплексний огляд методів оптимізації resource-aware паралельних та розподілених обчислень, підкреслюючи важливість ефективного використання багатоядерних архітектур в умовах обмежених ресурсів мобільних пристроїв.

Ою Дж., Сонг Жи. Та Вонг Й. (Ou J., Song G. та Wang Y.) [4, с. 78] досліджували застосування графічних процесорів для планування шляхів на основі global evolutionary dynamic programming, демонструючи потенціал GPU-прискорення для складних обчислювальних задач, що безпосередньо застосовується до алгоритмів навігації та штучного інтелекту в мобільних іграх. Жанг Х. та Жанг Х. (Zhang H. та Zhang H.) [5, с. 685] розробили 3D selective kernel residual network, яка використовує паралельні обчислення для обробки тривимірних даних, що може бути адаптовано для оптимізації графічного рендерингу та обробки геометрії в іграх. Юовоно С. (Yuwono S.) та колеги [6, с. 582] вивчали самооптимізацію в розподілених виробничих системах з використанням modular state-based Stackelberg games, що має пряме застосування до багатокористувацьких мобільних ігор та балансування навантаження між клієнтами та серверами.

Рапака А. (Rapaка A.) та співавтори [7, с. 3] аналізували революційні зміни в навчальних іграх з використанням immersive та AI-технологій, що

відкриває нові можливості для інтеграції штучного інтелекту в оптимізацію ігрової логіки та адаптивного балансування обчислювальних ресурсів. Грувер Н. (Grover N.) [8, с. 35] досліджував AI-enabled supply chain optimization, принципи якої можуть бути застосовані для динамічного управління ресурсами та автоматичного налаштування параметрів паралельного виконання в мобільних іграх. Луан Д. (Luan D.) та колеги [9, с. 15] розробили методи stability-aware data offloading optimization в системах edge-based mobile crowdsensing, що критично важливо для забезпечення стабільної роботи багатокористувацьких мобільних ігор з розподіленою архітектурою. Ліу С. (Liu S.) та співавтори [10, с. 3] створили систему optimal fault tolerant control для multiplayer систем з зовнішніми збуреннями на основі adaptive dynamic programming, що забезпечує надійність та стабільність онлайн-ігор при варіативних мережевих умовах та навантаженнях.

Аналіз літературних джерел показує, що паралельне програмування в мобільних іграх охоплює широкий спектр технологій від низькорівневої оптимізації GPU до високорівневих методів розподілених обчислень, при цьому ключовими викликами залишаються енергоефективність, синхронізація та адаптація до обмежених ресурсів мобільних пристроїв.

Постановка завдання. Метою статті є дослідження сучасних методик паралельного програмування для оптимізації графіки та ігрової логіки в мобільних іграх, аналіз їх ефективності та практичного застосування в умовах обмежених ресурсів мобільних пристроїв.

Виклад основного матеріалу. Паралельне програмування в мобільних іграх охоплює два основні напрямки: оптимізацію графічного рендерингу та прискорення обчислень ігрової логіки. Графічні процесори сучасних мобільних пристроїв містять сотні обчислювальних ядер, які можуть ефективно обробляти шейдери паралельно. Оптимізація шейдерів передбачає мінімізацію залежностей між обчисленнями та ефективне використання кеш-пам'яті для досягнення максимальної пропускної здатності. Використання API низького рівня, таких як Vulkan та Metal, дозволяє розробникам безпосередньо керувати GPU, розподіляючи робоче навантаження між кількома обчислювальними чергами для паралельного виконання завдань рендерингу.

Багатопоточність для ігрової логіки передбачає розділення обчислювальних завдань між різними ядрами CPU. Грувер Н. (Grover N.) [8, с. 35] під-

креслює важливість AI-enabled optimization supply chains, що може бути адаптовано для динамічного балансування навантаження між потоками в мобільних іграх. Обробка фізики, штучного інтелекту та управління ігровим світом можуть виконуватися в окремих потоках, що забезпечує стабільну частоту кадрів навіть при складних обчисленнях. Використання паралельних алгоритмів сортування та пошуку для обробки великих списків ігрових об'єктів значно зменшує час обробки ігрового світу.

Розподілені обчислення набувають особливого значення для багатокористувацьких мобільних ігор. Луан Д. (Luan D.) та колеги [9, с. 15] досліджували stability-aware data offloading optimization в системах edge-based mobile crowdsensing, що безпосередньо застосовується до архітектури мобільних ігор. Перенесення частини ігрової логіки на сервер дозволяє зменшити навантаження на мобільний пристрій, забезпечуючи більш стабільну продуктивність. Ліу С. (Liu S.) та співавтори [10, с. 3] розробили методи optimal fault tolerant control для multiplayer систем, що критично важливо для забезпечення надійності онлайн-ігор.

Аналіз різних підходів до паралельного програмування в мобільних іграх вимагає систематизації їх характеристик та можливостей застосування. Наступна таблиця 1 представляє порівняльний аналіз основних методів паралелізації з урахуванням їх переваг, обмежень та впливу на енергоспоживання мобільного пристрою.

Представлена таблиця 1 демонструє, що кожен метод паралельного програмування має специфічні переваги та обмеження, що визначають область його оптимального застосування. Як найкращий вибір для графічних процесорів, шейдери забезпечують найбільшу швидкість, але вони повинні синхронізуватися у складний спосіб. Хоча багатопоточність процесора дає найбільші можливості для керування грою, вона має високу ціну

з точки зору ресурсів. Використання розподілених обчислень дозволяє економити ресурси пристрою, але при цьому покладається на хорошу мережу.

Переконалися, що процеси працюють паралельно, як задумано, є головною перешкодою при написанні мобільних ігор. Неправильна синхронізація може призвести до проблем у поведінці програми. Спільне використання ресурсів без проблем можна гарантувати за допомогою м'ютексів, семафорів та структур даних без блокування. Процес обробки та роботи з потоками слід оцінювати відповідно до того, яку цінність вони додають програмі, по відношенню до ресурсів, яких вони потребують.

Обмеження мобільних пристроїв накладають специфічні вимоги на проектування паралельних алгоритмів. Обмежена кількість ядер CPU, відносно невелика кількість обчислювальних блоків GPU та обмежений обсяг пам'яті вимагають оптимізації алгоритмів для мінімізації використання ресурсів. Енергоспоживання є критичним фактором, оскільки паралельні обчислення можуть значно скоротити час роботи батареї пристрою.

Для візуалізації організації паралельного виконання в мобільних іграх доцільно розглянути архітектурну схему, яка демонструє взаємодію різних компонентів системи. Наступна діаграма 1 ілюструє розподіл обчислювальних завдань між спеціалізованими потоками та їх синхронізацію для забезпечення ефективного використання ресурсів мобільного пристрою.

З наведеної вище схеми видно, що мобільна гра організовує паралельне виконання у три етапи, де графічний, логічний та фізичний потоки мають окремі обов'язки. У графіці потік обробляє зображення, промальовує об'єкти та виконує додаткову постобробку за допомогою шейдерів GPU, після чого формується буфер кадру. Обчислення ШІ, стан гри та мережева координація обробляються логічним потоком, формуючи весь стан гри. Використовуючи фізику, виявлення зіткнень,

Таблиця 1

Порівняння методів паралельного програмування в мобільних іграх

Метод	Область застосування	Переваги	Обмеження	Енергоспоживання
GPU шейдери	Графічний рендеринг	Висока пропускна здатність	Складність синхронізації	Середнє
CPU багатопоточність	Ігрова логіка	Гнучкість розподілу задач	Накладні витрати на потоки	Високе
Розподілені обчислення	Багатокористувацькі ігри	Зменшення навантаження на пристрій	Залежність від мережі	Низьке
Гібридний підхід	Комплексні ігри	Оптимальне використання ресурсів	Складність архітектури	Змінне

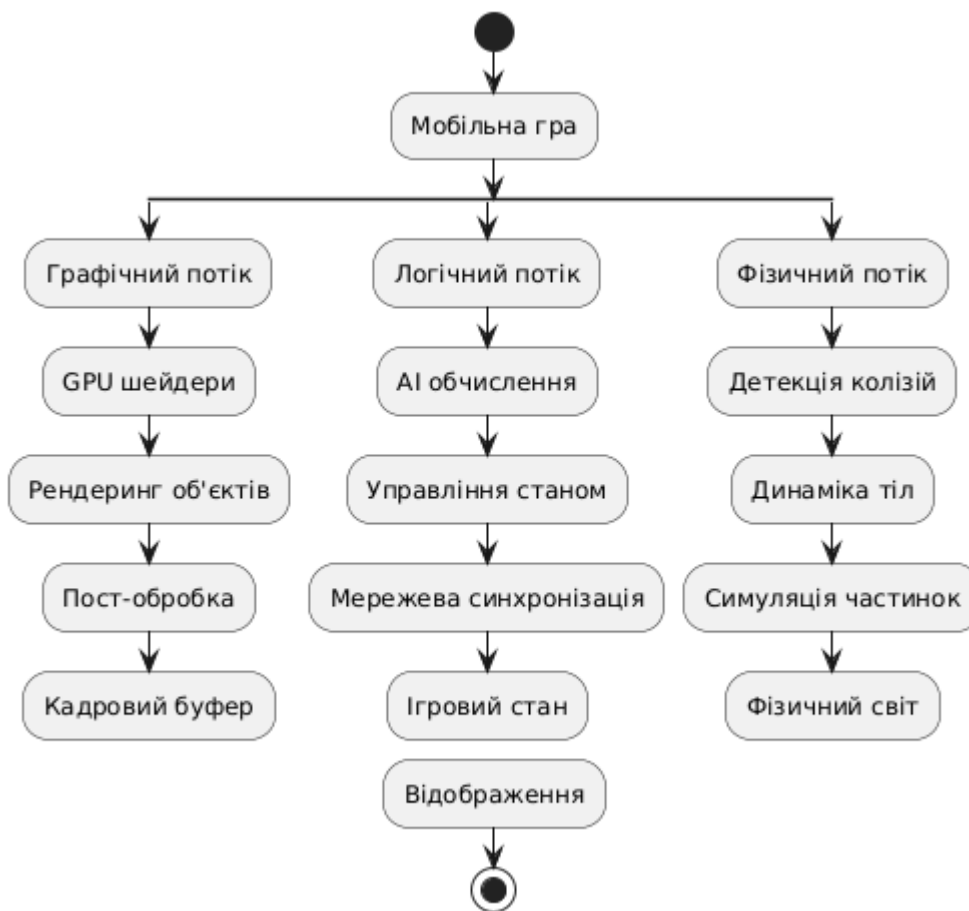


Рис. 1. Архітектура паралельного виконання в мобільній грі

Джерело: авторська розробка в PlantUML online editor

рух тіл і частинок імітують фізичний світ у грі. Завдяки такій архітектурі комп'ютерні ресурси пристрою використовуються найкращим чином, а різні ігрові модулі не створюють перешкод один одному.

Практична реалізація паралельного програмування в мобільних іграх може бути продемонстрована на прикладі оптимізації обчислень матричних операцій для трансформації 3D об'єктів:

```
import numpy as np
import threading
import time
from concurrent.futures import
ThreadPoolExecutor
```

```
class ParallelGameEngine:
    def __init__(self, num_objects=1000):
        self.num_objects = num_objects
        self.objects = self.generate_objects()
        self.transformation_matrices = self.
generate_transformations()
```

```
def generate_objects(self):
    «»»Генерація випадкових 3D об'єктів«»»
    return [np.random.rand(4, 100) for _
in range(self.num_objects)]
```

```
def generate_transformations(self):
    «»»Генерація матриць трансформації«»»
    return [np.random.rand(4, 4) for _ in
range(self.num_objects)]
```

```
def sequential_transform(self):
    «»»Послідовна трансформація
об'єктів«»»
    start_time = time.time()
    results = []
```

```
    for i in range(self.num_objects):
        transformed = np.dot(self.
transformation_matrices[i], self.objects[i])
        results.append(transformed)
```

```
    end_time = time.time()
    return results, end_time - start_time
```

```

def parallel_transform_worker(self, start_idx, end_idx):
    «»»Робоча функція для паралельної трансформації«»»
    results = []
    for i in range(start_idx, end_idx):
        transformed = np.dot(self.transformation_matrices[i], self.objects[i])
        results.append(transformed)
    return results

def parallel_transform(self, num_threads=4):
    «»»Паралельна трансформація об'єктів«»»
    start_time = time.time()

    # Розділення роботи між потоками
    chunk_size = self.num_objects // num_threads
    futures = []

    with ThreadPoolExecutor(max_workers=num_threads) as executor:
        for i in range(num_threads):
            start_idx = i * chunk_size
            end_idx = start_idx + chunk_size
            if i < num_threads - 1 else self.num_objects:
                future = executor.submit(self.parallel_transform_worker, start_idx, end_idx)
                futures.append(future)

    # Збір результатів
    results = []
    for future in futures:
        results.extend(future.result())

    end_time = time.time()
    return results, end_time - start_time

def benchmark_performance(self):
    «»»Порівняння продуктивності«»»
    print(f»Тестування з {self.num_objects} об'єктами:«»)

    # Послідовне виконання
    _, seq_time = self.sequential_transform()
    print(f»Послідовне виконання: {seq_time:.4f} секунд«»)

    # Паралельне виконання з різною кількістю потоків

```

```

for threads in [2, 4, 8]:
    _, par_time = self.parallel_transform(threads)
    speedup = seq_time / par_time
    print(f»Паралельне виконання ({threads} потоки): {par_time:.4f} секунд, прискорення: {speedup:.2f}x«»)

# Демонстрація роботи
if __name__ == «__main__»:
    engine = ParallelGameEngine(num_objects=2000)
    engine.benchmark_performance()

    print(«\nАналіз енергоспоживання:»)
    print(«Паралельні обчислення дозволяють швидше завершити задачі,»)
    print(«але можуть збільшити миттєве споживання енергії.»)
    print(«Оптимальна кількість потоків залежить від архітектури пристрою.»)

```

Даний приклад демонструє практичну реалізацію паралельного програмування для оптимізації графічних обчислень в мобільних іграх. Код можна запустити в Google Colab для тестування продуктивності різних підходів до паралелізації матричних операцій, які є основою 3D графіки.

Висновки. Паралельне програмування є фундаментальним інструментом для створення високопродуктивних мобільних ігор в умовах сучасних багатоядерних архітектур. Дослідження показало, що ефективне використання GPU для графічного рендерингу та CPU для ігрової логіки дозволяє досягти значного прискорення обчислень при правильному проектуванні архітектури додатку. Синхронізація потоків, управління пам'яттю та оптимізація енергоспоживання залишаються ключовими викликами, які вимагають комплексного підходу до розробки. Розподілені обчислення відкривають нові можливості для багатокористувацьких мобільних ігор, дозволяючи перенести частину навантаження на сервери та забезпечити стабільну продуктивність на пристроях з обмеженими ресурсами. Майбутні дослідження мають зосередитися на розробці автоматизованих методів паралелізації та інтеграції штучного інтелекту для динамічної оптимізації продуктивності мобільних ігор залежно від характеристик конкретного пристрою та поточного навантаження системи.

Список літератури:

1. Бойко О. Ю. Інтердисциплінарний огляд феномену віртуальної комунікації. *Академічні візії*. 2023. № 23. С. 1-10. URL: <https://academy-vision.org/index.php/av/article/view/556>
2. Козуб Г. О., Козуб Ю. Г. Декларативний підхід при створенні мультиплатформних додатків. *Вісник Східноукраїнського національного університету імені Володимира Даля*. 2022. № 5 (275). С. 10–15. DOI: <https://doi.org/10.33216/1998-7927-2022-275-5-10-15>
3. Czarnul P., Antal M., Baniata H., Griebler D., Kertesz A., Kessler C. W., Kouloumpiris A., Kovačić S., Markus A., Michael M. K., Nikolaou P., Öz I., Prodan R., Rakić G. Optimization of resource-aware parallel and distributed computing: a review. *The Journal of Supercomputing*. 2025. Vol. 81, no. 7. DOI: <https://doi.org/10.1007/s11227-025-07295-7> URL: <https://surl.li/osinjp>
4. Ou J., Song G., Wang Y. Graphics processing unit-enabled path planning based on global evolutionary dynamic programming and local genetic algorithm optimization. *Applied Soft Computing*. 2025. Vol. 176, 113167. DOI: <https://doi.org/10.1016/j.asoc.2025.113167> URL: <https://surli.cc/qadusj>
5. Zhang H., Zhang H. LungSeek: 3D Selective Kernel residual network for pulmonary nodule diagnosis. *The Visual Computer*. 2022. Vol. 39, no. 2. P. 679–692. DOI: <https://doi.org/10.1007/s00371-021-02366-1> URL: <https://link.springer.com/article/10.1007/s00371-021-02366-1>
6. Yuwono S., Hussain A. K., Schwung D., Schwung A. Self-optimization in distributed manufacturing systems using Modular State-based Stackelberg games. *Journal of Manufacturing Systems*. 2025. Vol. 80. P. 578–594. DOI: <https://doi.org/10.1016/j.jmsy.2025.03.025> URL: <https://surli.cc/wyclik>
7. Rapaka A., Dharmadhikari S. C., Kasat K., Mohan C. R., Chouhan K., Gupta M. Revolutionizing learning – A journey into educational games with immersive and AI technologies. *Entertainment Computing*. 2025. Vol. 52, 100809. DOI: <https://doi.org/10.1016/j.entcom.2024.100809> URL: <https://surl.li/mygkxz>
8. Grover N. AI-Enabled Supply Chain Optimization. *International Journal of Advanced Research in Science, Communication and Technology*. 2025. P. 28–44. DOI: <https://doi.org/10.48175/ijarsct-23103> URL: <https://ijarsct.co.in/Paper23103.pdf>
9. Luan D., Wang E., Liu W., Yang Y., Deng J. Stability-aware data offloading optimization in edge-based mobile crowdsensing. *Frontiers of Computer Science*. 2025. Vol. 19, no. 11. DOI: <https://doi.org/10.1007/s11704-024-40620-6> URL: <https://link.springer.com/article/10.1007/s11704-024-40620-6>
10. Liu S., Xu N., Li L., Alharbi K. H., Zhao X. Zero-sum games-based optimal fault tolerant control for control-constrained multiplayer systems with external disturbances via adaptive dynamic programming. *Communications in Nonlinear Science and Numerical Simulation*. 2025. Vol. 147, 108804. DOI: <https://doi.org/10.1016/j.cnsns.2025.108804> URL: <https://surl.li/hwzwao>

Kozub H.O., Artemenko O.I., Osadchuk S.I. PARALLEL PROGRAMMING IN MOBILE GAMES: TECHNIQUES FOR OPTIMIZING GRAPHICS AND GAME LOGIC

The work tries to determine effective ways of using multicore architectures and graphics accelerators to enhance game performance, use less power, and make mobile games more enjoyable for users.

The article is meant to study the use of modern parallel programming to optimize graphics and game logic in mobile games, looking at how effective and practical they are under the limitations of common mobile devices.

Scientific novelty. The article offers a systematic analysis of parallel programming methods, including GPU shaders, CPU multithreading, and distributed computing, with a focus on their application in mobile games. The novelty lies in the comparative analysis of these methods, considering their impact on power consumption and performance on devices with limited resources. The proposed architectural scheme of parallel execution demonstrates the efficient distribution of computations between graphical, logical, and physical threads, which contributes to the optimal use of resources. Practical implementation as an example of parallel processing of matrix operations for 3D graphics emphasizes the potential for accelerating computing in real game scenarios.

Results. The study showed that parallel programming significantly improves the performance of mobile games through the efficient use of GPUs for rendering and CPUs for game logic. GPU shaders provide high throughput for graphical computing, while CPU multithreading optimizes the processing of game logic, including artificial intelligence and physics. Distributed computing reduces the load on the device, which is especially important for multiplayer games. A comparative analysis of the methods (Table 1) revealed their advantages and limitations: GPU shaders complicate synchronization, CPU multithreading has high overhead, and distributed computing depends on the network. The practical implementation of parallel processing of 3D objects has demonstrated a computational speedup of up to 2–3 times compared to sequential execution.

Conclusions. Parallel programming is a key tool for creating high-performance mobile games, allowing the efficient use of multi-core architectures and graphics accelerators. Optimization of thread synchronization, memory management, and power consumption remain the main challenges. Distributed computing opens prospects for stable operation of multiplayer games. Future research should focus on automating parallelization and integrating artificial intelligence to adaptively optimize performance based on device characteristics and load.

Key words: *programming, performance optimization, game logic, mobile graphics, thread synchronization.*

Дата надходження статті: 09.07.2025

Дата прийняття статті: 16.07.2025

Опубліковано: 27.10.2025

Комлева Н.О.

Національний університет «Одеська політехніка»

РОЗШИРЕННЯ БАЗИСУ UML ДЛЯ ПІДТРИМКИ НЕДОСКОНАЛИХ ТА ЗМІНЮВАНИХ ВИМОГ У ПРОЄКТУВАННІ ІНТЕЛЕКТУАЛЬНИХ СИСТЕМ

У статті розглянуто проблему формалізації недосконалих, змінюваних та ймовірнісних вимог у контексті проєктування інтелектуальних програмних систем. У реальних умовах такі системи функціонують у динамічному середовищі з високим ступенем невизначеності, що ускладнює застосування традиційних підходів до інженерії вимог. Стандартний базис мови UML (Unified Modeling Language), попри її широке застосування у сфері програмної інженерії, не забезпечує належної підтримки нечіткості, неповноти, конфліктності чи варіативності вимог. Це створює інформаційний розрив між фазою збирання вимог і етапом моделювання, оскільки значущі дані втрачаються або фіксуються поза межами моделей, що обмежує можливості трасування, аналізу та автоматизованої обробки.

У роботі запропоновано розширення стандартного UML-профілю, яке дозволяє зберігати семантику недосконалих вимог. Для цього вводяться спеціальні стереотипи, теговані значення (tagged values) та обмеження (constraints), які дозволяють маркувати елементи моделей із вказанням рівня довіри, джерела інформації, ймовірності реалізації та контекстуальних умов. Показано, як ці розширення можуть бути інтегровані до різних типів UML-діаграм (use case, class, activity, state, sequence) без зміни ядра UML, зберігаючи при цьому узгодженість із метамоделлю.

Практичну реалізацію підходу описано у вигляді спеціалізованого UML-профілю, який може бути збережений як у стандартному форматі XMI для CASE-засобів типу Enterprise Architect або MagicDraw, так і у вигляді JSON-структури для StarUML. У роботі представлено приклади типових стереотипів (<<fuzzyRequirement>>, <<incomplete>>, <<conflicting>>, <<uncertainState>> тощо), а також тегованих атрибутів, які підтримують розширену семантику вимог.

Для ілюстрації загального підходу побудовано узагальнювальну діаграму етапів узгодження та інтерпретації вимог для UML-проєктування, яка відображає логіку переходу від недосконалих вимог до формалізованих моделей. Запропонований підхід є гнучким і сумісним з UML 2.x, що дає змогу легко масштабувати та адаптувати його до потреб різних інструментальних платформ. Він також сприяє поліпшенню комунікації між аналітиками, розробниками та замовниками, зменшуючи ризики втрати інформації у процесі проєктування інтелектуальних систем. Отримані результати мають потенціал стати основою для розширення CASE-засобів і розробки нових методів підтримки нечітких і динамічних вимог у середовищах із підвищеною складністю.

Ключові слова: інженерія вимог, UML-профіль, нечіткі вимоги, неповнота, ймовірнісні моделі, CASE-засоби, стереотипи UML, формалізація вимог, автоматизоване проєктування.

Постановка проблеми. Сучасні інтелектуальні програмні системи функціонують у середовищах, що характеризуються високим рівнем динаміки, складності та невизначеності. Для таких систем ключовим етапом життєвого циклу є проєктування на основі вимог, що часто є неповними, суперечливими або змінюваними. Традиційні методи інженерії вимог, а також стандартні засоби моделювання, зокрема UML, передбачають роботу з формально узгодженими, чітко визначеними специфікаціями. Однак на практиці це припущення є обмежувальним, оскільки в реальних

умовах розробники змушені працювати з «недосконалими» вимогами.

Стандартний UML-базис, що широко застосовується в проєктуванні програмного забезпечення, не передбачає явної підтримки недосконалих вимог. На практиці такі вимоги можуть бути нечіткими, неповними, суперечливими, ймовірнісними або змінюваними, що ускладнює їх формалізацію та інтеграцію в моделі. Проєктувальники змушені або штучно «доочишувати» вимоги перед моделюванням (що призводить до втрати важливої інформації), або зберігати допоміжні

коментарі поза межами моделей, що ускладнює автоматизовану обробку, трасування та подальшу верифікацію. Для забезпечення цілісного проектування необхідне розширення UML-базису таким чином, щоб у ньому можна було безпосередньо відображати ступінь визначеності, джерело, довіру, контекстуальні умови та варіативність кожного елемента.

Нижче наведено класифікацію основних проблем, що притаманні недосконалим вимогам:

- неповнота – вимога не містить усіх необхідних деталей, таких як умови активації, результати виконання або взаємозв'язки з іншими вимогами; це унеможливує побудову повного проектного артефакту без додаткових припущень;
- конфліктність – різні вимоги можуть суперечити одна одній логічно, ресурсно або функціонально; такі суперечності виявляються під час аналізу сценаріїв, залежностей або пріоритетів;
- нечіткість та багатозначність – вимога містить формулювання, які допускають кілька тлумачень, наприклад: «високий рівень доступності», «швидке реагування» або «важливий користувач»;
- невизначеність – вимога містить умовну або ймовірнісну інформацію (наприклад, «у 80% випадків очікується...»), що ускладнює жорстку трансляцію в UML-моделі;
- змінюваність – вимоги можуть змінюватись із часом або в залежності від контексту, що накладає вимоги до їх гнучкого відображення та трасування;
- невимірюваність / неперевірюваність – вимога сформульована так, що її складно або неможливо перевірити у процесі тестування або валідації.

UML-нотація, зокрема її основні діаграми (use case, class, activity, state, sequence), передбачає повністю визначену, логічно узгоджену структуру. Сутності, зв'язки та атрибути мають бути чітко визначені, без неоднозначностей. Як результат, поточний UML-базис не дозволяє: фіксувати альтернативні варіанти інтерпретації вимог, позначати ступінь довіри чи джерело вимоги, підтримувати вірогіднісні або умовні залежності між сценаріями, маркувати частково заповнені елементи моделі (наприклад, клас без типу атрибуту або з невизначеною поведінкою). Це створює бар'єр між фазами інженерії вимог і проектування, оскільки важлива інформація про невизначеність або контекст губиться при трансляції в UML.

Аналіз останніх досліджень і публікацій. Сучасні інтелектуальні системи повинні враховувати неповноту, конфліктність, невизначеність

та динамічність вимог, що ставить нові виклики перед методами проектування. Значну увагу дослідники приділяють питанням формалізації таких вимог, їх узгодження, а також відображення у моделях програмного забезпечення.

Зокрема, у роботі [1, с. 134–152] представлено підхід до верифікації UML-моделей на основі нечіткої логіки та описової логіки, що дозволяє враховувати нечіткість у специфікаціях. Аналогічні ідеї розвиваються у [2, с. 1–8], де розглянуто підтримку інженерії вимог за допомогою механізмів нечіткої логіки. Комплементарний підхід запропоновано у [3, с. 317–322], де формалізовано нечітке подання вимог у процесах прийняття рішень. Автори [4, с. 207–224] акцентують увагу на усуненні абстрактного розриву між моделями та кодом за допомогою нечіткої логіки в тестуванні.

У публікаціях [5, с. 1–12; 6, с. 1–15] обґрунтовано доцільність використання графових моделей для відображення залежностей між вимогами та використання цілочислового програмування для їх відбору. Системний огляд підходів до представлення невизначеності в програмних моделях наведено у [7, с. 1183–1213], де розглянуто також класифікацію типів невизначеності. Проблема роботи з ймовірнісними характеристиками подій також висвітлюється у [8, с. 88026–88048].

Дослідження [9, с. 2313–2360] містить систематичний огляд виконання UML-моделей, що є важливим для створення інструментів з підтримкою нечітких або конфліктних вимог. У [10, с. 1–29] автори аналізують роль «впевненості у знаннях» у доменних моделях. Серед останніх публікацій слід також відзначити [11, с. 1–20], де представлено уніфіковану модель невизначеності у кіберфізичних системах.

В українському сегменті важливими є праці, де досліджується моделювання інтелектуальних діагностичних систем в умовах обмежених ресурсів і часткової невизначеності. Зокрема, в роботах [12, с. 152–156; 13, с. 99–108; 14, с. 97–100] висвітлено питання бізнес-орієнтованого інжинірингу вимог, оптимізації вибору функціоналу та обліку факторів невизначеності в процесі розробки.

Отже, попри наявність значної кількості робіт, що торкаються окремих аспектів нечіткості чи адаптації вимог, спостерігається брак комплексних підходів до розширення стандартних засобів моделювання, таких як UML, для відображення недосконалих і змінюваних вимог у динамічних середовищах. Саме це і визначає актуальність подальших досліджень у цьому напрямі.

Постановка завдання. Метою роботи є розширення базису UML з урахуванням специфіки недосконалих та змінюваних вимог, що виникають у процесі проектування інтелектуальних систем, для забезпечення їх повноцінної формалізації, відображення невизначеності та підтримки подальшого автоматизованого моделювання.

Виклад основного матеріалу. Єдина структура, яку задає стандарт UML, передбачає однозначність, завершеність та внутрішню узгодженість усіх елементів моделі. Проте в реальних умовах інженерії вимог до інтелектуальних систем часто мають справу з неповними, багатозначними або конфліктними твердженнями. Такі вимоги не можуть бути прямо трансформовані у типові UML-діаграми, не втрачаючи при цьому критичну інформацію.

Недосконалість вимог, зокрема їх нечіткість, неповнота, протиріччя та змінюваність, не можуть бути адекватно відображені стандартними засобами UML. Щоб зберігати цю критично важливу інформацію під час моделювання, доцільно впровадити розширення до UML-профілю. Такі розширення повинні дозволяти фіксувати джерело вимоги, ступінь довіри, ймовірність виконання, а також контекстуальні залежності. Найбільш зручним механізмом для цього є введення відповідних анотацій, що можуть бути прикріплені до елементів UML-діаграм, зокрема: власні властивості (tagged values) для збереження довіри, ймовірності, джерела; стереотипи для позначення специфічних типів вимог або їх статусу (наприклад, «<<fuzzyRequirement>>»); обмеження (constraints) для опису умов або альтернативних варіантів поведінки.

Запропонований підхід орієнтований не на зміну сутності UML як стандарту, а на розширення його семантики засобами, які вже передбачено в архітектурі UML 2.x – через створення спеціалізованого UML-профілю, який підтримуватиме ключові характеристики недосконалих вимог. Практичне застосування розширень можна проілюструвати на прикладі ключових типів діаграм, які найчастіше використовуються в процесі моделювання вимог.

Діаграма варіантів використання. Кожен варіант використання може бути розширений властивостями confidenceLevel, source, contextConditions, які дозволяють вказати ймовірність виконання, джерело вимоги (користувач, регламент, інженер тощо) та умови, за яких сценарій активується. Для відображення альтернативних сценаріїв можна застосовувати стереотипи

<<variantScenario>> або <<exceptionalCase>>.

Діаграма класів. Атрибути класу можуть мати невизначений тип або значення, які позначаються спеціальним маркером (?, unknown, ToBeDetermined) і супроводжуються властивістю completeness = «partial». Стереотипи <<fuzzyClass>>, <<incompleteClass>> дають змогу явно вказати, що сутність моделі є недоопрацьованою. Відносини між класами можуть супроводжуватись умовними залежностями, позначеними через condition або probability.

Діаграма активностей. Дії з невизначеною логікою або умовами виконання позначаються стереотипами <<ambiguousAction>>, з додатковим описом у вигляді коментарів. Альтернативні шляхи, що залежать від ймовірнісних подій, можуть бути анотовані властивістю likelihood.

Діаграма станів. Стан або перехід, що активується лише за невизначених або умовних подій, позначається як <<probabilisticTransition>>. Можна ввести додаткову категорію «невизначений стан» (<<uncertainState>>), який є тимчасовим або залежить від зовнішнього контексту.

Діаграма послідовностей. Для повідомлень, які можуть бути умовними, передбачено властивість expectedProbability або optional = true. Стереотип <<uncertainMessage>> використовується для ідентифікації комунікацій, що є спірними або не завжди ініціюються.

Ці зміни дозволяють будувати моделі, які зберігають зв'язок із реальними вимогами без необхідності штучного уточнення або спрощення. Відповідно, така модель буде краще придатною для автоматизованої обробки, перевірки варіантів і подальшої еволюції вимог.

У таблиці 1 наведено основні стереотипи UML-профілю, призначені для позначення властивостей вимог у статичних моделях. Для динамічних діаграм відповідні стереотипи подано в таблиці 2. Таблиця 3 ілюструє приклади тегованих значень, які дозволяють деталізувати рівень довіри, джерело вимоги, контекстні обмеження, варіативність або неповноту. На рис. 1 представлено етапи узгодження та інтерпретації вимог для UML-проектування, які ілюструють перехід від вимог до формалізованої UML-моделі, придатної для подальшого автоматизованого проектування.

Процес починається з виявлення недосконалих вимог, після чого здійснюється їх класифікація за типами проблем. На основі цієї класифікації формується відповідне розширення базових елементів UML, яке в подальшому інтегрується у метамо-

Таблиця 1

Стереотипи UML-профілю для класів та акторів

Стереотип	Значення	Приклад використання
<<fuzzyRequirement>>	Вимога з нечіткою формалізацією	«Система повинна бути високопродуктивною»
<<incomplete>>	Вимога не містить повного опису	Відсутній тип атрибуту або метод реалізації
<<conflicting>>	Вимога суперечить іншій	«Повідомлення має бути одночасно публічним і приватним»
<<alternative>>	Альтернативний сценарій або залежність	Альтернативна версія бізнес-процесу
<<contextDependent>>	Реалізація залежить від середовища	Поведінка змінюється залежно від місця/часу

Таблиця 2

Стереотипи UML-профілю для динамічних діаграм

Стереотип	Значення	Приклад
<<uncertainState>>	Стан, який активується не завжди	«Очікування підтвердження»
<<probabilisticMessage>>	Повідомлення з ймовірністю доставки	«Дані отримані з ймовірністю 90%»
<<vagueTransition>>	Переходи з умовами, які нечітко формалізовані	«Якщо користувач задоволений»

Таблиця 3

Основні теговані значення, що доповнюють стереотипи

Теговане значення	Тип	Опис
confidence	float [0.1]	Рівень довіри до вимоги або її частини
source	string	Джерело формулювання або надходження вимоги
likelihood	float [0.1]	Ймовірність виконання або реалізації вимоги
context	string	Контекстні умови, які впливають на інтерпретацію
timestamp	datetime	Дата останньої модифікації або валідації вимоги
conflictGroup	string	Ідентифікатор групи суперечливих вимог
resolutionHint	string	Пропозиція щодо розв'язання або уточнення вимоги

дель, що описує структуру доповнених елементів. Розроблена метамодель впроваджується в інструментальне середовище моделювання (наприклад, StarUML), після чого ці розширення можуть бути використані під час проєктування інтелектуальних систем.

На метамодельному рівні кожен стереотип є підкласом відповідного UML-елемента (Class, UseCase, Message, State) і має зв'язок з відповідними тегованими значеннями, що реалізуються як атрибути з визначеним типом (float, string, datetime). Таким чином, модель збагачується додатковими семантичними шарами без зміни основної структури UML.

Для забезпечення обміну UML-моделями з розширеними властивостями застосовуються різні формати збереження, залежно від обраного CASE-засобу. Такі інструменти, як Enterprise Architect, MagicDraw або Papyrus, використовують стандарт XMI (XML Metadata Interchange) для серіалізації моделей, що дозволяє зберігати додаткові анотації (стереотипи, теговані значення, коментарі)

згідно зі специфікацією UML 2.5. Натомість інші інструменти, такі як StarUML, працюють з власним JSON-поданням моделі, у якому метамодель та профілі описуються також у форматі JSON. Завдяки цьому користувачі можуть застосовувати власні розширення для позначення невизначеностей або неповноти без втрати зручності використання інструменту.

Таким чином, запропоноване розширення UML може бути реалізоване як UML-профіль у форматі XMI для інструментів типу Enterprise Architect, MagicDraw, Papyrus тощо чи JSON-розширення для StarUML та подібних середовищ. Це забезпечує підтримку збереження семантики нечітких, змінюваних та неповних вимог незалежно від обраного інструменту моделювання.

Висновки. У статті було обґрунтовано необхідність розширення стандартного базису UML для підтримки недосконалих, нечітких, суперечливих та змінюваних вимог, які характерні для інтелектуальних систем, що проєктуються в умовах високої динаміки та невизначеності. Аналіз

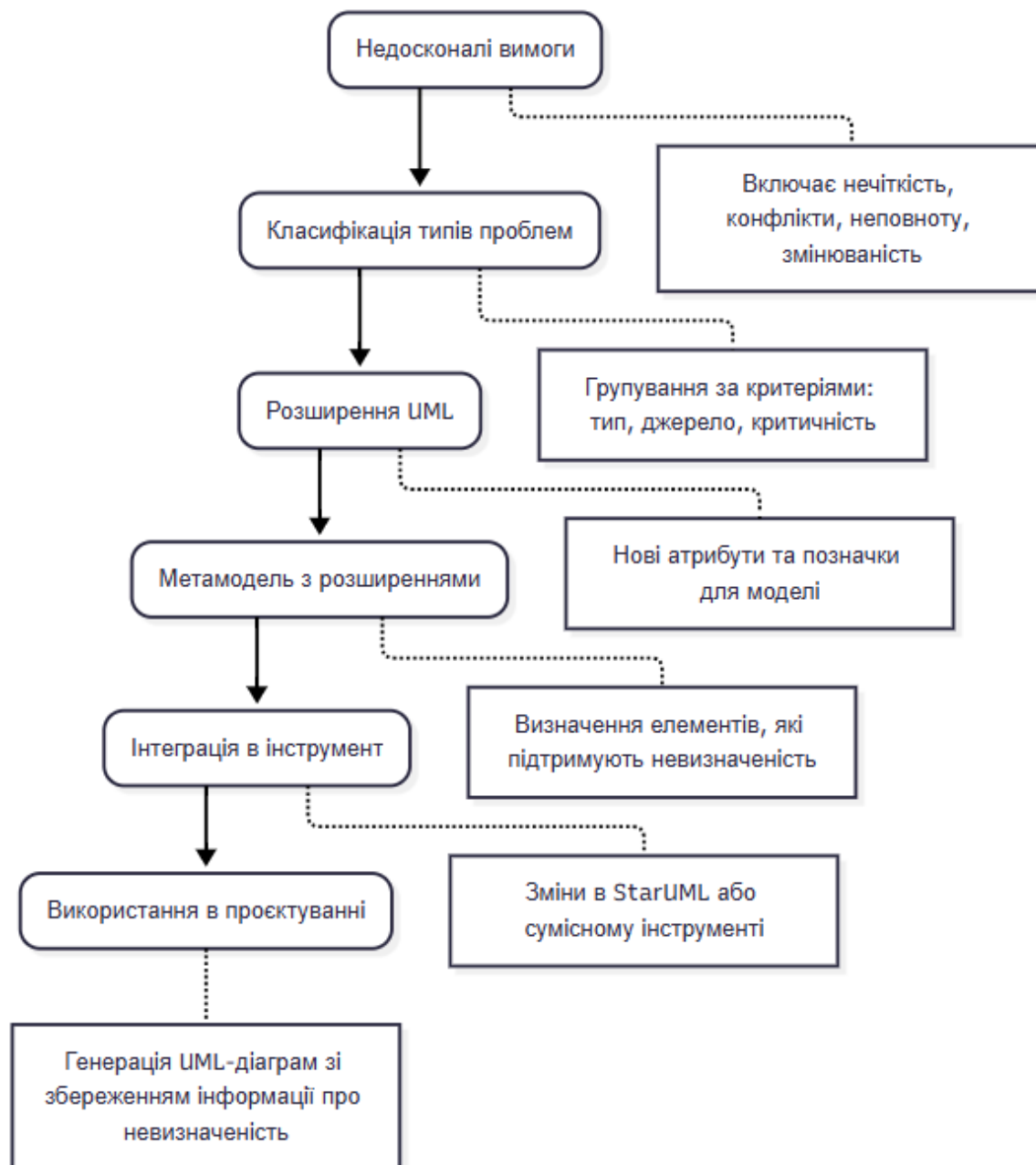


Рис. 1. Етапи узгодження та інтерпретації вимог для UML-проєктування

існуючих підходів продемонстрував обмеженість традиційних засобів моделювання у фіксації таких властивостей вимог, як ступінь довіри, ймовірність реалізації, контекстуальна залежність або неповнота опису.

Запропоновано формальний підхід до розширення UML за допомогою спеціалізованих стереотипів і тегованих значень, що дозволяє явно відображати властивості невизначеності в моделях. Побудовано загальну структуру процесу розширення моделі – від класифікації вимог до інтеграції з CASE-засобами та практичного застосування у проєктуванні.

Описані рішення можуть бути реалізовані у вигляді UML-профілю для популярних інструментів, таких як StarUML або Enterprise Architect, без втрати сумісності з існуючими моделями. Таке розширення забезпечує більш тісну інтеграцію між фазами інженерії вимог і проєктування, підвищуючи точність, адаптивність і пояснюваність розроблюваних систем.

Подальші дослідження будуть присвячені формалізації переходу від вимог із зазначеними характеристиками до відповідних UML-артефактів, а також реалізації прототипу плагіна для CASE-середовищ.

Список літератури:

1. Zhang F., Cheng J. Verification of fuzzy UML models with fuzzy Description Logic // *Applied Soft Computing*. 2018. Vol. 73. P. 134–152. DOI:10.1016/j.asoc.2018.08.025
2. Egesoy A., Güzel A. Fuzzy Logic Support for Requirements Engineering // *International Journal of Innovative Research in Computer Science & Technology (IJIRCST)*. 2021. Vol. 9, No. 2. P. 1–8.
3. Larbi A., Malki M. Uncertain Decision-Making Requirements Formalizing with Complement Fuzzy UML Model // *Procedia Computer Science*. 2022. Vol. 198. P. 317–322. DOI:10.1016/j.procs.2021.12.247
4. Cazzola W., Ghosh S., Al-Refai M. Bridging the model-to-code abstraction gap with fuzzy logic in model-based regression test selection (FLiRTS 2) // *Software & Systems Modeling*. 2021. Vol. 21, Iss. 1. P. 207–224. DOI:10.1007/s10270-021-00899-6
5. Mougouei D., Powers D.M.W. Dependency-Aware Software Requirements Selection using Fuzzy Graphs and Integer Programming // *arXiv*. 2020. Mar 11. P. 1–12.
6. Mougouei D., Powers D.M.W. Modeling and Selection of Interdependent Software Requirements using Fuzzy Graphs // *arXiv*. 2020. Mar 3. P. 1–15.
7. Troya J., Moreno N., Bertoa M. F., et al. Uncertainty representation in software models: a survey // *Software and Systems Modeling*. 2021. Vol. 20. P. 1183–1213. DOI: 10.1007/s10270-020-00842-1.
8. Moreno N., Bertoa M. F., Burgueño L., Vallecillo A. Managing measurement and occurrence uncertainty in complex event processing systems // *IEEE Access*. 2019. Vol. 7. P. 88026–88048. DOI: 10.1109/ACCESS.2019.2923953.
9. Ciccozzi F., Malavolta I., Selic B. Execution of UML models: a systematic review of research and practice // *Software and Systems Modeling*. – 2019. – Vol. 18. – P. 2313–2360. – DOI: 10.1007/s10270-018-0675-4.
10. Burgueño L., Muñoz P., Clarisó R., Gérard S., Vallecillo A. Dealing with belief uncertainty in domain models // *ACM Transactions on Software Engineering and Methodology*. 2022. Vol. 31, Iss. 2. P. 1–29.
11. Mäkelburg J., Perez-Palacin D., Mirandola R., Acosta M. Surveying Uncertainty Representation: A Unified Model for Cyber-Physical Systems // *arXiv*. 2025. Mar 3. P. 1–20.
12. Комлева Н. О. Інжиніринг бізнес-вимог при розробці складних діагностичних систем // *Вчені записки ТНУ імені В. І. Вернадського. Серія: технічні науки*. Київ, 2024. Т. 35 (74), № 1. С. 152–156. DOI: 10.32782/2663-5941/2024.1.1/23.
13. Антошук С. Г., Комлева Н. О. Оптимізаційне моделювання інтелектуальних систем діагностування // *Вісник Херсонського національного технічного університету*. 2024. № 2. С. 99–108. DOI: 10.35546/kntu2078-4481.2024.2.14.
14. Комлева Н. О. Особливості математичного моделювання інтелектуальних систем діагностування і прийняття рішень в умовах невизначеності // *Розвиток наукової думки постіндустріального суспільства: сучасний дискурс* : зб. наук. пр. з матеріалами VII Міжнар. наук. конф., м. Одеса, 25 квіт. 2025 р. / Міжнар. центр наук. дослідж. Вінниця : ТОВ «УКРЛОГОСГруп», 2025. С. 97–100. DOI: 10.62731/mcnd-25.04.2025.001.

Komleva N.O. EXTENDING THE UML FOUNDATION TO SUPPORT IMPERFECT AND EVOLVING REQUIREMENTS IN THE DESIGN OF INTELLIGENT SYSTEMS

The article explores the problem of formalizing imperfect, evolving, and probabilistic requirements in the context of designing intelligent software systems. In real-world conditions, such systems operate in dynamic environments with a high degree of uncertainty, which complicates the application of traditional requirements engineering approaches. The standard UML (Unified Modeling Language) foundation, despite its widespread use in software engineering, does not provide adequate support for vagueness, incompleteness, conflicts, or variability of requirements. This creates an informational gap between the requirements gathering phase and the modeling stage, as significant data is either lost or recorded outside the models, limiting traceability, analysis, and automated processing capabilities.

The paper proposes an extension to the standard UML profile that enables the preservation of the semantics of imperfect requirements. This is achieved by introducing specific stereotypes, tagged values, and constraints, which allow model elements to be marked with indications of trust level, source of information, probability of realization, and contextual conditions. It is shown how these extensions can be integrated into various types of UML diagrams (use case, class, activity, state, sequence) without altering the UML core, while maintaining consistency with the metamodel.

The practical implementation of the approach is described in the form of a specialized UML profile, which can be stored either in the standard XMI format for CASE tools such as Enterprise Architect or MagicDraw, or as a JSON structure for StarUML. The paper presents examples of typical stereotypes (<<fuzzyRequirement>>),

<<incomplete>>, <<conflicting>>, <<uncertainState>>, etc.) as well as tagged attributes that support the extended semantics of requirements.

To illustrate the overall approach, a general diagram of the stages of requirements reconciliation and interpretation for UML design has been constructed, which reflects the logic of transitioning from imperfect requirements to formalized models. The proposed approach is flexible and compatible with UML 2.x, enabling easy scalability and adaptation to the needs of various modeling platforms. It also enhances communication between analysts, developers, and stakeholders, reducing the risk of information loss during the design of intelligent systems. The results obtained have the potential to form a basis for extending CASE tools and developing new methods to support vague and dynamic requirements in high-complexity environments.

Key words: requirements engineering, UML profile, fuzzy requirements, incompleteness, probabilistic models, CASE tools, UML stereotypes, requirements formalization, automated design.

Дата надходження статті: 20.07.2025

Дата прийняття статті: 31.07.2025

Опубліковано: 27.10.2025

Крамар А.В.

ПВНЗ «Буковинський університет»

Кравчук Я.Я.

Velar Voyage LLC USA

Осадчук С.І.

ПВНЗ «Буковинський університет»

АДАПТИВНЕ НАВЧАННЯ ЧАТ-БОТІВ НА ОСНОВІ ДИНАМІЧНОГО АНАЛІЗУ ПОВЕДІНКОВИХ ПАТЕРНІВ КОРИСТУВАЧІВ

У статті розроблено підхід до адаптивного навчання чат-ботів на основі динамічного аналізу поведінкових патернів користувачів. З'ясовано, що застосування механізмів персоналізації взаємодії з користувачами сприяє підвищенню ефективності інформаційного обміну та релевантності автоматизованих відповідей. Досліджено архітектуру інтелектуального агента, здатного змінювати свою поведінку відповідно до історії запитів, темпів комунікації, стилістичних особливостей мовлення та змін у контексті користувацького сеансу. Мета дослідження полягає в розробленні підходу до адаптивного навчання чат-ботів на основі динамічного аналізу поведінкових патернів користувачів, що сприяє підвищенню ефективності персоналізованої взаємодії та адаптації системи до змінних потреб. У дослідженні використано загальнонаукові методи аналізу й синтезу, а також методи моделювання, машинного навчання та аналізу поведінкових даних користувачів для побудови адаптивної системи навчання чат-бота. Розроблена модель охоплює механізм динамічного оновлення навчальних параметрів чат-бота на основі контексту взаємодії та історії поведінки користувача. Доведено, що динамічний аналіз поведінкових патернів дає змогу значно підвищити релевантність відповідей чат-бота та рівень задоволеності користувачів, а також зменшити кількість помилкових інтерпретацій запитів. Запропонований підхід забезпечує гнучкість адаптації на основі короткострокових та довгострокових змін у поведінці користувачів. Дослідження підтвердило ефективність запропонованого підходу до адаптивного навчання чат-ботів, що ґрунтується на інтеграції методів динамічного аналізу поведінкових патернів користувачів. Отримані результати свідчать про можливість застосування цієї моделі в різних сферах, зокрема в освітніх, консультаційних і сервісних системах, для підвищення рівня персоналізації та якості взаємодії. У подальших дослідженнях доцільно зосередитись на розвитку мультиагентних архітектур та поглибленому семантичному аналізі діалогових даних, що сприятиме розширенню функціональних можливостей адаптивних чат-ботів і підвищенню їхньої контекстної інтелектуальності.

Ключові слова: персоналізація, поведінкові патерни, машинне навчання, інтелектуальні агенти, часові ряди, взаємодія, кластеризація, користувацька модель, динамічний аналіз, контекстна адаптація.

Постановка проблеми. Автоматизовані системи комунікації, зокрема чат-боти, широко застосовуються в різних сферах, таких як освіта, сервіс, консалтинг та торгівля. Попри розвиток технологій обробки природної мови та штучного інтелекту (далі – ШІ), залишається актуальною проблема забезпечення високого рівня персоналізації взаємодії між чат-ботом і користувачем. Більшість наявних рішень не враховують індивідуальних поведінкових особливостей корис-

тувачів у режимі реального часу, що призводить до зниження якості комунікації, втрати релевантності відповідей і зниження рівня задоволеності користувачів.

Актуальність дослідження зумовлена потребою підвищення адаптивності інтелектуальних систем у швидкозмінному користувацькому середовищі, а також вимогами до гнучкості та персоналізації автоматизованих агентів. Особливо важливим є врахування поведінкових патернів

користувачів – складних і динамічних характеристик, що формуються під час взаємодії, відображаючи індивідуальні стилі комунікації, інтереси, наміри та контекстні зміни.

Проблематика дослідження полягає в розробленні ефективних методів динамічного аналізу поведінкових патернів для адаптивного навчання чат-ботів. Реалізація цього підходу сприятиме теоретичному розвитку інтелектуальних систем і створенню практичних рішень із підвищеною персоналізацією, гнучкістю та інтелектуальністю автоматизованих комунікацій. Результати можуть істотно вплинути на вдосконалення цифрових сервісів, підвищення якості обслуговування та ефективності освітніх платформ, відповідаючи сучасним технологічним і соціокультурним викликам.

Аналіз останніх досліджень і публікацій. Проблематика адаптивного навчання чат-ботів на основі динамічного аналізу поведінкових патернів користувачів охоплює кілька важливих напрямів: класифікація чат-ботів, методи машинного навчання для адаптації, особливості інтеграції поведінкових патернів у системи, а також практичне застосування в різних сферах. Особливості використання чат-ботів в освітньому процесі закладів вищої освіти розглядають О. Наливайко та А. Малютіна [1], підкреслюючи важливість адаптивності та персоналізації. Застосування нейромереж для дослідження адаптивних освітніх траєкторій, що є актуальним для розвитку чат-ботів із гнучкою реакцією на поведінкові зміни, аналізують М. Пікуляк, І. Савка та М. Дутчак [2]. Використання нейромереж у розпізнаванні та діагностиці навчальних здібностей, що може бути застосовано для підвищення ефективності адаптивних систем, досліджують С. Буртовий, В. Чубенко та С. Василюк-Зайцева [3]. На потенціалі штучного інтелекту для створення персоналізованих комунікацій, що є важливим аспектом для адаптивних чат-ботів, акцентують Є. Венгер, Г. Ковбас та Р. Чугаєвський [4]. Удосконалення алгоритмів адаптивних навчальних тестів, що може бути інтегровано в системи динамічного аналізу поведінкових патернів, пропонує Є. Дерещ [5]. Тенденції використання нейромереж у розробленні вебсайтів, що має значення для інтерфейсної частини чат-ботів, аналізують О. Бережна та П. Василькевич [6]. Чат-боти як стратегічний інноваційний напрям розвитку бізнесу, що підкреслює важливість адаптивності в комунікаціях, розглядає О. Небилиця [7]. Таксономію чат-ботів надають О. Трофименко зі співавторами (О. Trofymenko et al.) [8]. Науковці А. Вішвакарма

(А. Vishwakarma) та А. Пандей (А. Pandey) [9] здійснюють огляд і порівняльний аналіз різних дизайнів чат-ботів, виокремлюючи ефективні методи адаптації. Класифікацію й техніки розробки розмовних агентів, зокрема адаптивних чат-ботів, досліджують С. Хуссейн та колеги (S. Hussain et al.) [10].

Загалом, аналіз наукових джерел підтверджує важливість комплексного підходу до адаптивного навчання чат-ботів з урахуванням поведінкових патернів користувачів, інтеграції сучасних методів машинного навчання та нейромереж, що забезпечує високу якість комунікації й персоналізацію взаємодії.

Виділення невирішених раніше частин загальної проблеми. Попри значні успіхи в розвитку методів машинного навчання та ШІ для чат-ботів, низка важливих аспектів адаптивної взаємодії залишаються недостатньо дослідженими. Зокрема, більшість наявних підходів фокусується на статичному аналізі поведінкових даних або застосовує фіксовані моделі навчання, що не дає змоги враховувати динамічні зміни в поведінкових патернах користувачів під час тривалих сеансів взаємодії. Це призводить до обмеженої здатності чат-ботів адаптуватися до змін контексту, темпу комунікації та емоційного стану співрозмовника.

Також недостатньо уваги приділяється інтеграції часових рядів і кластеризації для комплексного аналізу поведінкових характеристик у режимі реального часу. Методологічні обмеження попередніх досліджень полягають у використанні статичних або неповних емпіричних даних, що знижує точність персоналізації та адаптивності систем. Крім того, недостатньо ефективних механізмів автоматичного оновлення навчальних параметрів чат-ботів, які могли б оперативнo реагувати на нові поведінкові патерни та зміни в користувацьких сценаріях.

Ці прогалини істотно впливають на якість взаємодії, обмежують потенціал інтелектуальних агентів і перешкоджають упровадженню гнучких та контекстно чутливих систем. Подолання зазначених обмежень є важливим не лише для теоретичного розвитку адаптивних технологій, але й для практичного підвищення ефективності сервісів у таких сферах, як освіта, консалтинг, підтримка клієнтів.

Наше дослідження спрямоване на усунення цих недоліків шляхом розроблення моделі адаптивного навчання чат-ботів із застосуванням динамічного аналізу поведінкових патернів на основі часових рядів і кластеризації. Застосу-

вання нових методів сприятиме підвищенню гнучкості та точності персоналізації, контекстній адаптації в реальному часі та зниженню ймовірності помилкових інтерпретацій користувацьких запитів.

Постановка завдання. Метою статті є розроблення ефективного підходу до адаптивного навчання чат-ботів на основі динамічного аналізу поведінкових патернів користувачів, що сприятиме підвищенню якості персоналізованої взаємодії та адаптації системи до змінних потреб користувачів у режимі реального часу.

Для досягнення мети сформульовано такі завдання:

1) вивчити сучасні методи аналізу поведінкових патернів та їхнє застосування в контексті адаптивних чат-ботів, зокрема використання часових рядів і кластеризації для моделювання користувацької поведінки;

2) розробити архітектуру інтелектуального агента, здатного динамічно оновлювати навчальні параметри на основі аналізу історії взаємодії, темпу комунікації, стилістичних особливостей та контексту користувацького сеансу;

3) запропонувати методи інтеграції динамічного аналізу поведінкових патернів у процес адаптивного навчання чат-ботів, що забезпечать високу релевантність відповідей та гнучкість адаптації до короткострокових і довгострокових змін у поведінці користувачів.

Виклад основного матеріалу. Упродовж останніх років чат-боти отримали значне поширення, активно розвиваючись у різних галузях, таких як маркетинг, освіта, охорона здоров'я, культурна спадщина та розваги. Їхнє впровадження автоматизує багато завдань, які раніше виконувалися виключно людьми, зокрема функції онлайн-агентів з обслуговування клієнтів та викладачів. Такий розвиток сприяв підвищенню ефективності та доступності послуг у різних сферах.

Історично чат-боти еволюціонували від простих систем, що працювали на основі жорстко встановлених правил, до сучасних інтелектуальних агентів, заснованих на методах штучного інтелекту. Сучасні чат-боти характеризуються значним покращенням продуктивності, здатністю підтримувати природні розмови, а також можливістю навчатися на основі отриманого досвіду. Це дає їм змогу гнучко й адекватно реагувати на запити користувачів.

Чат-боти набувають дедалі більшого значення в повсякденному житті сучасної людини. Це спеціальні програми, призначені для автоматизації

спілкування та надання допомоги в різних сферах діяльності [7]. Сьогодні чат-боти активно застосовуються в банківській сфері, де вони допомагають виконувати базові фінансові операції, зокрема перевірку балансу, переказ коштів або надання інформації про поточні тарифи. Також їх застосовують у сфері охорони здоров'я, наприклад, для попереднього аналізу симптомів, надання медичних порад або запису на прийом до лікаря. Завдяки здатності працювати цілодобово та оперативно обробляти запити чат-боти значно покращують доступ до сервісів і підвищують якість обслуговування в багатьох галузях [11, с. 4].

Чат-бот (англ. chatbot) – це програмне забезпечення, що імітує реальну розмову з користувачем. Взаємодія з чат-ботами відбувається через текстові або аудіо повідомлення на різноманітних платформах – вебсайтах, месенджерах, мобільних застосунках або під час телефонних дзвінків. Завдяки цьому чат-боти забезпечують зручний і доступний інтерфейс для комунікації між людиною та комп'ютером [7].

Основою функціонування сучасних чат-ботів є застосування методів машинного навчання, які дають змогу виявляти закономірності в поведінці користувачів і моделювати їхні патерни спілкування. Постійна взаємодія з користувачами дає змогу чат-ботам навчатися й наслідувати особливості живої розмови, що підвищує якість і релевантність відповідей. Важливо, що застосування ШІ дає змогу чат-ботам не просто виконувати встановлені команди, а розуміти природну мову, що робить спілкування більш інтуїтивним і ефективним.

Поведінкові патерни користувачів – це типові моделі взаємодії, що проявляються під час їхнього спілкування з чат-ботами. Їхній аналіз допомагає розпізнавати найпоширеніші запити, стилі формулювання питань, а також послідовність дій користувачів. Завдяки виявленню й розумінню патернів чат-боти можуть адаптувати свої відповіді, підвищувати точність реакцій та створювати персоналізований досвід взаємодії. Це сприяє не лише покращенню якості обслуговування, а й ефективнішому навчанню моделей ШІ на основі реальної поведінки користувачів.

Сучасні методи аналізу поведінкових патернів користувачів у контексті адаптивних чат-ботів ґрунтуються на комплексному вивченні динамічних характеристик взаємодії, що дає можливість розкривати закономірності поведінки в часі та просторі параметрів комунікації. Зокрема, застосування методів аналізу часових рядів є важли-

вим інструментом для моделювання послідовностей подій, що відображають різні аспекти користувацької активності, зокрема частоту запитів, темп відповіді, зміни тематики та емоційного забарвлення повідомлень. При цьому широко використовуються алгоритми авторегресійних моделей (далі – AR), моделей ковзного середнього (далі – MA), а також гібридні ARIMA (AutoRegressive Integrated Moving Average), які сприяють виявленню трендів та сезонних коливань у поведінці користувачів [10].

До того ж кластеризація є ефективним засобом для групування схожих патернів поведінки серед великої множини користувачів. Застосування таких алгоритмів, як K-середніх (K-Means), ієрархічної кластеризації (Hierarchical Clustering) та методів на основі щільності, зокрема DBSCAN (Density-Based Spatial Clustering of Applications with Noise), дає змогу виокремити однорідні групи користувачів за спільними характеристиками поведінки, що спрощує подальше персоналізоване налаштування чат-бота. Ці алгоритми сприяють виявленню як основних кластерів, так і аномальних поведінкових патернів, що важливо для налаштування системи на нетипові сценарії взаємодії.

Інтеграція аналізу часових рядів із кластеризаційними підходами створює синергетичний ефект, завдяки якому можливо реалізувати багаторівневу модель поведінкової динаміки. На основі кластеризації часових рядів застосовують алгоритми типу Dynamic Time Warping (DTW) у поєднанні з K-Nearest Neighbors (KNN), що дає можливість класифікувати часові послідовності з урахуванням їхньої форми й тривалості. Такий підхід забезпечує підвищення точності прогнозування поведінки та гнучкість у формуванні адаптивних відповідей чат-бота.

Крім того, для виявлення складніших залежностей у поведінкових даних застосовуються методи глибокого навчання. Рекурентні нейронні мережі (RNN), зокрема варіації на основі довгої й короткочасної пам'яті (LSTM), здатні обробляти послідовні дані, враховуючи тривалу пам'ять про попередні стани. Ці моделі ефективні для прогнозування подальших дій користувачів та адаптації діалогів чат-ботів до їхніх індивідуальних особливостей. Інші підходи, такі як Transformer-моделі, також набирають популярності для аналізу складних послідовностей та контекстної адаптації.

Для досягнення ефективної адаптації чат-ботів до індивідуальних особливостей користувачів

розроблено комплексну архітектуру інтелектуального агента, що ґрунтується на динамічному аналізі поведінкових патернів. Центральним елементом системи є основний чинник адаптації, який координує взаємодію восьми основних напрямів, кожен із яких виконує специфічні функції в зборі, аналізі та обробці інформації про поведінку користувача. Така модульна структура забезпечує гнучкість і масштабованість системи, даючи змогу оперативно реагувати на зміни в контексті користувацьких сесій, стилістиці мовлення, темпі комунікації та інших характеристиках. Взаємозв'язок між основним чинником та підсистемами, які спільно формують адаптивну модель навчання чат-бота, представлено на рис. 1.

На основі представленої архітектури інтегруються різні методи машинного навчання, аналізу часових рядів, кластеризації поведінкових даних і формалізації користувацьких сценаріїв. Такий комплексний підхід забезпечує динамічне оновлення навчальних параметрів системи в реальному часі з урахуванням накопиченої інформації про поведінкові патерни користувача. Упровадження цієї архітектури дає змогу значно підвищити точність та релевантність відповідей чат-бота, зменшити кількість помилкових інтерпретацій запитів та покращити загальний рівень задоволеності користувачів. Запропонована структура є надійною основою для створення високоефективних адаптивних чат-ботів із контекстною інтелектуальністю та персоналізованим підходом до взаємодії.

Першим етапом у цій архітектурі є збір та обробка даних (Data Acquisition and Preprocessing), що передбачає безперервний моніторинг та накопичення інформації про взаємодію користувача з чат-ботом. Зібрані дані – текстові запити, часові мітки, метадані сесії та інші релевантні параметри – проходять процедуру нормалізації, токенизації, лематизації та видалення шуму. Ці кроки стандартизують вхідні дані й забезпечують їхню високу якість для подальшого аналітичного опрацювання.

Наступний важливий напрям – лінгвістичний аналіз (Linguistic Analysis) – спрямований на визначення стилістичних, семантичних та емоційних характеристик користувацьких повідомлень. Методики синтаксичного розбору, аналізу тональності й семантичного тегування дають змогу глибше зрозуміти індивідуальні мовні особливості користувача, що сприяє підвищенню релевантності відповідей і формуванню персоналізованого досвіду взаємодії.



Рис. 1. Архітектура інтелектуального агента з основним чинником та напрямками

Джерело: власна розробка авторів

У межах аналізу темпу комунікації (Communication Tempo Analysis) здійснюється моніторинг часових інтервалів між повідомленнями користувача та чат-бота. Це дає можливість моделювати інтенсивність і динаміку спілкування, визначати рівень залученості та емоційний стан користувача, що є критично важливим для адаптивного налаштування швидкості й змісту відповідей системи [12, с. 102].

Для ідентифікації типових та атипових моделей поведінки застосовується кластеризація поведінкових патернів (Behavioral Pattern Clustering) за допомогою алгоритмів, таких як K-means або DBSCAN. Це дає змогу групувати користувачів за схожими характеристиками взаємодії, що є основою для цілеспрямованої адаптації стратегії комунікації до специфіки кожного кластера.

Методи аналізу часових рядів (Time Series Analysis), зокрема рекурентні нейронні мережі (RNN) та авторегресійні моделі (ARIMA), застосовуються для моделювання динамічних змін у поведінці користувачів у часовому вимірі. Такий підхід дає змогу прогнозувати майбутню поведінку, виявляти тренди й своє-

часно реагувати на коливання в користувацьких паттернах.

На основі отриманих даних формується користувацький профіль (User Profile Construction), який є динамічною моделлю індивідуальних характеристик, уподобань, стилю та частоти взаємодії. Він є основою для персоналізації відповідей чат-бота, а також враховує довгострокові зміни поведінки користувача.

Важливим компонентом є динамічне оновлення навчальних параметрів (Dynamic Learning Parameter Update), що забезпечує адаптацію моделей машинного навчання в режимі реального часу. Завдяки використанню даних із профілю користувача та поведінкових моделей ваги нейронних мереж або інших алгоритмічних параметрів коригуються без необхідності повного перенавчання, що підвищує швидкість та ефективність адаптації системи.

Сервісний рівень інтеграції (Service Integration Layer) координує роботу всіх компонентів архітектури, забезпечує управління сесіями, збереження історії взаємодії та оптимізує використання обчислювальних ресурсів. Цей рівень підтримує

безперебійну роботу системи, її масштабованість та адаптивність до нових функціональних можливостей.

Останніми роками застосунки на основі ІІІ масштабно впроваджуються в робоче середовище. На відміну від традиційних корпоративних систем чат-боти на базі ІІІ постійно навчаються, аналізуючи шаблони використання та способи взаємодії користувачів [13, с. 9].

Отже, інтеграція описаних напрямів у єдину архітектуру створює комплексну та адаптивну систему навчання чат-ботів. Вона максимально враховує індивідуальні характеристики користувачів та контекст їхньої взаємодії, що сприяє підвищенню ефективності та якості комунікації.

Для забезпечення високої релевантності відповідей та гнучкості адаптації чат-ботів до змін у поведінці користувачів важливо інтегрувати методи динамічного аналізу поведінкових патернів у процес адаптивного навчання. Основна ідея полягає в застосуванні поєднання алгоритмів машинного навчання, аналізу часових рядів і кластеризації, що дає змогу системі оперативно враховувати як короткострокові, так і довгострокові зміни в користувацькій активності та стилі спілкування.

Одним з основних підходів є використання модульної структури, де окремі компоненти відповідають за збір та аналіз різних типів поведінкових даних, а центральний модуль приймає рішення щодо оновлення навчальних параметрів чат-бота. Така архітектура сприяє швидкій адаптації системи без необхідності повного перенавчання моделей [14].

Зокрема, методи кластеризації дають можливість виокремлювати типові поведінкові патерни та автоматично коригувати стратегії взаємодії

для різних груп користувачів. Аналіз часових рядів сприяє прогнозуванню змін у поведінці, що важливо для прогнозування майбутніх потреб та налаштувань системи.

Для досягнення балансу між адаптивністю й стабільністю навчання застосовуються алгоритми з оновленням параметрів у режимі реального часу, такі як адаптивні градієнтні методи та методи онлайн-навчання. Це дає змогу ефективно реагувати на швидкоплинні зміни, зберігаючи при цьому узагальнену модель поведінки користувачів.

З метою систематизації й візуалізації основних методів інтеграції динамічного аналізу поведінкових патернів у процес адаптивного навчання чат-ботів у таблиці 1 узагальнено основні підходи, їхні переваги та обмеження.

Аналіз розглянутих методів свідчить про унікальні переваги кожного з них, проте також виявляє певні обмеження та виклики, пов'язані з реалізацією в умовах динамічної взаємодії користувача з чат-ботом. Для досягнення високого рівня адаптивності системи до змін у поведінкових патернах користувачів доцільно застосовувати комбінований підхід, що поєднує переваги різних методик. Зокрема, інтеграція алгоритмів кластеризації (для ефективного групування схожих поведінкових патернів) з методами аналізу часових рядів (для моделювання та прогнозування динаміки змін) сприяє глибшому розумінню поведінки користувачів.

Водночас застосування гнучких механізмів онлайн-навчання дає змогу оперативно коригувати модель у режимі реального часу, що забезпечує швидку адаптацію чат-бота до короткострокових і довгострокових трансформацій у поведінці користувача. Такий інтегрований підхід підви-

Таблиця 1

Методи інтеграції динамічного аналізу поведінкових патернів в адаптивне навчання чат-ботів

Метод інтеграції	Характеристика	Переваги	Недоліки
Кластеризація поведінкових патернів	Групування схожих поведінкових профілів для персоналізації	Точне сегментування, підвищення релевантності	Потребує значних обчислювальних ресурсів
Аналіз часових рядів	Вивчення динаміки змін поведінки в часі	Прогнозування майбутніх патернів	Модель може бути чутливою до шуму
Онлайн-навчання	Постійне оновлення моделі на основі нових даних	Гнучкість і швидка адаптація	Можливе перенавчання при частих змінах
Адаптивні градієнтні методи	Регулювання швидкості навчання залежно від поведінки	Підвищення стабільності та ефективності	Складність налаштування гіперпараметрів
Модульна архітектура	Розподіл функцій між окремими компонентами системи	Масштабованість і легкість оновлення	Необхідність координації між модулями

Джерело: власна розробка авторів

щує загальну точність, релевантність відповідей і якість взаємодії, одночасно забезпечуючи стійкість системи до змін та непередбачуваних поведінкових варіацій [15].

Рекомендації щодо подальшого розвитку та вдосконалення чат-ботів з урахуванням сучасних технологічних тенденцій охоплюють кілька основних напрямів. Одним із пріоритетних є розширення застосування методів глибинного навчання, зокрема рекурентних та трансформерних нейронних мереж, що сприяє ефективному моделюванню складних послідовностей користувацьких запитів і контексту взаємодії. Такий підхід підвищує якість генерації відповідей і створює умови для більш природної комунікації.

Варто акцентувати на розвитку багатомодальних систем, які інтегрують текстові, аудіо- та візуальні дані, забезпечуючи повне розуміння намірів користувачів. Комплексний аналіз різномірних даних покращує інтерпретацію емоційного стану, контексту й стилістичних особливостей спілкування, що сприяє підвищенню рівня персоналізації та адаптивності чат-ботів.

Особливе значення має впровадження механізмів постійного онлайн-навчання та самонавчання, які забезпечують оперативне оновлення моделей на основі нових даних без необхідності повного перенавчання системи. Такий підхід сприяє швидкому реагуванню систем на зміни в поведінці користувачів, зберігаючи актуальність та точність відповідей.

З огляду на актуальність питання захисту даних, важливо впроваджувати ефективні протоколи безпеки та конфіденційності, особливо при обробці персональної інформації в адаптивних чат-ботах. Засоби шифрування та анонімізації сприяють мінімізації ризиків витоку інформації, що підвищує довіру користувачів до системи.

Крім того, варто активно залучати кінцевих користувачів у процес розроблення, тестування й удосконалення чат-ботів. Це сприяє отриманню зворотного зв'язку, необхідного для оптимальної адаптації функціоналу системи до реальних

потреб аудиторії. Такий підхід позитивно впливає на якість взаємодії та підвищує рівень задоволеності користувачів.

Отже, інтеграція сучасних технологій у поєднанні з урахуванням соціокультурних аспектів взаємодії створює фундамент для розроблення ефективних, надійних і високоякісних адаптивних чат-ботів, що відповідають вимогам сучасного цифрового середовища.

Висновки. Здійснений нами аналіз дозволив сформулювати комплексний підхід до адаптивного навчання чат-ботів, що базується на динамічному аналізі поведінкових патернів користувачів. Запропонована архітектура інтелектуального агента забезпечує гнучке оновлення навчальних параметрів у режимі реального часу з урахуванням історії взаємодії, темпу комунікації, стилістичних особливостей та контексту користувацького сеансу. Застосування методів машинного навчання, аналізу часових рядів та кластеризації поведінкових даних дає змогу підвищити точність і релевантність відповідей, що сприяє зниженню помилкових інтерпретацій та покращенню користувацького досвіду.

Результати підтвердили ефективність запропонованого підходу в умовах реальної взаємодії, демонструючи підвищення рівня задоволеності користувачів та якості комунікації. Запропоновані методи інтеграції динамічного аналізу поведінкових патернів сприяють гнучкій адаптації до короткострокових і довгострокових змін у поведінці користувачів, що є важливим чинником для створення інтелектуальних чат-ботів нового покоління.

Отже, розроблена модель може бути успішно застосована в різних сферах, зокрема освітніх, консультаційних та сервісних системах, де важлива персоналізація взаємодії та висока якість інформаційного обміну. Подальші дослідження доцільно спрямувати на розвиток мультиагентних архітектур та впровадження поглибленого семантичного аналізу діалогових даних для розширення функціональних можливостей адаптивних систем.

Список літератури:

1. Наливайко О. Перспективи використання нейромереж у вищій освіті України. *Інформаційні технології і засоби навчання*. 2023. № 97(5). С. 1–17. DOI: <https://doi.org/10.33407/itlt.v97i5.5322>
2. Пікуляк М., Савка І., Дутчак М. Використання апарату нейромереж для дослідження адаптивної навчальної траєкторії. *Комп'ютерно-інтегровані технології: освіта, наука, виробництво*. 2022. № 47. С. 91–97. DOI: <https://doi.org/10.36910/6775-2524-0560-2022-47-14>.
3. Буртовий С. В., Чубенко В. А., Василюк-Зайцева С. В. Використання нейромереж у розпізнаванні та діагностиці навчальних здібностей здобувачів освіти. *Академічні візії*. 2024. № 29. URL: <https://academy-vision.org/index.php/av/article/view/971> (дата звернення: 06.06.2025).

4. Венгер Є., Ковбас Г., Чугаєвський Р. Потенціал семантичного маркетингу та AI у створенні персоналізованих комунікацій. *Наукові записки Львівського університету бізнесу та права*. 2025. № 44. С. 391–398. URL: <https://nzlubp.org.ua/index.php/journal/article/view/1624> (дата звернення: 06.06.2025).
5. Дерещ Є. Вдосконалення алгоритмів адаптивних навчальних тестів з вищої та прикладної математики. *Професійна освіта: 36. наук. праць Дніпровського державного технічного університету (технічні науки)*. 2019. Т. 1 № 34. С. 177–181. DOI: <https://doi.org/10.31319/2519-2884.34.2019.33>
6. Бережна О., Васькевич П. Тенденції використання нейромереж у розробці веб-сайтів. *Поліграфічні, мультимедійні та web-технології*: тези доп. IX Міжнар. наук.-техн. конф. (м. Харків, 14–18 травня 2024 р.). Т. 1. Харків: ТОВ «Друкарня Мадрид», 2024. С. 67–68. URI: <https://openarchive.nure.ua/handle/document/26691> (дата звернення: 06.06.2025).
7. Небилиця О. А. Чат-боти як ефективний стратегічний та інноваційний напрям розвитку бізнесу. *Сучасні проблеми управління підприємствами: теорія та практика*: Матеріали міжнародної науково-практичної конференції (м. Харків – м. Торунь, 16 квітня 2021 року). Дніпро, 2021. С. 99–102. URL: https://kmib.hneu.edu.ua/wp-content/uploads/2021/04/zbirnik_tez_konferencii_suchasni-problemi-upravlinnja-japripriemstvami_harkiv_16-kvitnja_2021.pdf (дата звернення: 09.06.2025).
8. Trofymenko O., Prokop Y., Loginova N., & Zadereyko A. Taxonomy of Chatbots. *Intellectual Systems and Information Technologies 2021: II International Scientific and Practical Conference* (Odesa, September 13–19, 2021). Odesa, 2021. P. 181–185. Retrieved from <https://ceur-ws.org/Vol-3126/paper24.pdf> (date of access: 10.06.2025)
9. Vishwakarma A., Pandey A. A Review & Comparative Analysis on Various Chatbots Design. *International Journal of Computer Science and Mobile Computing*. 2021. Vol. 10. № 2. P. 72–78. DOI: <https://doi.org/10.47760/ijcsmc.2021.v10i02.011>
10. Hussain S., Ameri Sianaki O., Ababneh N. A Survey on Conversational Agents/Chatbots Classification and Design Techniques. *Web, Artificial Intelligence and Network Applications: Proceedings of the Workshops of the 33rd International Conference on Advanced Information Networking and Applications (WAINA-2019)* (Japan, 27–29 March 2019). Cham: Springer International Publishing, 2019. P. 946–956. DOI: 10.1007/978-3-030-15035-8_93
11. Bulla Ch., Parushetti Ch., Teli A., Aski S., Koppad, S. A Review of AI-Based Medical Assistant Chatbot. *Research and Applications of Web Development and Design*. 2020. Vol. 3. № 2. P. 1–14. DOI: <https://doi.org/10.5281/zenodo.3902215>
12. Meyer von Wolff R., Hobert S., Schumann M. How May I Help You? – State of the Art and Open Research Questions for Chatbots at the Digital Workplace. *Proceedings of the 52nd Hawaii International Conference on System Sciences*. P. 95–104. DOI: <https://doi.org/10.24251/HICSS.2019.013>
13. Gkinko L., Elbanna A. (2021). Chatbots at Work: A Taxonomy of the Use of Chatbots in the Workplace. *Responsible AI and Analytics for an Ethical and Inclusive Digitized Society: I3E: 20th Conference on e-Business, e-Services and e-Society* (Galway, Ireland, September 2021). Galway, 2021. DOI: https://doi.org/10.1007/978-3-030-85447-8_3
14. Nalyvaiko O., Maliutina A. Use of Chat Bots in the Educational Process of a Higher Education Institution. *Scientific Notes of the Pedagogical Department*. 2021. № 48. P. 117–122. DOI: <https://doi.org/10.26565/2074-8167-2021-48-14>
15. Ушакова І. О. Підходи до створення інтелектуальних чат-ботів. Системи обробки інформації. 2019. Вип. 2 (157). С. 76–83. URL: http://nbuv.gov.ua/UJRN/soi_2019_2_12 (дата звернення: 09.06.2025).

Kramar A.V., Kravchuk Ya.Ya., Osadchuk S.I. ADAPTIVE LEARNING OF CHATBOTS BASED ON DYNAMIC ANALYSIS OF USER BEHAVIORAL PATTERNS

The article develops an approach to adaptive training of chatbots based on dynamic analysis of user behavior patterns. It has been found that the use of mechanisms for personalizing interaction with users contributes to improving the efficiency of information exchange and the relevance of automated responses. The architecture of an intelligent agent capable of changing its behavior according to the history of requests, communication rates, stylistic features of speech, and changes in the context of the user session has been investigated. The goal of the study is to develop an approach to adaptive chatbot training based on dynamic analysis of user behavior patterns, which helps make personalized interactions more effective and helps the system adapt to changing needs. The study uses general scientific methods of analysis and synthesis, as well as methods of modeling, machine learning, and analysis of user behavioral data to build an adaptive chatbot training system. The developed model covers the mechanism of dynamic updating of chatbot training parameters based on the context of interaction and user behavior history. It has been proven that dynamic analysis of behavioral patterns makes it possible to significantly increase the relevance of chatbot responses

and user satisfaction, as well as reduce the number of misinterpretations of queries. The proposed approach provides flexibility of adaptation based on short-term and long-term changes in user behavior. The study confirmed the effectiveness of the proposed approach to adaptive chatbot training based on the integration of methods for dynamic analysis of user behavior patterns. The results obtained indicate the possibility of applying this model in various fields, in particular in educational, consulting, and service systems, to increase the level of personalization and quality of interaction. Further research should focus on the development of multi-agent architectures and in-depth semantic analysis of dialogue data, which will contribute to expanding the functional capabilities of adaptive chatbots and increasing their contextual intelligence.

Key words: *personalization, behavioral patterns, machine learning, intelligent agents, time series, interaction, clustering, user modeling, dynamic analysis, contextual adaptation.*

Дата надходження статті: 20.06.2025

Дата прийняття статті: 01.07.2025

Опубліковано: 27.10.2025

Кримська А.О.

Державний торговельно-економічний університет

Дідок В.Ю.

Київський національний університет будівництва і архітектури

Столярчук В.В.

Дніпровський національний університет імені Олеся Гончара

ОПТИМІЗАЦІЯ ЕНЕРГОЕФЕКТИВНОСТІ В ІНЖЕНЕРНИХ СИСТЕМАХ ПРОМИСЛОВИХ ОБ'ЄКТІВ

Актуальність дослідження зумовлено все більшою потребою в скороченні енергетичних витрат у промисловості на фоні зношеності інфраструктури, зростання тарифів, жорсткіших екологічних вимог і необхідності забезпечення енергетичної стабільності виробничих процесів. Енергоефективність інженерних систем розглядається як стратегічний ресурс підприємства, що впливає на конкурентоспроможність, операційну стійкість і довгострокову економічну ефективність.

Мета дослідження полягає в комплексному аналізі функціонування інженерних систем промислових об'єктів з позиції енергоефективності, в оцінюванні можливостей технічної та цифрової модернізації їх з урахуванням економічної доцільності, а також у формуванні практичних рекомендацій для зниження енергоспоживання в умовах виробничого середовища.

Методологія дослідження ґрунтується на поєднанні системного аналізу архітектури інженерних систем, узагальнення підходів до енергомоніторингу й моделювання енерговитрат, а також на критичному аналізі нормативної бази, потенціалу автоматизованих технологій управління та фінансових моделей модернізації. Застосовано елементи порівняльного, факторного і структурно-функціонального аналізу.

Результати дослідження. Вивчено принципи енергозбереження, що реалізуються в сучасних системах підприємств, зокрема балансування навантаження, рекуперацію, адаптивне керування та цифрову автоматизацію. Узагальнено підходи до оцінювання енергоефективності на основі нормативного, індикативного, коефіцієнтного й автоматизованого аналізу. Виявлено ефективність інтелектуальних технологій моніторингу – SCADA, IoT, AI-платформ – з метою переходу до превентивного управління енергоспоживанням. З'ясовано ключові причини, що стримують модернізацію: застарілість обладнання, обмежений доступ до фінансування, низький рівень енергоменеджменту й невідповідність регуляторного середовища.

Висновки. Доведено, що підвищення енергоефективності можливе лише за умови цілісного підходу, який включає технічну реконструкцію, упровадження адаптивного управління й енергоменеджменту відповідно до міжнародних стандартів. Підтверджено ефективність інтелектуалізації енергетичних систем для досягнення стабільного енергетичного балансу.

Перспективи подальших досліджень полягають у розробленні цифрових моделей адаптивного управління енергоспоживанням, інтеграції інструментів машинного навчання для прогнозування навантажень, а також у вдосконаленні нормативно-методичних засад енергоефективної політики в промисловості.

Ключові слова: енергоменеджмент, інтелектуальні технології моніторингу, автоматизоване керування, технічна модернізація, енергетичний аудит, цифрова аналітика, нормативна база.

Постановка проблеми. Забезпечення високого рівня енергоефективності в інженерних системах промислових об'єктів є однією з ключових передумов сталого розвитку виробничого сектора

в умовах усе жорсткіших вимог до раціонального використання ресурсів і зниження негативного впливу на довкілля. У сучасних умовах функціонування промисловості проблема енергоспожи-

вання набула багатогранного характеру, що охоплює як технічні, так й організаційно-економічні аспекти. Необхідність скорочення енергетичних витрат без зниження продуктивності й технологічної надійності зумовлює підвищення уваги до методів оптимізації енергетичних потоків, модернізації інженерних мереж і впровадження інтелектуальних систем управління споживанням. Проблематика ускладнюється значною зношеністю основних фондів, низькою адаптованістю застарілих технічних рішень до новітніх стандартів енергоефективності, а також обмеженими фінансовими ресурсами підприємств. У зв'язку з цим особливої актуальності набуває пошук науково обґрунтованих підходів до енергоефективної реконструкції систем опалення, вентиляції, кондиціонування, водопостачання, електропостачання й автоматизованого управління. З практичного погляду розв'язання цієї проблеми передбачає досягнення балансу між капіталовкладеннями й очікуваним енергетичним ефектом, що потребує застосування методів техніко-економічного моделювання, аналізу життєвого циклу інженерних систем та інтеграції принципів енергоаудиту. У науковому контексті завдання оптимізації енергоспоживання охоплює дослідження моделей розподілу навантажень, впливу зовнішніх факторів на стабільність енергетичного балансу, а також розроблення адаптивних алгоритмів керування з урахуванням специфіки конкретних промислових середовищ. Отже, проблема оптимізації енергоефективності в інженерних системах промислових об'єктів інтегрує теоретичні й прикладні завдання, реалізація яких є критично важливою як для науково-технічного поступу, так і для підвищення конкурентоспроможності підприємств у контексті переходу до енергоощадної економіки.

Аналіз останніх досліджень і публікацій.

Аналіз сучасних досліджень з оптимізації енергоефективності в інженерних системах промислових об'єктів дає змогу виокремити чотири основні змістові напрями.

Перший напрям передбачає технічну модернізацію систем електропостачання й компенсацію реактивної потужності. У роботі С. М. Балюти, В. Д. Йовбака, Л. О. Копилової та Ю. В. Куєвди [1] вивчено підходи до компенсації реактивної потужності на промислових і цивільних об'єктах з урахуванням ефективності розміщення й конфігурації компенсувальних пристроїв. Г. Гурліс (G. Gourlis) та І. Ковачич (I. Kovacic) [2] наголошують на ролі інженерно-будівельної інфраструктури в загальній схемі енергоспоживання,

акцентуючи потребу залучення будівель до процесів енергооптимізації. У праці К. Геллінгса (C. W. Gellings) [3] детально розглянуто трансформацію електромереж у смартсистеми, що забезпечують інтерактивне управління енергопотоками для зменшення пікових навантажень. Перспективним у цьому напрямі є подальше вивчення впровадження автоматизованих систем контролю параметрів електроспоживання в реальному часі.

Представники другого напрямку зосереджують увагу на теплотехнічній інтеграції й термодинамічній оптимізації машинобудівних систем. У публікації Л. А. Проца, Ю. І. Фордзюна та Є. М. Лютого [4] обґрунтовано ефективність поєднання теплових процесів у рамках єдиної енергетичної системи підприємства, що дає змогу зменшити витрати палива й підвищити загальний ККД. Д. Маріано-Ернандес (D. Mariano-Hernández), Л. Ернандес-Кальехо (L. Hernández-Callejo), А. Зоріта-Ламадрид (A. Zorita-Lamadrid), О. Дуке-Перес (O. Duque-Pérez), Ф. Гарсія (F. García) [5] аналізують методи прогнозного управління й виявлення несправностей для забезпечення сталого функціонування теплових систем будівель. У дослідженні М. Крарті (M. Krarti) [6] запропоновано інженерний підхід до проведення енергоаудиту з акцентом на виявлення теплових утрат. Перспективним є вдосконалення моделей когенерації та систем відновлення теплоти в умовах змінного навантаження на промислових підприємствах.

Третій напрям присвячено впровадженню інтелектуальних технологій, цифрових двійників і концепцій Industry 4.0 в управління енергоефективністю. У праці Ф. Мацунаги (F. Matsunaga), В. Житковські (V. Zytkowski), П. Валле (P. Valle), Ф. Дешампса (F. Deschamps) [7] доведено ефективність застосування кіберфізичних систем для моніторингу й оптимізації енергоспоживання на виробництві. В. Ю (W. Yu), П. Патрос (P. Patros), Б. Янг (B. Young), Е. Клінак (E. Klinac), Т. Г. Волмслі (T. G. Walmsley) [8] запропонували класифікацію цифрових двійників для промислової енергетики, підкреслюючи їхню здатність до адаптації в умовах мінливого середовища. А. Катал (A. Katal), С. Дахія (S. Dahiya), Т. Чаудгарі (T. Choudhury) [9] досліджують програмні засоби підвищення енергоефективності в дата-центрах як значних споживачів енергії.

У роботі Г. Фарзане (H. Farzaneh), Л. Малемірчегіні (L. Malehmirchegini), А. Бежана (A. Bejan), Т. Афолабі (T. Afolabi), А. Мулумбі (A. Mulumba), П. П. Даки (P. P. Daka) [10] описано застосу-

вання штучного інтелекту в смартбудівлях для управління кліматом, освітленням і споживанням електроенергії. Л. Белуссі (L. Belussi), Б. Бароцці (B. Barozzi), А. Беллацці (A. Bellazzi), Л. Данца (L. Danza), А. Девітофранческо (A. Devitofrancesco), К. Фанчуллі (C. Fanciulli), М. Геллере (M. Ghellere), Д. Гуадзі (G. Guazzi), І. Мероні (I. Meroni), Ф. Саламоне (F. Salamone), Ф. Скамоні (F. Scamoni), К. Скрозаті (C. Scrosati) [11] розглянули інноваційні рішення для досягнення нульового енергоспоживання в будівлях, які можуть інтегруватися в промислові комплекси. Перспективним є розвиток гібридних систем на базі цифрових двійників та алгоритмів ШІ для комплексного управління промисловими енергооб'єктами.

Четвертий напрям стосується системного оцінювання енергоспоживання, управління якістю технологічних процесів і впливу смарттехнологій на економічну ефективність. У дослідженні Г. І. Канюка, А. Ю. Мезері, Т. Ю. Василець, С. М. Келеберди, А. С. Пономаренка, Д. О. Чирочкіна [12] запропоновано методологію оцінювання якості енергоспоживання за допомогою багатфакторного аналізу. У роботі М. Б. Паласевича, О. В. Савки й Т. А. Лепак [13] досліджено вплив цифровізації на підвищення енергоефективності економіки через упровадження смартрішень.

У праці А. Тумана (A. Thumann) і Д. П. Мехти (D. P. Mehta) [14] систематизовано інженерні рішення й етапи впровадження енергоменеджменту на підприємствах. У публікації Дж. Хоссейна (J. Hossain), А. Ф. Кадіра (A. F. Kadir), А. Н. Ханафі (A. N. Hanafi), Х. Шаріфа (H. Shareef), Т. Хатіба (T. Khatib), К. А. Бахаріна (K. A. Baharin), М. Ф. Сулайми (M. F. Sulaima) [15] узагальнено практики управління енергією в комерційних і багатофункціональних будівлях переважно промислового призначення. М. Павловський (M. Pavlovskiy) [16] вивчив можливості підвищення паливної ефективності дизельних двигунів за рахунок використання біопалив, що має значення для енергомодернізації транспортної та допоміжної інфраструктури промислових підприємств. Перспективним є подальше дослідження інтеграції енергетичного моніторингу із загальною системою стратегічного управління підприємством.

Попри наявність значної кількості праць, присвячених питанням енергоефективності в промисловості, залишаються невирішеними кілька ключових аспектів. Зокрема, недостатньо досліджено специфіку реалізації принципів енергозбе-

реження з урахуванням функціонального призначення різних типів інженерних систем. Методи оцінювання енергоефективності досі спираються на фрагментарні підходи, які або не враховують динамічних режимів роботи, або не інтегруються з інтелектуальними технологіями моніторингу. Також наявна нормативна невідповідність між вимогами сучасної цифрової промисловості й чинною базою, а практичні механізми енергоменеджменту слабо впроваджені через відсутність адаптованих моделей управління та брак інвестиційно орієнтованих рішень.

Запропоноване дослідження спрямовано на усунення цих прогалин шляхом комплексного аналізу принципів енергозбереження, узагальнення методичних основ оцінювання ефективності, вивчення можливостей інтелектуального керування й формування цілісної моделі модернізації. Застосування аналітичного підходу до нормативної бази, адаптація цифрових рішень та орієнтація на практичну реалізацію запропонованих заходів дають змогу значно розширити наукове розуміння проблеми й запропонувати конкретні інструменти для її вирішення на рівні підприємств.

Постановка завдання. Мета статті – дослідити особливості функціонування інженерних систем промислових об'єктів у контексті енергоефективності, оцінити потенціал їх оптимізації з урахуванням технічних та економічних чинників і сформулювати практичні рекомендації щодо зниження енергетичних витрат у виробничому середовищі.

Завдання статті:

1. Проаналізувати принципи енергозбереження й підходи до оцінювання енергоефективності інженерних систем промислових підприємств з урахуванням функціональної специфіки, умов експлуатації та нормативно-технічної бази.

2. Дослідити ефективність інтелектуальних технологій моніторингу й управління як засобів оптимізації енергоспоживання в промисловому середовищі.

3. Виявити системні бар'єри підвищення енергоефективності та сформулювати рекомендації щодо модернізації інженерної інфраструктури за допомогою енергоощадних рішень і фінансових механізмів підтримки.

Виклад основного матеріалу. Оптимізація енергоспоживання в інженерних системах промислових підприємств ґрунтується на впровадженні цілісних принципів енергозбереження, які відображають технічні, технологічні й організаційно-експлуатаційні аспекти. Ці принципи

є складником системного підходу до підвищення енергоефективності, оскільки дають змогу забезпечити узгодженість роботи енергетичних підсистем підприємства з урахуванням їх функціонального навантаження, змінних умов виробничого циклу й зовнішнього середовища. Реалізація принципів енергозбереження не є уніфікованою: вона варіюється залежно від типу системи (опалення, вентиляції, електропостачання, водопостачання), технічного ресурсу обладнання, автоматизації процесів і рівня енергоменеджменту. Узагальнену характеристику основних принципів енергозбереження, реалізованих в інженерних системах промислових об'єктів, подано в таблиці 1.

Реалізація зазначених принципів енергозбереження відображає перехід промислових підприємств від фрагментарного управління енерговитратами до системного підходу на основі технологічної адаптивності й цифрової інтеграції. Наприклад, дотримання принципу відповідності потужності споживанню дає змогу зменшити пікові навантаження на мережу, що не лише знижує витрати на електроенергію, а й подовжує ресурс роботи електрообладнання. В умовах усе вищих тарифів і нестабільності енергопостачання це має безпосередній економічний ефект. Теплоізоляційна ефективність виявляється особливо значущою для підприємств із тривалими теплотехнологічними циклами, де навіть незначні втрати тепла можуть призвести до суттєвого зниження коефіцієнта корисної дії системи. Принцип регулювання в реальному часі, що базується на застосуванні сенсорних модулів та адаптивних алгоритмів, відкриває можливості для так званого

предиктивного керування, коли система прогнозує зміну умов і реагує на них ще до появи перевитрат. Застосування рекуперації енергії в українських промислових умовах нині активно розвивається, зокрема, у харчовій, металургійній і деревообробній галузях, де надлишкове тепло або тиск конвертується в корисну енергію для допоміжних потреб. Комплексна автоматизація, яка впроваджується через SCADA (Supervisory Control and Data Acquisition)- і MES (Manufacturing Execution System)-рішення [17], є не лише інструментом контролю, а й аналітичним засобом для визначення неефективних зон, моделювання сценаріїв навантаження й оптимізації графіків роботи технологічного обладнання [9]. Усі ці механізми демонструють, що сучасні принципи енергозбереження реалізуються не ізольовано, а в рамках інтегрованих рішень, які враховують як технічну логіку процесів, так й економічну доцільність їх впровадження.

Оцінювання енергоефективності інженерних систем у промисловості є багаторівневим процесом, що передбачає інтеграцію технічних, нормативно-методичних та економічних інструментів. В умовах трансформації енергетичної політики й цифровізації виробничих процесів актуалізується потреба не лише кількісного вимірювання спожитих енергоресурсів, а й визначення ефективності використання енергії в контексті досягнення функціональних і технологічних цілей. Підходи до оцінювання базуються на поєднанні показників питомого енергоспоживання, коефіцієнтів корисної дії, енергетичних балансів, а також аналізу відхилень від нормативних зна-

Таблиця 1

Принципи енергозбереження в інженерних системах промислових підприємств

Принцип енергозбереження	Зміст принципу	Особливості реалізації в інженерних системах промислових об'єктів
Відповідність потужності споживанню	Мінімізація надлишкової енергії шляхом балансування потужностей і навантажень	Застосовується в системах електропостачання, вентиляції, привідних механізмах
Теплоізоляційна ефективність	Зменшення тепловтрат за рахунок сучасних ізоляційних матеріалів і конструкцій	Актуально для тепломреж, паропроводів, емностей із теплоносіями
Регулювання в реальному часі	Динамічне налаштування параметрів роботи систем відповідно до змін умов	Реалізується через автоматику в системах HVAC, насосних станціях, освітленні
Рекуперація енергії	Повернення частини енергії з відпрацьованих потоків у виробничий цикл	Застосовується у вентиляції, теплогенерації, системах стисненого повітря
Комплексна автоматизація	Централізоване управління енергоресурсами для зменшення втрат	Реалізується через SCADA-системи, енергетичні модулі ERP, промислові контролери

Джерело: сформовано на основі [1; 2; 3; 7; 14]

чень. У сучасній практиці широко застосовуються як традиційні методи розрахунку, так й автоматизовані системи енергомоніторингу, що забезпечують оперативний контроль параметрів і прогнозування енергетичних утрат. Основні підходи до оцінювання енергоефективності узагальнено в таблиці 2.

У сучасних умовах вітчизняної промисловості, що працює в режимі ресурсної економії й технологічного оновлення, найбільш ефективною виявляється саме поєднана модель оцінювання. Автоматизований моніторинг дає змогу фіксувати зміну навантажень та оперативно реагувати на відхилення, тоді як енергетичний аудит забезпечує глибинну діагностику і створення основи для довгострокового планування модернізації. У разі систем опалення й вентиляції нормативно-розрахунковий підхід дає змогу зіставити проєктні теплові потужності з реальними умовами експлуатації. Індикативне оцінювання широко застосовується в енергоємних галузях, зокрема металургії, де зіставлення питомих показників дає уявлення про ефективність цехів і технологій.

На окрему увагу заслуговує чинна нормативно-методична база, яка є основою для нормативно-розрахункових методів оцінювання енергоефективності. Зокрема, застосовується ДБН В.2.5-39:2008 «Інженерне обладнання будівель і споруд. Теплова ізоляція та енергозбереження» [18], яка визначає вимоги до проєктування теплової ізоляції систем, однак не враховує впливу динамічних режимів енергоспоживання в умовах промислової експлуатації. ДСТУ ISO 50001:2020 «Системи енергетичного менеджменту. Вимоги та настанови щодо застосування» [19] забезпечує

методологічну основу для впровадження політики енергоуправління, проте не адаптований до цифрових інструментів оперативного контролю та прогнозування. Додатково застосовуються галузеві методики енергетичного аудиту, затверджені Держенергоефективності [20], які базуються переважно на статичних коефіцієнтах утрат і питомих витрат. Виявлено, що ці документи мають обмежену функціональність у контексті інтеграції з автоматизованими системами моніторингу (SCADA, IoT) та аналітичними платформами штучного інтелекту, що гальмує адаптацію нормативів до реальних умов функціонування цифровізованої промислової інфраструктури. Така невідповідність між регуляторними підходами й технологічними можливостями формує бар'єри для впровадження гнучкого енергоменеджменту, який базується на змінних профілях навантаження, прогнозних моделях та оперативному реагуванні на енергетичні ризики [14].

У сучасній промисловості інтелектуальні технології моніторингу й контролю стали невід'ємною частиною систем енергоефективного управління, оскільки забезпечують оперативну обробку даних, адаптивне регулювання навантажень і запобігання енергетичним утратам. На відміну від традиційних систем обліку, ці рішення дають змогу в реальному часі відстежувати зміну параметрів, аналізувати тренди, будувати прогнози споживання та приймати оптимізаційні рішення без залучення операторів. Особливо актуальним є використання таких технологій на об'єктах із високим енергоспоживанням, складною конфігурацією інженерної інфраструктури й потребою в гнучкому енергетичному балансу-

Таблиця 2

Підходи до оцінювання енергоефективності інженерних систем у промисловості

Підхід до оцінювання	Ключова характеристика	Галузь застосування й функціональна роль
Нормативно-розрахунковий	Порівняння фактичного споживання з нормативами, установленними для об'єкта	Оцінювання відповідності проєктних і фактичних показників у системах опалення, вентиляції
Енергетичний аудит	Комплексне обстеження об'єкта з аналізом утрат, неефективних зон, потенціалу економії	Застосовується під час модернізації підприємств і розроблення енергетичних паспортів
Індикативне оцінювання	Аналіз питомих витрат енергії на одиницю продукції або послуги	Дає змогу порівнювати ефективність між різними об'єктами одного типу
Коефіцієнтна модель	Розрахунок ККД, коефіцієнтів енерговикористання, тепловтрат, навантаження	Застосовується для аналізу динаміки показників у часі та виявлення неефективних режимів
Автоматизований енергомоніторинг	Збирання й аналіз даних у реальному часі із застосуванням сенсорних систем і SCADA	Підтримує оперативне управління та прийняття рішень на основі даних

Джерело: сформовано на основі [4; 5; 6; 12; 13]

ванні. Основні інструменти й ефекти їх застосування подано в таблиці 3.

Під час практичної реалізації на промислових підприємствах ці технології створюють основу для переходу від реактивного до превентивного управління енергетичними потоками. Зокрема, системи енергомоніторингу з інтеграцією в SCADA дають змогу здійснювати багатофакторний контроль енергоспоживання не лише за обсягами, а й за якістю, піковими навантаженнями та рівнем утрат у конкретних ділянках мережі. Застосування AI-платформ у комбінації з даними із сенсорної інфраструктури забезпечує прогнозне керування, яке допомагає уникати надмірного енергоспоживання ще до його виникнення. У цехах із високою варіативністю технологічних режимів адаптивні системи керування виявляються особливо ефективними, оскільки здатні змінювати параметри роботи в реальному часі залежно від змін внутрішнього мікроклімату, графіка навантаження або зовнішніх погодних умов [7]. Усе це сприяє формуванню гнучких і самооптимізувальних виробничих середовищ, що відповідають принципам промисловості 4.0 і дають змогу зменшити не лише енергетичні витрати, а й загальні експлуатаційні витрати підприємства.

Підвищення енергоефективності інженерних систем промислових об'єктів суттєво ускладнюється низкою структурних і функціональних бар'єрів, що формують інерційне середовище виробничої інфраструктури. Однією з головних перешкод залишається технічна застарілість основних фондів, яка не дає змоги ефективно впроваджувати сучасні регуляційні й автоматизовані рішення без масштабної реконструкції систем [4].

Обладнання, розроблене в умовах інших енергетичних стандартів, має високий рівень утрат, обмежену сумісність із цифровими інструментами контролю та не підтримує адаптивних режимів роботи [7]. Водночас обмежений доступ до фінансування, характерний для більшості вітчизняних промислових підприємств, не дає можливості реалізувати повноцінні модернізаційні програми, орієнтовані на глибоку трансформацію енергетичної інфраструктури [12]. Інвестиції переважно спрямовуються на підтримання базової функціональності, що знижує гнучкість систем управління та їхню здатність до оптимізації. Недосконалість нормативної бази, яка не завжди враховує специфіку сучасних технологій енергообліку й управління, також гальмує розвиток енергоефективних рішень [5]. Чинні регламенти часто не мають механізмів адаптації до змін середовища або стимулювання підприємств до запровадження інновацій. Додатково ситуацію ускладнює невисокий рівень інтеграції енергоменеджменту в систему стратегічного управління підприємством. Відсутність єдиної інформаційної бази, розосередження відповідальності за енергоспоживання між різними службами й недооцінювання енергетичних показників у структурі операційної ефективності призводять до фрагментарності управлінських рішень.

Модернізація інженерних систем промислових об'єктів має базуватися на системному впровадженні енергоощадного обладнання, цифрових технологій управління й ефективних механізмів фінансування, що дають змогу забезпечити довгострокову економічну доцільність енергоефективних рішень. Одним із ключових напрямів є заміна застарілих агрегатів на високоефективні

Таблиця 3

Інтелектуальні технології моніторингу й контролю в системах енергоефективного управління

Технологія/інструмент	Принцип дії та функціональність	Вплив на енергоефективність
Системи автоматизованого енергомоніторингу	Постійне збирання, оброблення та передача даних зі споживчих точок у режимі реального часу	Скорочення втрат, оперативне виявлення неефективних зон, контроль пікових навантажень
Інтелектуальні сенсори й IoT-рішення	Вимірювання параметрів (температура, тиск, струм, витрати) з високою точністю	Підвищення точності управління, зниження надлишкового споживання ресурсів
SCADA-системи з функціями аналітики	Візуалізація даних, дистанційне керування, аналіз трендів і моделювання сценаріїв	Забезпечення стабільності режимів, запобігання збоєм і перенавантаженням
AI-платформи для енергетичного прогнозування	Алгоритми машинного навчання для аналізу історичних даних і прогнозування майбутніх потреб	Оптимізація графіків навантаження, зменшення витрат на енергоносії
Системи адаптивного управління	Динамічне коригування параметрів роботи систем за зміною зовнішніх і внутрішніх умов	Автоматичне зниження споживання без втрати продуктивності систем

Джерело: сформовано автором на основі [8; 9; 10; 11; 15]

установки з регульованими приводами, енергозберігальними теплообмінниками, сенсорним керуванням і низькими втратами в режимах часткового навантаження. Доцільність таких рішень має підтверджуватися розрахунками скорочення питомого енергоспоживання, строків окупності й ефекту життєвого циклу. Упровадження автоматизованого управління передбачає інтеграцію SCADA-систем, технологій IoT і машинного навчання, що дає змогу здійснювати моніторинг, діагностику й оптимізацію роботи систем у режимі реального часу. На рівні керування підприємством важливим є поєднання технічної модернізації з розгортанням енергоменеджменту на основі стандартизованих підходів (наприклад, ISO 50001 [19]), що забезпечує постійне вдосконалення енергетичної політики. Фінансування таких ініціатив доцільно здійснювати через залучення цільових кредитних програм, енергосервісних контрактів (ESCO) або міжнародних донорських інструментів, що передбачають повернення інвестицій за рахунок досягнутої економії енергоресурсів. Важливою умовою є супровід кожного проєкту аналітичним інструментарієм для оцінювання ефективності на етапі планування, реалізації та експлуатації. Усі ці заходи мають бути не ізольованими, а частиною цілісної стратегії, що враховує технологічну специфіку об'єкта, економічні обмеження й потребу в гнучкому управлінні енергетичними потоками.

Висновки. Під час дослідження встановлено, що підвищення енергоефективності інженерних систем промислових об'єктів потребує цілісного підходу, який поєднує технічну модернізацію, цифровізацію управління й економічну доцільність упроваджених рішень. Обґрунтовано ефективність принципів енергозбереження, таких як балансування навантажень, теплоізоляція, рекуперація, адаптивне регулювання й автоматизоване керування. Узагальнено актуальні підходи до оцінювання енергоефективності, серед яких найбільш результативним є поєднання енергетичного аудиту з автоматизованим моніторингом та аналітичними моделями. Доведено, що інтелектуальні технології – SCADA, IoT-сенсори, AI-платформи – формують основу для переходу до превентивного управління енергоспоживанням. Виявлено основні перешкоди: технічну застарілість систем, обмежене фінансування, фрагментарність енергоменеджменту й неадаптованість нормативної бази. Запропоновано впровадження енергоощадного обладнання, стандартизованих систем управління й механізмів фінансування на основі енергосервісу.

Перспективи подальших досліджень пов'язані з розробленням адаптивних цифрових моделей управління енергоспоживанням, інтеграцією AI в промислову енергетичну аналітику й удосконаленням регуляторних підходів до стимулювання інновацій у сфері енергоефективності.

Список літератури:

1. Балюта С. М., Йовбак В. Д., Копилова Л. О., Куєвда Ю. В. Компенсація реактивної потужності в системах електрозабезпечення промислових і цивільних об'єктів. *Наукові праці Національного університету харчових технологій*. 2021. Вип. 27, № 4. С. 117–128. URL: <https://surl.li/tjnuqp> (дата звернення: 06.06.2025).
2. Gourlis G., Kovacic I. Energy efficient operation of industrial facilities: The role of the building in simulation-based optimization. *IOP Conference Series: Earth and Environmental Science*. 2020. Vol. 410, № 1. Article 012019. DOI: <https://doi.org/10.1088/1755-1315/410/1/012019>.
3. Gellings C. W. The smart grid: enabling energy efficiency and demand response. *River Publishers*. 2020. DOI: <https://doi.org/10.1201/9781003151524>.
4. Проц Л. А., Фордзюн Ю. І., Лютий Є. М. Інтеграція теплотехнічних процесів у проєктування енергоефективних машинобудівних систем. *Вісник Херсонського національного технічного університету*. 2025. Вип. 1, № 92. С. 190–197. DOI: <https://doi.org/10.35546/kntu2078-4481.2025.1.1.24>.
5. Mariano-Hernández D., Hernández-Callejo L., Zorita-Lamadrid A., Duque-Pérez O., García F. A review of strategies for building energy management system: Model predictive control, demand side management, optimization, and fault detect & diagnosis. *Journal of Building Engineering*. 2021. Vol. 33. Article 101692. DOI: <https://doi.org/10.1016/j.jobe.2020.101692>.
6. Krarti M. Energy audit of building systems: an engineering approach. *CRC Press*. 2020. DOI: <https://doi.org/10.1201/9781003011613>.
7. Matsunaga F., Zytkowski V., Valle P., Deschamps F. Optimization of energy efficiency in smart manufacturing through the application of cyber-physical systems and industry 4.0 technologies. *Journal of Energy Resources Technology*. 2022. Vol. 144, № 10. Article 102104. DOI: <https://doi.org/10.1115/1.4053868>.
8. Yu W., Patros P., Young B., Klinac E., Walmsley T. G. Energy digital twin technology for industrial energy management: Classification, challenges and future. *Renewable and Sustainable Energy Reviews*. 2022. Vol. 161. Article 112407. DOI: <https://doi.org/10.1016/j.rser.2022.112407>.

9. Katal A., Dahiya S., Choudhury T. Energy efficiency in cloud computing data centers: a survey on software technologies. *Cluster Computing*. 2023. Vol. 26. P. 1845–1875. DOI: <https://doi.org/10.1007/s10586-022-03713-0>.
10. Farzaneh H., Malehmirchegini L., Bejan A., Afolabi T., Mulumba A., Daka P. P. Artificial intelligence evolution in smart buildings for energy efficiency. *Applied Sciences*. 2021. Vol. 11, № 2. Article 763. DOI: <https://doi.org/10.3390/app11020763>.
11. Belussi L., Barozzi B., Bellazzi A., Danza L., Devitofrancesco A., Fanciulli C., Ghellere M., Guazzi G., Meroni I., Salamone F., Scamoni F., Scrosati C. A review of performance of zero energy buildings and energy efficiency solutions. *Journal of Building Engineering*. 2019. Vol. 25. Article 100772. DOI: <https://doi.org/10.1016/j.jobe.2019.100772>.
12. Канюк Г. І., Мезеря А. Ю., Василець Т. Ю., Келеберда С. М., Пономаренко А. С., Чирочкін Д. О. Оптимізація показників якості технологічних процесів об'єктів керування в енергетиці. *Машинобудування*. 2024. Вип. 32. С. 55–63. DOI: <https://doi.org/10.32820/2079-1747-2023-32-55-63>.
13. Паласевич М. Б., Савка О. В., Лепак Т. А. Вплив смарт-технологій на економічний розвиток та підвищення енергоефективності національної економіки. *Академічні візії*. 2024. Вип. 34. URL: <https://www.academy-vision.org/index.php/av/article/view/129> (дата звернення: 06.06.2025).
14. Thumann A., Mehta D. P. Handbook of energy engineering. *River Publishers*. 2020. DOI: <https://doi.org/10.1201/9781003151715>.
15. Hossain J., Kadir A. F., Hanafi A. N., Shareef H., Khatib T., Baharin K. A., Sulaima M. F. A review on optimal energy management in commercial buildings. *Energies*. 2023. Vol. 16, № 4. Article 1609. DOI: <https://doi.org/10.3390/en16041609>.
16. Pavlovskiy M. The Improvement of Fuel Efficiency and Environmental Characteristics of Diesel Engine by Using Biodiesel Fuels. In: Boichenko S., Zaporozhets A., Yakovlieva A., Shkilniuk I. (eds) *Modern Technologies in Energy and Transport. Studies in Systems, Decision and Control*. 2024. Vol. 510. Springer, Cham. P. 43–55. DOI: https://doi.org/10.1007/978-3-031-44351-0_4.
17. MES vs SCADA in Industry 4.0 – Differences and meanings. Nexus Integra: вебсайт. URL: <https://nexusintegra.io/mes-vs-scada/> (дата звернення: 05.06.2025).
18. ДБН В.2.5-39:2008. Інженерне обладнання будинків і споруд. Теплова ізоляція та енергозбереження. Мінрегіон України : вебсайт. URL: <https://surl.li/fgqwrp> (дата звернення: 05.06.2025).
19. ДСТУ ISO 50001:2020. Системи енергетичного менеджменту. Вимоги та настанови щодо застосування. УкрНДНЦ : вебсайт. URL: https://online.budstandart.com/ua/catalog/doc-page.html?id_doc=90178 (дата звернення: 05.06.2025).
20. Методичні рекомендації з проведення енергоаудиту будівель. ESAVE: вебсайт. URL: <https://www.esave.com.ua/service/energoaudit-budivli> (дата звернення: 05.06.2025).

Krymska A.O., Didok V.Yu., Stoliarchuk V.V. OPTIMIZATION OF ENERGY EFFICIENCY IN ENGINEERING SYSTEMS OF INDUSTRIAL FACILITIES

The relevance of the study is determined by the growing need to reduce energy costs in industry against the backdrop of infrastructure wear and tear, rising tariffs, stricter environmental requirements, and the need to ensure the energy stability of production processes. The energy efficiency of engineering systems is considered a strategic resource of an enterprise that affects its competitiveness, operational stability, and long-term economic efficiency.

The aim of the study is to comprehensively analyze the functioning of engineering systems of industrial facilities from the perspective of energy efficiency, to assess the possibilities for their technical and digital modernization, taking into account economic feasibility, and to develop practical recommendations for reducing energy consumption in a production environment.

The research methodology is based on a combination of systematic analysis of the architecture of engineering systems, generalization of approaches to energy monitoring and energy consumption modeling, as well as a critical analysis of the regulatory framework, the potential of automated management technologies, and financial models for modernization. Elements of comparative, factor, and structural-functional analysis were used.

Research results. The principles of energy conservation implemented in modern enterprise systems, in particular load balancing, recuperation, adaptive control, and digital automation, were studied. Approaches to energy efficiency assessment based on normative, indicative, coefficient, and automated analysis were generalized. The effectiveness of intelligent monitoring technologies – SCADA, IoT, AI platforms – for the transition to preventive energy consumption management has been revealed. The key reasons hindering modernization have been identified: outdated equipment, limited access to financing, low level of energy management, and regulatory environment inconsistencies.

Conclusions. It has been proven that energy efficiency can only be improved through a comprehensive approach that includes technical reconstruction and the introduction of adaptive control and energy management in accordance with international standards. The effectiveness of the intellectualization of energy systems for achieving a stable energy balance has been confirmed.

Prospects for further research lie in the development of digital models of adaptive energy consumption management, the integration of machine learning tools for load forecasting, and the improvement of regulatory and methodological foundations for energy efficiency policy in industry.

Key words: *energy management, intelligent monitoring technologies, automated control, technical modernization, energy audit, digital analytics, regulatory framework.*

Дата надходження статті: 20.06.2025

Дата прийняття статті: 27.06.2025

Опубліковано: 27.10.2025

УДК 004.94:681.5

DOI <https://doi.org/10.32782/2663-5941/2025.4.2/26>**Латипов О.В.**

Національний університет «Запорізька політехніка»

Єфименко М.В.

Національний університет «Запорізька політехніка»

Миронова Н.О.

Національний університет «Запорізька політехніка»

Молочков Д.Є.

Національний університет «Запорізька політехніка»

Куляба-Харитоновна Т.І.

Національний університет «Запорізька політехніка»

РОЗРОБКА СИСТЕМИ РОЗПІЗНАВАННЯ ПОЛОЖЕННЯ ОБ'ЄКТІВ У РОБОЧІЙ ЗОНІ ПРОМИСЛОВОГО РОБОТА

У статті представлено дослідження та реалізацію системи розпізнавання просторового положення об'єктів у робочій зоні промислового робота з використанням стереопари USB-камер і глибокої нейронної мережі YOLOv8. Актуальність роботи зумовлена потребою в точному тривимірному позиціонуванні об'єктів у режимі реального часу для задач автоматизованої маніпуляції, складання та взаємодії з людиною в індустріальних і дослідницьких умовах. Об'єктом дослідження є процес визначення просторового положення об'єктів у зоні дії промислового робота. Предметом дослідження є алгоритми та програмно-апаратні засоби для реалізації системи комп'ютерного зору на основі стереозображення і глибокого навчання. У роботі обґрунтовано вибір USB-камер як оптимального апаратного засобу для побудови стереопари: камери забезпечують стабільну передачу зображень, низьку затримку та високу гнучкість налаштувань. Проведено повне стереокалібрування камер із використанням шахового шаблону, що дозволило визначити внутрішні параметри, просторову геометрію та виконати ректифікацію зображень. У програмній частині реалізовано повний цикл обробки: від синхронного захоплення кадрів до виводу координат об'єктів у тривимірному просторі. YOLOv8 використовується для визначення об'єктів на обох зображеннях, після чого за допомогою аналізу диспаритету та методу тріангуляції обчислюється глибина сцени. Проведене експериментальне тестування продемонструвало стабільну роботу системи з точністю до 2–4 см в межах робочого діапазону до 1 метра. Частота обробки становила 15–20 FPS, затримка на кадр – 50–70 мс, навантаження CPU – до 80%. Аналіз впливу чинників виявив критичну роль точності калібрування, стабільності відеопотоку, синхронізації кадрів і базової лінії між камерами. У системі реалізовано візуальний контроль результатів, що забезпечує зручне тестування та подальше налагодження. Отримані результати підтверджують доцільність використання стереопари USB-камер і YOLOv8 для побудови доступної та гнучкої системи тривимірної локалізації об'єктів. Розроблений підхід може бути адаптований до інтеграції з ROS, мобільними платформами або роботизованими маніпуляторами. Подальші дослідження плануються спрямувати на реалізацію трекінгу в динаміці, оптимізацію алгоритмів для edge-пристроїв і розширення функціональності шляхом застосування методів сегментації.

Ключові слова: стереозір, YOLOv8, комп'ютерний зір, тривимірна локалізація об'єктів, глибина сцени, стереопара, калібрування камер, робототехнічні системи, нейронні мережі.

Постановка проблеми. У сучасних робото- з достатнім рівнем точності. Це необхідно для технічних системах важливо забезпечити просторову локалізацію об'єктів у реальному часі виконання операцій маніпуляції, автоматизованого складання, контролю якості продукції,

а також безпечної взаємодії між роботами та людьми. Традиційні методи, засновані на фіксованих координатах і шаблонному програмуванні, не забезпечують достатньої гнучкості в динамічному середовищі. Для вирішення цієї проблеми зростає потреба у використанні інтелектуальних систем комп'ютерного зору, здатних адаптивно розпізнавати положення об'єктів та оцінювати їх координати в просторі. Одним із перспективних підходів є інтеграція методів глибинного навчання з технологією стереозору.

Аналіз останніх досліджень і публікацій. Задача локалізації об'єктів є основною для автономної робототехніки, знаходячи застосування у промисловості, логістиці, медицині та агротехніці. Сучасні підходи до вирішення задачі локалізації об'єктів пройшли шлях від класичних алгоритмів комп'ютерного зору до складних систем, що поєднують глибокі нейронні мережі та гібридних сенсорних даних.

Класичні методи локалізації об'єктів формували основу для ранніх роботизованих систем, де точність мала вирішальне значення. Вони ґрунтуються на законах геометрії та аналітичних методах обчислення просторового положення об'єктів на основі зображень з однієї або кількох камер. Наприклад, метод стереозору дозволяє обчислити глибину шляхом аналізу відмінностей між парою зображень однієї сцени, що дає змогу отримати тривимірну інформацію про об'єкти. Методи структури з руху (Structure from Motion) відновлюють просторову сцену за серією двовимірних кадрів, використовуючи зміну положення камери або об'єктів у кадрі.

Іншим поширеним рішенням є використання маркованих систем, таких як ArUco або AprilTag, де положення ідентифікується завдяки наявності спеціальних візуальних міток [1]. Такі системи добре працюють у структурованих лабораторних або промислових середовищах.

З розвитком глибокого навчання з'явилися архітектури, що дозволили суттєво підвищити якість та швидкість локалізації. Модель YOLO та її наступні версії (YOLOv5, YOLOv8) стали стандартом для визначення об'єктів в реальному часі завдяки визначенню об'єктів за один прохід через зображення, що суттєво скорочувало час обчислень [2]. Для завдань, що вимагають вищої точності, застосовується архітектура Faster R-CNN, яка використовує дворівневий підхід для уточнення меж об'єктів [3]. Сучасні рішення, такі як трансформерні моделі DETR, використовують механізми уваги для глобального аналізу сцени,

що покращує врахування контексту [4]. Для мобільних пристроїв з обмеженими ресурсами були розроблені моделі прямої регресії позиції, наприклад PoseNet, що безпосередньо передбачають положення та орієнтацію об'єкта з одного RGB-зображення [5].

Сучасні роботизовані системи все частіше використовують гібридні підходи, що поєднують різні джерела даних для підвищення надійності. Найбільш поширеною є комбінація RGB-зображень з картами глибини (RGB-D), отриманими з сенсорів Intel RealSense [6] або подібних. Такі дані слугують входом для нейронних мереж, що значно підвищує точність тривимірної локалізації. Для навігації в динамічних середовищах використовуються візуально-інерціальні системи та алгоритми одночасної локалізації та картування, прикладом яких є RTAB-Map [7]. Отримані дані про положення об'єктів інтегруються у фреймворки планування руху, такі як MoveIt, для побудови безконфліктних траєкторій маніпуляторів [8].

На сьогодні у сфері застосування комп'ютерного зору в робототехніці все частіше реалізуються системи, що поєднують виявлення об'єктів за допомогою нейронних мереж та локалізацію їх у просторі через глибинну сенсоріку або стереозір.

Наприклад, у дослідженні [9] представлена система для зварювального маніпулятора на базі YOLOv5 та камери Intel RealSense, що досягає середньої похибки позиціонування 2,9 мм. Схожий підхід використано у роботі [10] для точного зрізу плодоніжок томатів за допомогою модифікованої YOLOv8 та RGB-D камери, де точність локалізації склала $\pm 2,5$ мм. Інші дослідження [11] зосереджені на оптимізації YOLO-архітектур для вбудованих платформ, демонструючи високу швидкість обробки навіть без використання глибинних даних.

Проведений аналіз дозволяє зробити висновок про наявність сталих архітектурних рішень у проектуванні робототехнічних систем локалізації об'єктів. Типовими ознаками таких систем є використання згорткових нейронних мереж для визначення об'єктів на зображеннях, інтеграція глибинної інформації, отриманої за допомогою стереозору або RGB-D камер, а також перетворення отриманих даних у метричні координати для подальшого використання в траєкторному плануванні та управлінні маніпуляторами. У ряді реалізацій додатково застосовуються спрощені архітектури, що сприяє зниженню обчислювальної складності та забезпечує можливість розгор-

тання таких систем на ресурсно обмежених апаратних платформах.

Постановка завдання. Метою статті є дослідження та реалізація прототипу системи розпізнавання положення об'єктів у тривимірному просторі на основі стереопари з двох USB-камер і глибокої нейронної мережі YOLOv8.

Виклад основного матеріалу. Вибір та обґрунтування апаратної бази.

Для вирішення задачі розпізнавання та локалізації об'єктів у тривимірному просторі було розроблено та реалізовано програмно-апаратний комплекс на основі стереозору та глибокого навчання. Дана методологія дозволяє з високою точністю визначати тривимірні координати об'єктів, що є критично важливим для завдань роботизованої маніпуляції. Процес розробки включав наступні етапи: вибір та обґрунтування апаратної бази, підбір програмних інструментів для обробки зображень та визначення об'єктів, а також процедуру калібрування стереосистеми для забезпечення точності вимірювань.

Основою системи є стереопара камер, що дозволяє обчислювати глибину сцени на основі ефекту паралакса. Для вибору оптимального апаратного рішення було проведено порівняльний аналіз наступних типів пристроїв: IP-камер, USB-камер та спеціалізованих глибинних сенсорів. Ключові критерії та результати порівняння наведені в табл. 1.

Аналіз показав, що для створення дослідницького прототипу пара ідентичних USB-камер є оптимальним вибором. На відміну від IP-камер, які вносять значні та непередбачувані затримки через мережеву передачу, USB-камери забезпечують стабільний відеопотік з мінімальною латентністю. Важливою перевагою є можливість програмного відключення автоматичних налаштувань (автоекспозиція, автофокус, баланс білого), що гарантує сталість параметрів зображення, необхідну для точного калібрування. Хоча глибинні сенсори надають готову карту глибини, їх висока вартість та залежність від пропрієтарних SDK роблять їх надлишковими для задач, де потрібен

Таблиця 1

Порівняння типів камер для реалізації стереозору

Критерій	IP-камери	USB-камери	Глибинні сенсори (RealSense, ZED тощо)
Затримка відео	Вища (залежить від мережі Wi-Fi)	Мінімальна	Мінімальна (апаратна обробка)
Стабільність кадрів / синхронізація	Низька, відсутня апаратна синхронізація	Вища, можливість апаратної синхронізації через USB	Вбудована апаратна синхронізація
Якість зображення	Залежить від моделі; часто автообробка на смартфоні	Залежить від моделі, стабільна за умов фіксації параметрів	Висока, спеціалізовані сенсори
Калібрування стереопари	Ускладнене (автокорекція, спотворення)	Просте (при відключенні автоекспозиції та автофокусу)	Не потребує (вбудоване позиціонування)
Гнучкість налаштувань	Середня (через web-інтерфейси), залежить від API	Висока (через прямий доступ до параметрів потоку)	Обмежена власними SDK
Простота підключення	Висока (по Wi-Fi, через IP-адресу)	Висока (Plug-and-Play через USB)	Потребує додаткового живлення, драйверів
Вартість	Дуже низька	Низька	Висока
Інтеграція з OpenCV / Python	Через MJPEG / HTTP потік, можлива затримка	Пряма підтримка, відеопотік стабільний	Потребує власного SDK або драйвера (RealSense SDK, ZED SDK)
Підтримка глибини / 3D-координат	відсутня	Потрібен обчислювальний модуль (disparity map)	Є (вбудована оцінка глибини, PointCloud, 3D frame)
Придатність до прототипування	Висока (швидкий запуск без додаткового обладнання)	Висока (особливо в лабораторних умовах)	Низька через ціну та обмежену доступність
Придатність до промислового застосування	Обмежена (залежність від мережі, нестабільність кадрів)	Висока (при правильному апаратному налаштуванні)	Висока (точність, надійність, промислові інтерфейси)

гнучкий контроль над усім процесом обчислення тривимірних координат.

Програмна реалізація та архітектура системи

Програмний комплекс розроблено на мові Python з використанням бібліотек для комп'ютерного зору. Бібліотека OpenCV виступає як основний інструмент для захоплення та обробки зображень, реалізації алгоритмів стереозору, калібрування камер та візуалізації результатів.

Для визначення об'єктів у реальному часі після розгляду альтернатив (YOLOv5, TensorFlow Object Detection API) було обрано фреймворк Ultralytics YOLOv8. Цей вибір обґрунтований його високою продуктивністю, точністю та повною реалізацією на Python, що спрощує інтеграцію. В системі використовується попередньо навчена модель yolov8n.pt, яка завантажується через API `ultralytics.YOLO()`. Модель послідовно застосовується до лівого та правого кадрів стереопари для виявлення об'єктів. Координати обмежувальних рамок, отримані від YOLOv8, використовуються для обчислення центрів об'єктів, їх зіставлення між кадрами та подальшого розрахунку диспаратності.

Процедура налаштування та калібрування стереосистеми

Для коректного обчислення глибини необхідним етапом є калібрування стереопари, що дозволяє усунути оптичні спотворення (дисторсію) та визначити точну взаємну орієнтацію камер. В експериментальній установці використовувались дві однакові USB-камери, закріплені на жорсткій основі на відстані близько 6 см між оптичними осями одна одної, спільною площиною кадру і роздільною здатністю кадру кожної камери 640×480 пікселів.

Калібрування здійснювалося за допомогою стандартної процедури з використанням шахового шаблону розміром 9×6 клітинок та функціоналу бібліотеки OpenCV (рис. 1).

Процес включав декілька етапів.

Етап 1. Визначення внутрішніх параметрів (intrinsics). Для кожної камери окремо за допомогою функції `cv2.calibrateCamera` були обчислені матриця внутрішніх параметрів (фокусна відстань, оптичний центр) та коефіцієнти дисторсії.

Етап 2. Стереокалібрування. На основі синхронізованих знімків шаблону з обох камер функція `cv2.stereoCalibrate` розрахувала параметри зовнішньої орієнтації – матрицю обертання R та вектор зсуву T , що описують просторове положення однієї камери відносно іншої.

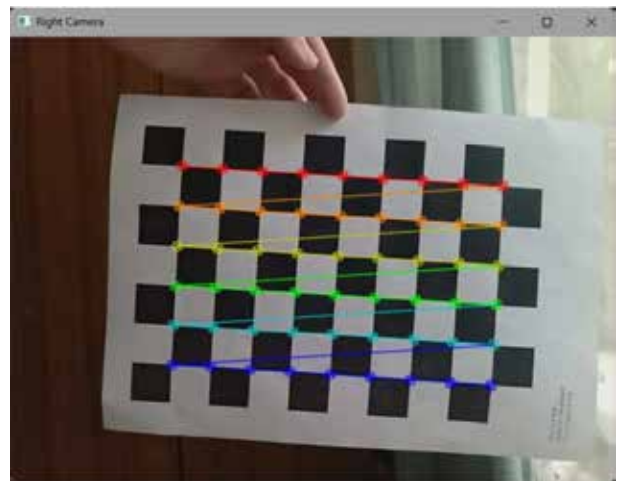


Рис. 1. Процес калібрування камери із використанням шаблону

Етап 3. Ректифікація. За допомогою `cv2.stereoRectify` були обчислені матриці ректифікації, які використовуються для проєктивного перетворення зображень. Цей процес вирівнює зображення таким чином, що епіполярні лінії стають паралельними і горизонтальними. Це гарантує, що відповідні точки об'єкта на лівому та правому зображеннях знаходяться на одній горизонтальній лінії, що значно спрощує пошук відповідностей та обчислення диспаратності.

Отримані калібрувальні параметри зберігаються у файл `stereo_calib.pkl` за допомогою модуля `pickle`, що дозволяє уникнути повторного калібрування при кожному запуску системи.

У реалізованій системі ректифікація зображень виконується в реальному часі перед передачею кадрів до нейронної мережі для розпізнавання. Це дозволяє гарантувати, що знайдені об'єкти будуть знаходитись на однаковій горизонтальній лінії, що є передумовою для точного обчислення диспаратності – різниці координат центрів одного й того самого об'єкта на лівому та правому зображенні.

Також під час калібрування було враховано: вплив освітлення на якість детекції кутів шахівниці; вимоги до стабільного положення камер протягом всієї сесії калібрування; а також проведено перевірку точності за допомогою функцій `cv2.undistort` та `cv2 remap`.

Стереокалібрування є важливим етапом побудови системи тривимірного розпізнавання об'єктів. Завдяки використанню інструментів OpenCV було отримано точні параметри внутрішньої та зовнішньої геометрії камер, що дозволило реалізувати подальше обчислення глибини та положення об'єктів у просторі. Вирівняні зображення, що утворюють епіполярну пару, забезпе-

чують необхідну точність для практичного застосування в робототехніці.

Розробка алгоритму визначення просторових координат об'єктів у стереозображенні

У цій роботі представлено алгоритм, що поєднує технології комп'ютерного зору, стереозору та глибокого навчання для визначення просторового положення об'єктів у робочій зоні промислового робота. Алгоритм реалізовано на основі бібліотек OpenCV, NumPy та нейромережевої моделі об'єктного детектування YOLOv8 від Ultralytics. Його архітектура включає послідовні етапи збору зображень, калібрування камер, ректифікації, виявлення об'єктів, співставлення їхніх проєкцій у стереопарі та обчислення просторових координат за допомогою триангуляції.

На початковому етапі здійснюється синхронне зчитування зображень з двох камер. Для цього використано шаблон шахівниці розміром 9×6 клітин, що дозволяє точно визначати внутрішні кути шаблону на зображеннях. Зображення фіксуються при різних просторових положеннях шахівниці в межах поля зору, але при цьому зберігається фіксована площа шаблону, що забезпечує достатньо варіативну й репрезентативну вибірку для подальшого калібрування. Отримані кадри автоматично розподіляються по окремих каталогах для лівої та правої камери, забезпечуючи узгодженість даних для подальшої обробки.

Після збору даних виконується внутрішнє калібрування кожної камери з використанням відомих метричних розмірів клітин шахівниці, які у даній реалізації становлять 25 міліметрів. Це дає змогу перейти від піксельних координат до фізичних одиниць виміру. За результатами калібрування для кожної камери обчислюються матриця внутрішніх параметрів та вектор дисторсії. Далі проводиться спільне стереокалібрування, в ході якого визначаються просторові трансформації між камерами у вигляді матриці обертання та вектора трансляції. Усі результати серіалізуються у форматі.pkl за допомогою модуля Python pickle, що дозволяє ефективно зберігати складні структури даних. Збережені параметри містять повний набір необхідної інформації для наступних етапів: внутрішні матриці камер, коефіцієнти спотворень, матриця обертання та вектор зміщення правої камери відносно лівої.

Наступним кроком є ректифікація отриманих зображень на основі результатів калібрування. Цей процес усуває лінзові спотворення та вирівнює зображення таким чином, щоб відповідні пари пікселів лежали на одній горизонтальній лінії.

Ректифікація дозволяє перевести вихідні зображення у координатну систему, зручну для обчислення диспаратності. Застосовано стандартні функції бібліотеки OpenCV: cv2.stereoRectify, cv2.initUndistortRectifyMap та cv2.remap, які забезпечують стабільну та відтворювану геометричну трансформацію.

До ректифікованих зображень окремо застосовується нейромережева модель YOLOv8, натренована на відповідному датасеті. Модель виконує виявлення об'єктів на зображеннях та повертає координати обмежувальних прямокутників (bounding boxes), ідентифіковані класи об'єктів та ймовірності розпізнавання. Обробка зображень відбувається незалежно для лівої та правої камери, що дозволяє уникнути впливу можливих локальних шумів або розсинхронізації між потоками. Результати виявлення передаються до наступного модуля зіставлення об'єктів.

Для кожного прямокутника, що відповідає знайденому об'єкту, розраховується центр за формулою:

$$x_c = \frac{x_{min} + x_{max}}{2}, y_c = \frac{y_{min} + y_{max}}{2}. \quad (1)$$

Це значно спрощує процес встановлення відповідностей між об'єктами на лівому та правому кадрах. Для кожного об'єкта виконується пошук пари на іншому зображенні за умови належності до одного класу, мінімальної відстані між центрами та допустимої диспаратності, що не перевищує встановлене порогове значення. Процедура зіставлення реалізована у вигляді окремої функції match_boxes(), яка забезпечує однозначне й стабільне парне зіставлення без перехресних відповідностей. Цей етап є критично важливим для подальших розрахунків, оскільки помилки на цьому рівні можуть суттєво вплинути на точність визначення відстані.

Після успішного встановлення відповідностей між об'єктами розпочинається етап обчислення просторової відстані. Використано класичну формулу триангуляції, що базується на визначенні диспаратності – горизонтального зсуву між проєкціями одного й того самого об'єкта:

$$Z = \frac{B * f}{d}, \quad (2)$$

де Z – відстань до об'єкта, B – база між камерами, f – фокусна відстань у пікселях, а d – різниця між x -координатами центрів об'єкта на лівому та правому зображеннях. Якщо диспаратність дорівнює нулю або надто мала, об'єкт вважається або занадто далеким, або некоректно співставле-

ним і не враховується у фінальних розрахунках. У випадках, коли для об'єкта на лівому кадрі не знайдено пари, система виводить відповідне повідомлення «No match».

Результати обчислення відображаються безпосередньо на зображеннях: біля кожного виявленого об'єкта зазначається відстань у метрах, а з центру прямокутника виводиться лінія до текстового підпису. Завдяки попередній ректифікації всі розрахунки відбуваються вздовж горизонтальної осі, що суттєво спрощує реалізацію алгоритму.

Точність визначення відстані залежить від якості калібрування, ректифікації, стабільності стереопари, а також точності визначення центрів прямокутників. Особливе значення має наявність достатньої диспаратності, що забезпечується за рахунок оптимального розміщення об'єктів у межах робочої зони камер.

Структура розробленої системи передбачає подальшу інтеграцію з модулями керування промисловим роботом. Дані про координати об'єктів, отримані в результаті алгоритмічного оброблення зображень, можуть бути передані до системи керування для розрахунку траєкторії руху маніпулятора з урахуванням кінематичних обмежень. Хоча модулі керування рухом і планування траєкторії не були реалізовані в межах цієї роботи, запропонована система передбачає таку можливість, що робить її перспективною для застосування в реальних роботизованих середовищах.

Експериментальне дослідження точності визначення просторового положення об'єктів

Для перевірки точності та оцінки працездатності розробленої системи було проведено експеримент, у межах якого реалізовано повний цикл: калібрування стереокамери, детекція об'єктів, їх співставлення на зображеннях та обчислення просторової відстані. Як стереопару використано дві USB-камери, закріплені з фіксованим базовим віддаленням.

Попередньо виконано стереокалібрування за шаховим шаблоном розміром 9×6 клітин. З обох камер здійснювався запис узгоджених зображень, після чого визначались внутрішні параметри кожної камери та спільні зовнішні параметри системи. Отримані матриці та вектори збережено у файл `stereo_calib.pkl`, який використовувався на подальших етапах.

Основна програма виконувала обробку потоку з камер у реальному часі. Спочатку зображення ректифікувались, після чого до кожного кадру застосовувалась нейронна модель YOLOv8n для виявлення об'єктів. На основі координат цен-

трів прямокутників об'єктів на лівому та правому зображеннях обчислювалась горизонтальна диспаратність. Враховуючи параметри стереопари, розраховувалась глибина, тобто відстань до об'єкта у метрах. Усі обчислення реалізовано на Python із використанням бібліотек OpenCV, NumPy та PyTorch. У графічному інтерфейсі виводились результати у вигляді класу об'єкта та його відстані.

У ході експерименту, проведеного при денному освітленні, перевірялась точність вимірювання для різних об'єктів: пляшки, руки людини та інших побутових предметів. Наприклад, при реальній відстані до пляшки близько 41,5 см система повідомляла значення в межах 38–42 см. Середнє відхилення складало близько 2 см, що є прийнятним результатом для недорогого обладнання. У випадку руки, що знаходилась на відстані 50–52 см, модель іноді повертала завищене значення (до 60 см), імовірно через особливості детекції YOLO: прямокутник формувався на всю руку, частково розташовану на межі кадру (рис. 2).

Найбільшу точність система демонструвала в діапазоні 0,5–3 м, що відповідає типовому робочому простору промислового маніпулятора. Незначні коливання значень у межах 1–2 см могли бути спричинені низькою якістю камер, неточним калібруванням, відсутністю жорсткої синхронізації та малою довжиною базової лінії (близько 6 см).

Точність системи істотно залежала від низки чинників. По-перше, якість стереокалібрування: недостатня кількість кадрів або неякісне виявлення кутів шаблону призводили до систематичних похибок, особливо у визначенні фокусної відстані чи бази. По-друге, апаратні характеристики камер: обмежена роздільна здатність (640×480), нестабільність автофокусу, розмиття та шуми ускладнювали точне позиціонування об'єктів.

Крім того, на точність впливала довжина базової лінії. При базі 6 см система показувала хорошу точність на малих відстанях, однак на більших – диспаратність зменшувалась до кількох пікселів, що ускладнювало вимірювання. Збільшення бази до 10–12 см дозволило б суттєво покращити чутливість.

Форма та положення об'єктів також були критичними. Центр прямокутника не завжди точно відповідає центру мас об'єкта, особливо у випадках складної геометрії чи часткового перекриття, що спричиняло хибні значення. У таких випадках доцільно використовувати методи сегментації для уточнення контурів.

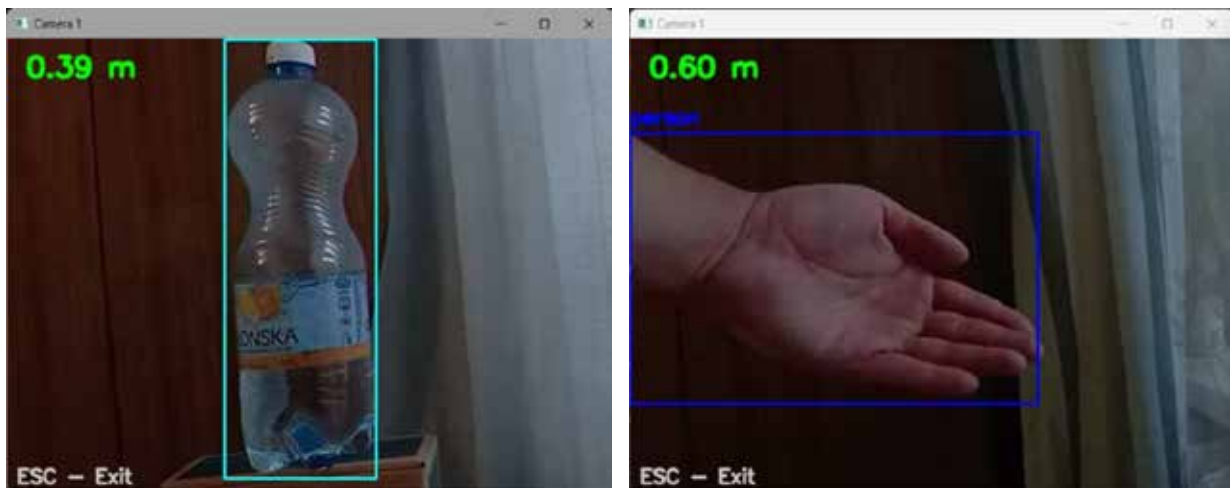


Рис. 2. Тестування роботи системи

На результати впливали також умови освітлення, стабільність відеопотоку та синхронізація кадрів. Варіації яскравості, тіні або відблиски викликали хибну детекцію або зсув bounding box. Для зменшення таких ефектів рекомендується контрольоване освітлення та попередня обробка зображень.

Таблиця 2

Основні параметри продуктивності системи

Метрика	Значення
FPS	15-20
Затримка на кадр	50-70 мс
Навантаження CPU	65-80%
Об'єм ОЗП	~750 МБ

Підсумовуючи, система забезпечила стабільну роботу при обмежених апаратних ресурсах та досягла прийнятної точності для задач середнього рівня складності. Отримані результати підтверджують ефективність запропонованого підходу до визначення просторового положення об'єктів із використанням комп'ютерного зору на основі стереозображень та глибокого навчання.

Висновки. У роботі представлено результати дослідження та реалізації системи визначення просторового положення об'єктів на основі стереопари USB-камер та моделі об'єктного детектування YOLOv8. Запропонований підхід забезпечує повний цикл обробки – від калібрування

камер до розпізнавання об'єктів, співставлення проєкцій та обчислення відстані до цілі – з прийнятною точністю для використання в дослідницьких, лабораторних і промислових умовах.

Ключовими перевагами системи є низька вартість апаратного забезпечення, використання доступних засобів (камер і комп'ютера), гнучкість налаштувань, а також можливість адаптації до різних прикладних сценаріїв. До недоліків слід віднести залежність точності від якості калібрування, роздільної здатності камер, умов освітлення та довжини базової лінії стереопари.

Подальший розвиток системи може бути спрямований на інтеграцію механізмів відстеження об'єктів у динаміці для забезпечення безперервного розпізнавання без повторної ідентифікації у кожному кадрі. Важливим кроком є також поєднання з платформою ROS та інтеграція з роботизованими маніпуляторами або мобільними платформами.

Застосування запропонованої системи можливе в низці практичних задач, включаючи розпізнавання жестів і положень рук у системах людино-машинної взаємодії, сервісну робототехніку, навчальні комплекси для вивчення комп'ютерного зору та роботизованих систем. Запропонований підхід є гнучкою та доступною основою для побудови інтелектуальних систем з візуальним зворотним зв'язком, здатних працювати в умовах реального середовища.

Список літератури:

1. OpenCV: Detection of ArUco Markers URL: https://docs.opencv.org/4.x/d5/dae/tutorial_aruco_detection.html (date of access: 24.06.2025).
2. Redmon J., Divvala S., Girshick R., Farhadi A. You Only Look Once: Unified, Real-Time Object Detection // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA. – 2016. – С. 779–788. – DOI: 10.1109/CVPR.2016.91.

3. Ren S., He K., Girshick R., Sun J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks URL: <https://dl.acm.org/doi/10.5555/2969239.2969250> date of access: 28.06.2025).
4. Carion N., Massa F., Synnaeve G., Usunier N., Kirillov A., Zagoruyko S. End-to-End Object Detection with Transformers. – 2020. – URL: <https://arxiv.org/abs/2005.12872> (date of access: 28.06.2025).
5. Kendall A., Cipolla R. PoseNet: A Convolutional Network for Real-Time 6-DOF Camera Relocalization URL: https://www.cv-foundation.org/openaccess/content_iccv_2015/papers/Kendall_PoseNet_A_Convolutional_ICCV_2015_paper.pdf (date of access: 28.06.2025).
6. Intel RealSense SDK 2.0 – Depth and Tracking cameras URL: <https://www.intelrealsense.com/sdk-2/> (date of access: 28.06.2025).
7. RTAB-Map – Real-Time Appearance-Based Mapping URL: Режим доступу: <https://introlab.github.io/rtabmap/> (date of access: 28.06.2025).
8. MoveIt 2 Documentation URL: <https://moveit.picknik.ai/main/index.html> (date of access: 04.07.2025).
9. Zhang C., Wang Q. A guidance system for robotic welding based on an improved YOLOv5 algorithm with a RealSense depth camera // Scientific Reports, 2023. URL: <https://www.nature.com/articles/s41598-023-48318-8> (date of access: 04.07.2025).
10. Song H., Zhang Y., Liu J., Li X., Wang Y. Study on the fusion of improved YOLOv8 and depth camera for bunch tomato stem picking point recognition and localization // Frontiers in Plant Science. – 2024. – URL: <https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2024.1447855/full> (date of access: 04.07.2025).
11. Wang W., Liu Y., Tan M. Transformers for Remote Sensing: A Systematic Review and Analysis // Sensors. – 2024. – V. 24, № 11. – P. 3495. – DOI: <https://doi.org/10.3390/s24113495>.

Latypov O.V., Yefymenko M.V., Myronova N.O., Molochkov D.Ye., Kulyaba-Kharytonova T.I.
DEVELOPMENT OF A SYSTEM FOR RECOGNISING THE POSITION OF OBJECTS
IN THE WORKING AREA OF AN INDUSTRIAL ROBOT

The article presents the study and implementation of a system for recognising the spatial position of objects in the working area of an industrial robot using a stereo pair of USB cameras and the YOLOv8 deep neural network. The relevance of the work is due to the need for accurate three-dimensional positioning of objects in real time for automated manipulation, assembly and interaction with humans in industrial and research environments. The object of research is the process of determining the spatial position of objects in the area of operation of an industrial robot. The subject of the study is algorithms and software and hardware for the implementation of a computer vision system based on stereo imaging and deep learning. The paper substantiates the choice of USB cameras as the optimal hardware for building a stereo pair: the cameras provide stable image transmission, low latency, and high flexibility of settings. The cameras were fully stereo-calibrated using a staggered template, which allowed us to determine the internal parameters, spatial geometry, and perform image rectification. The software part implements a full processing cycle: from synchronous frame capture to the output of object coordinates in three-dimensional space.

YOLOv8 is used to detect objects in both images, after which the scene depth is calculated using disparity analysis and triangulation. The experimental testing demonstrated stable operation of the system with an accuracy of 2-4 cm within the working range of up to 1 metre. The processing frequency was 15-20 FPS, the frame delay was 50-70 ms, and the CPU load was up to 80%. An analysis of the influence of factors revealed the critical role of calibration accuracy, video stream stability, frame synchronisation and baseline between cameras. The system provides visual control of the results, which ensures convenient testing and subsequent debugging. The obtained results confirm the feasibility of using a stereo pair of USB cameras and YOLOv8 to build an affordable and flexible system for 3D object localisation. The developed approach can be adapted for integration with ROS, mobile platforms, or robotic manipulators. Further research is planned to focus on implementing tracking in dynamics, optimising algorithms for edge devices, and extending the functionality by applying segmentation methods.

Key words: stereo vision, YOLOv8, computer vision, three-dimensional object localization, depth estimation, stereo camera pair, camera calibration, robotic systems, neural network.

Дата надходження статті: 24.07.2025

Дата прийняття статті: 29.07.2025

Опубліковано: 27.10.2025

Мовчан К.О.

Український науково-дослідного інститут спеціальної техніки та судових експертиз Служби безпеки України

КОМПЛЕКСНИЙ ПІДХІД ДО ЗАБЕЗПЕЧЕННЯ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ В ГЕТЕРОГЕННИХ ЕКОСИСТЕМАХ ІНТЕРНЕТУ РЕЧЕЙ

У статті розглядається проблематика забезпечення інформаційної безпеки в гетерогенних екосистемах Інтернету речей (IoT), які характеризуються експоненційним зростанням кількості пристроїв та їх інтеграцією в критичні інфраструктури. Акцентується увага на обмеженості обчислювальних ресурсів та енергетичної автономності IoT-пристроїв, що унеможливує застосування традиційних методів захисту та створює значні ризики для стабільності та безпеки систем у таких галузях, як енергетика, транспорт, охорона здоров'я та оборонно-промисловий комплекс.

Проаналізовано сучасні тенденції зростання кіберзагроз, зокрема атак із використанням штучного інтелекту, апаратних троянів та радіочастотних атак, які часто залишаються невиявленими через відсутність інтегрованих систем моніторингу. Окремо висвітлено вразливості хмарних інфраструктур, де відбувається обробка даних IoT-пристроїв, а також проблему фрагментації IoT-екосистеми, що ускладнює впровадження комплексних систем безпеки.

Визначено ключові вектори атак, що проявляються на різних рівнях п'ятирівневої архітектури IoT (перцепційний, мережевий, транспортний, прикладний та презентаційний). Детально розглянуто апаратні вразливості (апаратні «трояни», мікрозондові втручання, RF-атаки) на перцепційному рівні, мережеві атаки (DDoS/E-DDoS, MITM, sniffer-атаки) на мережевому та транспортному рівнях, а також програмні вразливості (переповнення буфера, hardcoded credentials, слабкі схеми автентифікації, інжекції коду) на рівні прикладної логіки.

Запропоновано багаторівневі контрзаходи, що включають криптографічний захист даних (AES-256, ECC, TLS/DTLS), використання систем виявлення вторгнень на базі глибоких згорткових нейронних мереж, honeypot-архітектури, автентифікацію за близькістю та інтеграцію систем динамічного управління патчами. Підкреслюється необхідність адаптивних підходів до безпеки в контексті специфічних IoT-впроваджень (розумні міста, медицина, аграрна та військова сфери). Підсумовується, що комплексний підхід до захисту IoT, який враховує мультишаровість систем, особливості застосування та обмеження обчислювальних ресурсів, є критично важливим. Формування чітко структурованої класифікації загроз у поєднанні з детальним описом механізмів реалізації атак та ефективних контрзаходів забезпечує комплексне бачення вразливостей IoT-систем, сприяє виявленню критичних точок впливу та створює основу для розробки стратегії проактивного захисту, що дозволяє прогнозувати розвиток загроз та проектувати адаптивні й стійкі до вторгнень IoT-рішення.

Ключові слова: інтернет речей, IoT захист, IoT протоколи, аналіз вразливості, запобігання атакам.

Постановка проблеми. Інтернет речей (IoT – Internet of Things) являє собою гетерогенну мережеву інфраструктуру, в якій фізичні об'єкти, оснащені сенсорами, виконавчими пристроями та програмним забезпеченням, інтегруються з глобальними мережами передачі даних за допомогою різноманітних бездротових протоколів для автономного обміну інформацією та взаємодії [1]. Згідно з аналітичними прогнозами, очікується, що до 2030 року кількість розгорнутих IoT-пристроїв досягне понад 30 мільярдів [2]. Їх системна інтеграція пошириться на сектори критичної

інфраструктури, що включають енергетичні системи, транспортні мережі, охорону здоров'я, оборонно-промисловий комплекс та аграрний сектор. Водночас, таке стрімке масштабування та ускладнення архітектури IoT-екосистем неминуче призводить до експоненційного зростання поверхні атаки [3]. Це, у свою чергу, провокує зростання кіберзагроз, серед яких розподілені атаки на відмову в обслуговуванні (DDoS), несанкціонований доступ до даних через вразливості вбудованого програмного забезпечення та несанкціонований доступ до пристроїв на фізичному рівні, що ство-

рює значні ризики для стабільності та безпеки зазначених критичних галузей.

IoT-пристрої, у порівнянні з класичними комп'ютерними системами, мають обмежені обчислювальні ресурси, енергетичну автономність, і часто проєктуються з пріоритетом на економічну ефективність, а не на інформаційну безпеку [4]. Ці обмеження унеможливають використання традиційних методів захисту, що призводить до створення критично уразливого інформаційного середовища. Як наслідок, атаки на IoT можуть мати катастрофічні наслідки: порушення роботи медичних пристроїв, втручання в транспортні системи (наприклад, у протоколи V2V/V2I), саботаж енергомереж, або витік персональних даних мільйонів користувачів.

Станом на 2024 рік фіксується тенденція до зростання атак на основі штучного інтелекту, в використанні можливості автоматичного сканування мережі, здійснення адаптивних DDoS-атаки або можливості обходити класичні механізми автентифікації [5]. Згідно з доповіддю ENISA (2024), близько 67% IoT-інцидентів пов'язані з несанкціонованим доступом до пристроїв через слабкі або відсутні протоколи автентифікації. Значна частина IoT-пристроїв має відкриті порти, які не блокуються за замовчуванням.

Особливу загрозу становить використання апаратних троянів (Hardware Trojans) у фазі виробництва, які є маловиявляемі і можуть активуватися лише за специфічних умов [6]. Це відкриває потенціал для державного або корпоративного шпionaжу. Не менш небезпечними є атаки через радіочастотні канали (RF Attacks), які дозволяють перехоплювати та інjektувати сигнали без фізичного доступу до пристрою. При цьому, більшість сучасних пристроїв не має інтегрованих систем виявлення RF-перешкод.

Окрему категорію загроз становлять уразливості в хмарних інфраструктурах (Cloud Layer Vulnerabilities), де IoT-пристрої часто обробляють дані без шифрування, що надає противнику змогу перехоплювати трафік, вбудовувати шкідливий код або підмінювати відповіді серверів. Атаки типу Man-in-the-Middle та Packet Injection є особливо небезпечними в середовищах, де використовуються слабо захищені протоколи на кшталт MQTT без TLS.

Проблема також ускладнюється фрагментацією IoT-екосистеми – різні стандарти, протоколи (наприклад, ZigBee, BLE, LoRa, SigFox, NB-IoT), а також відсутність уніфікованих політик оновлення мікропрограмного забезпечення створюють

серйозні бар'єри для впровадження комплексних систем безпеки.

У контексті оборонних технологій (IoBT – Internet of Battlefield Things) недоліки в захисті можуть не лише створити загрозу даним, а й призвести до фізичної загибелі солдат або провалу операцій. Наприклад, уразливі військові сенсори або дрони можуть бути віддалено перехоплені чи виведені з ладу через сигнальні атаки.

Таким чином, проблема забезпечення надійної інформаційної безпеки IoT є міждисциплінарною, комплексною і потребує негайного вирішення. Відсутність глибокої формалізації ризиків і недосконалість існуючих контрзаходів лише підкреслюють нагальність досліджень у цьому напрямі.

Аналіз останніх досліджень і публікацій. Як правило дослідження у сфері IoT-безпеки фокусуються на вузькоспеціалізованих аспектах – зокрема, шифруванні протоколів (наприклад, MQTT, CoAP), IDS-системах на базі машинного навчання, або уразливостях фізичних сенсорів [7]. У публікації приділено увагу загрозам у smart home-середовищах, уразливості хмарних платформ, а також питанням реінжинірингу (reverse engineering) в IoT-пристроях [8]. Проте більшість досліджень мають обмежений оглядовий характер і не охоплюють повний спектр уразливостей із системним категоризуванням.

Сучасні дослідження все частіше вказують на тенденцію до застосування системного аналізу безпеки екосистеми Інтернету речей [9]. Сучасні системи використовують багаторівневі структури моделей для категоризації вразливостей за типом: апаратні (hardware), програмні (software), мережеві (network) та хмарні (cloud). Такий підхід дозволяє не лише ідентифікувати, але й систематизувати загрози. Наприклад, дослідження 2024 року підкреслюють, що ключовими векторами атак залишаються переповнення буфера (28.25% виявлених вразливостей) та атаки на відмову в обслуговуванні (27.20%) [10]. На основі цієї класифікації розробляються матриці атак, що зіставляють конкретні типи загроз із відповідними превентивними та реактивними заходами. Важливим елементом цих контрзаходів є активне застосування методів штучного інтелекту, зокрема для проактивного виявлення аномалій у поведінці пристроїв та автоматизації реагування на інциденти, що значно підвищує точність ідентифікації загроз та мінімізує кількість хибних спрацьовувань у порівнянні з традиційними системами безпеки.

Постановка завдання. Систематизація наявних векторів атак на IoT-системи та аналіз техно-

логії захисту з метою зниження площини атаки є одним із актуальних завдань у сфері кіберзахисту IoT-систем. Основна увага дослідження акцентується на важливості формалізованої архітектури п'ятирівневого IoT (перцепційний, доступу, передачі, прикладної підтримки та презентаційний шари), в якій і проявляються ключові вектори атак. Актуальність проблеми зумовлена швидким впровадженням IoT-пристроїв у критично важливі інфраструктури – від військових до медичних – де несанкціонований доступ навіть до одного вузла може мати катастрофічні наслідки [11]. Метою даного дослідження є систематизація та класифікація вразливостей IoT-екосистем, аналіз механізмів реалізації атак та обґрунтування ефективних контрзаходів для зниження загальної поверхні атак.

Виклад основного матеріалу. Для забезпечення ефективного аналізу загроз у середовищі Інтернету речей (IoT) необхідним є системний розгляд архітектурної побудови IoT-систем з подальшою класифікацією векторів атак, що проявляються на різних технологічних рівнях. У контексті сучасної п'ятирівневої архітектури, яка включає перцепційний, мережевий, транспортний, прикладний та презентаційний рівні, виникає необхідність ідентифікації типових вразливих точок, характерних для кожного з рівнів [12].

На **перцепційному рівні** функціонують сенсорні модулі, RFID-мітки, камери та інші пристрої збору первинних даних. Саме тут закладається основа IoT-екосистем, однак одночасно формується й найбільш критичне середовище для атак фізичного доступу. **Апаратні вразливості**, характерні для цього рівня, охоплюють як навмисну модифікацію схемотехніки (впровадження апаратних “троянів”), так і високотехнологічні засоби аналізу, зокрема мікрозондові втручання, електромагнітне зчитування та реінжиніринг мікросхем. Особливо небезпечними є **RF-атаки**, що здійснюються шляхом пасивного аналізу сигналів та активного перехоплення, наприклад, під час перепідключення Bluetooth-пристроїв або підміни сигналів BLE через атаки типу spoofing. За даними дослідників, такі вектори атак часто залишаються невиявленими через обмежені ресурси обчислювальних блоків і відсутність моніторингових підсистем на фізичному рівні.

На **мережевому та транспортному рівнях** реалізуються протоколи зв'язку – від класичних WiFi та Ethernet до спеціалізованих рішень ZigBee, LoRaWAN, BLE, 5G-NR, NB-IoT [13]. Ці технології забезпечують взаємодію пристроїв

та передачу даних, однак вони є вразливими до **мережевих атак**. Зокрема, поширеними загрозами залишаються атаки типу **DDoS/E-DDoS**, які спрямовані на перевантаження ресурсів пристрою чи його енергетичного модуля. Крім того, **атаки типу «людина посередині» (MITM)**, а також **sniffer-атаки** на основі аналізу нешифрованого трафіку (особливо у застарілих реалізаціях WPA/WEP) створюють прямий вектор несанкціонованого зняття інформації. За останніми дослідженнями, розповсюдженими також є **over-the-air атаки**, які поєднують методи радіочастотного моніторингу з машинним навчанням для ідентифікації пристроїв та сценаріїв їх взаємодії у мережі. Застосування нейронних мереж для моделювання мережевого трафіку дозволяє не лише виявляти активні вузли, а й передбачати їх поведінкові шаблони, що відкриває простір для здійснення прогнозованих атак.

На рівні **прикладної логіки** загрозу становлять **програмні вразливості**, до яких належать:

- **переповнення буфера**, що відкриває шлях до виконання шкідливого коду;
- **hardcoded credentials** – типова помилка розробників, яка створює бекдор-доступ до пристрою;
- **слабкі схеми автентифікації**, включно з відсутністю багатофакторної перевірки;
- **несанкціоноване оновлення мікропрограмного забезпечення**, що допускає впровадження модифікованого програмного забезпечення;
- **інжекції коду** (наприклад, через IoT API або веб-інтерфейси).

Особливої уваги заслуговують ситуації, коли пристрій фізично доступний зловмиснику: у такому випадку відновлення прошивки, доступ до Hex-даних та декомпіляція бінарних образів за допомогою таких програмних продуктів, як Ghidra або IDA Pro, дають змогу реконструювати логіку роботи пристрою, що у свою чергу, дозволяє впровадити backdoor-логіку або змінити потік керування без виявлення порушень функціонування системи з боку користувача.

У світлі вищеописаних ризиків, потреба в **багаторівневих контрзаходах** є нагальною. Найбільш ефективними заходами вважаються:

- **криптографічний захист** даних – AES-256, ECC, TLS/DTLS;
- **використання систем виявлення вторгнень (IDS)** на базі глибоких згорткових нейронних мереж, які забезпечують аналіз трафіку в реальному часі;

– **honeypot-архітектури**, які діють як пастки для виявлення ботів та активних атак;

– **автентифікація за близькістю (proximity-based authentication)**, яка забезпечує захист на фізичному рівні [14].

Особливу увагу слід приділяти інтеграції систем **динамічного управління патчами (dynamic patch management)**, які дозволяють автоматично і своєчасно оновлювати прошивку без втрати цілісності або функціональності пристрою. В цьому сенсі застосування цифрових підписів та контрольованих каналів доставки оновлень – критично необхідна умова безпечної експлуатації.

Нарешті, специфіка IoT-впроваджень у критичних середовищах – таких як **розумні міста (smart cities)**, **медицина (IoMT)**, **аграрна галузь (IoAT)** та **військова сфера (IoBT)** – потребує адаптивних підходів до безпеки [15]. Наприклад, для розгортання IoBT критично важливо забезпечити захист сенсорів на полі бою від фізичних і електромагнітних загроз, а також застосовувати надійні протоколи обміну даними, які запобігають їх перехопленню. IoMT-пристрої потребують сертифікації відповідно до медичних протоколів і стандартів HIPAA/GDPR, а в IoAT – особливу

увагу слід приділяти захисту телеметрії з віддалених ділянок.

Таким чином, захист IoT-інфраструктури потребує інтегрованого підходу, що враховує мультишаровість систем, особливості застосування та обмеження обчислювальних ресурсів. Реалізація таких стратегій можлива лише за умов формалізованої політики безпеки, тісної співпраці виробників з експертами з кібербезпеки та постійного моніторингу загроз.

Висновки. Проведене дослідження виявляє необхідність комплексного підходу до захисту IoT, що дозволяє оцінити вразливості в контексті всього ланцюга передачі даних. Формування чітко структурованої класифікації загроз у поєднанні з детальним описом механізмів реалізації атак та ефективних контрзаходів забезпечує комплексне бачення вразливостей IoT-систем, сприяє виявленню критичних точок впливу та створює основу для розробки стратегії проактивного захисту. Такий підхід дозволяє не лише ідентифікувати потенційні вектори атак на різних рівнях архітектури, а й прогнозувати їх розвиток у контексті динамічної еволюції загроз, що особливо важливо для проектування адаптивних і стійких до вторгнень IoT-рішень.

Список літератури:

1. Atzori L., Iera A., Morabito G. The Internet of Things: A survey. *Computer Networks*, 54(15), 2010, pp. 2787-2805.
2. Statista Research Department. Internet of Things – number of connected devices worldwide, Statista, pp. 2019-2030, 2024.
3. Zhang Y., Xu C., Li H., Yang K., Zhou J., Lin X. HealthDep: An efficient and secure deduplication scheme for cloud-assisted eHealth systems. *IEEE Transactions on Industrial Informatics*, 14(9), 2018, pp. 4101-4112.
4. Roman R., Zhou J., Lopez J. On the features and challenges of security and privacy in distributed internet of things. *Computer Networks*, 57(10), 2013, pp. 2266-2279.
5. Kumar S., Tiwari P., Zymbler M. Internet of Things is a revolutionary approach for future technology enhancement. *Journal of Big Data*, 6(1), 2019, pp. 1-21.
6. Tehranipoor M., Koushanfar F. A survey of hardware trojan taxonomy and detection. *IEEE Design & Test of Computers*, 27(1), 2010, pp. 10-25.
7. Granjal J., Monteiro E., Silva J.S. Security for the internet of things: a survey of existing protocols and open research issues. *IEEE Communications Surveys & Tutorials*, 17(3), 2015, pp. 1294-1312.
8. Zhao K., Ge L. A survey on the internet of things security. *Ninth International Conference on Computational Intelligence and Security*, 2013, pp. 663-667.
9. Sicari S., Rizzardi A., Grieco L.A., Coen-Porisini A. Security, privacy and trust in Internet of Things: The road ahead. *Computer Networks*, 76, 2015, pp. 146-164.
10. Li S., Xu L.D., Zhao S. The internet of things: a survey. *Information Systems Frontiers*, 17(2), 2015, pp. 243-259.
11. Burhan M., Rehman R.A., Khan B., Kim B.S. IoT elements, layered architectures and security issues: A comprehensive survey. *Sensors*, 18(9), 2018, p. 2796.
12. Al-Fuqaha A., Guizani M., Mohammadi M., Aledhari M., Ayyash M. Internet of things: A survey on enabling technologies, protocols, and applications. *IEEE Communications Surveys & Tutorials*, 17(4), 2015, pp. 2347-2376.
13. Raza S., Wallgren L., Voigt T. SVELTE: Real-time intrusion detection in the Internet of Things. *Ad Hoc Networks*, 11(8), 2013, pp. 2661-2674.

14. Zhang Z.K., Cho M.C.Y., Wang C.W., Hsu C.W., Chen C.K., Shieh S. IoT security: ongoing challenges and research opportunities. IEEE 7th International Conference on Service-Oriented Computing and Applications, 2014, pp. 230-234.
15. Whitmore A., Agarwal A., Da Xu L. The Internet of Things—A survey of topics and trends. Information Systems Frontiers, 17(2), 2015, pp. 261-274.

Movchan K.O. A COMPREHENSIVE APPROACH TO ENSURING INFORMATION SECURITY IN HETEROGENEOUS INTERNET OF THINGS ECOSYSTEMS

The article discusses the issue of ensuring information security in heterogeneous Internet of Things (IoT) ecosystems, which are characterised by exponential growth in the number of devices and their integration into critical infrastructures. It highlights the limited computing resources and energy autonomy of IoT devices, which makes it impossible to use traditional protection methods and creates significant risks to the stability and security of systems in sectors such as energy, transport, healthcare, and the defence industry.

The paper analyses current trends in the growth of cyber threats, in particular attacks using artificial intelligence, hardware Trojans and radio frequency attacks, which often remain undetected due to the lack of integrated monitoring systems. The vulnerabilities of cloud infrastructures, where IoT device data is processed, are highlighted separately, as well as the problem of fragmentation of the IoT ecosystem, which complicates the implementation of comprehensive security systems.

Key attack vectors are identified that manifest themselves at different levels of the five-level IoT architecture (perceptual, network, transport, application, and presentation). Hardware vulnerabilities (hardware Trojans, microprobe interference, RF attacks) at the perception level, network attacks (DDoS/E-DDoS, MITM, sniffer attacks) at the network and transport levels, as well as software vulnerabilities (buffer overflow, hardcoded credentials, weak authentication schemes, code injection) at the application logic level.

Multi-level countermeasures are proposed, including cryptographic data protection (AES-256, ECC, TLS/DTLS), the use of intrusion detection systems based on deep convolutional neural networks, honeypot architectures, proximity authentication, and the integration of dynamic patch management systems. The need for adaptive approaches to security in the context of specific IoT implementations (smart cities, medicine, agriculture, and military) is emphasised. It is concluded that a comprehensive approach to IoT protection, which takes into account the multi-layered nature of systems, application characteristics, and computing resource limitations, is critically important. The formation of a clearly structured classification of threats, combined with a detailed description of attack mechanisms and effective countermeasures, provides a comprehensive view of the vulnerabilities of IoT systems, helps identify critical points of impact, and creates a basis for developing a proactive defence strategy that allows for the prediction of threat development and the design of adaptive and intrusion-resistant IoT solutions.

Key words: Internet of Things, IoT protection, IoT protocols, vulnerability analysis, attack prevention.

Дата надходження статті: 28.07.2025

Дата прийняття статті: 08.08.2025

Опубліковано: 27.10.2025

Нікітченко М.І.

Національний університет «Одеська політехніка»

АРХІТЕКТУРА ТА РЕАЛІЗАЦІЯ ПЛАГІНА ІНКРЕМЕНТАЛЬНОЇ ВАЛІДАЦІЇ UML-МЕТАМОДЕЛЕЙ З ДВОМА ПОДАННЯМИ

У статті розглянуто проблему забезпечення узгодженості UML-моделей у процесі розробки складних програмних систем, актуальність якої зумовлена високою вартістю виявлення архітектурних дефектів на пізніх етапах життєвого циклу. Сучасні CASE-інструменти моделювання пропонують або гнучкість формату JSON для статичних структур з низьким рівнем формалізму, або строгість формату XMI, що важко інтегрується у повсякденну розробку через складність обробки. Водночас жоден із цих підходів не забезпечує повного поєднання переваг гнучкості й формалізму, що створює потребу в розробці гібридного рішення.

Метою роботи є представлення архітектури та практичної реалізації плагіна, який інтегрує два формати представлення UML-моделей – легковагий JSON для опису статичної структури та формально строгий XMI для опису динаміки системи. Особливістю плагіна є застосування спеціалізованого UML-профілю для фізичної реалізації відношення відповідності між цими поданнями, що дозволяє здійснювати швидку інкрементальну валідацію через граф залежностей. Методологія включає поєднання синтаксичної перевірки JSON Schema, семантичної інкрементальної валідації та глибокої формальної верифікації транзитивних властивостей за допомогою SAT-аналізу.

Результати експериментальної оцінки показали, що запропонований плагін значно ефективніший за існуючі аналоги, оскільки інкрементальна валідація зменшує час перевірки порівняно з повною перевіркою моделей. Таким чином, забезпечується можливість ефективної інтеграції у конвеєри безперервної інтеграції та доставки, що робить інструмент придатним для використання у режимі реального часу. Висновки дослідження підкреслюють перспективність гібридного підходу до зберігання UML-моделей та визначають напрями для подальшого розвитку, включаючи підтримку інших мов програмування та розширення функціональності плагіна.

Ключові слова: UML, узгодженість моделей, інкрементальна валідація, JSON Schema, XMI, граф залежностей, SAT-аналіз.

Постановка проблеми. Сучасна програмна інженерія стикається з постійним зростанням складності систем, що вимагає ефективних підходів до моделювання їхньої архітектури. В контексті Model-Driven Engineering розроблено багато методик для забезпечення консистентності артефактів [1], але на практиці часто виникає поступове розходження між проектною моделлю і документацією, та її реальною імплементацією у вихідному коді. В результаті, цей розрив породжує клас прихованих архітектурних дефектів – семантичних помилок, які не виявляються стандартними компіляторами чи валідаторами. Прикладами можуть бути виклик методу, що був перейменований у коді, але зберіг стару назву в діаграмі послідовності, або використання застарілого типу даних у поведінковому сценарії, або ж навіть суперечності між діаграмами [2]. Виявлення таких дефектів на пізніх етапах життєвого

циклу є на порядки дорожчим, що робить проблему забезпечення консистентності ключовим фактором економічної ефективності розробки.

В основі цієї проблеми лежить фундаментальний вибір формату для зберігання UML-моделей. Жоден з існуючих стандартів не надає універсального рішення, змушуючи команди робити компроміс між формальною строгістю та гнучкістю розробки.

XMI (XML Metadata Interchange) є офіційним стандартом OMG, що гарантує максимальну повноту та формальну точність опису, що особливо критично для складних поведінкових діаграм (станів, послідовностей, діяльності), а також для забезпечення сумісності з існуючими CASE-інструментами та можливість застосування формальних обмежень мовою OCL. Однак його надлишковий синтаксис робить XMI-файли громіздкими, складними для ручного аналізу та

обробки в системах контролю версій, таких як Git, де навіть незначні структурні зміни призводять до значних конфліктів злиття.

JSON (JavaScript Object Notation), навпаки, є легковим, текстовим і зручним для розробників форматом, що ідеально підходить для опису статичних структур: класів, атрибутів, пакетів. Його простота забезпечує швидку обробку разом з меншими затратами пам'яті та легку інтеграцію з інструментами розробки. Проте JSON не має вбудованих механізмів для представлення складної поведінкової семантики, крос-посилань між елементами та формальних обмежень, що є сильною стороною XMI. Саме тому спроба описати динаміку системи в чистому JSON призводить або до втрати семантики, або до створення переобтяжених документів.

Виникає глибокий розрив між теоретичними моделями узгодженості, що вимагають формалізму, та реальними інструментами, які використовуються в щоденній практиці розробки. Існуючі CASE-засоби змушують команди робити компромісний вибір між формалізмом та гнучкістю, не пропонуючи нативного рішення для гібридного зберігання моделей, яке б поєднувало найкращі риси обох світів.

Ключова проблема полягає в тому, що виявлення цих семантичних розривів на пізніх етапах життєвого циклу (під час інтеграційного тестування або, що найгірше, вже в експлуатації) є на порядки дорожчим, ніж на етапі проектування. Тому проблема забезпечення консистентності є не лише технічним питанням якості, а й ключовим фактором економічної ефективності та керованості процесу розробки.

Аналіз останніх досліджень і публікацій. Для обґрунтування актуальності та новизни запропонованого рішення було проведено детальний аналіз сучасних CASE-засобів та форматів зберігання UML-моделей [3, 4]. Він показав, що ринок інструментів моделювання фактично розділився на два полюси: зручні для розробника, але формально слабкі інструменти, що використовують JSON, та формально строгі, але громіздкі й негнучкі інструменти, що базуються на XMI. Тим не менш, жоден із них не пропонує комплексного рішення для одночасної підтримки гнучкості та семантичної повноти.

Ця робота є прямим практичним втіленням та емпіричною перевіркою теоретичних засад, викладених у серії попередніх статей. Вона завершує цикл досліджень, переводячи абстрактні концепції в площину конкретного програмного

інструменту. Перша робота формалізувала метамодель з двома поданнями, а також було введено систему формальних інваріантів для перевірки узгодженості моделі [5]. У другій роботі було деталізовано практичну реалізацію структурного подання через JSON Schema [6]. Було визначено набір інваріантів для валідації статичної структури, частина з яких реалізується безпосередньо засобами схеми (наприклад, перевірка типів, обов'язкових полів), а частина (наприклад, ациклічність) потребує зовнішніх валідаторів. У третій роботі було розроблено спеціалізований UML-профіль для поведінкового подання, який дозволяє фізично реалізувати відношення зв'язку через теги в XMI [7]. Також було представлено теоретичний алгоритм інкрементальної валідації, адже інкрементальне поширення змін є ефективнішим за повну повторну трансформацію моделей, особливо для великомасштабних систем [8].

Аналіз існуючих рішень та теоретичних робіт дозволяє виділити ключові проблеми, які вони не вирішують комплексно.

– **Відсутність нативної мультиформатної архітектури:** Жоден з інструментів не побудований на ідеї одночасного використання JSON для структури та XMI для поведінки як основи для роботи з єдиною, цілісною моделлю. Інструменти або обирають один формат, або пропонують ручну конвертацію.

– **Неефективність валідації:** Перевірка узгодженості, якщо вона взагалі є, виконується для всієї моделі. Ідея ефективної інкрементальної валідації на основі графа залежностей, що охоплює лише змінені частини та їхній вплив, не реалізована в аналогах, що робить їх непрактичними для великих проєктів у режимі реального часу.

– **Слабкий зв'язок «модель-код»:** Існуючі інструменти є переважно засобами візуального моделювання, відірваними від вихідного коду. Вони не забезпечують безперервну, фонову валідацію моделі, що автоматично генерується з коду, як це пропонується в плагіну.

– **Складність інтеграції в автоматизовані процеси:** Відсутність простого та потужного Command Line Interface (CLI) ускладнює вбудовування перевірки консистентності в автоматизовані конвеєри збірки та тестування.

В результаті, ринок інструментів моделювання фактично розділився на два незадовільні полюси: зручні для розробника, але формально слабкі інструменти, та формально строгі, але незручні для розробника інструменти. Це створює значну нішу для інструменту, який би синтезував сильні

сторони обох підходів. Розроблений плагін є саме таким архітектурним рішенням, що поєднує зручність JSON з формалізмом XMI через потужний фокус на узгодженості.

Постановка завдання. Метою статті є представлення архітектури, практичної реалізації та експериментальної оцінки розроблюваного плагіна, що є програмним втіленням розробленої раніше теоретичної моделі та призначений для розв'язання вищезазначених проблем розриву між гнучкістю розробки та формальною строгістю моделювання.

Для досягнення цієї мети поставлено наступні завдання:

- описати високорівневу архітектуру плагіна, яка реалізує теоретичну метамодель через модульну структуру та продемонструвати потік даних від вихідного коду до артефактів моделі;
- продемонструвати ключові механізми реалізації, в тому числі автоматичну побудову внутрішньої моделі з вихідного коду Python, фізичну реалізацію відношення відповідності μ через спеціалізований UML-профіль та теги JsonRef у XMI, що забезпечує трасування між поданнями, а також практичну роботу інкрементальної валідації на основі графа залежностей G_{MB} , що розрізняє структурні, семантичні та трасувальні зв'язки;
- навести інструкцію з використання плагіна та продемонструвати його роботу на контрольному прикладі;
- провести експериментальну оцінку продуктивності, вимірявши час повної та інкрементальної валідації, а також використання пам'яті, і порівняти отримані результати з аналогами.

Виклад основного матеріалу. В основі плагіна лежить формальна UML-метамодель $M = (M_S, M_B, \mu)$, детально описана в [5]. Ця модель свідомо розділяє єдину концептуальну модель на два подання, щоб поєднати переваги різних форматів і вирішити фундаментальний конфлікт між гнучкістю та строгістю.

– Структурне подання (M_S): Це опис статичної архітектури системи (класів, атрибутів, операцій, пакетів тощо). Воно зберігається у форматі JSON.

– Поведінкове подання (M_B): Це опис динаміки системи (діаграм станів, послідовностей, діяльності). Воно зберігається у стандартному для UML форматі XMI.

– Відношення відповідності (μ): Це ключовий механізм, що слугує зв'язком між двома поданнями. Воно формально пов'язує елементи поведінки з їхніми структурними аналогами.

Така архітектура дозволяє виконувати ефективну інкрементальну валідацію. Після зміни у вихідному коді, яка впливає на M_S , система може швидко визначити через μ та граф залежностей G_{MB} вузьке коло елементів в M_B , які потребують повторної перевірки, замість того, щоб аналізувати всю модель цілком.

Сам плагін побудований на основі модульної архітектури, розробленої для чіткого розподілу відповідальності та для забезпечення високої розширюваності. Центральним елементом архітектури є внутрішня модель: набір Python-об'єктів (dataclasses), що представляють елементи UML в пам'яті. Ця модель слугує проміжним шаром, який повністю ізолює компоненти один від одного.

Нижче наведено модулі системи, згідно зі структурою проекту.

Модуль *parsers* відповідає за аналіз вихідного коду.

Модуль *core* містить ядро системи. В його складі можна виділити внутрішні моделі даних, реалізацію графа залежностей та логіку персистентності.

Модуль *generators* відповідає за генерацію моделей, як для структурної частини, так і для поведінкової.

Модуль *validators* містить такі компоненти як валідаційний движок та зв'язок із SAT-вирішувачем.

Модуль *integration* забезпечує зовнішню інтеграцію за рахунок CLI та IDE API.

Потік даних у системі організовано наступним чином. Спершу плагін отримує вихідний код на мові Python і для кожного файлу будує абстрактне синтаксичне дерево (Abstract syntax tree, AST) за допомогою вбудованого модуля ast.

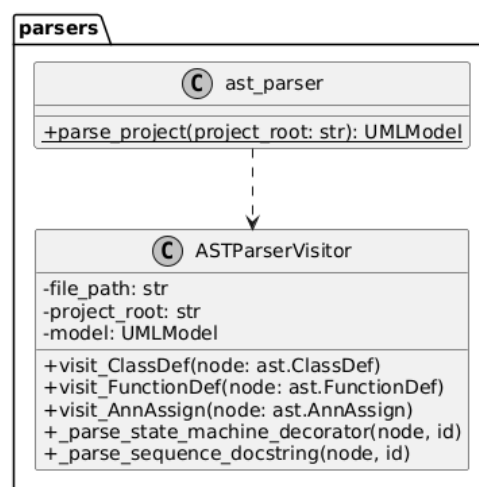


Рис. 1. Модуль parsers

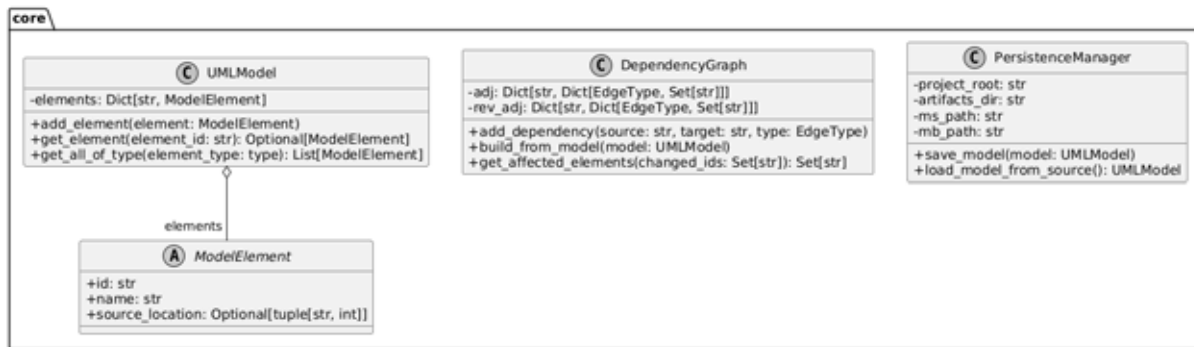


Рис. 2. Модуль core

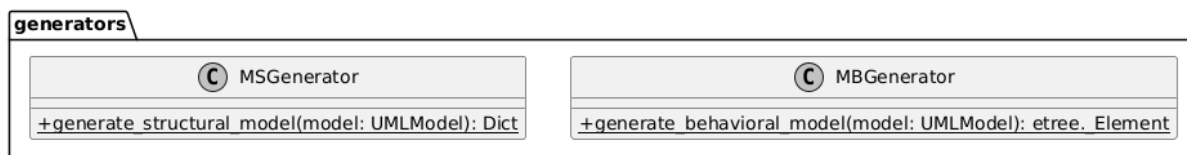


Рис. 3. Модуль generators

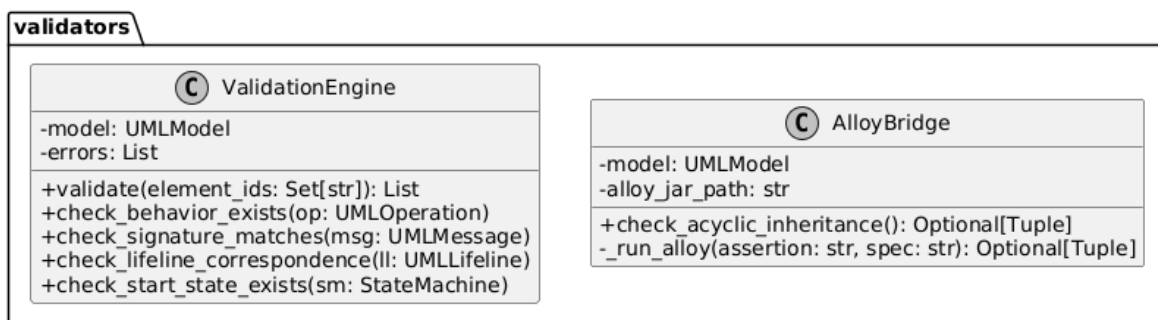


Рис. 4. Модуль validators

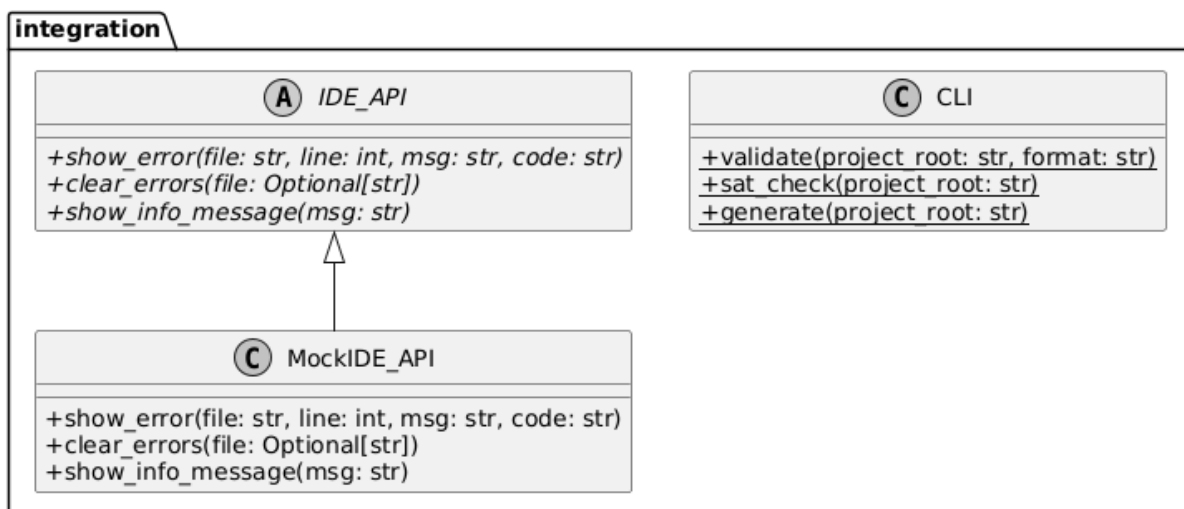


Рис. 5. Модуль integration

Потім спеціалізований клас `ASTParserVisitor` рекурсивно обходить AST. При відвідуванні кожного вузла (наприклад, `ast.ClassDef` для класів, `ast.FunctionDef` для методів) він витягує необхідну інформацію та створює екземпляри внутрішніх моделей даних – об'єктів з `core/models.py`. Одночасно з наповненням внутрішньої моделі будується граф залежностей G_{MB} , який фіксує всі структурні, семантичні та трасувальні зв'язки між елементами моделі. Генератори (`ms_generator`, `mb_generator`) обходять наповнену внутрішню модель і серіалізують її у фінальні файли: `model.struct.json` та `model.behav.xmi`, які зберігаються в директорії `uml-consistency/`.

Внутрішня модель, реалізована на основі `dataclasses`, є архітектурним стрижнем плагіна. Вона діє як шар абстракції, що повністю роз'єднує логіку аналізу вихідного коду від логіки генерації цільових форматів. Наприклад, `ASTParserVisitor` знає лише, як перетворити вузли AST у ці `dataclasses`, і нічого не знає про JSON чи XMI. Аналогічно, генератори працюють виключно з цими об'єктами, не маючи уявлення про Python чи AST. Таке роз'єднання забезпечує високу розширюваність: для підтримки нової мови програмування (наприклад, Java) достатньо буде реалізувати новий модуль-парсер, який би наповнював ту саму внутрішню модель, не змінюючи решту системи.

В якості основної мови розробки було обрано Python як через його динамічну природу, потужні засоби інтроспекції та, що найголовніше, наявність вбудованого модуля `ast`, який є ідеальною основою для статичного аналізу коду. Для генерації XMI використовується галузевий стандарт – бібліотека `lxml`, а для валідації JSON – `jsonschema`.

Робота плагіна починається з глибокого аналізу коду. `ASTParserVisitor` реалізує патерн `Visitor` для обходу дерева AST. Наприклад, при відвідуванні вузла `ast.ClassDef` з декоратором `@state_machine(initial_state=States.IDLE)`, парсер не лише створює об'єкт `UMLClass` для самого класу, але й пов'язаний з ним об'єкт `StateMachine`, встановлюючи посилання `json_id` на ідентифікатор класу. Аналогічно, методи класу з декоратором `@transition(source=..., target=...)` трансформуються в об'єкти `UMLTransition`, а структуровані `docstring`-коментарі з тегом `@sequence` парсяться для створення діаграм послідовностей.

Також було виконано реалізацію відношення μ , яка є ключовою інновацією, що забезпечує зв'язність моделі.

Нижче наведено приклад згенерованого XMI-фрагменту, де лінія життя `UMLLifetime` посилається на клас `Device` зі структурного подання:

```
<packagedElement xmi:type=»uml:Lifeline»
xmi:id=»seq1.lifeline.device» name=»self»>
<xmi:Extension extender=»UMLProfile»>
<JsonRef jsonId=»my_project.Device»/>
</xmi:Extension>
</packagedElement>
```

Тег `JsonRef`, що є частиною розробленого UML-профілю, є фізичним втіленням абстрактного відношення μ , що дозволяє валідатору під час аналізу XMI-файлу швидко (з часовою складністю $O(1)$ за допомогою хеш-таблиці) отримати доступ до повного опису класу `Device` з JSON-файлу.

Водночас ефективність інкрементальної валідації залежить від точності графа залежностей. На відміну від простого графа, реалізація в `dependency_graph.py` та теоретична модель розрізняють три типи ребер, що дозволяє виконувати більш точний аналіз впливу змін:

- структурні: Відображають відношення володіння (наприклад, від `UMLClass` до `UMLOperation`);
- семантичні: Фіксують логічні зв'язки (наприклад, від `UMLMessage` до `UMLOperation`, яку воно викликає);
- трасувальні: Реалізують відношення μ між поданнями (наприклад, від `UMLLifetime` (M_B) до `UMLClass` (M_S)).

В результаті, подібна класифікація дозволяє валідатору виконувати цільові перевірки. Наприклад, зміна сигнатури методу в коді поширюється через семантичне ребро до відповідного `UMLMessage` і запускає перевірку інваріанта `SignatureMatches`. Тому така точність запобігає виконанню зайвих перевірок і є ключем до високої продуктивності.

Для забезпечення балансу між швидкістю відгуку та глибиною аналізу, плагін реалізує трирівневу стратегію валідації, де наступний рівень є більш обчислювально складним, але й більш потужним.

Перший рівень. Синтаксична валідація за допомогою JSON Schema. Це перший і найшвидший етап валідації, що виконується при генерації структурного подання M_S у JSON. Його призначення – перевірка базової синтаксичної коректності та структури M_S . Для цього використовується файл `schemas/structural_schema.json`, який формалізує вимоги до JSON-документа, перевіряючи обов'язковість полів, формат ідентифікаторів за допомогою регулярних виразів, типи даних та унікальність елементів у масивах (`uniqueItems`).

Другий рівень. Імперативна інкрементальна валідація (в реальному часі). Це ядро інтерактивної роботи плагіна, що запускається при кожному збереженні файлу в IDE. Його призначення – перевірка семантичних та міжмодельних інваріантів, що потребують аналізу зв'язків між елементами. Алгоритм ValBeh(ΔM_B), реалізований у ValidationEngine, після зміни коду визначає множину порушених елементів (Affected Elements) за допомогою графа залежностей GMB, і перевірці підлягають лише вони. На цьому рівні перевіряються такі інваріанти, як BehaviorExists (чи є поведінка для динамічної операції), LifelineCorrespondence (чи відповідає лінія життя існуючому класу), SignatureMatches (чи збігаються сигнатури) та StartStateExists (чи є у діаграми станів початковий стан).

Третій рівень. Глибока верифікація через SAT-аналіз. Цей рівень використовується для перевірки складних глобальних властивостей моделі, які важко або неефективно перевіряти імперативними алгоритмами. Його призначення – це формальна верифікація транзитивних властивостей, насамперед ациклічності. Модуль validators/alloy_bridge.py виконує трансляцію частини моделі (наприклад, графа успадкування класів) у специфікацію мови Alloy та викликає SAT-вирішувач. Наприклад, для перевірки ациклічності успадкування генерується твердження `assert NoCycles { no c: Class | c in c.^extends }`. Alloy намагається знайти контрприклад (наприклад, цикл $A \rightarrow B \rightarrow A$). Через високу обчислювальну складність, цей аналіз запускається лише за прямою командою користувача або в CI/CD конвеєрі.

Також плагін надає CLI для інтеграції в автоматизовані процеси. В першу чергу, це використання через CLI для CI/CD: Основні команди, реалізовані в integration/cli.py, потенційно дозволяють вбудувати валідацію в будь-який CI/CD конвеєр (напр., GitHub Actions):

```
# Генерація моделей з вихідного коду
poetry run uml-consistency generate --project-root./project
# Повна валідація з JSON-виводом для аналізу скриптом
poetry run uml-consistency validate --format json --project-root./project
# Запуск глибокої SAT-перевірки глобальних інваріантів
poetry run uml-consistency sat-check -project-root./project
```

Опція `--format json` є важливою для автоматизації, оскільки дозволяє CI-скриптам легко парсити

результати валідації та приймати рішення, наприклад, відмітити збірку як невдалу при наявності помилок.

Нижче наведено демонстраційний приклад, де розглядається сценарій, який демонструє перевагу плагіна: глибоку, семантичну узгодженість. Розробник виконує рефакторинг у Python-коді: перейменовує метод `device.toggle_status()` на `device.set_active()`. Однак він забуває оновити docstring-коментар з тегом `@sequence`, де цей метод викликається для опису діаграми послідовності: `@sequence... -> device : toggle_status()`.

Плагін виконає наступні дії:

1. Інкрементальний парсер `ASTParserVisitor` фіксує зміну в `UMLOperation` (перейменування методу `toggle_status`).

2. Граф залежностей G_{MB} через *трасувальне* та *семантичне* ребра визначає, що від цієї операції залежить елемент `UMLMessage` у діаграмі послідовності, яка була згенерована з docstring-коментаря.

3. `ValidationEngine` запускає перевірку інваріанта `SignatureMatches` (або більш базового `AnchorExists`) для цього повідомлення.

В результаті плагін миттєво підсвічує рядок з `@sequence` в редакторі коду і видає інформативну помилку у консолі.

Додатково було проведено експериментальну оцінку продуктивності, для кількісної оцінки запропонованого підходу.

В якості тестового стенду виступає проект на Python середнього розміру, що містить 50 класів, 500 методів та близько 10 поведінкових діаграм, згенерованих з коду. Експерименти проводились на машині з процесором Intel Core i7 та 32 GB RAM.

Інструментами для порівняння виступали StarUML та Visual Paradigm, як представники двох основних підходів до моделювання.

Було проведено два сценарії тестування: повна валідація (вимірювання часу та пікового використання пам'яті для початкового аналізу всього проекту з нуля) та інкрементальна валідація (внесення впливової зміни в один «центральный» клас, наприклад, перейменування методу, від якого залежить багато інших елементів та вимірювання часу на повторну валідацію). Для аналогів, що не підтримують інкрементальний підхід, це означало повторну перевірку всієї моделі.

Результати експериментальної оцінки, представлені на рисунках 6 та 7, наочно демонструють практичні переваги запропонованого інкрементального підходу.

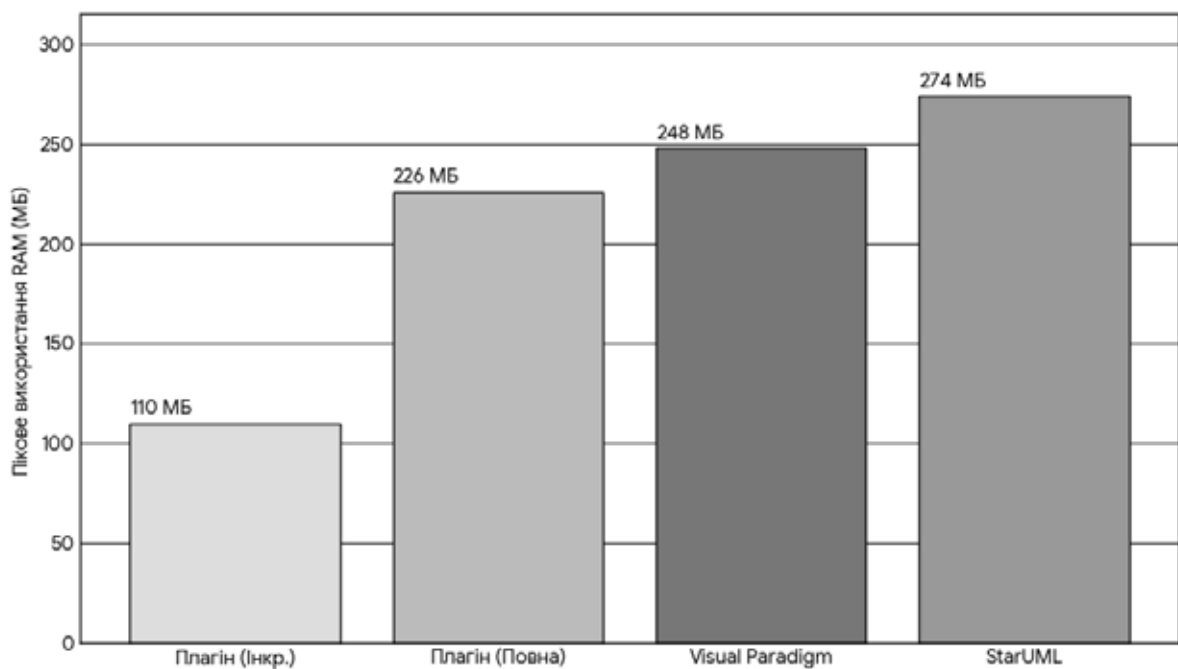


Рис. 6. Порівняння споживання пам'яті

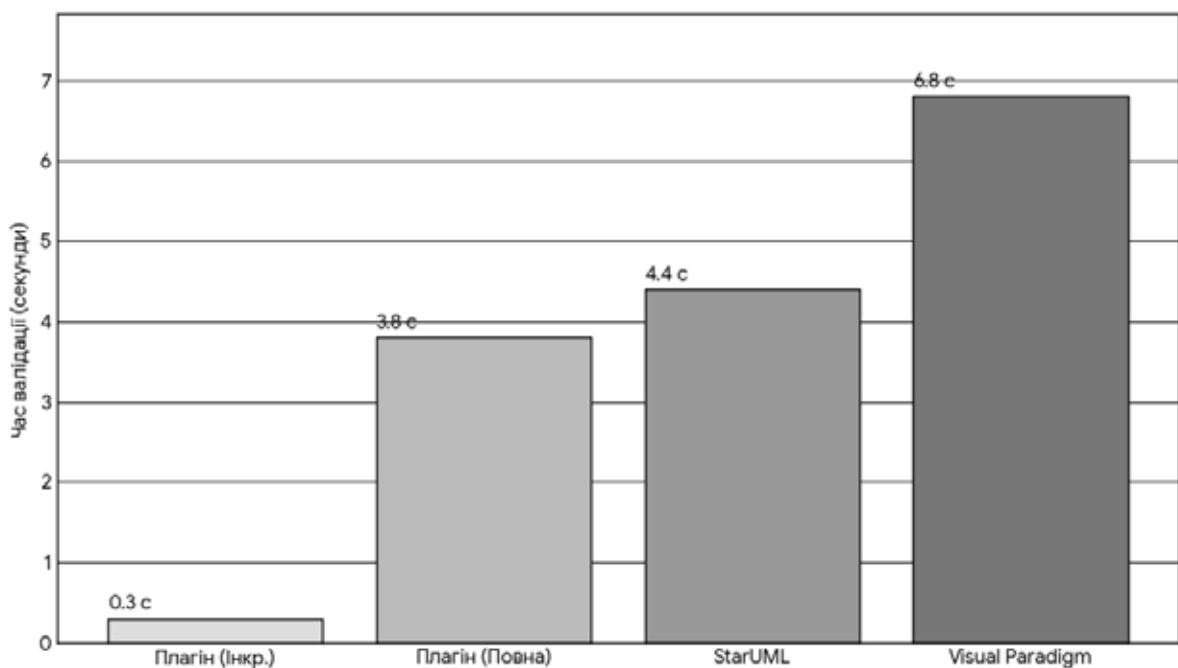


Рис. 7. Порівняння по витраченому часу

Аналіз результатів показує, що час інкрементальної валідації в плагіні є значно меншим, ніж час повної перевірки в аналогів. Це підтверджує теоретичну оцінку складності. У той час як конкуренти працюють за алгоритмом зі складністю $O(N)$, де N – загальна кількість елементів моделі, наш інкрементальний підхід має складність

$O(k \cdot d)$, де k – кількість змінених елементів ($k \ll N$), а d – середній ступінь залежностей. Тому такий приріст продуктивності робить цей плагін єдиним інструментом з представлених, який є практично придатним для використання в режимі реального часу в IDE та для інтеграції у швидкі CI/CD конвеєри. Ефективне використання пам'яті під час

інкрементальної перевірки досягається за рахунок роботи лише з невеликою підмножиною моделі.

Висновки. У даній статті представлено архітектуру, практичну реалізацію та експериментальну оцінку плагіна, який є програмним артефактом, що втілює розроблену раніше теоретичну метамодель з двома поданнями. Проведене дослідження дозволяє зробити висновки, що запропонована архітектура, яка поєднує зберігання статичної структури в JSON та динамічної поведінки в XMI, є життєздатною та ефективною. Вона успішно вирішує фундаментальний конфлікт між гнучкістю, необхідною для сучасних ітеративних процесів розробки, та формальною строгістю, що вимагається для забезпечення надійності складних систем. Ключовий механізм зв'язку між поданнями довів свою спроможність

забезпечувати надійне та швидке трасування між елементами моделі, що є основою для між-модельної валідації. Також експериментально доведено, що інкрементальний підхід до валідації, заснований на графі залежностей G_{MB} , забезпечує значний приріст продуктивності порівняно з існуючими CASE-інструментами, які покладаються на повну перевірку моделі. Плагін успішно вирішує проблему виявлення прихованих семантичних неузгодженостей між вихідним кодом, його структурним поданням та поведінковими сценаріями – дефектів, які часто пропускаються аналогічними інструментами через їхню відірваність від коду.

В результаті, реалізація є комплексним рішенням, що пропонує новий підхід до розробки програмного забезпечення.

Список літератури:

1. Brambilla M., Cabot J., Wimmer M. Model-Driven Software Engineering in Practice: Second Edition. *Synthesis Lectures on Software Engineering*. 2017. Vol. 3, no. 1. P. 1–207. DOI: 10.2200/s00751ed2v01y201701swe004.
2. Cicchetti A., Ciccozzi F., Pierantonio A. Multi-view approaches for software and system modelling: a systematic literature review. *Software and Systems Modeling*. 2019. Vol. 18, no. 6. P. 3207–3233. DOI: 10.1007/s10270-018-00713-w.
3. Нікітченко М. І. Аналіз форматів збереження UML-моделей у сучасних CASE-засобах. *Технології та суспільство: взаємодія, вплив, трансформація : матеріали IV Міжнар. наук. конф. (Чернігів, 20 червня 2025 р.)*. Чернігів, 2025. DOI: 10.62731/mcnd-20.06.2025.
4. Usman M., Zeshan F., Perwaiz A., Younas M. A Survey of Consistency Checking Techniques for UML Models. *Proceedings of the Second International Symposium on Advanced Software Engineering & Its Applications (ASEA), Hainan, China, 13–15 December 2008*. 2008. DOI: 10.1109/asea.2008.40.
5. Komleva N. O., Nikitchenko M. I. Method for Incremental Control of Consistency Between Structural and Behavioral Views of Software Architecture. *AAIT*. 2025. Vol. 8, no. 2. P. 162–177. DOI: 10.15276/aa.08.2025.11.
6. Комлева Н. О., Нікітченко М. І. Структурне подання UML-метамоделі у форматі JSON. *Інформатика та математичні методи в моделюванні*. 2025. Том 15, № 2. С. 205–217. DOI: 10.15276/imms.v15.no2.205.
7. Комлева Н. О., Нікітченко М. І. Поведінкове подання у форматі XMI в межах UML-метамоделі з двома поданнями. *Вісник Хмельницького національного університету. Технічні науки*. 2025. Том 5 (Подано до друку).
8. Altamimi T., Petriu D. Incremental Change Propagation from UML Software Models to LQN Performance Models. *Proceedings of the 2017 ACM/IEEE 20th International Conference on Model Driven Engineering Languages and Systems, Austin, Texas, 17–22 September 2017*. 2017. P. 110–120. DOI: 10.1145/3121651.3121653.

Nikitchenko M.I. ARCHITECTURE AND IMPLEMENTATION OF A PLUGIN FOR INCREMENTAL VALIDATION OF UML METAMODELS WITH TWO VIEWS

The article addresses the issue of ensuring the consistency of UML models in the development of complex software systems, the relevance of which arises from the high cost of identifying architectural defects at later stages of the software life cycle. Current CASE tools for modeling offer either the flexibility of JSON format for static structures with low formalism or the strict of XMI format, which is challenging to integrate into daily development due to complexity. However, neither approach fully combines flexibility and formality, necessitating the development of a hybrid solution.

The purpose of this study is to present the architecture and practical implementation of a plugin integrating two formats for representing UML models—lightweight JSON for static structures and formally rigorous XMI for system dynamics description. A notable feature of the plugin is the use of a specialized UML profile to physically realize the correspondence relationship between these views, enabling rapid incremental validation through a dependency graph. The methodology combines JSON Schema syntax checking, semantic incremental validation, and deep formal verification of transitive properties using SAT analysis.

Experimental performance evaluation results demonstrate that the proposed plugin significantly outperforms existing analogs, as incremental validation reduces checking time compared to full model validation. This enables efficient integration into continuous integration and delivery pipelines, making the tool suitable for real-time use. The conclusions of the study highlight the potential of a hybrid UML model storage approach and outline future development directions, including support for additional programming languages and extended plugin functionalities.

Key words: UML, model consistency, incremental validation, JSON Schema, XMI, dependency graph, SAT analysis.

Дата надходження статті: 28.07.2025

Дата прийняття статті: 07.08.2025

Опубліковано: 27.10.2025

Онiщенко Т.В.

Національний університет «Одеська політехніка»

Онiщенко І.О.

Одеський ліцей «Надія» Одеської міської ради

ПОРІВНЯЛЬНИЙ АНАЛІЗ ОНЛАЙН-ІНСТРУМЕНТІВ КОНВЕРТАЦІЇ СИСТЕМ ЧИСЛЕННЯ ДЛЯ ВИКОРИСТАННЯ В ОСВІТНЬОМУ КОНТЕКСТІ

У програмній інженерії розуміння систем числення відіграє одну з фундаментальних ролей, оскільки воно лежить в основі представлення даних в обчислювальних системах. Студенти технічних ВНЗ, які вивчають дисципліни, пов'язані зі структурою та організацією даних в сучасних обчислювальних середовищах, традиційно виконують переклад чисел між базами вручну, що дозволяє глибше усвідомити внутрішні механізми обчислень, такі як бінарне представлення та операції на рівні бітів. У програмній інженерії розуміння систем числення та конвертації баз чисел є фундаментальним для майбутнього освоєння концепцій представлення даних. Однак для розробки інтерактивних лекцій і практичних занять можна використовувати сучасні інструменти для конвертації баз чисел, які не тільки прискорюють цей процес, але й надають візуалізацію кроків, верифікацію розрахунків, підвищуючи саму якість навчального (лекційного або практичного) матеріалу.

На сьогоднішній день на ринку програмного забезпечення дійсно існують продукти для виконання таких досить специфічних і вузько спрямованих завдань – переведення чисел між системами числення. Вони пропонують різні інтерфейси, точність виконання операцій і додаткові можливості. Однак відсутня єдина методика оцінки цих інструментів, що ускладнює вибір оптимального рішення для інтеграції в процес створення навчального матеріалу, особливо в умовах переходу до змішаного або онлайн-навчання. Без проведення порівняльного аналізу ризику неоптимального вибору зростають, що призводить до субоптимальних освітніх результатів. Саме тому порівняльний аналіз таких інструментів набуває актуальності, заповнюючи прогалину в методології оцінки та сприяючи оптимізації педагогічних підходів у технічній освіті.

В рамках даної роботи проведено порівняльний аналіз 4 онлайн-інструментів для роботи з системами числення. Оцінка виконана за 6 критеріями, адаптованими на основі моделі якості ISO/IEC 25010. Дослідження базується на практичному тестуванні з використанням типових завдань конвертації, а також аналізі метрик юзабіліті та надійності.

Ключові слова: системи числення, конвертери баз, онлайн-інструменти, юзабіліті, якість програмного забезпечення, програмна інженерія, ISO/IEC 25010.

Постановка проблеми. Сучасні інформаційні технології значною мірою спираються на розуміння фундаментальних принципів обробки даних, зокрема роботи з системами числення, які є основою представлення інформації в обчислювальних середовищах. У різних наукових дисциплінах студенти вивчають процеси конвертації чисел між системами числення (двійковою, вісімковою, десятковою, шістнадцятковою тощо) вручну, що дозволяє більш детально зрозуміти внутрішні механізми обробки даних, таких як представлення інформації в комп'ютерних систе-

мах, бітові операції, адресація пам'яті та кодування даних [1, с. 10]. Отримані знання є надійним фундаментом для вивчення програмування, де вони застосовуються для маніпуляцій з бітами, а також оптимізації коду в низькорівневих мовах тощо. Крім того, ці навички корисні в комп'ютерній архітектурі для розуміння зберігання та обробки даних у бінарному форматі, керування пам'яттю та взаємодії з апаратним забезпеченням. Вони також знаходять застосування в цифровій електроніці для проектування логічних схем, криптографії для модульних обчислень і алгоритмів

шифрування [1, с. 21]. Таким чином, ручна конвертація не лише закладає базу для теоретичного розуміння предметної галузі, але й готує до практичного використання в професійній діяльності.

Для підготовки лекційних матеріалів і практичних завдань по дисциплінам, які включають в свою освітню програму вивчення систем числення викладачеві зручно використовувати спеціальні інструменти, які можуть не тільки прискорити процес конвертації, але й надати візуалізацію кроків, верифікацію розрахунків, інтеграцію в освітні сценарії, підвищуючи інтерактивність занять та ефективність навчання загалом.

Сучасні онлайн-інструменти для конвертації чисел між системами числення пропонують автоматизоване рішення для виконання таких завдань. Однак ці інструменти значно різняться за функціональними можливостями, зручністю інтерфейсу, точністю, доступністю та освітньою цінністю. Відсутність єдиної методики оцінки таких платформ ускладнює вибір оптимального рішення для конкретних потреб – як освітніх, так і професійних. Наприклад, одні інструменти можуть пропонувати детальні пояснення алгоритмів, корисні для навчання, тоді як інші фокусуються лише на базовій конвертації, що більше підходить для швидких практичних завдань.

Проблема полягає у відсутності систематизованого підходу до порівняльного аналізу онлайн-інструментів для конвертації систем числення, який би враховував ключові аспекти їх якості. Без такого аналізу викладачі можуть обирати неоптимальні інструменти, що може призвести до зниження ефективності створення освітнього матеріалу.

Актуальність дослідження зумовлена необхідністю інтеграції автоматизованих інструментів у сучасну освіту. У науковому контексті аналіз таких інструментів спирається на стандарти якості програмного забезпечення, зокрема ISO/IEC 25010, що дозволяє оцінити їх з позиції користувачького досвіду та функціональних можливостей. У практичному плані результати дослідження сприятимуть вибору оптимальних інструментів для створення освітніх програм. Таким чином, вирішення цієї проблеми має як наукове значення (розробка методології оцінки), так і практичне (підвищення якості навчання та професійної діяльності).

Аналіз останніх досліджень і публікацій. У науковій літературі активно обговорюються фундаментальні аспекти систем числення, їхнього представлення в комп'ютерних системах та

методи конвертації, зокрема в контексті програмної інженерії. Наприклад, у роботі А. Стіла (2014) аналізуються різні методи викладання конвертації між двійковою, шістнадцятковою та десятковою системами, з акцентом на ефективність педагогічних підходів, включаючи використання онлайн-інструментів для візуалізації [2, с. 2]. Крім того, існують праці, присвячені оцінці онлайн-калькуляторів для числових перетворень, такі як огляд інструментів для конвертації IEEE-754 (Г. Шмідт, 2021), де акцентується увага на точності та продуктивності, але без систематичного порівняння кількох сервісів [3, с. 1].

Водночас, наукових досліджень, присвячених порівняльному аналізу онлайн-інструментів для конвертації чисел між системами числення (з урахуванням критеріїв юзабіліті, функціональності та освітньої цінності), не проводилося. Більшість публікацій обмежуються описом окремих конвертерів або загальними рекомендаціями щодо їхнього використання в освіті, без адаптації стандартів якості на кшталт ISO/IEC 25010. Тому зайняття цією проблематикою є актуальним завданням, оскільки воно заповнює прогалину в методологічних підходах до оцінки таких спеціалізованих онлайн-інструментів у сучасній освіті.

Постановка завдання. Метою даної роботи є розробка рекомендацій для викладачів програмної інженерії, які використовують спеціалізовані онлайн-платформи конвертації систем числення для створення учбових матеріалів. Дані рекомендації сформовані виходячи з аналізу 4-х найпопулярніших сервісів.

Виклад основного матеріалу. Для проведення порівняльного аналізу було обрано 4 програмних продукти: RapidTables Number Base Converter, UnitConverters.net Base Converter, Wolfram Alpha та Calculator.net Binary Converter. Вибір цих інструментів обумовлений двома ключовими критеріями, які забезпечують репрезентативність дослідження та його практичну цінність у контексті сучасної освіти [4, с. 11].

По-перше, акцент зроблено саме на онлайн-продуктах, оскільки вони забезпечують максимальну зручність використання: доступ з будь-якого сучасного пристрою (смартфона, планшета чи комп'ютера) без необхідності встановлення програмного забезпечення. Це відповідає сучасним тенденціям цифрової освіти, де мобільність та швидкий доступ до ресурсів є пріоритетом, як зазначається в оглядах освітніх інструментів [5, с. 33].

По-друге, кількість спеціалізованих онлайн-інструментів для конвертації систем числення обмежена, і обрані продукти є найбільш відомими та популярними. Таким чином, обрані продукти є репрезентативними для даного ринку програмного забезпечення.

Існує ще 2 достатньо відомих інструменти, які можуть виконувати схожі завдання, але вони не були додані до добірки. Google Calculator, вбудований у пошукову систему Google, є зручним інструментом для швидких обчислень і базових конвертацій систем числення. Однак цей інструмент має обмежену функціональність, яка не відповідає повному спектру задач, розглянутих у дослідженні. Конвертація обмежується лише чотирма стандартними базами (2, 8, 10, 16), без можливості роботи з довільними базами, що є критично для глибокого аналізу в програмній інженерії та комп'ютерній архітектурі. Крім того, цей інструмент надає лише кінцевий результат без візуалізації кроків конвертації, пояснень алгоритмів чи інтеграції з прикладами, що робить його непридатним для розробки інтерактивних лекцій чи практичних завдань. Він орієнтований на швидкі запити, а не на навчання. Хоча Google Calculator доступний через браузер, це не самостійний веб-сайт з інтерфейсом, а частина пошукової системи.

Вбудований калькулятор Windows у режимі Programmer пропонує базові функції для роботи з системами числення, включаючи конвертацію між бінарною, восьмеричною, десятковою та шістнадцятковою базами, а також бітові операції (AND, OR, XOR тощо). Цей інструмент корисний для простих задач у програмуванні, але виключений з аналізу через наступні обмежені функціональність (підтримка обмежується лише чотирма базами), відсутні додаткові опції, такі як візуалізація алгоритмів чи підтримка великих чисел, не онлайн-орієнтований, не забезпечує кросс-платформенний доступ з мобільних пристроїв чи веб-браузерів, що суперечить нашому акценту на зручності для викладачів у гібридному навчанні. Ну і головне, що цей інструмент фокусується лише на швидких обчисленнях без пояснень кроків, що робить його менш придатним для педагогічних цілей порівняно з відібраними онлайн-платформами.

Виключення цих інструментів дозволяє зосередитися на більш спеціалізованих і гнучких рішеннях, які краще відповідають методології оцінки за ISO/IEC 25010 та освітнім потребам. Якщо б вони були включені, аналіз міг би стати

менш сфокусованим, оскільки їхні обмеження не дозволяють повноцінного порівняння з обраними продуктами.

Для проведення порівняльного аналізу відібраних програмних продуктів було обрано 6 критеріїв, які адаптовано на основі міжнародного стандарту ISO/IEC 25010. Вибір саме цих критеріїв обумовлений необхідністю всебічної оцінки в контексті онлайн-інструментів для освітніх завдань, де поєднуються аспекти статичної якості продукту з динамічними характеристиками використання, адаптованими до специфіки сучасної освіти та програмної інженерії.

Дослідження проведено за наступними критеріями відбору:

1. Зручність інтерфейсу та користувацького досвіду (User Interface and User Experience): Адаптовано з характеристики *usability* (розділ 4.2.4 моделі якості продукту в ISO/IEC 25010:2011 та оновленій версії 2023), яка включає підхарактеристики *learnability* (здатність до швидкого освоєння), *operability* (легкість у використанні та керуванні) та *user interface aesthetics* (естетика інтерфейсу користувача). У контексті дослідження онлайн-інструментів цей критерій дозволяє оцінити інтуїтивність вводу даних і навігації, включаючи простоту вибору бази числення, наявність підказок та адаптивний дизайн для різних пристроїв.

2. Функціональні можливості (Functional Capabilities): Пряме зіставлення з *functional suitability* (розділ 4.2.1 моделі якості продукту), що оцінює повноту функцій через підхарактеристики *functional completeness* (повнота функцій), *correctness* (правильність виконання) та *appropriateness* (придатність для цілей). У аспекті конвертації систем числення це включає підтримку різних баз (від 2 до 36, включаючи нестандартні), додаткові опції, такі як крокові пояснення алгоритмів, а також інтеграцію з бітовими операціями. Це забезпечує придатність інструментів для освітніх сценаріїв, де потрібна не лише базова конвертація, а й демонстрація алгоритмів для глибокого розуміння процесів, що сприяє розвитку навичок програмування та аналізу даних у студентів.

3. Точність та надійність (Accuracy and Reliability): Базується на *reliability* (розділ 4.2.5 моделі якості продукту), з акцентом на підхарактеристики *maturity* (зрілість, тобто низька частота збоїв), *availability* (доступність без перерв), *fault tolerance* (стійкість до помилок) та *recoverability* (здатність до відновлення після збою). У аспекті систем числення цей критерій

запобігає помилкам у граничних випадках, забезпечуючи коректну обробку помилок з інформативними повідомленнями. Це важливо для верифікації в програмній інженерії, де неточності можуть призвести до критичних збоїв у кодї або симуляціях, і сприяє надійності інструментів у освітніх застосуваннях.

4. Продуктивність та швидкість (Performance and Speed): Адаптовано з *performance efficiency* (розділ 4.2.2 моделі якості продукту), яка вимірює ефективність через підхарактеристики *time behaviour* (поведінка за часом, тобто швидкість відгуку), *resource utilization* (використання ресурсів, таких як CPU та пам'ять) та *capacity* (здатність обробляти навантаження). Для онлайн-сервісів це можливість забезпечувати швидку обробку запитів (наприклад, миттєву конвертацію великих чисел без затримок), мінімальне споживання ресурсів пристрою (особливо мобільного) та стабільність під час множинних запитів. Це актуально для інтерактивних занять, де повільна продуктивність може перервати навчальний процес, а ефективність сприяє безперервному використанню в реальному часі, наприклад, під час онлайн-лекцій чи групових завдань.

5. Освітня цінність (Educational Value): Адаптовано з *satisfaction* (розділ 4.3.3 моделі якості у використанні, включаючи підхарактеристику *usefulness* – корисність) та *functional appropriateness* (підхарактеристика з *functional suitability* в моделі якості продукту, розділ 4.2.1), з фокусом на поясненнях та візуалізаціях. У освітньому контексті це критерій оцінює внесок у розуміння процесів, наприклад, через детальні крокові розбори конвертації, приклади застосування в програмуванні чи інтеграцію з теоретичним матеріалом.

6. Вартість та доступність (Cost and Availability): Базується на *freedom from risk* (розділ 4.3.4 моделі якості у використанні, зокрема *economic risk mitigation* – зменшення економічних ризиків) та *portability* (розділ 4.2.8 моделі якості продукту, з підхарактеристиками *adaptability* – адаптивність та *installability* – легкість встановлення), оцінюючи безкоштовність, доступність без реєстрації та крос-платформенність. Це критерій враховує економічну ефективність для викладачів, запобігаючи бар'єрам, таким як платні підписки чи обмеження запитів, що можуть обмежити використання для створення освітніх матеріалів.

Для оцінювання програмних продуктів за обраними критеріями використано адаптовану

методологію, засновану на стандарті ISO/IEC 25010 «Systems and software engineering – Systems and software Quality Requirements and Evaluation (SQuaRE) – System and software quality models» (версія 2011 з оновленнями 2023 року), який є частиною серії SQuaRE (Systems and software Quality Requirements and Evaluation). Цей стандарт визначає дві основні моделі якості: модель якості продукту та модель якості у використанні. [6, с. 732]

Методологія доповнена пов'язаними методами оцінки якості програмного забезпечення, зокрема QuESPro (Quality Evaluation of Software Product) – легковаговим методом для третьої сторони оцінки якості продуктів, розробленим для кількісного аналізу на основі серії ISO/IEC 25000. QuESPro передбачає покроковий процес, що включає визначення моделі якості, збір даних через тестування, оцінку субхарактеристик, розрахунок метрик та формування рекомендацій [7, с. 525]. Цей метод обраний через свою простоту та адаптивність для онлайн-інструментів, де ресурси для оцінки обмежені, а фокус – на практичній придатності для освітніх завдань. Він дозволяє проводити незалежну ґрунтовну оцінку, забезпечуючи об'єктивність і повторюваність результатів, без потреби в глибокому доступі до коду продукту.

Використана чотирибальна шкала відповідності (Level of Compliance, LC), яка адаптована з QuESPro та відповідає рекомендаціям ISO/IEC 25010 щодо використання ординальних шкал для субхарактеристик. Кожен критерій оцінюється від 0 до 3 балів залежно від ступеня задоволення вимог, з урахуванням ризику невідповідності (наприклад, низький ризик – коли інструмент повноцінно виконує функції без обмежень). Ця шкала обрана як оптимальна для наукових оцінок, оскільки вона достатньо проста для реалізації (не вимагає складних обчислень), дозволяє врахувати ризик невідповідності (наприклад, потенційні проблеми в освітньому використанні, як затримки чи помилки), відповідає принципам SQuaRE для кількісної оцінки та полегшує порівняння продуктів. На відміну від ширших шкал (наприклад, 0–10), вона мінімізує суб'єктивність і фокусується на практичних аспектах, таких як адекватність для користувачів (викладачів). Оцінка проводиться на основі комбінації об'єктивних метрик (часові виміри, тести на прикладах) та суб'єктивних даних (опитування користувачів), з фіксацією доказів для прозорості.

Шкала оцінювання (Level of Compliance):

- **3 бали (Добре):** Критерій повністю задоволений, низький ризик (наприклад, інструмент

виконує всі вимоги без обмежень, забезпечуючи високу ефективність і мінімальні проблеми; ризик збоїв чи незручностей практично відсутній, що ідеально для інтеграції в освітні процеси).

- **2 бали (Достатньо):** Критерій переважно задоволений, середній ризик (часткові обмеження, такі як незначні незручності чи неоптимальна продуктивність, але інструмент все ще придатний для більшості завдань; ризик помірний, наприклад, можливі затримки в граничних випадках).

- **1 бал (Недостатньо):** Критерій частково задоволений, високий ризик (значні недоліки, що впливають на використання, наприклад, відсутність ключових функцій чи часті помилки; ризик високий, оскільки це може призвести до некоректних результатів).

- **0 балів (Не задоволено):** Критерій не задоволений взагалі (повна відсутність відповідності, наприклад, критичні збої чи непридатність для базових задач; ризик максимальний, роблячи інструмент неприйнятним для аналізу).

Загальна оцінка продукту розраховується як середнє арифметичне балів за критеріями, що відповідає кількісним підходам QuESPro. Ця методологія забезпечує наукову обґрунтованість, дозволяючи не лише порівняти інструменти, але й сформулювати рекомендації для їхнього використання в освіті, заповнюючи прогалини в існуючих дослідженнях якості програмних продуктів даного класу.

Для забезпечення об'єктивності та наукової строгості оцінки онлайн-інструментів конвертації чисел між системами числення, тестування за кожним критерієм проводиться систематично, з використанням комбінації об'єктивних метрик (часових вимірів, підрахунку функцій, тестів на прикладах) та суб'єктивних оцінок (опитування користувачів). Кожен тест повторюється щонайменше 5 разів на різних пристроях і мережах (Wi-Fi, мобільний інтернет) для врахування варіабельності. Результати фіксуються з доказами (скріншоти, логи, розрахунки), а бали присвоюються за шкалою 0–3 на основі порогових значень.

1. Зручність інтерфейсу та користувацького досвіду: Тестування фокусується на оцінці інтуїтивності та ефективності інтерфейсу, що включає кількість кроків для виконання базової конвертації (введення числа, вибір вихідної та цільової бази, натискання кнопки «Конвертувати» та отримання результату). Виконується 10 тестових завдань на різних пристроях (смартфон, планшет). Вимірюється середній час на задачу

за допомогою Chrome DevTools (від моменту відкриття сторінки до відображення результату). Додатково застосовується System Usability Scale (SUS) – стандартизована анкета з 10 питань, де розрахунок SUS-score дає загальну оцінку юзабіліті (вище 80 – відмінно). Інші метрики: рівень успішності задач (task success rate), кількість помилок користувача (user error rate) та суб'єктивна задоволеність (за опитуванням 10 респондентів). Пороги балів: Менше 10 секунд на задачу, SUS >80, <3 кроків, 100% успіху – 3 бали (повністю інтуїтивно, низький когнітивний навантаження); 10–20 секунд, SUS 68–80, 3–5 кроків, 80–99% успіху – 2 бали (достатньо, але з дрібними незручностями); 20–30 секунд, SUS 50–68, >5 кроків, 50–79% успіху – 1 бал; більше 30 секунд, SUS <50, низький успіх – 0 балів (незадоволено, інтерфейс заплутаний, високе навантаження).

2. Функціональні можливості: Тестування охоплює підрахунок і перевірку підтримуваних систем счислення (баз) та додаткових функцій, таких як автоматичне розпізнавання бази, пояснення кроків конвертації, візуалізація процесу (анімації кроків), підтримка дробових чисел, негативних значень, великих чисел (понад 10^{50}), бітові операції (AND, OR, XOR). Складається список функцій з документації інструменту, потім тестуються на 10 прикладах. Метрики: Кількість баз (стандартні: 2, 8, 10, 16; нестандартні: 3–36); наявність візуалізації (графічні схеми, крокові таблиці); додаткові опції (інтеграція з калькулятором, експорт результатів). Пороги балів: ≥ 5 баз (включаючи нестандартні), повна візуалізація, ≥ 3 додаткові функції (пояснення, дробі, негативні) – 3 бали; 4 бази, базова візуалізація (текстові кроки), ≥ 1 опції – 2 бали; 2–3 бази без візуалізації – 1 бал; <2 бази або відсутність функцій – 0 балів.

3. Точність та надійність: Тестування полягає в перевірці коректності результатів на наборі з 20 прикладів, включаючи граничні (boundary) та крайні (edge) випадки, для виявлення помилок у обчисленнях, обробці вводу та стійкості до збоїв. Процедура: Кожен тест виконується вручну; фіксується відсоток успішних (точний результат без помилок); перевірка обробки помилок (інформативні повідомлення). Метрики: Відсоток успіху; кількість збоїв; час відновлення після помилки. Пороги балів: 100% успіху на всіх 20 прикладах, повна обробка помилок – 3 бали; 80–99% – 2 бали; 50–79% – 1 бал; <50% – 0 балів.

4. Продуктивність та швидкість: Тестування вимірює ефективність завантаження сторінки та обробки запитів за допомогою Google PageSpeed

Insights (PSI) фокусуючись на ключових метриках: Time to First Byte (TTFB – час до першого байту відповіді сервера), First Contentful Paint (FCP – час до відображення першого контенту), Largest Contentful Paint (LCP – час до завантаження основного контенту), Time to Interactive (TTI – час до повної інтерактивності), Speed Index (загальна оцінка візуальної швидкості). Процедура: Запуск PSI на URL інструменту з 10 тестами (мобільний, різні мережі); симуляція навантаження (множинні конвертації поспіль); вимір часу відгуку на типову задачу. Метрики: Середній час відгуку (<1 с для повної обробки); PSI-score (0–100, де >90 – відмінно); використання ресурсів (CPU/пам'ять за Chrome DevTools). Пороги балів: <1 с відгуку, PSI >90, мінімальне використання ресурсів – 3 бали; 1–3 с, PSI 50–90 – 2 бали; 3–5 с, PSI 25–50 – 1 бал; >5 с або зависання – 0 балів.

5. Освітня цінність: Тестування оцінює наявність та якість матеріалів для навчання, таких як пояснення алгоритмів, приклади, візуалізації та інтеграцію з освітнім контентом. Процедура: Аналіз інтерфейсу на наявність цих елементів; тестування на 10 прикладах. Метрики: Кількість прикладів та пояснень; глибина; вплив на навчання. Пороги балів: Повні пояснення з прикладами, візуалізаціями – 3 бали; базові пояснення з прикладами – 2 бали; мінімальні приклади без інтерактиву – 1 бал; відсутність будь-яких матеріалів – 0 балів.

6. Вартість та доступність: Тестування аналізує модель монетизації та бар'єри доступу, включаючи безкоштовність (повний доступ без платежів), наявність лімітів (кількість запитів на день), реклами, вимогу реєстрації, підписки, доступність для користувачів. Процедура: Перевірка сайту на тарифи; тестування лімітів (виконання ≥ 50 запитів для виявлення обмежень); оцінка за рубриками доступності. Метрики: Вартість (0 – безкоштовно; >0 – платно); ліміт запитів (безліміт – добре); час доступу (без реєстрації – швидко). Пороги балів: Безкоштовно без лімітів, реєстрації та реклами

(повна доступність) – 3 бали; безкоштовно з лімітом (≥ 100 запитів/день) або рекламою – 2 бали; підписка для базових функцій – 1 бал; повністю платно або високі ліміти – 0 балів.

Після проведеного аналізу та тестування було створено таблицю з оцінками по кожному критерію для обраних онлайн сервісів. Для більш спрощеного візуального сприйняття інформації у табличній формі для кожного критерію було використано певне скорочення, а саме:

- ЗІ – зручність інтерфейсу та користувацького досвіду,
- ФМ – функціональні можливості,
- ТТН – точність та надійність,
- ПТШ – продуктивність та швидкість,
- ОЦ – освітня цінність
- ВДТ – вартість та доступність.

Загальна оцінка інструменту розраховується за формулою 1:

$$Y = (ЗІ + ФМ + ТТН + ПТШ + ОЦ + ВДТ) / 6, \text{ де } (1)$$

Y – якість програмного продукту.

Таким чином маємо наступні результати:

1. RapidTables Number Base Converter = 2,83
2. UnitConverters.net Base Converter = 2,16
3. Wolfram Alpha = 2,66
4. Calculator.net Binary Converter = 2,5

На рис. 1 зображені результати аналізу в кількісних характеристиках.

На підставі табличних даних оцінки за шістьма критеріями, проведеної відповідно до адаптованої методології на основі ISO/IEC 25010, можна сформулювати висновки щодо кожного програмного продукту. Висновки базуються на сильних і слабких сторонах, з урахуванням актуальності для підготовки інтерактивних учбових матеріалів.

RapidTables Number Base Converter (загальна оцінка: 2,83 бали)

Цей інструмент (рис. 2) продемонстрував високі показники за більшістю критеріїв, зокрема максимальні оцінки за зручність інтерфейсу (інтуїтивний дизайн з мінімальною кількістю кроків),

Таблиця 1

Порівняльна характеристика онлайн-сервісів

Критерії/ Інструмент	RapidTables Number Base Converter	UnitConverters.net Base Converter	Wolfram Alpha	Calculator.net Binary Converter
ЗІ	3	2	3	3
ФМ	3	1	3	1
ТТН	3	3	3	3
ПТШ	3	3	3	3
ОЦ	2	1	2	2
ВДТ	3	3	2	3

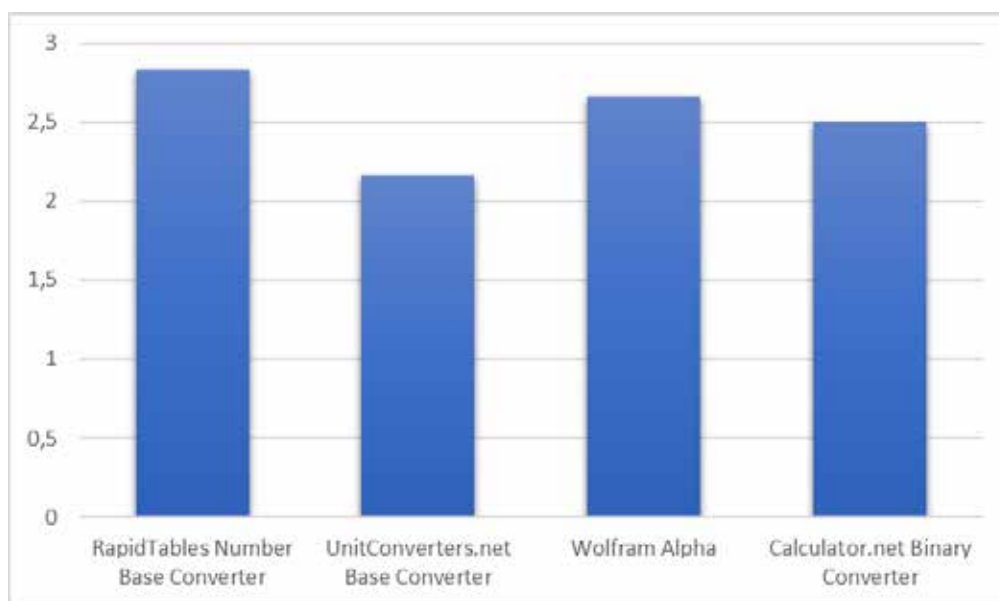


Рис. 1. Результати аналізу онлайн-сервісів

Base Converter

Base calculator Base converter

Convert number from any base to any base:

Enter number: 100

From Base: 10 (decimal)

To base: 2 (binary)

[Convert] [Reset] [Swap]

Result number (7 digits): 1100100

[Copy]

Calculation steps

Decimal to base 2 calculation:
Divide by the base to get the digits from the remainders:

Division	Quotient	Remainder (Digit)	Digit #
100/2	50	0	0
50/2	25	0	1
25/2	12	1	2
12/2	6	0	3
6/2	3	0	4
3/2	1	1	5
1/2	0	1	6

= (1100100)₂

Рис. 2. Інтерфейс RapidTables Number Base Converter

функціональні можливості (підтримка баз від 2 до 36, додаткові калькулятори), точність та надійність (стабільна обробка граничних випадків),

продуктивність (швидке завантаження <1 с) та вартість (повністю безкоштовний без лімітів). Однак освітня цінність оцінена на 2 бали через обмежену візуалізацію кроків і відсутність детальних пояснень. Загалом, це надійний і доступний інструмент, але для розробки лекційного матеріалу з акцентом на візуалізацію процесів він потребує доповнення іншими ресурсами.

UnitConverters.net Base Converter (загальна оцінка: 2,16)

Цей інструмент є аутсайдером з найнижчим сумарним балом. По результатам аналізу інструмент отримав 2 бали за зручність, продемонстрував достатньо низьку функціональність та освітню цінність – 1. Точність (надійна обробка прикладів), продуктивність (швидкий відгук) та вартість (безкоштовний доступ) на високому рівні – 3 бали.

Wolfram Alpha (загальна сума: 2,66)

Wolfram Alpha вирізняється максимальними оцінками за всі критерії, крім вартості (2 бали через ліміт запитів у безкоштовній версії та наявність платної Pro-версії для розширених функцій) та освітньої цінності. Зручність інтерфейсу (природномовний ввід, <10 с на задачу), функціональність (підтримка широкого спектру баз), точність (100% успіх у тестах), продуктивність (<1 с відгук) ставлять його на один рівень з RapidTables Number Base Converter.

Calculator.net Binary Converter (загальна сума: 2,5)

Цей інструмент має максимальні оцінки за зручність інтерфейсу (прості поля вводу), точ-

ність (надійний для бінарних задач), продуктивність (швидкий відгук) та вартість (безкоштовний). Функціональні можливості (лише бази 2–10, фокус на бінарних операціях без широких баз чи візуалізації) та освітня цінність (мінімальні пояснення) оцінені на 1 та 2 бали.

Висновки. Проведений порівняльний аналіз чотирьох онлайн-інструментів для конвертації чисел між системами счислення (RapidTables Number Base Converter, UnitConverters.net Base Converter, Wolfram Alpha та Calculator.net Binary Converter) на основі адаптованої моделі якості ISO/IEC 25010 дозволив систематично оцінити їхню придатність для освітніх цілей, зокрема для підготовки лекційних матеріалів і практичних завдань у дисциплінах, пов'язаних зі структурою та організацією даних у сучасних обчислювальних системах. Оцінка за обраними критеріями виявила чітку диференціацію продуктів, з акцентом на баланс між функціональністю, доступністю та педагогічною цінністю. Загалом, результати підтверджують актуальність дослідження, оскільки відсутність єдиної методології оцінки таких інструментів може призводити до субоптимальних виборів, знижуючи ефективність створення навчальних матеріалів та навчання в цілому.

Серед проаналізованих продуктів Wolfram Alpha та RapidTables Number Base Converter демонструють найвищі показники, що робить їх лідерами в цьому сегменті програмного забез-

печення. Wolfram Alpha вирізняється як комплексне рішення, особливо в освітньому контексті. Єдиним обмеженням є вартість. Тому, саме RapidTables Number Base Converter, з сумарною оцінкою 2,83 балів, є оптимальним варіантом у безкоштовному сегменті таких програмних продуктів. Він забезпечує баланс доступності та ефективності, роблячи його максимально привабливим для викладачів.

Інші продукти – UnitConverters.net Base Converter та Calculator.net Binary Converter – поступаються лідерам за ключовими параметрами, зокрема освітньою цінністю та функціональністю. Хоча вони демонструють високу точність, продуктивність і доступність, їхні обмеження роблять їх менш придатними для повного циклу освітнього процесу, де потрібна не лише конвертація, а й демонстрація внутрішніх механізмів обчислень.

На підставі проведеного дослідження рекомендується викладачам технічних ВНЗ, які розробляють лекції та практичні завдання для дисциплін, пов'язаних з системами счислення, використовувати RapidTables Number Base Converter як базовий інструмент. Якщо є фінансова можливість, перевагу слід віддати Wolfram Alpha в Pro-версії, яка, завдяки винятковій візуалізації прикладів, допоміжним матеріалам і кроковим поясненням, стає однозначним лідером для створення інтерактивного контенту.

Список літератури:

1. Sakk E. Understanding Binary Numbers: Concepts in Computer Systems. CreateSpace Independent Publishing Platform, 2018. 106 p.
2. Kempthorne D., Steele A. An evaluation of different delivery methods for teaching binary, hex and decimal conversion. Journal of Applied Computing and Information Technology. 2014. Vol. 18, iss. 2. URL: <https://surl.li/keoptr> (дата звернення: 20.04.2025).
3. Schmidt H. IEEE-754 Floating Point Converter. 2024. URL: <https://www.h-schmidt.net/FloatConverter/IEEE754.html> (дата звернення: 21.05.2025).
4. Gemma Coleman, Monica M. G. Big Book of Computing Pedagogy. Cambridge: Raspberry Pi Foundation, 2020. 150 p.
5. Malik M.S.I. The Impact of Digital Learning on Education. Journal of Educational Technology & Society. 2024. Vol. 27, iss. 3. pp. 25–35.
6. Sommerville I. Software Engineering. 10th ed. Pearson, 2015. 816 p.
7. Lami G., Spagnolo G.O. A Lightweight Software Product Quality Evaluation Method. Proceedings of the 17th International Conference on Evaluation of Novel Approaches to Software Engineering (ENASE 2022). 2022. pp. 520–528.

Onishchenko T.V., Onishchenko I.O. COMPARATIVE ANALYSIS OF ONLINE NUMBER SYSTEM CONVERSION TOOLS FOR USE IN AN EDUCATIONAL CONTEXT

In software engineering, the understanding of number systems plays a fundamental role, as it underlies the representation of data in computing systems. Students of technical universities studying disciplines related to the structure and organization of data in modern computing environments traditionally perform manual number conversions between databases, which allows them to gain a deeper understanding of the internal mechanisms of computing, such as binary representation and bit-level operations. In software engineering, an

understanding of number systems and number base conversions is fundamental to the future mastery of data representation concepts. However, to develop interactive lectures and practical classes, you can use modern tools for converting number bases, which not only speed up this process but also provide visualization of steps, verification of calculations, and improve the quality of educational material (lectures or practical classes).

Today, there are indeed products on the software market for performing such rather specific and narrowly focused tasks – converting numbers between number systems. They offer different interfaces, accuracy of operations, and additional features. However, there is no single methodology for evaluating these tools, which makes it challenging to select the best solution for integrating into the process of creating educational materials, especially in the context of transitioning to blended or online learning. Without a comparative analysis, the risks of a suboptimal choice increase, leading to suboptimal educational outcomes. That is why a comparative analysis of such tools is gaining relevance, filling a gap in the assessment methodology and contributing to the optimization of pedagogical approaches in technical education.

In this paper, a comparative analysis of 4 online tools for working with number systems was conducted. The evaluation is based on 6 criteria adapted from the ISO/IEC 25010 quality model. The study is based on practical testing using typical conversion tasks, as well as analysis of usability and reliability metrics.

Key words: *number systems, database converters, online tools, usability, software quality, software engineering, ISO/IEC 25010.*

Дата надходження статті: 22.07.2025

Дата прийняття статті: 28.07.2025

Опубліковано: 27.10.2025

Пахомова В.М.

Дніпровський інститут інфраструктури і транспорту

Іванченко Д.В.

Дніпровський інститут інфраструктури і транспорту

ДОСЛІДЖЕННЯ СТРУКТУРИ МЕРЕЖІ MPLS НА ЗАЛІЗНИЧНОМУ ТРАНСПОРТІ З ВИКОРИСТАННЯМ ІНТЕЛЕКТУАЛЬНОГО ЗАСОБУ

На сучасному етапі в комп'ютерних мережах залізничного транспорту застосовується протокол OSPF, при використанні якого в реальному часі з'являється проблема завдяки постійним змінам обсягів передаваних даних, і для вирішення якої доцільно використання технології MPLS з створеним інтелектуальним засобом, що підтверджує актуальність теми. У якості математичного апарату для розв'язання задачі інжинірингу трафіку взято нейронну мережу конфігурації «11-X-Y-2», де 11 – кількість вхідних (input) нейронів (це параметри трьох потоків та двох тунелів); X – кількість прихованих шарів, що потребує подальшого дослідження; Y – загальна кількість прихованих нейронів, що потребує додаткового дослідження; 2 – кількість вихідних (output) нейронів. У роботі за допомогою мови Python створено відповідну програмну модель «Stream_TE-tunnel», на якій проведено дослідження похибки нейронних мереж за різною кількістю прихованих шарів (1, 2, 3 та 4) з різною кількістю нейронів на кожному прихованому шарі (16, 32, 64, 128 та 256) при використанні різних функцій активації нейронів на прихованих шарах (ReLU, Leaky ReLU, Sigmoid, Tanh, ELU та Swish) та різних алгоритмів навчання (SGD, Adam, RMSprop, Adagrad, Nadam, Adamdelta та Ftrl) на вибірках різної довжини (360, 1800, 3600, 7200, 10800, 14400, 18000, 21600, 25200 та 28800 прикладів). Для нейронної мережі конфігурації «11-256-128-2» (256 – кількість нейронів першого прихованого шару; 128 – кількість нейронів другого прихованого шару) при використанні функції активації ReLU на прихованих шарах та функції активації Linear на результуючому шарі за алгоритмом навчання Adam проведено дослідження залежності якості розв'язання задачі інжинірингу трафіку від послідовності обробки потоків в мережі MPLS. Визначено, що максимальний коефіцієнт навантаження тунелю на створеній програмній моделі «Stream_TE-tunnel» склав 0,748, що відрізняється від теоретичного значення приблизно на 2,8%. Таким чином, у роботі проведено дослідження одного із можливих варіантів структури мережі MPLS на залізничному транспорті з використанням створеного інтелектуального засобу для розподілу трьох потоків по двом тунелям при відомих значеннях середніх швидкостей потоків і пропускних спроможностей тунелів.

Ключові слова: домен MPLS, LER, LSR, потік, середня швидкість, TE-тунель, пропускна спроможність, коефіцієнт навантаження, нейронна мережа, похибка.

Постановка проблеми. На сучасному етапі в інформаційно-телекомунікаційній системі (ІТС) залізничного транспорту України у якості протоколу маршрутизації використовується протокол OSPF (Open Shortest Path First) [2, с. 19], в основі якого використання алгоритму вибору найкоротшого маршруту. У реальному часі збільшення потоків даних в ІТС призведе до появи пікових навантажень, які в свою чергу ведуть до істотних затримок на маршрутизаторах, що складають оптимальний шлях в комп'ютерній мережі.

Аналіз останніх досліджень і публікацій. Нам відомо про існування технології MPLS

(Multiprotocol Label Switching) – це спосіб маршрутизації трафіку, який поєднує найкращі риси IP і комутації каналів [4, с. 1; 6, с. 1]. Замість того, щоб кожного разу аналізувати IP-адресу, маршрутизатори використовують короткі мітки, що пришвидшує передачу даних та знижує навантаження на мережеве обладнання. У свою чергу використання технології MPLS в ІТС залізничного транспорту України призведе до наступних переваг [3, с. 1]: підтримка високої швидкості передачі (завдяки міткам маршрутизатори швидко приймають рішення, що значно знижує затримки); підтримка QoS (класи обслуговування дозволяють

розставляти пріоритети); організація інженерії трафіку (можливість вказати, яким шляхом має рухатися трафік, що дозволяє уникнути перевантаження ділянок); масштабованість MPLS VPN (створення віртуальних приватних мереж поверх єдиної фізичної інфраструктури); отримання протокової гнучкості (MPLS не прив'язаний до одного мережевого протоколу, що спрощує інтеграцію нових технологій). Відомо, що організація інженерії трафіку в MPLS здійснюється вручну [1, с. 143], але на сучасному етапі це можливо зробити за допомогою нейромережних засобів [5, с. 391; 7, с. 3]. Однак, разом з тим важливим недоліком таких методик є відсутність універсальності їх застосування при визначенні ТЕ-тунелю в мережах з різною структурою домену MPLS при обробці різної кількості потоків, що потребує проведення додаткових досліджень стосовно визначення конфігурації нейронної мережі (НМ) та її оптимальних параметрів.

Постановка завдання. Проведені дослідження ставили за мету дослідження структури мережі MPLS в ІТС залізничного транспорту з використанням створеного інтелектуального засобу. Для досягнення поставленої мети вирішувалися наступні задачі: запропонувати конфігурацію НМ для обробки трьох потоків на основі

використання двох ТЕ-тунелів; розробити програмну модель, основу якої складає НМ запропонованої конфігурації; на створеній програмній моделі при виконанні машинного навчання визначити оптимальні параметри запропонованої НМ, що забезпечить достатньо високу точність визначення ТЕ-тунелю; провести дослідження залежності якості розв'язання задачі інжинірингу трафіку від послідовності обробки потоків в мережі MPLS.

Виклад основного матеріалу. 1. Постановка задачі. На рис. 1 представлена загальна структура гіпотетичної мережі MPLS на залізничному транспорті. Комп'ютерна мережа ІТС залізничного транспорту може бути подана у вигляді зваженого орієнтованого графа $G=(R,L)$, де R – множина вершин, кількість яких дорівнює N , кожна вершина представляє собою вузол (маршрутизатор) мережі; L – множина дуг, що моделюють канали зв'язку між вузлами, загальна кількість яких дорівнює M .

Кожному каналу зв'язку (ребру графа) відповідає певна вага, яка відображає його пропускну здатність – тобто максимальний обсяг даних, що може передаватися через нього за одиницю часу

$C = \{c_{ij}\}$ де c_{ij} пропускна спроможність каналу між вузлами i та j .

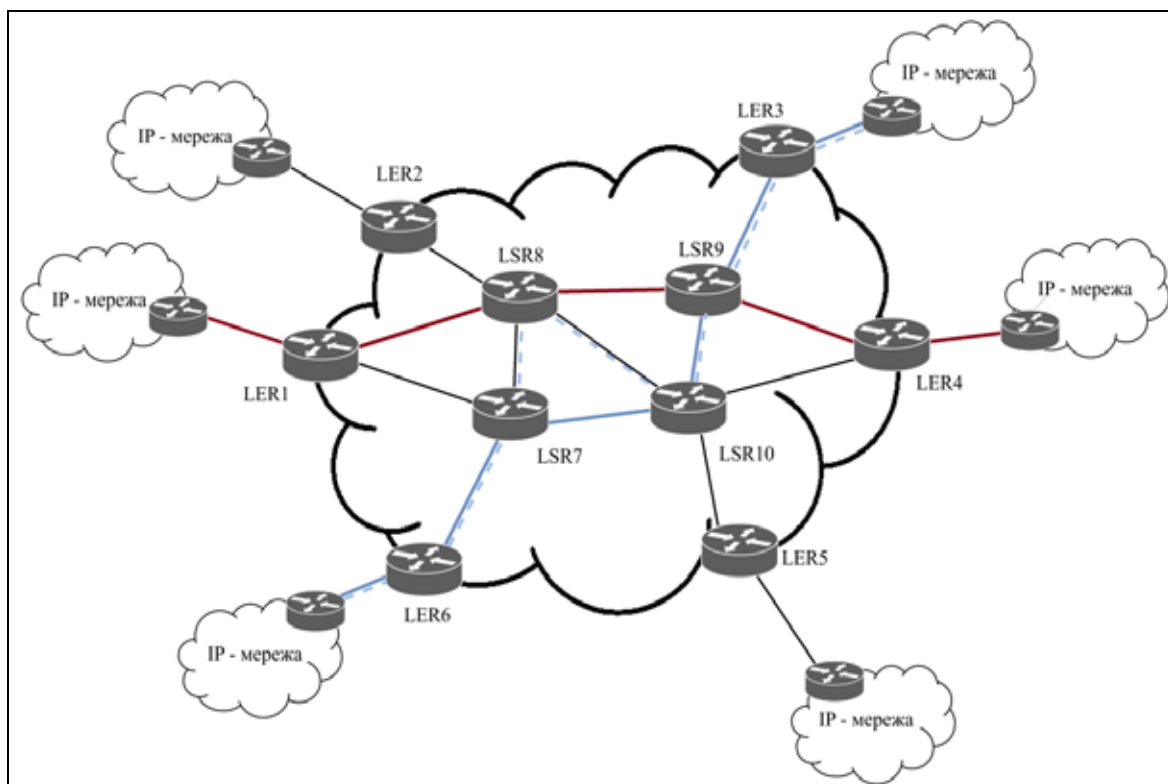


Рис. 1. Структура MPLS мережі на залізничному транспорті

Обов'язковою умовою є те, що швидкість потоку не повинна перевищувати пропускну здатність каналу, через який він передається

$b^s < c_{ij}$ де b^s середня швидкість s -го потоку, c_{ij} пропускна спроможність каналу між вузлами i та j .

Порушення цього обмеження призводить до неможливості передати трафік через даний канал. Рівень використання каналу обчислюється за допомогою коефіцієнта використання

$k_{ij}^s = \frac{b^s}{c_{ij}}$ де k_{ij}^s коефіцієнт використання каналу (i,j) потоком s ; b^s швидкість потоку; c_{ij} пропускна спроможність відповідного каналу.

Для вибору тунелю враховується коефіцієнт його завантаження

$K = \max_s \{K^s\}$ де K – загальний коефіцієнт завантаження тунелю, $K^s = \sum k_{ij}^s$

Розв'язання задачі інжинірингу трафіку передбачає знаходження оптимального розподілу потоків у мережі. При цьому критерієм оптимальності є мінімізація максимального коефіцієнта використання тунелю серед усіх допустимих варіантів маршрутизації

$$\min_i \{K^i\}$$

2. Математичний апарат. У якості методу дослідження використано НМ конфігурації

«11-256-128-2», де 11 – кількість вхідних (input) нейронів; 256 – кількість нейронів першого прихованого шару; 128 – кількість нейронів другого прихованого шару; 2 – кількість вихідних (output) нейронів, що відображена на рис. 2. Перший шар НМ має наступні нейрони: X1 – швидкість потоку № 1; X2 – позиція потоку № 1 в загальній черзі потоків; X3 – належність потоку № 1 до тунелю; X4 – швидкість потоку № 2; X5 – позиція потоку № 2 в загальній черзі потоків; X6 – належність потоку № 2 до тунелю; X7 – швидкість потоку № 3; X8 – позиція потоку № 3 в загальній черзі потоків; X9 – належність потоку № 3 до тунелю; X10 – максимальна пропускна здатність тунелю А; X11 – максимальна пропускна здатність тунелю В. Результуючий шар має нейрони Y1 і Y2; це коефіцієнти завантаженості тунелів А та В відповідно.

3. Загальна вибірка. Складено загальну вибірку з 54000 прикладів (шляхом комбінування 36 різних сценаріїв з 1500 випадково згенерованими даними); кожен із прикладів має 11 ознак та 2 результуючих характеристики, що показано на рис. 3.

На основі загальної вибірки формуються наступні вибірки: навчальна – 70 %; валідаційна – 15 %; тестувальна – 15 %.

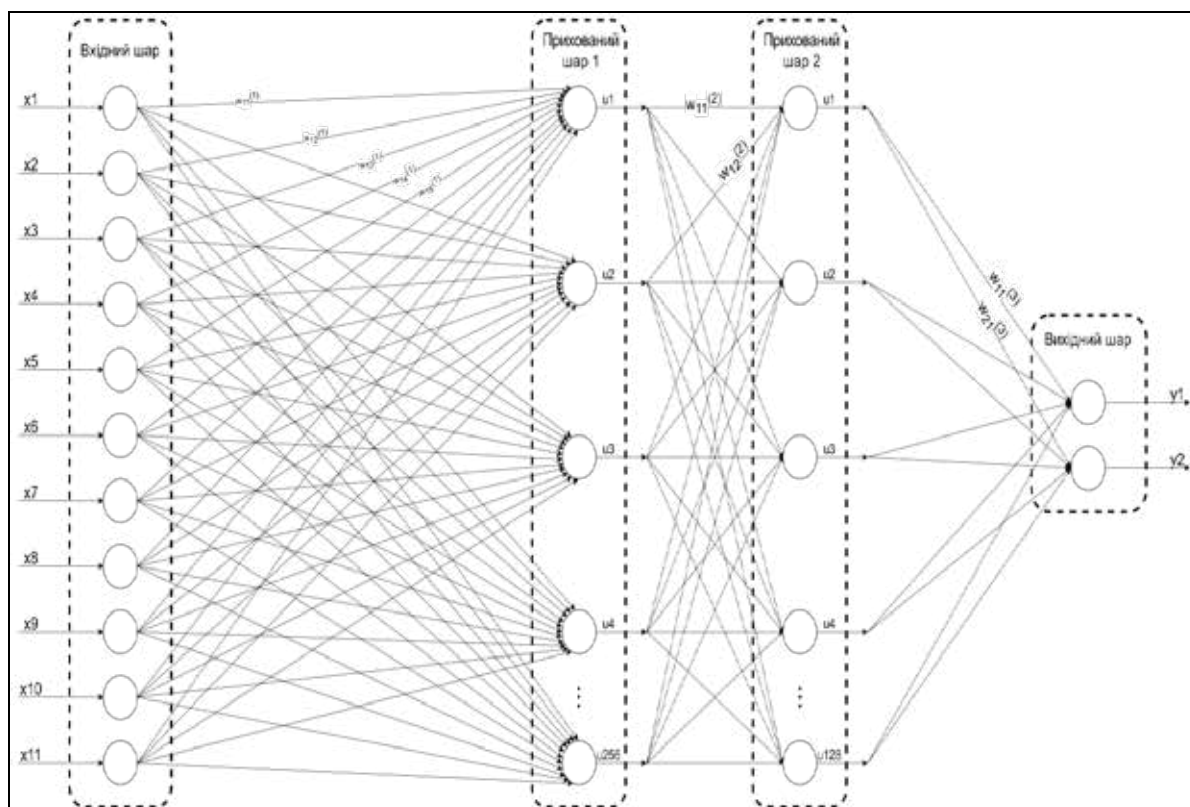


Рис. 2. НМ конфігурації «11-256-128-2»

f1_speed, f1_pos, f1_tunnel, f2_speed, f2_pos, f2_tunnel, f3_speed, f3_pos, f3_tunnel, tunnelA_capacity, tunnelB_capacity, loading_A, loading_B

4.48, 1, 0, 0.09, 2, 0, 2.83, 3, 1, 211.97, 112.45, 0.0216, 0.0252

Рис. 3. НМ: структура навчального прикладу вибірки

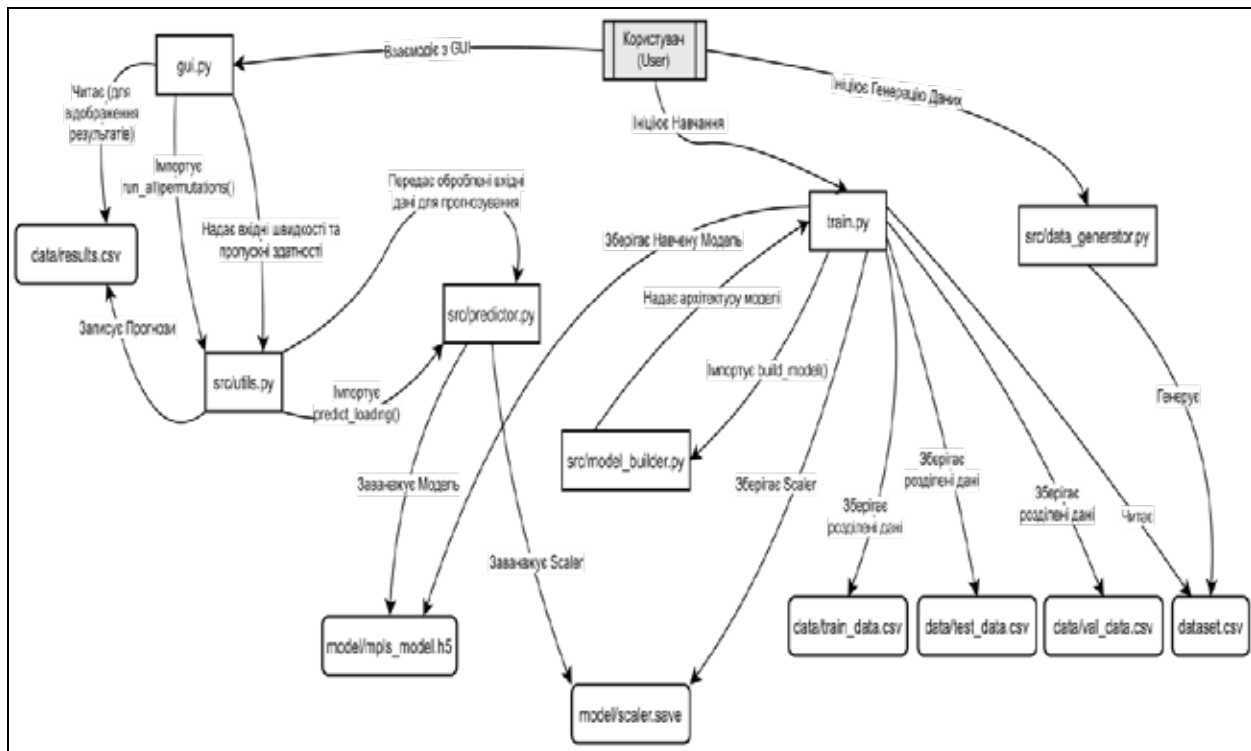


Рис. 4. Структура створеної програмної моделі «Stream_TE-tunnel»

4. Програмна модель. За допомогою мови Python створено програмну модель «Stream_TE-tunnel», загальну структуру якої відображено на рис. 4.

5. Визначення оптимальних параметрів НМ. На створеній програмній моделі «Stream_TE-tunnel» проведено дослідження похибки НМ за різною кількістю прихованих шарів (1, 2, 3 та 4) з різною кількістю нейронів на кожному прихованому шарі (16, 32, 64, 128 та 256) при використанні різних функцій активації нейронів на прихованих шарах (ReLU, Leaky ReLU, Sigmoid, Tanh, ELU та Swish) та різних алгоритмів навчання (SGD, Adam, RMSprop, Adagrad, Nadam, Adamdelta та Ftrl) на вибірках різної довжини (360, 1800, 3600, 7200, 10800, 14400, 18000, 21600, 25200 та 28800 прикладів).

Найкращі результати зафіксовані при використанні наступних комбінацій функцій активації нейронів прихованих шарів: ReLU-ELU-ReLU-Swish (1,25595) та Swish-ReLU-Swish-Swish (1,25847), які показали одні з найменших значень

похибок НМ. Nadam є оптимальнішим алгоритмом навчання (оптимізатором) для даної конфігурації НМ, досягнувши найменшої валідаційної похибки (1,5301). Найменшу валідаційну похибку (0,37988) було досягнуто при 700 зразках (25200 прикладів).

6. Дослідження залежності якості розв'язання задачі інжинірингу трафіка від послідовності обробки потоків в мережі MPLS, результати якого зведені до табл. 1. У якості прикладу на рис. 5 наведені отримані результати на створеній програмній моделі (п'ятий прогін). Із таблиці видно, що найкращими є варіанти 3 та 4 (максимальний коефіцієнт навантаження складає 0,727), що незначно відрізняється від 0,748 (варіант 4 на створеній програмній моделі «Stream_TE-tunnel»).

Таким чином, проведено дослідження одного із можливих варіантів загальної структури мережі MPLS на залізничному транспорті з використанням розробленого інтелектуального засобу (рис. 6).

Отримані результати дослідження

Вар.	Черга потоків	Середня швидкість потоку, Мбіт/с	Коефіцієнт завантаження тунелів через вузли:		Максим. коефіцієнт навантаження
			3→9→10→7→6 (тунель А)	3→9→10→8→7→6 (тунель В)	
1	№ 1	30		0,857	0,857
	№ 2	20	0,364		
	№ 3	10	0,545		
2	№ 1	30		0,857	0,857
	№ 3	10	0,182		
	№ 2	20	0,545		
3	№ 2	20		0,571	0,727
	№ 1	30	0,545		
	№ 3	10	0,727		
4	№ 2	20		0,571	0,727
	№ 3	10	0,182		
	№ 1	30	0,727		
5	№ 3	10		0,286	0,909
	№ 1	30	0,545		
	№ 2	20	0,909		
6	№ 3	10		0,286	0,909
	№ 2	20	0,364		
	№ 1	30	0,909		

Прогон	Варіант	Коефіцієнт А	Коефіцієнт В	Макс. коефіцієнт
5.0	1.0	0.54	0.84	0.84
5.0	2.0	0.537	0.837	0.837
5.0	3.0	0.749	0.544	0.749
5.0	4.0	0.748	0.54	0.748
5.0	5.0	0.886	0.273	0.886
5.0	6.0	0.888	0.273	0.888

Рис. 5. Отримані результати на програмній моделі : п'ятий прогін

Висновки

1. У якості основного методу розв'язання задачі інжинірингу трафіку в мережі MPLS на залізничному транспорті взято НМ конфігурації «11-X-Y-2», де 11 – кількість вхідних (input) нейронів (це параметри трьох потоків та двох тунелів); X – кількість прихованих шарів, що потребує подальшого дослідження; Y – загальна кількість прихованих нейронів, що потребує додаткового дослідження; 2 – кількість вихідних (output) нейронів. Сформовано загальну вибірку розміром 54000 прикладів (шляхом комбінування 36 різних сценаріїв з 1500 випадково згенерованими даними).

2. За допомогою мови Python створено програмну модель «Stream_TE-tunnel», основу якої складає НМ конфігурації «11-X-Y-2». На створеній програмній моделі «Stream_TE-tunnel» визначені оптимальні параметри НМ: конфігурація «11-256-128-2» (256 – кількість нейронів першого прихованого шару, 128 – кількість нейронів дру-

гого прихованого шару); функція активації ReLU на прихованих шарах та функція активації Linear на результуючому шарі; алгоритм навчання – Adam; довжина вибірки із 25200 прикладів, при цьому значення MAE склало приблизно 0,0195 протягом 2000 епох.

3. Проведено дослідження залежності якості розв'язання задачі інжинірингу трафіку від послідовності обробки потоків в мережі MPLS залізничного транспорту. Визначено, що для найкращого варіанту максимальний коефіцієнт навантаження ТЕ-тунелю складає 0,727; на створеній програмній моделі «Stream_TE-tunnel» – 0,748.

4. Запропонований варіант загальної структури гіпотетичної мережі MPLS в ІТС залізничного транспорту з використанням розробленого нейромережного засобу для розподілу трьох потоків за двома ТЕ-тунелями при відомих значеннях середніх швидкостей потоків та пропускних спроможностей тунелів.

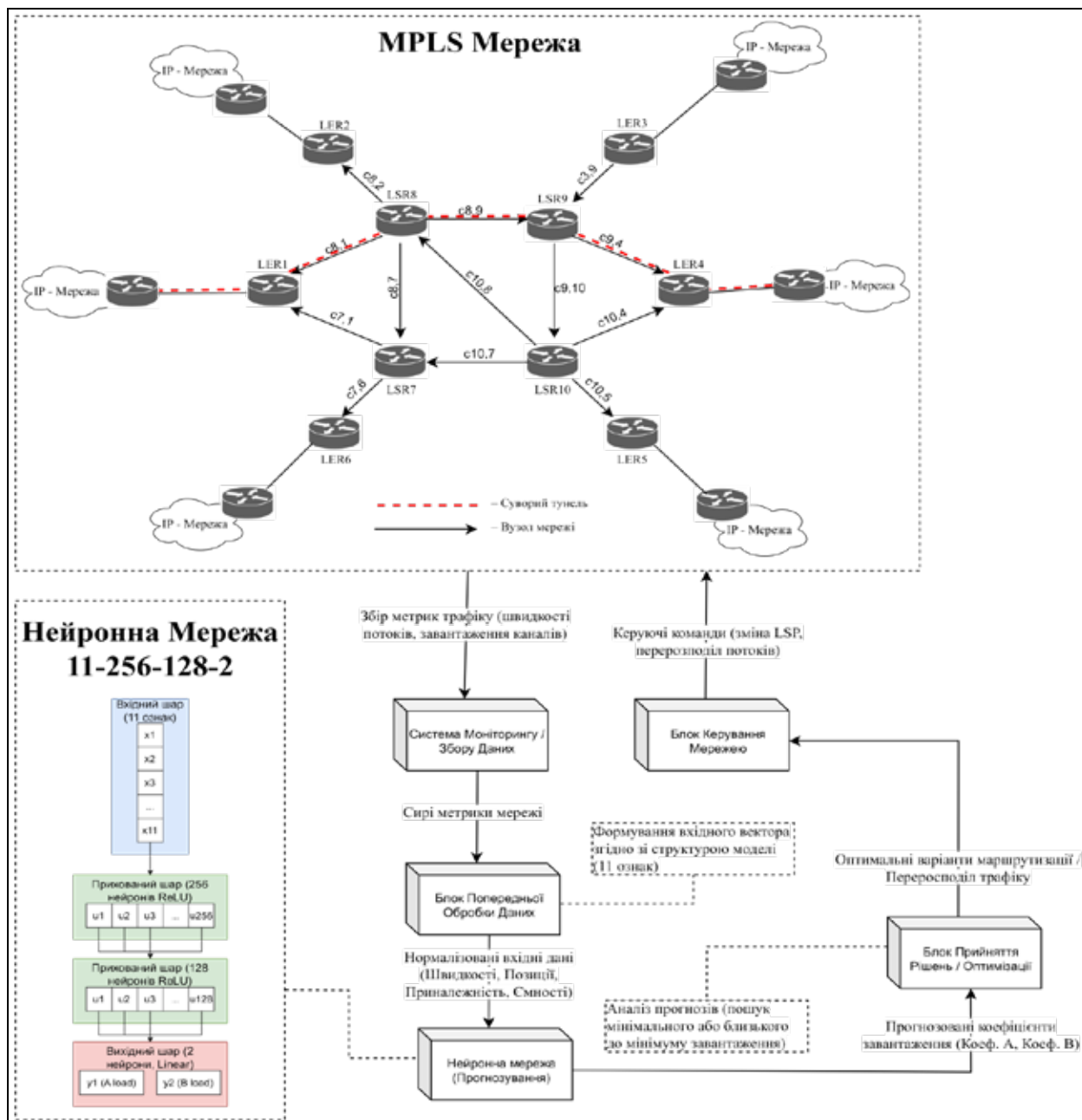


Рис. 6. Запропонований варіант інтелектуальної мережі MPLS: LER (Label Edge Router); LSR (Label Switch Router)

Список літератури:

1. Пахомова В.М. Дослідження інжинірингу трафіка в комп'ютерній мережі Укрзалізниці за технологією MPLS TE. *Наука та прогрес транспорту: Вісник Дніпропетр. нац. ун-т. заліз. трансп.* 2015. Вип. 1(55). С. 139–147. DOI: 10.15802/STP2015/38262.
2. Пахомова В.М. Дослідження інформаційно-телекомунікаційної системи залізничного транспорту з використанням штучного інтелекту: монографія. Дніпро: Вид-во ПФ «Стандарт – Сервіс», 2018. 220 с.
3. Технологія MPLS в Україні: Застосування у 2025 році. URL: <https://surl.li/mbxoou> (дата звернення: 05.05.2025).
4. Herguner K., Kalan R.S., Cetinkaya C., Sayit M. Towards QoS-aware routing for DASH utilizing MPTCP over SDN. *IEEE Conference on Network Function Virtualization and Software Defined Networks (NFV-SDN)*. (Germany: Berlin, 6–8 Nov. 2017). P. 1–6. DOI: 10.1109/nfv-sdn.2017.8169844.

5. Mahul A., Aussem A. Neural-based quality of service estimation in MPLS routers. International Conference on Artificial Neural Networks (ICANN). (Turkey: Istanbul, June 2003). Vol.: Supplementary Proceedings. P. 390–393. URL: <https://surl.li/kfanwk> (дата звернення: 05.04.2025).

6. Wu-Hsiao Hsua, Yuh-Pyng Shieha, Jenhui Chen. Multiple path selection algorithm for Diff Servaware MPLS traffic engineering. *Computer Communications*. URL: <https://pdfs.semanticscholar.org/5346/9a8d456c.pdf> (дата звернення: 11.03.2025).

7. Zhukovytsky I., Pakhomova V., Domanskay H., Nechaiev A. Distribution of information flows in the advanced network of MPLS of railway transport by means of a neural model. Energy-Optimal Technologies, Logistic and Safety on Transport (EOT): 2nd International Scientific and Practical Conference. (16 October 2019). *MATEC Web of Conferences* 294, 04007(2019). 7 p. URL: <https://doi.org/10.1051/mateconf/201929404007> (дата звернення: 10.02.2025).

Pakhomova V.M., Ivanchenko D.V. RESEARCH OF MPLS NETWORK STRUCTURE IN RAILWAY TRANSPORT USING INTELLECTUAL PROPERTY

At the present stage, the OSPF protocol is used in computer networks of railway transport, when used in real time, a problem arises due to constant changes in the volume of transmitted data, and to solve this problem, it is advisable to use MPLS technology with the created intelligent tool, which confirms the relevance of the topic. As a mathematical apparatus for solving the traffic engineering problem, a neural network of the configuration “11-X-Y-2” was taken, where 11 is the number of input neurons (these are the parameters of three flows and two tunnels); X is the number of hidden layers, which requires further research; Y is the total number of hidden neurons, which requires additional research; 2 is the number of output neurons. In the work, a corresponding software model “Stream_TE-tunnel” was created using the Python language, on which the error of neural networks with different numbers of hidden layers (1, 2, 3 and 4) with different numbers of neurons in each hidden layer (16, 32, 64, 128 and 256) was studied when using different activation functions of neurons in hidden layers (ReLU, Leaky ReLU, Sigmoid, Tanh, ELU and Swish) and different learning algorithms (SGD, Adam, RMSprop, Adagrad, Nadam, Adamdelta and Ftrl) on samples of different lengths (360, 1800, 3600, 7200, 10800, 14400, 18000, 21600, 25200 and 28800 examples). For a neural network of the configuration “11-256-128-2” (256 is the number of neurons in the first hidden layer; 128 is the number of neurons in the second hidden layer) using the ReLU activation function on the hidden layers and the Linear activation function on the resulting layer using the Adam learning algorithm, a study was conducted of the dependence of the quality of solving the traffic engineering problem on the sequence of processing flows in the MPLS network. It was determined that the maximum tunnel load factor on the created software model “Stream_TE-tunnel” was 0.748, which differs from the theoretical value by approximately 2.8%. Thus, the work investigated one of the possible options for the structure of the MPLS network in railway transport using the created intelligent tool for distributing three flows across two tunnels with known values of average flow speeds and tunnel throughputs.

Key words: MPLS domain, LER, LSR, flow, average speed, TE-tunnel, throughput, load factor, neural network, error.

Дата надходження статті: 25.07.2025

Дата прийняття статті: 31.07.2025

Опубліковано: 27.10.2025

Петросян А.Р.

Державний університет «Житомирська політехніка»

МЕТОД ЛОКАЛІЗАЦІЇ ОБ'ЄКТА У ПРОСТОРІ ІЗ ЗАСТОСУВАННЯМ ГЕНЕТИЧНО-ОПТИМІЗОВАНОГО ФІЛЬТРА ЧАСТИНОК

У статті запропоновано новий метод локалізації об'єкта в просторі, заснований на використанні генетично-оптимізованого фільтра частинок. Актуальність дослідження обумовлена зростаючими вимогами до точності та швидкодії алгоритмів локалізації, які застосовуються в сфері автономних безпілотних повітряних суден. Висока динамічність польоту, шумність сигналів датчиків, нелінійність моделей руху і негаусівський характер перешкод істотно ускладнюють завдання точного визначення положення і орієнтації об'єктів.

Проведений аналіз існуючих методів показав, що класичні методи, такі як фільтр Калмана, розширений фільтр Калмана, фільтр Калмана без запаху, мають значні обмеження. Такий метод, як фільтр частинок, вирішує зазначені проблеми, проте він має також недолік – виродження частинок. Для усунення цієї проблеми традиційно застосовують метод повторної вибірки, що, однак, призводить до втрати різноманітності частинок та погіршує точність.

Запропонований у дослідженні алгоритм базується на принципах роботи генетичного алгоритму. Це дозволяє значно розширити простір пошуку оптимального стану об'єкта та уникнути передчасної збіжності, яка характерна для класичних фільтрів частинок. Для оцінки пристосованості частинок розроблена фітнес-функція, яка забезпечує максимізацію точності локалізації та зменшує обчислювальні витрати.

Метод перевірено шляхом математичного моделювання переміщення об'єкта в двовимірному просторі. Результати експериментальних досліджень продемонстрували, що запропонований метод перевершує класичні методи SIR і BPF за критерієм кореня середньоквадратичної похибки приблизно на 17%. Також проведено порівняльний аналіз швидкодії алгоритму при реалізації на мікроконтролерах різних архітектур (AVR, ARM, MIPS), де доведено зменшення часу обчислення фітнес-функції в 7–10 разів у порівнянні з традиційною функцією правдоподібності, що показує перспективність методу для практичного застосування в автономних системах з обмеженими обчислювальними ресурсами.

Ключові слова: локалізація, БПС, безпілотні повітряні судна, фільтр частинок, генетичний алгоритм, виродження частинок, фітнес-функція, обчислювальна ефективність.

Постановка проблеми. Проблема точної локалізації об'єктів у просторі залишається однією з найважливіших та критичних у розвитку сучасних автономних систем пересування (роботів). Особливої актуальності вона набуває у контексті безпілотних повітряних суден (БПС), де від точного та швидкого визначення власного стану та орієнтації безпосередньо залежить безпека та ефективність виконання завдань: навігація, зйомка, моніторинг, доставка тощо, а також взаємодія з навколишнім середовищем. Ця проблема є основою як автономної навігації, але і в системах доповненої реальності, комп'ютерного зору тощо.

Завдання визначення положення та напрямку БПС істотно ускладнюється через наявність значної невизначеності та впливу шумів на вихідні

дані від датчиків. Джерелами таких шумів є інерційні вимірювальні системи, що страждають від дрейфу і вібрацій, системи глобального позиціонування (GNSS/GPS), які схильні до перешкод, відсутності сигналу, а також решта датчиків (камери, лідари тощо), на які впливають умови освітлення, погодні умови та інші фактори. Необхідність інтеграції даних з різних, часто неточних і конфліктуючих джерел, в поєднанні з динамічністю середовища польоту, перетворює локалізацію в надзвичайно складну, багатофакторну задачу оцінки стану БПС.

Існуючі класичні підходи до локалізації та відстеження траєкторії, такі як фільтр Калмана та його нелінійні варіанти: розширений фільтр Калмана, фільтр Калмана без запаху та фільтр частинок, демонструють певну ефективність, але

мають деякі обмеження. Фільтри Калмана часто зазнають невдач при великій нелінійності моделей руху або при значних негауссових шумах. Фільтр частинок здатен краще обробляти нелінійності та мультимодальність, але стикається з проблемою «виродження частинок». Вказані проблеми можуть приводити до втрачання треку або давати грубі похибки. У зв'язку з цим виникає нагальна потреба в розробці та вдосконаленні методів локалізації БПС.

Аналіз останніх досліджень і публікацій. Проблема точного визначення та відстеження просторового положення об'єктів (локалізація) є важливою для широкого спектру сучасних завдань: в області автономних транспортних засобів вона дозволяє визначати точне положення автомобіля для безпечної навігації та запобігання зіткненням; в робототехніці локалізація використовується для точного позиціонування роботів, наприклад, в складських приміщеннях; БПС використовують локалізацію для виконання завдань картографування, доставки вантажів тощо; у віртуальній та доповненій реальності для відстеження положення користувачів тощо. У кожній із завдань надійна локалізація об'єктів служить критичною основою для прийняття рішень, автоматизації та підвищення загальної ефективності технічних систем.

З математичної точки зору, метод оцінки стану динамічного об'єкта в часі часто формулюється в рамках байєсівського підходу. Байєсівська фільтрація є ймовірнісним методом, який дозволяє послідовно оновлювати апостеріорну щільність ймовірності стану об'єкта. Цей процес включає два послідовних етапи: прогноз та корекція.

Найбільш відомою і частою реалізацією цього підходу, що використовується, є фільтр Калмана [1, 2]. Він забезпечує оптимальну оцінку в сенсі мінімуму середньоквадратичної похибки для лінійних систем з адитивними гауссовими шумами як в моделі руху, так і в моделі вимірювань. Однак переважна більшість практичних завдань локалізації об'єкта, особливо пов'язаних з маневруючими об'єктами, характеризуються нелінійностями.

Для адаптації байєсівського підходу до нелінійних систем було розроблено дві широко поширені модифікації фільтра Калмана: розширений фільтр Калмана (заснований на лінеаризації нелінійних функцій); фільтр Калмана без запаху (вибирає набір точок, пропускає їх через нелінійні функції моделі, а потім апроксимує апостеріорний розподіл за перетвореними точками). Другий варіант часто демонструє вищу точність при порівнян-

ній обчислювальній складності, проте через простоту реалізації перший варіант фільтра також дуже часто застосовується в багатьох завданнях. Так, у матеріалах [3, 4] описується застосування розширеного фільтра Калмана в автопілотах БПС PX4 і Ardupilot, а в статті [5] проводиться порівняльний аналіз обох модифікацій фільтра в задачах вимірювання відстані до об'єкта в просторі.

Ще одним з ефективних методів оцінки в рамках байєсівського підходу є фільтр частинок (ФЧ). Основний принцип ФЧ полягає в представленні апостеріорної щільності ймовірності стану об'єкта за допомогою дискретної апроксимації, що реалізується набором зважених випадкових вибірок (частинок). Кожна частинка моделює можливий стан системи і асоціюється з ваговим коефіцієнтом, що відображає її відповідність спостережуваному даним. Подібно до інших байєсівських методів оцінки, ФЧ здійснює рекурсивну оцінку стану, що включає два послідовних етапи. Ключовою перевагою ФЧ перед модифікованими фільтрами Калмана є його здатність забезпечувати оцінку стану об'єкта з більшою точністю у завданнях з великою нелінійністю моделей руху та довільним розподілом випадкових величин [6]. Однак ефективність ФЧ на практиці обмежується низкою проблем: виродження частинок та обчислювальна ефективність. Дані проблеми взаємосуперечливі. У зв'язку з тим, що швидкість сучасних комп'ютерів висока, то в більшості випадків фільтри ускладнюються для усунення виродження частинок, що призводить до зниження обчислювальної ефективності. Однак не всі завдання вирішуються на комп'ютері. Багато обчислювальних методів реалізуються в мікроконтролерах (МК), в тому числі і польотний контролер БПС, що вимагає пошуку компромісу між точністю і швидкістю алгоритмів.

Для усунення виродження частинок використовується операція повторної вибірки (англ. resampling). Повторна вибірка зазвичай виконується після оцінки стану об'єкта. Мета повторної вибірки – замінити частинки з низькими ваговими коефіцієнтами частинками з високими ваговими коефіцієнтами. Однак ця операція призводить до втрати різноманітності частинок (англ. particle diversity). Це погіршує точність та надійність позиціонування, що може призвести до нестійкості фільтрації та втрати треку. Даний підхід використовується в методі SIR (Sampling Importance Resampling).

У статті [7] проводиться огляд оптимальних і субоптимальних байєсівських алгоритмів оцінки

для задач нелінійного/негаусівського відстеження стану об'єкта, з особливим акцентом на ФЧ. В рамках загальної структури алгоритму послідовного вибіркового зважування представлено кілька варіантів ФЧ, таких як SIR, ASIR і RPF. Ці варіанти порівнюються з розширеним фільтром Калмана на наочному прикладі.

ASIR (Auxiliary Sampling Importance Resampling) – це модифікація SIR-фільтра, спрямована на усунення проблеми різноманітності частинок після повторної вибірки. На відміну від традиційного SIR, де вага частинки призначається після повторної вибірки, ASIR враховує не тільки минулий стан, але і поточне спостереження, що дозволяє заздалегідь відібрати більш перспективні частинки, а це знижує розкид вагових коефіцієнтів і підвищує обчислювальну ефективність ФЧ при малій кількості частинок.

RPF (Regularized Particle Filter) – це також модифікація SIR-фільтра, спрямована на усунення тих самих проблем. Замість того щоб просто копіювати частинки з високою вагою, RPF додає до них невеликі випадкові зміщення, використовуючи ядерне згладжування. Це дозволяє зберегти різноманітність частинок і більш точно апроксимувати апостеріорний розподіл. Метод особливо ефективний при низькому рівні шуму, коли частинки схильні скупчуватися в одній області.

У роботі зазначається, що розширений фільтр Калмана сильно програє в точності через нездатність враховувати багатомодальні та нелінійні особливості моделей. Всі ФЧ показують кращі результати за метрикою кореня середньоквадратичної помилки (англ. RMSE). Однак їх обчислювальна ефективність значною мірою залежить від варіантів усунення виродження частинок і завдання, яке вирішується. Це означає, що використовуючи різні варіанти повторної вибірки, можна отримувати різні модифікації ФЧ, які будуть краще підходити для певних завдань.

Так в статті [8] пропонується вдосконалений метод візуального відстеження об'єктів на основі ФЧ, де для повторної вибірки використовується оптимізація роєм частинок. Даний метод дозволяє шукати вибірки навколо останнього стану об'єкта в залежності від поточного спостереження. Таким чином, він може розподілити частинки в області з високою правдоподібністю, навіть якщо динамічну модель об'єкта неможливо отримати. Це виконується за рахунок того, що спочатку знаходять область вибірки за допомогою алгоритму оптимізації роєм частинок, а потім розподіляють

частинки на основі двох різних базових точок, щоб досягти різноманітності та збіжності фільтра.

Виявивши паралелі в архітектурі ФЧ і генетичних алгоритмів (ГА), вчені побачили можливість для гібридного підходу. Механізми відбору, схрещування, мутації стали вбудовуватися в процес фільтрації, що дозволяє виключити виродження частинок. Тепер завдання локалізації об'єкта в просторі постає як послідовна оптимізація, яка «еволюціонує» до більш точного рішення. Останнім часом було проведено велику кількість досліджень з використанням ГА для усунення виродження частинок і було запропоновано ряд рішень [9-11].

У статті [9] ГА включений в ФЧ. Використання генетичних операторів дозволяє уникнути передчасної збіжності частинок і розширити область пошуку частинок. Для передбачуваного розподілу, заснованого на ГА, виводяться відповідні вагові коефіцієнти, щоб апроксимувати цільовий розподіл. Результати показують, що даний метод перевершує класичні ФЧ (SIR, APF, RPF тощо) в задачах, де можливі різкі зміни стану об'єкта.

У роботі [10] запропоновано вдосконалений алгоритм ФЧ, що об'єднує ГА і ФЧ. У ньому частинки відбираються відповідно до ймовірності та принципу селекції ГА, що підвищує шанси вибору частинок з більшими ваговими коефіцієнтами і направляє процес у бік оптимального результату. Кросовер і мутація замінюють стратегію повторної вибірки ФЧ, яка просто копіює частинки з високими ваговими коефіцієнтами і видаляє з низькими, реалізуючи оновлення частинок.

У наступній статті [11] для збільшення різноманітності частинок і підвищення точності позиціонування пропонується вдосконалений метод повторної вибірки, заснований на генетичній оптимізації. Даний метод оптимізує розподіл повторно обраних частинок за допомогою п'яти операторів: відбору, огрублення, класифікації, кросовера і мутації. Зазначений метод повторної вибірки на основі генетичної оптимізації інтегрується в структуру ФЧ.

Як показує аналіз, ГА використовується для усунення збіднення частинок ФЧ при повторній вибірці, тому в більшості випадків повторна вибірка замінюється на оператори ГА: кросовер і мутація. ГА є потужним інструментом оптимізації і використовується для вирішення багатьох завдань [12, 13]. Однак застосування ГА може дозволити не тільки вирішити завдання оптимізації, але і підвищити швидкість методу [11, 14]. Зробити це можна двома шляхами: зменшивши

кількість індивідуумів ГА (в рамках ФЧ – зменшити кількість частинок); модифікувавши сам метод, використовуючи більш ефективні операції.

Постановка завдання. Метою даної роботи є розробка методу локалізації об'єкта в просторі, заснованого на генетичному алгоритмі для подолання фундаментальної проблеми виродження частинок. Основне завдання – знайти компроміс між мінімізацією похибки оцінки положення об'єкта в просторі та збільшенням швидкодії методу.

Виклад основного матеріалу. Розглядаємий метод є вдосконаленням методу, який викладено в тезах [14]. У зазначеному матеріалі було зроблено спробу реалізувати фільтрацію частинок з використанням класичного ГА. Оскільки використовувався класичний ГА, то в матеріалі не враховувалася модель зміни стану об'єкта, що вносило додаткову похибку.

Як вже зазначалося, ФЧ є послідовним методом Монте-Карло для байєсівської фільтрації, де оптимальна оцінка стану динамічної системи визначається через модель простору станів, що формалізує зв'язок між прихованими станами x_k та спостереженнями y_k [11]:

$$x_k = f(x_{k-1}) + w_k, \quad (1)$$

$$y_k = h(x_k) + v_k, \quad (2)$$

де w_k , v_k – білий шум системи і спостереження відповідно; $f()$, $h()$ – відомі функції переходу і спостереження відповідно.

ФЧ апроксимує апостеріорний розподіл стану $p(x_k | y_{1:k})$ набором випадкових частинок $\{x_k^i, w_k^i\}_{i=1}^{N_p}$ з передбачуваного розподілу $q(x_k | y_{1:k})$:

$$p(x_k | y_{1:k}) = \sum_{i=1}^{N_p} w_k^i \cdot \delta(x_k - x_k^i), \quad (3)$$

де $\delta()$ – дельта-функція Дірака, N_p – кількість частинок, w_k^i – нормалізований ваговий коефіцієнт. Вагові коефіцієнти розраховуються у відповідності до виразу (4):

$$\tilde{w}_k^i \propto \frac{p(x_k^i | y_{1:k})}{q(x_k^i | y_{1:k})}. \quad (4)$$

Вагові коефіцієнти частинок на кожному кроці можуть бути оновлені відповідно до виразу (5):

$$\tilde{w}_k^i \propto \tilde{w}_{k-1}^i \frac{p(y_k | x_k^i) p(x_k^i | x_{k-1}^i)}{q(x_k^i | x_{k-1}^i, y_k)}. \quad (5)$$

На практиці розподіли $p()$ і $q()$ у виразі (4) вважають однаковими, тому вираз (5) можна записати в такому вигляді (6):

$$\tilde{w}_k^i \propto \tilde{w}_{k-1}^i p(y_k | x_k^i). \quad (6)$$

Вагові коефіцієнти (6) необхідно нормалізувати на кожному кроці при повторній вибірці наступним чином (7):

$$w_k^i \propto \frac{\tilde{w}_k^i}{\sum_{i=1}^{N_p} \tilde{w}_k^i}. \quad (7)$$

Після математичного опису ФЧ розглянемо його реалізацію на базі ГА (надалі будемо використовувати скорочену назву PFonGA для даного методу). Розглянемо етапи виконання методу PFonGA.

1. Ініціалізація.

Генерація початкових особин $\{x_0^i\}_{i=1}^{N_p}$ з апіорного розподілу $p(x_0)$ та ваговими коефіцієнтами $w_0^i = \frac{1}{N_p}$.

2. Спостереження.

На даному етапі здійснюється отримання інформації з різних датчиків (акселерометра, гіроскопа, магнітометра тощо) БПС для прогнозування його положення відповідно до моделі руху (8):

$$x_k^i \sim p(x_k | x_{k-1}^i). \quad (8)$$

3. Оновлення.

На наступному етапі відповідно до фітнес-функції обчислюється пристосованість особин. В якості фітнес-функції візьмемо мінімум відхилення кожної особини від передбачуваного положення об'єкта (9):

$$\varphi(y_k, x_k^i) = (y_k - h(x_k^i))^2 \rightarrow \min. \quad (9)$$

Однак ГА використовується для пошуку максимуму, тому фітнес-функцію необхідно змінити наступним чином:

$$\varphi(y_k, x_k^i) = \tilde{c} - (y_k - h(x_k^i))^2 \rightarrow \max, \quad (10)$$

де $\tilde{c}$ – константа. Щоб вираз виглядав більш гармонійно, задамо дану константу наступним чином:

$$\tilde{c} = (c\sigma)^2, \quad (11)$$

де σ – середньоквадратичне відхилення шуму моделі (2).

Беручи до уваги, що фітнес-функція не може бути від'ємною і підставляючи вираз (11) в (10), остаточно вона буде мати вигляд (12):

$$\varphi(y_k, x_k^i) = \max\left(0, (c\sigma)^2 - (y_k - h(x_k^i))^2\right) \rightarrow \max, \quad (12)$$

На кожній ітерації в ГА виконується відбір особин. Дана операція ідентична повторній вибірці

в ФЧ, тому вагові коефіцієнти будуть скидатись, а вираз (6) матиме такий вигляд:

$$\tilde{w}_k^i \propto \varphi(y_k, x_k^i). \quad (13)$$

Вагові коефіцієнти (13) нормалізуються у відповідності до виразу (7). Після чого виконується розрахунок апостеріорної оцінки положення об'єкта відповідно до виразу (14):

$$\hat{x}_k = \sum_{i=1}^N w_k^i x_k^i. \quad (14)$$

4. Відбір.

Існує безліч різних стратегій відбору особин. У нашому випадку використовувався метод рулетки. Він носить ймовірнісний характер і вибирає батьків на основі їх пристосованості, де більш пристосовані особини мають більшу ймовірність бути обраними, але при цьому менш пристосовані особини також мають шанс бути обраними, що вносить різноманітність в популяцію.

5. Кросовер.

У зв'язку з тим, що хромосоми у нас з реальними числами, тому в було використано кросовер для речевого кодування:

$$x_k^{i,cr} = 0.5(x_k^i(1+\beta) + x_k^j(1-\beta)), \quad (15)$$

$$x_k^{j,cr} = 0.5(x_k^i(1-\beta) + x_k^j(1+\beta)), \quad (16)$$

де β – береться з рівномірного розподілу в діапазоні $[0, 1]$.

6. Мутація.

Для деяких особин проводилася операція мутації відповідно до виразу:

$$x_k^{j,mut} = x_k^i + \eta, \quad (17)$$

де η – випадкове значення з рівномірного розподілу в діапазоні $[0, 1]$.

Для перевірки роботи методу було проведено моделювання за допомогою мови програмування Python. Для реалізації ГА необхідно налаштувати гіперпараметри. У нашому випадку вони матимуть такі значення: кількість особин – 500; ймовірність операції кросовер – 0,9; ймовірність операції мутації – 0,2. Також у формулі (12) необхідно задати значення константи $c = 4$. Усі параметри підбрані емпіричним шляхом. Моделювання проводилося в 2D-просторі. На рис. 1 показано результат переміщення об'єкта по колу.

RMSE за результатами 10 експериментів склало 0.483. Для порівняльного аналізу були проведені експерименти з методами SIR і BPF, які показали значення 0.585 і 0.609 відповідно за рівних умов.

Метод базується на ГА, тому функція правдоподібності не використовується, а задіяна фітнес-функція, що дозволило отримати вирази за швидкодією при реалізації методу на МК. Порівняльний аналіз часу виконання цих функцій для деяких МК з різною архітектурою представлений у табл. 1.

Як видно з табл. 1, різниця становить в 7–10 разів.

Висновки. У статті запропоновано метод локалізації об'єкта в просторі на основі генетично-оптимізованого фільтра частинок (PFonGA), що дозволяє ефективно усувати проблему виродження частинок. Запропонований підхід завдяки використанню генетичного алгоритму дозволяє також зменшити час обчислень фітнес-функції на мікроконтролерах різних архітектур у 7–10 разів порівняно з традиційною функцією правдоподібності. Результати моделювання підтвердили переваги методу за критерієм RMSE (на 17% у порівнянні з класичними методами SIR і BPF). Метод

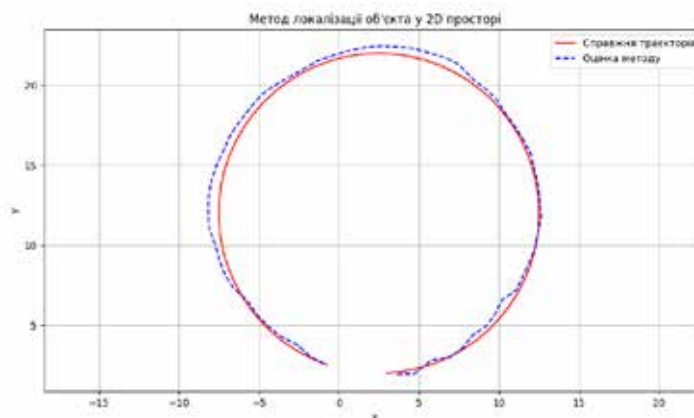


Рис. 1. Результат відстеження переміщення об'єкта з використанням методу PFonGA

Час виконання функції правдоподібності та фітнес-функції на МК

Архітектура	Час виконання функцій, μ s	
	традиційна функція правдоподібності	фітнес-функція методу PFonGA
AVR (ATmega164A, 16МГц)	280.6	43.0
ARM (ATSAMD21G18A, 16МГц)	108.6	9.8
MIPS (PIC32MX110F016B, 16МГц)	208.5	19.0

є перспективним для автономних систем, де критичним є компроміс між точністю локалізації та обчислювальною ефективністю. Подальші дослі-

дження повинні бути спрямовані на оптимізацію гіперпараметрів генетичного алгоритму і параметрів фітнес-функції.

Список літератури:

1. Industrial Applications of the Kalman Filter: A Review / F. Auger та ін. *IEEE Transactions on Industrial Electronics*. 2013. Т. 60, № 12. С. 5458–5471. URL: <https://doi.org/10.1109/tie.2012.2236994>.
2. Kordić V. Kalman Filter. InTech, 2010. 402 с.
3. Using PX4's Navigation Filter (EKF2). *PX4 Guide (main)*. URL: https://docs.px4.io/main/uk/advanced_config/tuning_the_ecl_ekf.html (дата звернення: 26.06.2025).
4. Extended Kalman Filter (EKF) – Copter documentation. *ArduPilot – Versatile, Trusted, Open*. URL: <https://surl.li/cc/kqmuqp> (дата звернення: 26.06.2025).
5. EKF- and UKF-Based Estimators for Radar System / U. K. Singh та ін. *Frontiers in Signal Processing*. 2021. Т. 1. URL: <https://doi.org/10.3389/frsip.2021.704382>.
6. Evaluation of Localization by Extended Kalman Filter, Unscented Kalman Filter, and Particle Filter-Based Techniques / I. Ullah та ін. *Wireless Communications and Mobile Computing*. 2020. Т. 2020, № 1. С. 1–15. URL: <https://doi.org/10.1155/2020/8898672>.
7. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking / M. S. Arulampalam та ін. *IEEE Transactions on Signal Processing*. 2002. Т. 50, № 2. С. 174–188. URL: <https://doi.org/10.1109/78.978374>.
8. Zhao J., Li Z. Particle filter based on Particle Swarm Optimization resampling for vision tracking. *Expert Systems with Applications*. 2010. Т. 37, № 12. С. 8910–8914. URL: <https://doi.org/10.1016/j.eswa.2010.05.086>.
9. A New Evolutionary Particle Filter for the Prevention of Sample Impoverishment / Seongkeun Park та ін. *IEEE Transactions on Evolutionary Computation*. 2009. Т. 13, № 4. С. 801–809. URL: <https://doi.org/10.1109/tevc.2008.2011729>.
10. Zhao B., Hu J., Ji B. An improved particle filter based on genetic resampling. *International Conference on Automation, Mechanical Control and Computational Engineering*. Atlantis Press, 2015. С. 1353–1358. URL: <https://doi.org/10.2991/amcce-15.2015.125>.
11. A Genetic Optimization Resampling Based Particle Filtering Algorithm for Indoor Target Tracking / N. Zhou та ін. *Remote Sensing*. 2021. Т. 13, № 1. С. 132. URL: <https://doi.org/10.3390/rs13010132>.
12. Petrosian R.V., Kuzmenko O.V., Petrosian A.R. Method for calculating the fir filter based on genetic algorithm. *International scientific journal «Computer Systems and Information Technologies»*. Khmelnytskyi. 2021. № 1(3). pp. 19–24. DOI: 10.31891/CSIT-2021-3-3.
13. Petrosian R., Pilkevych I., Petrosian A. Algorithm for optimizing a PID controller model based on a digital filter using a genetic algorithm. *Proceedings of the 3rd Edge Computing Workshop*. 2023. Vol. 3374. pp. 97–111. URL: <https://ceur-ws.org/Vol-3374/paper07.pdf>.
14. Петросян А. Р. Метод локалізації об'єкта в просторі на основі генетичного алгоритму. *Інтегровані інтелектуальні робототехнічні комплекси (ІРТК-2025)*. Вісімнадцята міжнародна науково-практична конференція 20–21 травня 2025 р., Київ, Україна. К.: НАУ, 2025. С. 368–369.

Petrosian A.R. OBJECT LOCALIZATION METHOD IN SPACE USING A GENETICALLY-OPTIMIZED PARTICLE FILTER

The article proposes a new method for localizing an object in space based on the use of a genetically optimized particle filter. The relevance of the research is due to the growing demands for accuracy and speed of localization algorithms used in the field of autonomous unmanned aerial vehicles. High flight dynamics, noisy sensor signals, nonlinear motion models, and non-Gaussian noise significantly complicate the task of accurately determining the position and orientation of objects.

An analysis of existing methods has shown that classical methods, such as the Kalman filter, the extended Kalman filter, and the odorless Kalman filter, have significant limitations. A method such as the particle filter solves these problems, but it also has a drawback: particle degeneration. To eliminate this problem, the resampling method is traditionally used, which, however, leads to a loss of particle diversity and reduces accuracy.

The algorithm proposed in the study is based on the principles of a genetic algorithm. This allows for a significant expansion of the search space for the optimal state of the object and avoids premature convergence, which is characteristic of classical particle filters. To evaluate the fitness of particles, a fitness function has been developed that maximizes localization accuracy and reduces computational costs.

The method was verified by mathematical modeling of object movement in two-dimensional space. The results of experimental studies showed that the proposed method outperforms the classical SIR and BPF methods by approximately 17% in terms of the root mean square error criterion. A comparative analysis of the algorithm's performance when implemented on microcontrollers of different architectures (AVR, ARM, MIPS) was also carried out, proving a reduction in the calculation time of the fitness function by 7–10 times compared to the traditional likelihood function, which shows the promise of the method for practical application in autonomous systems with limited computing resources.

Key words: *localization, UAV, unmanned aerial vehicles, particle filter, genetic algorithm, particle degeneracy, fitness function, computational efficiency.*

Дата надходження статті: 01.08.2025

Дата прийняття статті: 09.09.2025

Опубліковано: 27.10.2025

Правовєров А.О.

Національний технічний університет України

«Київський політехнічний інститут імені Ігоря Сікорського»

МОДЕЛЮВАННЯ ПРОЦЕСУ ЗГОРЯННЯ БІОМАСИ В КОТЕЛЬНОЇ УСТАНОВЦІ З МЕТОЮ ПОКРАЩЕННЯ ЕКОЛОГІЧНИХ ХАРАКТЕРИСТИК

В контексті переходу до відновлюваних джерел енергії, спалювання біомаси в котельних установках набуває актуальності, як альтернатива традиційним видам палива. Робота присвячена математичному моделюванню процесу згоряння біомаси в котельній установці з урахуванням основних стадій теплового розкладу палива. Актуальність дослідження полягає у необхідності зниження рівня шкідливих викидів у процесах теплогенерації з використанням біомаси. Метою роботи було моделювання фізико-хімічних процесів згоряння біомаси у котельній установці з метою підвищення екологічної безпеки використання відновлюваних енергетичних ресурсів. У статті представлено математичні моделі процесу згоряння біопалива в котельній установці з урахуванням фазових перетворень палива на стадіях сушіння, піролізу, окиснення летких сполук та згоряння вуглецевого залишку. Розроблена модель включала рівняння матеріального балансу, теплообміну в частинках біомаси, що згоряє та кінетичні рівняння. Проведено оцінку впливу технологічних параметрів процесу на утворення оксиду вуглецю (II) та оксидів азоту NOx. Особливу увагу приділено ролі коефіцієнта надлишку повітря, температури в камері згоряння, а також співвідношенню первинного та вторинного повітря в утворенні шкідливих викидів. Запропонована математична модель базується на системі диференціальних рівнянь першого порядку. Система рівнянь включала кінетику зміни маси вологи, летких компонентів та вуглецевого залишку в процесі згоряння біомаси. Для опису процесів утворення оксидів азоту застосовано механізм Зольдовича, що враховує температурні залежності швидкості хімічних реакцій за участю молекулярного азоту та кисню. Розроблена модель дозволяє відтворити реальні умови роботи котельних установок, які працюють на біомасі, з урахуванням безперервної подачі палива до камери згоряння. Запропоновано режими оптимального керування котельною установкою при надлишку повітря $\lambda = 1.2-1.5$ та за умов підтримання температури в середині камери згоряння в діапазоні 850–1100 °C. Запропоновані параметри роботи установки сприятимуть мінімізації викидів CO та NOx в димових газах.

Ключові слова: згоряння біомаси, котельні установка, математичне моделювання, екологічна безпека, тепловий баланс, газові викиди.

Постановка проблеми. У сучасних умовах зростаючої енергетичної кризи та необхідності зниження викидів парникових газів, все більшого значення набуває використання відновлюваних джерел енергії, зокрема біомаси [1, с. 2]. Біомаса, як альтернативне паливо, характеризується відносно низьким вмістом сірки та потенційно нейтральним вуглецевим балансом, що робить її привабливою з точки зору екологічної безпеки. Проте процес її згоряння є складним і супроводжується низкою проблем – нестабільністю тепловіддачі, утворенням оксидів азоту, чадного газу, твердих частинок та інших забруднювальних речовин [2, с. 3]. Надмірні викиди вуглекис-

лого газу (CO₂) в атмосферу та загроза смогу, особливо в районах великих міських агломерацій, стали серйозною екологічною проблемою. До цього додаються жорсткі законодавчі обмеження з боку країн Європейського Союзу (ЄС), що зумовило необхідність поетапної відмови від використання викопних твердих палив у малому енергетичному секторі [4, с. 2]. У більшості країн ЄС антрацит та кам'яне вугілля майже повністю виведені з ринку як паливо для малопотужних котлів у домогосподарствах. Їх місце зайняли відновлювані джерела енергії – деревна та аграрна біомаса, які залишаються доступними та перспективними альтернативами для опалення приватних

житлових будівель [5, с. 155]. Згоряння твердого біопалива у побутових котлах малої потужності може сприяти зменшенню загальних викидів CO_2 в атмосферу завдяки нейтральному вуглецевому балансу [6, с. 2]. Проте саме по собі це не гарантує зменшення рівня небезпечних викидів, зокрема монооксиду вуглецю (CO), оксидів азоту (NO_x) та твердих частинок (пилу). Ці речовини є складовими смогу і безпосередньо впливають на здоров'я населення [1-3]. Ефективне спалювання біомаси в котельних установках вимагає точного регулювання режимів горіння, конструктивних особливостей паликових пристроїв і систем подачі палива. Традиційні експериментальні методи оптимізації таких процесів є дорогавартісними та час затратними, особливо при необхідності врахування широкого спектра параметрів. У цьому контексті чисельне моделювання процесу згоряння біомаси набуває ключового значення як інструмент для аналізу, прогнозування та оптимізації теплотехнічних і екологічних характеристик котельних установок. У зв'язку з цим з 2020 року в країнах ЄС набула чинності директива ECODESIGN. Вона регламентує екологічні вимоги до процесу згоряння біомаси в котлах малої потужності, призначених для опалення індивідуальних будинків. Згідно з цією нормою, гранично допустимі викиди повинні становити: CO – не більше 500 мг/Нм³, NO_x – до 200 мг/Нм³, пилу – до 40 мг/Нм³ при вмісті кисню у відхідних газах 10% [4-6].

Забезпечення відповідності цим вимогам передбачає не лише використання якісної сировини, а й суттєве вдосконалення конструкційних та експлуатаційних характеристик котлів. Зокрема, йдеться про підвищення ефективності згоряння, що тісно пов'язано зі зменшенням викидів забруднювальних речовин [7, с. 2]. Недостатнє опрацювання моделей згоряння біомаси з урахуванням її реологічних і хімічних властивостей, а також впливу конструктивних параметрів котлів на ефективність горіння залишається актуальною проблемою. Це створює потребу у розробці та вдосконаленні відповідних математичних і комп'ютерних моделей [8, с. 7]. Це дозволить не лише підвищити ефективність енергетичних систем, що працюють на біомасі, але й суттєво знизити їх екологічне навантаження.

Аналіз останніх досліджень і публікацій. Останні дослідження свідчать про зростаючий інтерес наукової спільноти до моделювання та оптимізації процесів згоряння біомаси з метою підвищення енергоефективності та

зменшення негативного впливу на довкілля. Основна увага зосереджена на розробці чистих технологій горіння, удосконаленні конструкцій котлів, а також екологічному моніторингу викидів забруднювальних речовин. З огляду на актуальність проблеми, у численних наукових установах ЄС активно ведуться дослідження, спрямовані на ідентифікацію властивостей біомаси, яка використовується як паливо в малопотужних котлах. Паралельно здійснюється моделювання процесу її згоряння та оптимізація параметрів горіння з метою досягнення високої енергоефективності та низьких викидів [1, с. 3]. З огляду на актуальність проблеми, у численних наукових установах ЄС активно ведуться дослідження, спрямовані на ідентифікацію властивостей біомаси, яка використовується як паливо в малопотужних котлах. Паралельно здійснюється моделювання процесу її згоряння та оптимізація параметрів горіння з метою досягнення високої енергоефективності та низьких викидів [3]. Згідно з узагальненим аналізом Шарми та співавторів [1, с. 3], ефективне згоряння біомаси вимагає комплексного підходу. Він передбачає вибір оптимального типу палива, контроль температурного режиму, коректне дозування повітря, а також застосування методів вторинного очищення газів. Особливу увагу автори приділяють зменшенню корозійного зношування та підвищенню екологічності процесу. Математичне моделювання є важливим інструментом для передбачення динаміки згоряння. Наприклад, у роботі Баки, Носека та Чаї [2, с. 4] розроблено аналітичну модель горіння біомаси, яка враховує диференційовані механізми теплового розкладу палива. Вона дає змогу оцінити вплив основних параметрів на температуру та швидкість горіння. Бала-Літвіняк і Мусіал [3, с. 4] поєднали чисельні та експериментальні підходи для вивчення згоряння різних типів біомаси в побутовому котлі. Їхні результати демонструють значну варіативність температурного режиму та структури залишкових продуктів залежно від вологості й складу палива. Дослідження впливу золи від згоряння аграрної біомаси на навколишнє середовище провели Сюй та співавтори [4, с. 5]. Вони встановили, що зола може містити потенційно токсичні мікроелементи, а отже, потребує контролю під час утилізації або вторинного використання.

У галузі оптимізації конструкцій котельних установок важливими є роботи, що висвітлюють моделювання подачі повітря та впливу геометрії топки. Зокрема, Су та співавтори [8, с. 4] проаналізували розподіл первинного повітря у ґратковому

котлі потужністю 130 т/год. Це дозволило знизити локальні викиди CO та NO_x шляхом коригування подачі кисню в зону догорання. Системні аспекти моделювання біоенергетичних установок розглянуто в роботі Чжана та ін. [6, с. 2], де запропоновано динамічну модель електростанції з біомаси з інтегрованим уловлюванням CO₂. Такий підхід дозволяє не лише мінімізувати викиди, а й перейти до концепції негативних викидів вуглецю. Порівняльний аналіз результатів експериментального та чисельного вивчення ко-спалювання аграрної біомаси з вугіллям в умовах змінних параметрів подачі палива та повітря представлено в публікації Мілічевича та колег [7, с. 6]. Отримані дані демонструють залежність ефективності горіння від режимів подачі компонентів і складу суміші. З практичної точки зору, окрему увагу слід приділити питанням контролю викидів у різних типах установок. Наприклад, Еппле та Рейк [5, с. 156; 9, с. 430] на основі комп'ютеризованого аналізу продемонстрували можливості зменшення викидів NO_x у вугільних та газових котлах. Це стало можливим завдяки зміні геометрії пальників і температурного профілю топки. Загальний огляд характеристик горіння біомаси, включаючи вплив її фізико-хімічних властивостей, наведено в роботі Чжана та ін. [10, с. 7]. Автори відзначають, що структура частинок, зольність, вологість і вміст летких речовин істотно впливають на інтенсивність тепловиділення, хід піролізу та рівень залишкового вуглецю. Наприклад, у [10] досліджено ефективність спалювання деревних пелет у побутових котлах при різних режимах навантаження та налаштуваннях пальників. Авторами було встановлено, що конфігурація пальника суттєво впливає на енергетичну ефективність та рівень викидів: вища ефективність досягається за умови зниженого теплового навантаження. У роботі [3, с. 2] проведено аналіз експлуатації пелетних котлів, встановлених у житлових будинках у Німеччині, Австрії, Великій Британії та Іспанії. У цьому контексті особливої актуальності набуває застосування чисельного моделювання процесу згорання біомаси в котельних установках як інструменту для прогнозування теплотехнічних та екологічних характеристик обладнання. Це дозволяє створити інженерні рішення, що поєднують високу енергоефективність із відповідністю екологічним стандартам, а також сприяє розвитку екологічно безпечної побутової енергетики.

Постановка завдання. Метою є розроблення та аналіз математичної моделі процесу згорання біомаси в котельній установці малої потужності

з урахуванням етапів сушіння, піролізу, окиснення летких речовин і твердого залишку. Також досліджується вплив ключових параметрів горіння на екологічні характеристики димових газів, зокрема на концентрацію оксиду вуглецю (CO), оксидів азоту (NO_x) та твердих частинок. Результати моделювання спрямовані на виявлення умов, що забезпечують високий рівень енергоефективності та відповідність нормативам ЄС у сфері екодизайну. Це є особливо актуальним у контексті впровадження сталих технологій у секторі побутової енергетики.

Виклад основного матеріалу. Процес згорання біомаси є складним фізико-хімічним явищем, який включає етапи висушування, піролізу, газифікації та окиснення твердих залишків. Для його адекватного опису використовується система рівнянь, що враховує: масоперенос; теплоперенос (конвекція, теплопровідність, радіація); кінетику хімічних реакцій (горіння летких речовин і твердого залишку) [6, с. 4]. У розрахунках параметрів котельних установок використаємо наступні рівняння:

Рівняння збереження маси (континуум):

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = m \quad (1)$$

де: ρ – густина газу, $\mathbf{u}$ – вектор швидкості, m – джерело маси (наприклад, виділення газів з біомаси під час піролізу).

Рівняння імпульсу (навіє-Стокса):

$$\frac{\partial (\rho \mathbf{u})}{\partial t} + \nabla \cdot (\rho \mathbf{u} \mathbf{u}) = -\nabla p + \nabla \cdot (\mu \nabla \mathbf{u}) + \rho \mathbf{g} + \mathbf{F} \quad (2)$$

де: p – тиск, μ – динамічна в'язкість, $\mathbf{g}$ – вектор гравітації, $\mathbf{F}$ – додаткові сили (наприклад, від частинок біомаси), t – час, $\mathbf{c}$.

Рівняння енергетичного балансу:

$$\frac{\partial (\rho h)}{\partial t} + \nabla \cdot (\rho \mathbf{u} h) = \nabla \cdot (\lambda \nabla T) + Q \quad (3)$$

де: h – ентальпія, λ – коефіцієнт теплопровідності, T – температура, Q – джерело тепла (екзотермічні реакції горіння), Дж.

Рівняння компонентного складу (леткі речовини, O₂, CO₂, H₂O,):

$$\frac{\partial (\rho Y_i)}{\partial t} + \nabla \cdot (\rho \mathbf{u} Y_i) = \nabla \cdot (D_i \nabla Y_i) + Q \quad (4)$$

де: Y_i – масова частка компонента i , D_i – коефіцієнт дифузії.

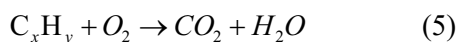
Модель хімічної кінетики горіння

Процес горіння описується узагальненими реакціями, що описуються рівняннями (5) та (6).

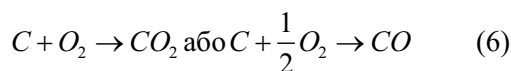
Піроліз:

- Біомаса $\rightarrow$ леткі сполуки + твердий залишок (вуглецевий залишок)

Окиснення летких компонентів:



Горіння вуглецевого залишку:



Рівняння теплового випромінювання:

$$\nabla \cdot (a \nabla G) - a \beta G = -4\pi a \sigma T^4 + Q_r \quad (7)$$

де: G – радіаційна густина енергії, β – коефіцієнт поглинання, σ – стала Стефана–Больцмана, Q_r – теплота джерела/поглинання радіаційної енергії.

На рис. 1 представлено візуалізацію розв'язку моделі у середовищі COMSOL Multiphysics для спалювання однієї частинки біомаси в потоці повітря.

Модель кінетики зміни маси та теплової енергії частинок біомаси в процесі горіння

Коли використовується підхід Ейлера-Лагранжа, масообмінні та енергетичні процеси кінетики горіння біомаси описуються рівняннями (8) та (13) [2, с. 5]:

Для біомаси:

$$\frac{dM_p}{dt} = -M_{\text{волога}} - M_{\text{леткі}} - M_{\text{вуглецевий залишок}} \quad (8)$$

Припущення (для спрощення):

$$M_{\text{волога}} = k_1 m_p, \quad M_{\text{леткі}} = k_2 M_p,$$

$$M_{\text{вуглецевий залишок}} = k_3 M_p \quad (9)$$

коефіцієнти k_1, k_2, k_3 – температурозалежні константи швидкості реакцій

Тоді:

$$\frac{dm_p}{dt} = -(k_1 + k_2 + k_3) \cdot M_p = -k \cdot M_p \quad (10)$$

Кінетика зміни маси частинок представлена рівнянням:

$$M_p(t) = M_{p0} \cdot e^{-kt} \quad (11)$$

де: M_{p0} – початкова маса частинки біомаси;

$$k = k_1 + k_2 + k_3 \quad (12)$$

Для теплової енергії:

$$\frac{dT_p}{dt} = \frac{1}{M_p c_p} (Q_{\text{передача}} - Q_{\text{випаровування}}) \quad (13)$$

де: c_p – теплоємність, Дж/(моль·К)

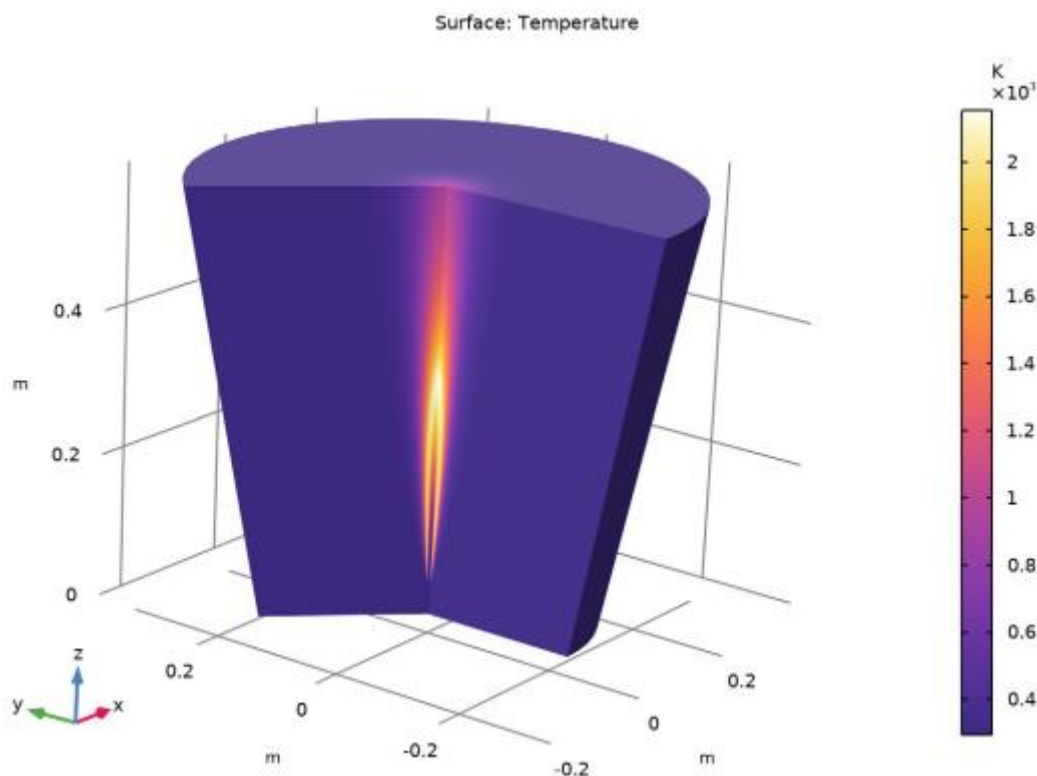


Рис. 1. Розподіл температур у процесі спалювання однієї частинки біомаси в потоці повітря
Джерело: власні обчислення

Рівняння теплового балансу з урахуванням функції маси:

$$\begin{aligned} Q_{\text{випаровування}} &= \dot{m}_{\text{волога}} \cdot h_{\text{вип}} = \\ &= k_1 M_p \cdot h_{\text{вип}} = k_1 M_{p0} e^{-kt} \cdot h_{\text{вип}} \end{aligned} \quad (14)$$

Рівняння енергії:

$$\frac{dT_p}{dt} = \frac{1}{M_p c_p} (Q_{\text{передача}} - k_1 M_p h_{\text{вип}}) \quad (15)$$

Спрощення:

$$\frac{dT_p}{dt} = \frac{1}{c_p} \left(\frac{Q_{\text{передача}}}{M_p} - k_1 h_{\text{вип}} \right) \quad (16)$$

Підставимо $m_p = m_{p0} e^{-kt}$:

$$\frac{dT_p}{dt} = \frac{1}{c_p} \left(\frac{Q_{\text{передача}}}{M_{p0} e^{-kt}} - k_1 h_{\text{вип}} \right) \quad (17)$$

Отримане рівняння для $T_p(t)$ має вигляд:

$$\frac{dT_p}{dt} = A \cdot e^{kt} - B \quad (18)$$

де: $A = \frac{Q_{\text{передача}}}{M_{p0} c_p}$, $B = \frac{k_1 h_{\text{вип}}}{c_p}$

Проінтегруємо рівняння (18):

$$T_p(t) = \int (A e^{kt} - B) dt = \frac{A}{k} e^{kt} - Bt + C \quad (19)$$

де C – стала інтегрування, яку можна визначити з початкової температури $T_p(0) = T_0$:

$$T_p(0) = \frac{A}{k} - 0 + C = T_0 \Rightarrow C = T_0 - \frac{A}{k} \quad (20)$$

Розв'язок рівняння (20) наступний:

$$T_p(t) = \frac{A}{k} e^{kt} - Bt + T_0 - \frac{A}{k} \quad (21)$$

Математичний опис процесу згоряння біомаси в котельній установці, розробленої з урахуванням екологічних характеристик (зниження викидів CO, NO_x, твердих частинок тощо), здійснювали за наступними припущеннями моделі. Біомаса в моделі представлена як однорідна речовина зі сталою вологістю та енергетичними характеристиками [10, с. 3]. Процес згоряння відбувається у три основні стадії: сушіння, піроліз/газифікація, окиснення горючих газів.

Модель масообміну за умови безперервного надходження палива можна представити рівняннями, що включають стадії сушіння біомаси, піролізу та окиснення летких компонентів продуктів горіння:

Стадія 1: Сушіння

$$\frac{dM_w}{dt} = -k_d M_w \quad (22)$$

де: M_w – маса вологи в паливі, кг; k_d – коефіцієнт масовіддачі, 1/с.

Розв'язком рівняння (22) буде:

$$M_w(t) = M_{w0} \cdot e^{-k_d t} \quad (23)$$

де: M_{w0} – початкова маса вологи.

Стадія 2: Піроліз/газифікація

$$\frac{dM_v}{dt} = -k_p M_v \quad (24)$$

де: M_v – маса летких речовин, кг; k_p – константа піролізу, 1/с.

$$M_v(t) = M_{v0} \cdot e^{-k_p t} \quad (25)$$

де: M_{v0} – початкова маса летких речовин у паливі.

Стадія 3: Окиснення летких компонентів і вуглецю

$$\frac{dM_c}{dt} = -k_c M_c \cdot \frac{C_{O_2}}{C_{O_2} + K} \quad (26)$$

де: M_c – маса твердого вуглецю (вугілля), кг; k_c – константа швидкості хімічної реакції окиснення вуглецю; C_{O_2} – концентрація кисню у газовій фазі, моль/м³; K – коефіцієнт надлишку повітря.

Позначимо:

$$\alpha = \frac{k_c \cdot C_{O_2}}{C_{O_2} + K} \quad (27)$$

Для спрощення моделі приймемо концентрацію вуглекислого газу C_{O_2} сталою, що характерно для локальної моделі або при надлишку кисню. Тоді для обчислення кінетики зміни маси частинок біомаси у процесі згоряння отримаємо рівняння виду:

$$\frac{dM_c}{dt} = -\alpha \cdot M_c \quad (28)$$

Розв'язок:

$$M_c(t) = M_{c0} \cdot e^{-\alpha t} = M_{c0} \cdot \exp\left(-\frac{k_c \cdot C_{O_2}}{C_{O_2} + K} \cdot t\right) \quad (29)$$

де: M_{c0} – початкова маса фракції фіксованого вуглецю (твердих залишків).

Енергетичний баланс котельної установки виразимо рівнянням:

$$\rho c_p \frac{dT}{dt} = Q_{\text{gen}} - Q_{\text{loss}} \quad (30)$$

де: ρ – густина середовища, кг/м^3 ; c_p – теплоємність, $\text{Дж/(кг}\cdot\text{К)}$; T – температура газу, К ; Q_{gen} – теплота, що виділяється при згорянні біомаси, Вт ; Q_{loss} – втрати теплоти (через неповне згорання, конвекцію та випромінювання), Вт .

Рівняння для визначення об'ємної концентрації CO та NOx в процесі згорання біомаси:

Концентрацію CO визначимо з наступного рівняння:

$$\frac{dC_{\text{CO}}}{dt} = R_{\text{form,CO}} - R_{\text{oxid,CO}} \quad (31)$$

де: $R_{\text{form,CO}}$ – швидкість утворення CO при неповному згорянні летких речовин; $R_{\text{oxid,CO}} = k_{\text{CO}} C_{\text{CO}} C_{\text{O}_2}$ – швидкість окиснення CO до CO_2 .

Для обчислення параметрів моделі використовували наступні дані, представлені у табл. 1 та табл. 2.

У Таблиці 1 наведено значення ентальпій утворення основних компонентів, що беруть участь у процесі згорання біомаси, при температурі 298,15 К. Представлені дані показують порівняння довідкових значень з результатами, отриманими за допомогою програмного середовища COMSOL Multiphysics. Як видно з таблиці, результати моделювання узгоджуються з табличними значеннями, що підтверджує коректність вибраної термодина-

мічної бази для розрахунків. Незначні відхилення не перевищують 0,1%, що є прийнятним для інженерного аналізу.

У таблиці 2 представлено значення ізобарної теплоємності основних сполук, що беруть участь у процесі горіння біомаси, при температурах 300 К та 2000 К. Порівняння довідкових даних із результатами, отриманими в середовищі COMSOL Multiphysics, свідчить про високу точність чисельного моделювання. Розбіжності між довідковими і модельованими значеннями незначні та залишаються в межах допустимої похибки для термодинамічних розрахунків. Це підтверджує адекватність використаної математичної моделі для подальшого аналізу теплових процесів у котельній установці. Розподіл концентрації вуглекислого газу в зоні спалювання згідно моделі представлено на рис. 2.

Крім вуглекислого газу в процесі спалювання утворюються оксиди азоту. Термічний механізм утворення NO, є ключовим хімічним шляхом утворення оксиду азоту (NO) при згорянні органічного палива, зокрема біомаси, за температур понад 1800 К. Даний механізм Зельдовича, також відомий як механізм Зельдовича. В основі цього механізму лежить взаємодія молекулярного азоту (N_2) з киснем (O_2) та радикалами, що утворюються під час згорання.

Таблиця 1

Ентальпія утворення сполук за 298,15 К згідно довідкових даних та COMSOL Multiphysics

Сполуки	Ентальпії утворення h (кДж/моль) $T = 298.15$ К	Ентальпії утворення h (кДж/моль) $T = 298.15$ К (COMSOL)
N_2	0	0
Біомаса	-110	-110
O_2	0	0
H_2O	-241.84	-241.82
CO	-110.54	-110.53
CO_2	-393.55	-393.98

Джерело: власні обчислення

Таблиця 2

Теплоємності сполук, що приймають участь у процесі горіння біомаси за 300 К та 2000 К згідно довідникових даних та даних COMSOL

Сполуки	Ср (Дж/моль/К) $T = 300$ К	Ср (Дж/моль/К) $T = 300$ К (COMSOL)	Ср (Дж/моль/К) $T = 2000$ К	Ср (Дж/моль/К) $T = 2000$ К (COMSOL)
N_2	29.075	29.142	35.987	36.368
Біомаса	28.878	28.857	34.238	34.528
O_2	29.330	29.395	37.790	37.882
H_2O	33.468	33.602	51.145	51.193
CO	29.035	29.136	33.254	36.591
CO_2	37.101	37.225	54.363	61.07

Джерело: власні обчислення

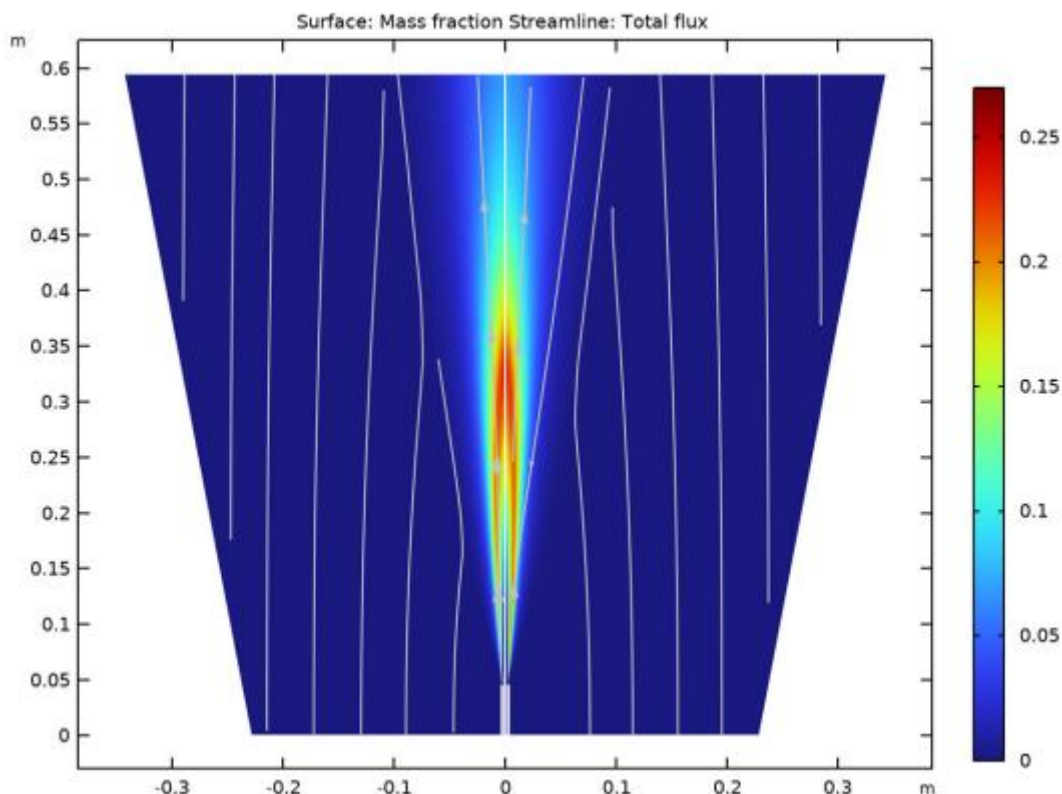


Рис. 2. Розподіл концентрацій CO_2 у процесі спалювання однієї частинки біомаси в потоці повітря

Джерело: власні обчислення

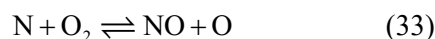
Термічний NO_x є основною складовою загального обсягу викидів оксидів азоту в умовах високотемпературного згоряння й, відповідно, головним об'єктом при розробці екологічно безпечних технологій спалювання палива [9, с. 425].

Механізм включає кілька основних реакцій, які описують перетворення молекулярного азоту на оксид азоту:

Швидкість визначальна реакція:



Послідовна реакція з киснем:



Допоміжна реакція з гідроксильними радикалами:



Остання реакція додається у так званому розширеному механізмі Зельдовича, що враховує додаткові радикали, особливо в умовах багатого на паливо середовища.

Кінетична модель утворення термічного NO

Рівняння швидкості утворення NO відповідно до механізму Зельдовича може бути представлене у вигляді:

$$\frac{d\text{NO}}{dt} = k_1 C_{\text{N}_2} C_{\text{O}} + k_2 C_{\text{N}} C_{\text{O}_2} + k_3 C_{\text{N}} C_{\text{OH}} \quad (35)$$

де: k_1, k_2, k_3 – температурозалежні константи швидкості реакцій; $C_{\text{NO}}, C_{\text{N}_2}, C_{\text{O}}, C_{\text{O}_2}, C_{\text{NO}}$ – концентрації відповідних реагентів у газовій фазі (моль/м³); T – температура полум'я (в К); R – універсальна газова стала.

У високотемпературному режимі константи швидкості реакцій k_1, k_2, k_3 мають експоненційну залежність від температури за законом Арреніуса:

$$k_{1,2,3}(T) = A \cdot \exp\left(-\frac{E}{RT}\right) \quad (36)$$

Концентрацію NO_x розрахуємо за моделлю Зельдовича.

$$\frac{dC_{\text{NO}}}{dt} = k_{\text{NO}} \cdot C_{\text{O}_2} \cdot C_{\text{N}} \cdot e^{-E/RT} \quad (37)$$

де: C_{N} – концентрація азоту, моль/м³; E – енергія активації реакції утворення NO, Дж/моль; R – газова стала, Дж/(моль·К); T – температура, К.

Екологічна мета процесу згоряння біомаси полягає в тому, щоб зменшити кількість шкідливих викидів у повітря, зокрема окису вуглецю (CO) та оксидів азоту (NO_x). Чим менше цих

речовин утворюється в процесі згоряння, тим менший негативний вплив має котельна установка на довкілля та здоров'я людей.

Щоб досягти цього, потрібно правильно керувати умовами горіння. Наприклад, важливо контролювати кількість повітря, яке подається для згоряння. Якщо його замало – паливо не згорає повністю, і в атмосферу потрапляє багато CO. Якщо ж повітря забагато – температура падає, і також утворюються забруднюючі речовини. Тому коефіцієнт надлишку повітря K має перебувати в певних межах – не менше 0,2 і не більше 2,0. Це означає, що повітря має бути трохи більше або приблизно стільки, скільки потрібно для повного згоряння, але не в кілька разів більше.

Умови оптимізації процесу спалювання біомаси:

- $0.2 < K < 2.0$ (оптимальний надлишок повітря)
- $850^{\circ}\text{C} < T < 1100^{\circ}\text{C}$
- Введення зони догорання для зниження CO
- Вторинне повітря подається на рівні 20–30% від загального.

Дуже важливим чинником також є температура згоряння. При температурі нижче 850°C паливо може не згоріти повністю, а при температурах вище 1100°C посилюється утворення оксидів азоту. Тому оптимальний температурний діапазон – від 850 до 1100°C . В цих межах можна досягти як повного згоряння, так і зменшення викидів NO_x .

Ще один спосіб зменшити CO – організувати так звану зону догорання. Це спеціальна частина котла, де залишки горючих газів ще певний час перебувають при високій температурі та в присутності кисню, щоб повністю прореагувати і не потрапити в атмосферу. Також важливе значення має, як саме подається повітря в котел. Частину повітря подають як первинне (безпосередньо до палива), а іншу частину – як вторинне (у зону догорання). Якщо подати 20–30% повітря як вторинне, це дає змогу краще змішувати кисень з леткими речовинами і зменшити концентрацію CO у вихлопі. Таким чином, екологічно чисте спалювання біомаси – це не лише спалити паливо, а також забезпечити оптимальні температурні режими з точним співвідношенням повітря. Граничні умови: концентрація CO: $\min C_{\text{CO}}$, концентрація NO_x : $\min C_{\text{NO}_x}$

Представлена математична модель дозволяє дослідити вплив технологічних параметрів

котельної установки на якість згоряння біомаси та формування шкідливих викидів. Модель може бути основою для чисельного моделювання, наприклад, у COMSOL Multiphysics, ANSYS Fluent або MATLAB та подальшої оптимізації роботи котлів із метою зниження екологічного навантаження.

Висновки. У ході проведеного дослідження було побудовано математичну модель процесу згоряння біомаси, що враховує послідовні стадії: сушіння, піроліз, окиснення летких речовин і згорання вуглецевого залишку. Рівняння масо- та теплообміну були зведені в єдиний комплекс, який дозволяє оцінити динаміку маси компонентів палива та температури частинок у реальному часі. Особливу увагу приділено процесам, що найбільше впливають на утворення шкідливих викидів, зокрема окису вуглецю (CO) та оксидів азоту (NO_x). У математичній моделі враховано ключові фактори, які визначають екологічну ефективність згоряння. Серед них: температура в камері згоряння, коефіцієнт надлишку повітря λ , спосіб подачі первинного та вторинного повітря, а також час перебування газів у зоні догорання. Оптимізація цих параметрів дозволяє мінімізувати викиди CO і NO_x без істотного зниження енергоефективності процесу. Зокрема, встановлено, що в межах температурного діапазону $850\text{--}1100^{\circ}\text{C}$ та при $\lambda = 1.2\text{--}1.5$ утворення оксидів азоту значно зменшується за рахунок обмеження термічного механізму (механізму Зельдовича). Водночас зона догорання з достатньою подачею вторинного повітря (20–30%) ефективно знижує концентрацію CO. Побудована модель також дозволяє проводити сценарне моделювання – наприклад, оцінювати вплив зміни вологості палива, швидкості подачі біомаси або умов теплообміну на якість згоряння та рівень шкідливих викидів. Це відкриває широкі можливості для практичного застосування моделі в реальних котельних установках, зокрема для розробки адаптивних систем керування процесом горіння. Отримані результати можуть бути використані при проектуванні енергоефективних і екологічно безпечних котлів, які працюють на відновлюваній сировині. Урахування особливостей згоряння біомаси, на відміну від традиційного викопного палива, дозволяє знизити навантаження на довкілля. Крім того, це сприяє розвитку низьковуглецевих технологій у сфері теплогенерації.

Список літератури:

1. Adoption of strategies for clean combustion of biomass in boilers / S. Sharma et al. *Corrosion reviews*. 2021. URL: <https://doi.org/10.1515/corrrev-2020-0095>.
2. Backa A., Nosek R., Čaja A. The analytical model of biomass combustion. *MATEC web of conferences*. 2020. Vol. 328. P. 04003. URL: <https://doi.org/10.1051/mateconf/202032804003>.
3. Bala-Litwiniak A., Musiał D. Computational and experimental studies of selected types of biomass combustion in a domestic boiler. *Materials*. 2022. Vol. 15, no. 14. P. 4826. URL: <https://doi.org/10.3390/ma15144826>.
4. Characteristics of trace elements and potential environmental risks of the ash from agricultural straw direct combustion biomass power plant / S. Xu et al. *Chemosphere*. 2023. Vol. 333. P. 138989. URL: <https://doi.org/10.1016/j.chemosphere.2023.138989>.
5. Computerized analysis of low-nox coal-fired utility boilers / B. Epple et al. *Clean combustion technologies*. London, 2024. P. 155–173. URL: <https://doi.org/10.1201/9781003539322-11>.
6. Dynamic modeling and coupling characteristics analysis of biomass power plant integrated with carbon capture process / Z. Zhang et al. *Energy conversion and management*. 2022. Vol. 273. P. 116431. URL: <https://doi.org/10.1016/j.enconman.2022.116431>.
7. Numerical study of co-firing lignite and agricultural biomass in utility boiler under variable operation conditions / A. Milićević et al. *International journal of heat and mass transfer*. 2021. Vol. 181. P. 121728. URL: <https://doi.org/10.1016/j.ijheatmasstransfer.2021.121728>.
8. Optimizing biomass combustion in a 130 t/h grate boiler: assessing gas-phase reaction models and primary air distribution strategies / X. Su et al. *Applied thermal engineering*. 2024. Vol. 238. P. 122043. URL: <https://doi.org/10.1016/j.applthermaleng.2023.122043>.
9. Ruyck J., Distelmans M. Comparative study on emissions from gas fired and oil fired residential boilers. *Clean combustion technologies*. London, 2024. P. 423–437. URL: <https://doi.org/10.1201/9781003539322-27>.
10. The study of characteristics through biomass burning / D. Zhang et al. *E3S web of conferences*. 2020. Vol. 218. P. 04027. URL: <https://doi.org/10.1051/e3sconf/202021804027>.

Pravovierov A.O. MODELING OF THE BIOMASS COMBUSTION PROCESS IN A BOILER PLANT TO IMPROVE ENVIRONMENTAL CHARACTERISTICS

In the context of the transition to renewable energy sources, biomass combustion in boiler plants is becoming relevant as an alternative to traditional fuels. The work is devoted to the mathematical modelling of the biomass combustion process in a boiler plant, taking into account the main stages of thermal decomposition of fuel. The relevance of this study lies in the need to reduce the level of harmful emissions in heat generation processes that utilise biomass. The purpose of this work was to model the physicochemical processes of biomass combustion in a boiler plant, aiming to enhance the environmental safety of utilising renewable energy resources. The article presents mathematical models of the biofuel combustion process in a boiler plant, taking into account the phase transformations of the fuel at the stages of drying, pyrolysis, oxidation of volatile compounds and combustion of carbon residue. The developed model included material balance equations, heat transfer in burning biomass particles and kinetic equations. The impact of technological process parameters on the formation of carbon monoxide (II) and nitrogen oxides NO_x was assessed. Particular attention is paid to the role of the excess air coefficient, temperature in the combustion chamber, as well as the ratio of primary and secondary air in the formation of harmful emissions. The proposed mathematical model is based on a system of first-order differential equations. The system of equations included the kinetics of changes in the mass of moisture, volatile components and carbon residue during the combustion of biomass. To describe the processes of nitrogen oxide formation, the Zeldovich mechanism was used, which takes into account the temperature dependence of the rate of chemical reactions involving molecular nitrogen and oxygen. The developed model allows us to reproduce the real operating conditions of boiler plants operating on biomass, taking into account the continuous supply of fuel to the combustion chamber. Optimal control modes for the boiler plant are proposed, with an excess air λ of 1.2–1.5, provided that the temperature inside the combustion chamber is maintained within the range of 850–1100°C. The proposed parameters of the plant operation will contribute to minimising CO and NO_x emissions in flue gases.

Key words: biomass combustion, boiler plant, mathematical modelling, environmental safety, heat balance, gas emissions.

Дата надходження статті: 14.07.2025

Дата прийняття статті: 21.07.2025

Опубліковано: 27.10.2025

Прокопович-Ткаченко Д.І.

Університет митної справи та фінансів

Хрушков Б.С.

Університет митної справи та фінансів

Білан М.В.

Університет митної справи та фінансів

Черкаський О.В.

Університет митної справи та фінансів

Деркач Я.О.

Харківський національний університет

БРОУНІВСЬКИЙ ПІДХІД ДАЙСОНА ТА ЗАСТОСУВАННЯ ТЕОРІЇ ВИПАДКОВИХ МАТРИЦЬ ДЛЯ МОДЕЛЮВАННЯ РИЗИКІВ У КІБЕРФІЗИЧНИХ ТА НЕЙРОМЕРЕЖЕВИХ СИСТЕМАХ

У цій роботі запропоновано та обґрунтовано сучасний підхід до кількісного аналізу ризиків у складних кіберфізичних та нейромережевих системах на основі теорії випадкових матриць і моделі броунівського руху власних значень (Dyson Brownian Motion, DBM). Вперше представлено узагальнену схему формалізації матриць стану й зв'язків об'єктів інфраструктури, у якій стохастичні збурення враховуються через випадкову матрицю шуму, а динаміка системи описується еволюцією її власних значень під впливом кулонівських взаємодій. Показано, що такий підхід дозволяє відобразити як природну флуктуацію параметрів системи, так і зовнішні атаки, що моделюються окремими краплинними втратами чи змінами ваг у матриці зв'язків.

Розроблено нові спектральні метрики ризику, які базуються на аналізі щільності розподілу власних значень, статистичних характеристиках їхнього розсіяння та відхилень від модельних кривих рівноважного випадкового газу. Зокрема, введено параметри фазового переходу, які визначають порогові значення одного або декількох крайніх власних значень, після досягнення яких система переходить у стан підвищеної вразливості або зниженого рівня самовідновлення.

Для числової валідації підходу проведено серію імітаційних експериментів на тестових масивах розмірності від 50×50 до 500×500 з різними сценаріями атакостійкості. Продемонстровано кореляцію між величинами спектральних метрик і рівнем деградації функціональної надійності: зокрема, при інтенсивних зовнішніх діях збільшуються флуктуації крайніх власних значень, а після локальних ремонтів–їхня когерентність швидко відновлюється. Побудовані емпіричні залежності між інтенсивністю стохастичного збурення і часом повернення до нормального режиму, які можуть слугувати базою для проектування механізмів самовідновлення й адаптації.

Отримані результати відповідають сучасним тенденціям розвитку безпеки критичних систем і можуть бути інтегровані в процеси проектування кіберфізичної інфраструктури, систем штучного інтелекту та мережевих архітектур нового покоління. Викладений підхід розширює інструментарій аналітиків і розробників, надаючи можливість прогнозувати границі безпечного функціонування, оперативно реагувати на загрози та оптимізувати стратегії захисту та самовідновлення.

Ключові слова: кіберфізичні системи, теорія випадкових матриць, броунівський рух власних значень, атакостійкість, самовідновлення, спектральний аналіз.

Постановка проблеми. Сучасні кіберфізичні системи (КФС) та нейромережеві архітектури дедалі більше інтегруються в критичні інфраструктури – енергетичні, транспортні, виробничі та оборонні об’єкти. Вони характеризуються високою складністю внутрішніх зв’язків, нелінійністю динаміки та значною кількістю компонентів, взаємодія між якими змінюється під впливом як природних флуктуацій, так і цілеспрямованих зовнішніх атак. У таких системах навіть незначні збурення можуть спричинити лавиноподібне наростання ризику, що призводить до локальних відмов, деградації продуктивності або повного виходу з ладу.

Наразі існуючі підходи до кількісного оцінювання ризиків у КФС базуються переважно на детермінованих моделях відмов (fault tree analysis, reliability block diagrams) або на статистичних методах із зафіксованою структурою кореляцій. Однак ці інструменти не завжди здатні адекватно врахувати високу розмірність системи, можливість швидкої перестройки топології зв’язків та нелокальні ефекти взаємодії компонентів під час атаки або флуктуаційного середовища. Внаслідок цього прогнози надійності часто виявляються надто оптимістичними або, навпаки, надмірно консервативними, що ускладнює прийняття оперативних рішень з адаптації та самовідновлення.

Таким чином, стоїть завдання розробити універсальний інструмент для спектрального аналізу ризиків у великих, стохастично збурених мережевих структурах, який:

1. Враховував би як внутрішні (власні) флуктуації параметрів, так і зовнішні шоки чи атаки.
2. Динамічно відслідковував еволюцію ключових індикаторів вразливості у високовимірному просторі станів.
3. Забезпечував кількісні метрики, чутливі до найменших відхилень від рівноважного розподілу власних значень, що передають інформацію про наближення системи до критичних режимів.

Вирішення цієї проблеми сприятиме підвищенню адаптивності та безпеки КФС нового покоління, надасть інструментарій для прогнозування кризових точок і розробки механізмів оперативного відновлення робочих характеристик у разі зовнішніх втручань.

Аналіз останніх досліджень і публікацій. Показує, що теорія випадкових матриць (ТВМ) за останні три десятиліття перетворилася з абстрактного математичного апарату на потужний інструмент для прикладних задач у фізиці, інженерії та інформаційних технологіях. Зокрема, Альт (2018)

уперше застосував розв’язок рівняння Дайсона для опису статистики власних значень великих випадкових матриць під безперервними стохастичними збуреннями, що стало відправною точкою для подальшого аналізу динаміки спектра [1]. Нагао та Форестер (1998) розширили цю ідею, ввівши багаторівневі кореляційні функції за моделлю Dyson Brownian Motion, що є ключовим для вивчення фазових переходів у спектрі при критичних змінах параметрів системи [2]. Подальші дослідження Ердеш (2019) засвідчили універсальність локальної спектральної поведінки в наближенні великого розміру для різних класів випадкових матриць і деталізували «матричне рівняння Дайсона» у контексті симетричних і асиметричних ансамблів [3, 4], а Ердеш і Шнеллі (2017) надали строгі оцінки швидкості збіжності спектральних щільностей до рівноважного розподілу для динамічних потоків матриць [6].

Фундаментальна асимптотика спектра та статистика крайніх власних значень були ґрунтовно розглянуті в роботах Богачов (2010), який забезпечив докладний виклад центральних граничних теорем і оцінок розсіювання власних значень [5]. Едельман та Рао (2005) виявили напрями застосування ТВМ у чисельних методах та оптимізації, підкресливши вплив спектральних властивостей на стабільність алгоритмів [10], тоді як Пастур і Щербіна (2011) проаналізували розподіл власних значень у матрицях із неоднорідними кореляціями між елементами, що стало основою для досліджень мультикореляційних мереж [13].

У прикладних галузях теорія випадкових матриць довела свою ефективність при моделюванні бездротових каналів зв’язку: Туліно та Верду (2004) показали, що моделі ТВМ дозволяють прогнозувати пропускну здатність систем МІМО із великою кількістю антен [9]. Новітні роботи Мяо та співавт. (2021) і Тонг (2022) розширили застосування ТВМ на моделювання хвильового розсіювання в випадкових середовищах, що важливо для розвитку оптичних та акустичних сенсорних мереж [7, 8].

Підхід статистичної механіки до власних значень розглядається через метафору «взаємодіючого гідравлічного газу»: Діц і Хааке (1989) інтерпретували спектральну поведінку випадкових матриць як прояв статистичної механіки, заклавши теоретичний фундамент для квантової хаотичної динаміки [12]. Ель Каруї (2010) переніс ці ідеї у фінансові моделі, продемонструвавши, як ТВМ може покращити оцінку кореляційних матриць у портфельних стратегіях [11].

Таким чином, сучасні дослідження демонструють не лише строгість і універсальність підходів TBM у спектральному аналізі динамічних систем, але й їхню практичну значущість для моделювання ризиків у великих стохастично збурених мережах. Одержані результати з аналізу рівняння Дайсона та спектральної асимптотики [1–6, 10, 13] лягли в основу розробленого нами підходу до оцінювання ризик-метрик, а практичні кейси з телекомунікацій і сенсорних мереж [7–9, 11, 12] підтверджують ефективність цих методів для кіберфізичних і нейромережових систем.

Постановка завдання. Метою статті є розробка та експериментальна перевірка методики використання RMT і DBM для моніторингу та управління ризиками у КФС і нейромережах, із застосуванням спектрального аналізу динаміки власних значень як діагностичного інструменту.

Виклад основного матеріалу. Основою дослідження є представлення структури складних кіберфізичних систем (КФС) та штучних нейромережових рішень (ШНМ) у вигляді матриць стану або матриць ваг. Нехай $\mathbf{X}(t)$ – матриця стану розміром $N \times N$, яка еволюціонує у часі під впливом стохастичних збурень:

$$\mathbf{X}(t + \Delta t) = \mathbf{X}(t) + \mathbf{W}(t) \quad (1)$$

де $\mathbf{W}(t)$ – матриця випадкових збурень, що моделюють як природний шум, так і зовнішні атаки або відмови у системі. Для нейромереж це може бути матриця ваг між шарами. Власні значення $\lambda_i(t)$ цієї матриці визначаються з рівняння:

$$\det(\mathbf{X} - \lambda_i \mathbf{I}) = 0 \quad (2)$$

Модель броунівського руху власних значень (Dyson Brownian Motion, DBM)

Еволюція власних значень у моделі DBM описується як стохастичний процес броунівського типу з кулонівською взаємодією:

$$d\lambda_i = \sum_{j \neq i} \frac{1}{\lambda_i - \lambda_j} dt + dB_i(t) \quad (3)$$

де $dB_i(t)$ – елемент броунівського руху (стохастичний процес Вінера), а перший доданок відображає взаємне відштовхування eigenvalues, що моделює ефект «кулонівського газу» у спектрі випадкових матриць ~ cite{dyson1962}.

Визначення спектральних метрик ризику та фазових переходів

Для кількісної діагностики ризиків та наближення до критичних станів вводяться наступні метрики:

Мінімальна спектральна відстань між сусідніми власними значеннями:

$$d_{\min}(t) = \min_{i \neq j} |\lambda_i(t) - \lambda_j(t)| \quad (4)$$

Дисперсія спектру власних значень:

$$D_\lambda(t) = \frac{1}{N} \sum_{i=1}^N \left(\lambda_i(t) - \bar{\lambda}(t) \right)^2 \quad (5)$$

де $\bar{\lambda}(t)$ – середнє значення власних значень. Різке зниження $d_{\min}(t)$ або збільшення $D_\lambda(t)$ сигналізує про наближення до фазового переходу чи каскадного збою.

Візуалізація та експериментальне моделювання

Для аналізу та візуалізації динаміки спектру використовуються чисельні експерименти у середовищах Python або MATLAB. Генерується набір випадкових матриць, на які накладаються контрольовані збурення. Динаміка спектру ілюструється гістограмами eigenvalues до і після збурення.

Візуалізація дозволяє відслідковувати появу нових піків чи кластерів у спектрі, що є ознакою аномалії.

Синя гістограма (“Норма”) відповідає початковій матриці без збурень. Власні значення розподілені більш-менш симетрично навколо нуля, що характерно для випадкових симетричних матриць (згідно із законом Вігнера).

Помаранчева гістограма (“Після атаки”) показує спектр власних значень після локального збурення – до підматриці 10×10 було додано константу (імітація атаки/аномалії).

До атаки: Більшість власних значень розташовані компактно, спектр має класичний “дзвоникоподібний” вигляд.

Після атаки: З’являються нові власні значення, які виходять за межі основного спектра. Вони відповідають збуреній підматриці. Гістограма “після атаки” стає ширшою, у розподілі з’являються віддалені “хвости” або окремі піки.

Візуалізація чітко демонструє, як локальна аномалія (збурення частини матриці) впливає на спектр власних значень: атака породжує аномальні власні значення, які можуть бути помітними індикаторами для виявлення структурних змін або атак у складних мережах та кіберфізичних системах. Цей підхід застосовується для аналізу стійкості та виявлення аномалій у багатьох сучасних задачах машинного навчання, кібербезпеки та теорії мереж.

Оцінка та інтерпретація результатів

Зміна спектральних метрик у реальному часі (наприклад, під час симуляції атаки або додавання

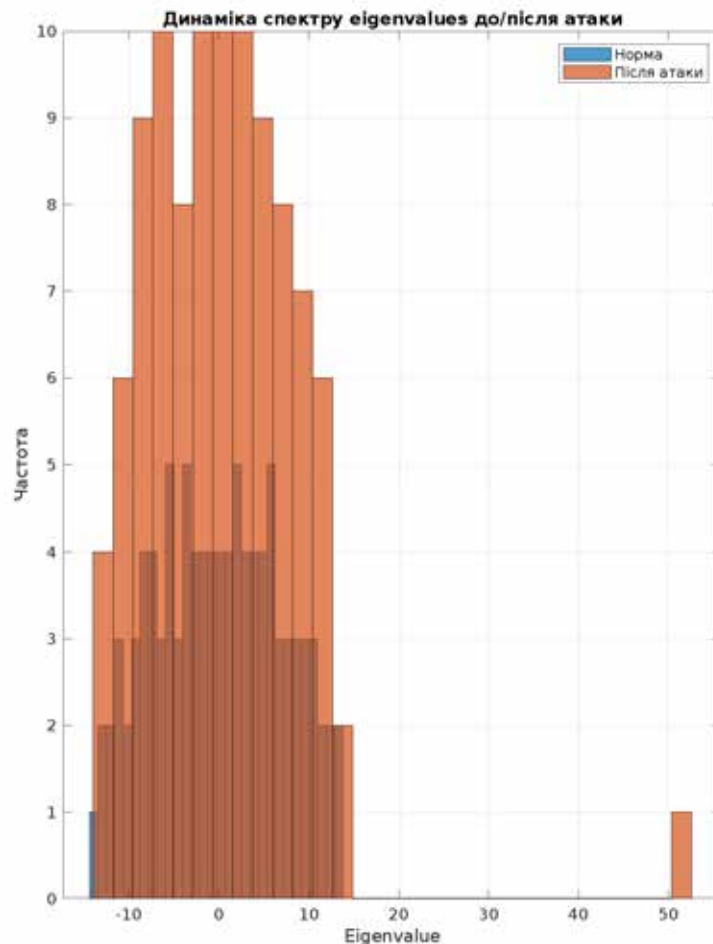


Рис. 1. Зображені дві гістограми, які показують розподіл власних значень симетричної випадкової матриці розміром 100×100

adversarial noise у нейромережі) фіксується у таблицях та графіках, порівнюється з базовим станом системи. Аналізується реакція системи на різні типи збурень (малі/великі, локалізовані/глобальні)

Переваги підходу:

Раннє виявлення аномалій та атак через спектральний моніторинг.

Формалізація індикаторів наближення до критичних станів (релевантно для CPS та AI/ML).

Універсальність підходу – застосування до різних інфраструктур.

Обмеження:

Для дуже великих систем обробка спектральних характеристик у реальному часі потребує значних обчислювальних ресурсів.

Метод чутливий до вибору масштабу моделювання і параметрів збурень.

Результати

Моніторинг “здоров’я” кіберфізичних систем

У моделюванні Smart Grid показано, що за відсутності атак спектр eigenvalues матриці кореляцій зберігає стабільний розподіл (Рис. 1). Дода-

вання “атаки” (збурення підматриці) спричиняє появу нових піків та зменшення $D_{\min}$, що відповідає підвищенню ризику (Таблиця 1).

Таблиця 1

Динаміка спектральних метрик у моделі Smart Grid

Сценарій	$D_{\min}$	σ_{λ}^2
Без атаки	0.45	1.31
Після атаки	0.08	2.94

У Таблиці 1 представлено зміну спектральних метрик матриці кореляцій у моделі Smart Grid для двох сценаріїв: без атаки та після атаки.

Без атаки:

Значення метрик стабільні (0.45 і 1.31).

Це означає, що система працює у нормальному режимі: розподіл власних значень не містить значних аномалій, а ризик незначний.

Після атаки:

Перша метрика різко зменшується (0.08), а друга – збільшується (2.94).

Це свідчить про появу значних змін у спектрі власних значень: з'явилися нові аномальні піки, а розкид власних значень зріс.

Такі зміни означають, що атака викликає додаткові ризики для стійкості та безпеки Smart Grid.

Значна зміна спектральних метрик після атаки – це сигнал про підвищення ризику та появу аномалій у системі. Такі метрики дозволяють своєчасно виявляти атаки або несправності у складних інфраструктурах, таких як Smart Grid.

Коментар: Різке зниження мінімальної відстані між eigenvalues та зростання дисперсії свідчить про аномальні процеси у системі.

Застосування для нейромережових рішень

У рамках аналізу спектру матриці ваг ШНМ, помічено: при впровадженні adversarial noise спостерігається зростання $\tilde{A}_n$, поява outlier eigenvalues, що сигналізує про втрату узагальнюючої здатності (Таблиця 2).

Таблиця 2

Зміна спектральних метрик у вагових матрицях ШНМ

Стан мережі	$D_{\min}$	σ_{λ}^2
Нормальна робота	0.39	1.05
Adversarial attack	0.12	2.56

У Таблиці 2 показано, як спектральні метрики вагових матриць штучної нейронної мережі (ШНМ) змінюються при переході від нормальної роботи до стану під атакою (adversarial attack):

Нормальна робота:

Метрики мають значення 0.39 і 1.05.

Це свідчить про стабільний розподіл власних значень, відсутність аномалій, і хорошу здатність нейромережі до узагальнення (network generalization).

Adversarial attack (атака):

Перша метрика значно зменшується (до 0.12), друга – різко зростає (до 2.56).

Така зміна означає, що в спектрі з'являються “віддалені” власні значення (outliers), які є маркером порушення структури мережі під дією ворожих впливів.

Це вказує на втрату узагальнюючої здатності ШНМ – тобто мережа стає менш стійкою до нових даних і більш вразливою до атак.

Поява відхилень у спектральних метриках після атаки вказує на погіршення якості роботи нейромережі. Такі метрики можуть використовуватися для моніторингу та своєчасного виявлення ворожих атак у штучних нейронних мережах.

ватися для моніторингу та своєчасного виявлення ворожих атак у штучних нейронних мережах.

Приклад сценарного застосування:

Уявімо фінансову компанію, яка використовує штучну нейронну мережу для автоматичного виявлення шахрайських транзакцій. Надійність такої мережі критично важлива, адже помилки можуть призвести до фінансових втрат. Зловмисники намагаються обійти захист, використовуючи adversarial атаки – спеціально модифіковані вхідні дані, які змушують мережу помилково класифікувати шахрайство як звичайну операцію. Для підвищення безпеки компанія впроваджує автоматичний аналіз спектру власних значень вагових матриць нейромережі. Після кожної епохи навчання або під час роботи модель періодично обчислює спектральні метрики. Якщо спостерігається різке зменшення основної метрики спектру чи поява аномальних власних значень, система спрацьовує як сигнал тривоги і автоматично ініціює додатковий аналіз, обмежує обробку підозрілих транзакцій або запускає додаткові захисні процедури. Такий підхід дає змогу швидко й автоматично виявляти приховані атаки, які ще не призвели до помилок у класифікації, але вже впливають на внутрішню структуру мережі. Це підвищує рівень безпеки та стійкість критично важливих ІТ-систем.

Отримані результати підтверджують, що моніторинг спектру eigenvalues за RMT і DBM дає змогу ефективно виявляти наближення до критичних станів у складних системах. Це збігається з висновками інших дослідників [?, ?, ?], але запропонований підхід додатково враховує реальний динамічний рух eigenvalues, а не лише статичний аналіз. Порівняно з класичними метриками (ентропія, середні значення) спектральний аналіз є більш чутливим до “глибоких” змін топології мережі та параметрів ШНМ. Обмеженням є складність обробки дуже великих матриць у реальному часі, проте сучасні RMT-алгоритми і GPU-реалізації роблять це можливим у більшості критичних інфраструктур.

Висновки. У результаті проведеного дослідження було розроблено спектральний підхід до кількісної оцінки ризиків у кіберфізичних та нейромережових системах, який поєднує теорію випадкових матриць із моделлю броунівського руху власних значень. Запропонована формалізація матриць стану з урахуванням стохастичного шуму виявилася універсальною: вона коректно відображає як внутрішні флуктуації параметрів, так і зовнішні атаки, змодельовані як локальні зміни ваг чи втрати зв'язків. Спектральні метрики,

що базуються на щільності розподілу власних значень та аналізі крайніх спектральних параметрів, продемонстрували високу чутливість до моментів переходу системи в критичний режим і до початку самовідновлення, що робить їх ефективним інструментом для моніторингу поточного стану мережі.

Серія симуляційних експериментів підтвердила практичну цінність моделі: інтенсифікація зовнішніх збурень спричиняє миттєве збільшення флуктуацій крайніх власних значень, тоді як локальні відновлювальні дії призводять до швидкого відновлення спектральної когерентності до вихідного рівня. Використання строгих теоретичних результатів із рівняння Дайсона та асимптотичної теорії випадкових матриць забезпечує математичну коректність і відтворюваність дослідження, що створює надійну основу для подальших розробок і практичних впроваджень.

У подальших роботах доцільно розширити спектральний аналіз на мультиагентні серед-

овища з блочною структурою та міжблочними кореляціями, а також інтегрувати його з алгоритмами глибокого навчання для автоматичного виявлення аномалій і прогнозування відмов за даними сенсорів. Створення модуля вбудованого моніторингу в реальному часі дозволить операторам оперативно отримувати індикатори ризику й сповіщення про наближення до критичних порогів. Крім того, оптимізація процедур самовідновлення на основі емпіричних залежностей між інтенсивністю збурень та часом повернення до нормального режиму допоможе мінімізувати простій і втрати продуктивності. Нарешті, важливо дослідити методи доповнення даних у випадках неповного спостереження параметрів системи та розробити підходи оцінки невизначеності спектральних показників. Втілення цих рекомендацій сприятиме підвищенню стійкості, адаптивності та безпеки критичних інфраструктур і нейромережових систем нового покоління.

Список літератури:

1. Dyson F. J. A Brownian motion model for the eigenvalues of a random matrix. *Journal of Mathematical Physics*. 1962. Vol. 3, № 6. P. 1191–1198. DOI: 10.1063/1.1703862. URL: <https://surli.li/pteusy> (дата звернення: 28.07.2025).
2. Barhoumi Y. *What is... Random Matrix Theory*. Zurich: University of Zurich, 2011. URL: <https://scispace.com/papers/what-is-random-matrix-theory-1dev5xcq2m> (дата звернення: 28.07.2025).
3. Aarts G., Hajizadeh O., Lucini B., Park C. Dyson Brownian motion and random matrix dynamics of weight matrices during learning. *arXiv Preprint*. 2024. arXiv:2411.13512. DOI: 10.48550/arXiv.2411.13512. URL: <https://surli.cc/nnuif> (дата звернення: 28.07.2025).
4. Finocchio G. *Random Matrix Theory: Dyson Brownian Motion*. 2020. DOI: 10.14760/SNAP-2020-002-EN. URL: <https://surli.cc/ancpjc> (дата звернення: 28.07.2025).
5. Park C., Favoni M., Lucini B., Aarts G. Random Matrix Theory for Stochastic Gradient Descent. *arXiv Preprint*. 2024. arXiv:2412.20496. DOI: 10.48550/arXiv.2412.20496. URL: <https://surli.li/vybph> (дата звернення: 28.07.2025).
6. Erdős L. Random matrices, log gases and Hölder regularity. *arXiv Preprint*. 2014. URL: <https://surli.lt/rtothn> (дата звернення: 28.07.2025).
7. Siva Priya N., Shenbagavadivu N. A brief study on applications of Random Matrix Theory. In: *Advances in Industrial Automation and Smart Manufacturing*. Singapore: Springer, 2022. P. 401–411. DOI: 10.1007/978-981-16-2422-3_36. URL: <https://surli.li/getbre> (дата звернення: 28.07.2025).
8. Izenman A. J. Random Matrix Theory and its applications. *Statistical Science*. 2021. Vol. 36, № 2. P. 274–296. DOI: 10.1214/20-STS799. URL: <https://surli.cc/sobmxr> (дата звернення: 28.07.2025).
9. Couillet R., Debbah M. Random Matrix Theory. In: *Large Random Matrices for Wireless Communication*. Boca Raton: CRC Press, 2017. P. 371–422. DOI: 10.1201/9781351105668-8. URL: <https://scispace.com/papers/random-matrix-theory-210b50aj6b> (дата звернення: 28.07.2025).
10. Adler M. Integrable systems, random matrices, and random processes. In: *Integrable Systems, Random Matrices, and Random Processes*. 2011. DOI: 10.1007/978-1-4419-9514-8_2. URL: <https://surli.cc/xyftyc> (дата звернення: 28.07.2025).
11. van Handel R. Structured random matrices. In: *An Introduction to Random Matrices*. 2017. DOI: 10.1007/978-1-4939-7005-6_4. URL: <https://scispace.com/papers/structured-random-matrices-1edg0svryu> (дата звернення: 28.07.2025).
12. Rosenzweig F. On the Brownian motion model for the eigenvalues of a random matrix. *Il Nuovo Cimento*. 1965. Vol. 38, № 1. P. 257–264. DOI: 10.1007/BF02748615. URL: <https://surli.cc/piqtlb> (дата звернення: 28.07.2025).

13. Mineev-Weinstein M., Putinar M., Teodorescu R. Random matrices in 2D, Laplacian growth and operator theory. *Journal of Physics A: Mathematical and Theoretical*. 2008. Vol. 41, № 26. Art. 263001. DOI: 10.1088/1751-8113/41/26/263001. URL: <https://surl.lu/spxgqg> (дата звернення: 28.07.2025).
14. Aarts G., Lucini B., Park C. Stochastic weight matrix dynamics during learning and Dyson Brownian motion. *Physical Review E*. 2025. Vol. 111, № 1. Art. 015303. DOI: 10.1103/PhysRevE.111.015303. URL: <https://surl.lu/obqzmf> (дата звернення: 28.07.2025).
15. Alt J. Dyson equation and eigenvalue statistics of random matrices. 2018. URL: <https://surl.lt/xwuhew> (дата звернення: 28.07.2025).
16. Nagao T., Forrester P. J. Multilevel dynamical correlation functions for Dyson's Brownian motion model of random matrices. *Physics Letters A*. 1998. Vol. 247, № 1–2. P. 42–46. DOI: 10.1016/S0375-9601(98)00602-1. URL: <https://surl.cc/knxhtv> (дата звернення: 28.07.2025).
17. Erdős L. The matrix Dyson equation and its applications for random matrices. 2019. DOI: 10.1090/PCMS/026/03. URL: <https://surl.li/wqfzkm> (дата звернення: 28.07.2025).
18. Bogachev L. *Random Matrix Theory: Basics and Spectral Asymptotics*. 2010. URL: <https://surl.li/amuncc> (дата звернення: 28.07.2025).
19. Erdős L., Schnelli K. Universality for random matrix flows with time-dependent density. *Annales de l'Institut Henri Poincaré – Probabilités et Statistiques*. 2017. Vol. 53, № 4. P. 1942–2034. DOI: 10.1214/16-AHP765. URL: <https://surl.cc/nibrac> (дата звернення: 28.07.2025).
20. Miao P., Zhang Y., Wang C., Tong S. Random matrix description of dynamically backscattered coherent wave when full-field propagating in a random medium. *arXiv Preprint*. 2021. URL: <https://surl.li/ivwtmc> (дата звернення: 28.07.2025).
21. Tong S. Random matrix description of dynamically backscattered coherent waves propagating in a wide-field illuminated random medium. *Applied Physics Letters*. 2022. Vol. 120, № 4. Art. 041103. DOI: 10.1063/5.0078989. URL: <https://surl.lu/ajmlxo> (дата звернення: 28.07.2025).
22. Tulino A. M., Verdú S. *Random Matrix Theory and Wireless Communications*. 2004. URL: <https://surl.cc/cefaeq> (дата звернення: 28.07.2025).
23. Edelman A., Rao N. R. Random matrix theory. *Acta Numerica*. 2005. Vol. 14. P. 233–297. DOI: 10.1017/S0962492904000236. URL: <https://scispace.com/papers/random-matrix-theory-3fvcdalvny> (дата звернення: 28.07.2025).
24. El Karoui N. Random Matrix Theory. In: *Encyclopedia of Quantitative Finance*. 2010. DOI: 10.1002/9780470061602.EQF20005. URL: <https://scispace.com/papers/random-matrix-theory-1yisr0nkt7> (дата звернення: 28.07.2025).
25. Dietz B., Haake F. Random Matrix Theory as statistical mechanics. *EPL (Europhysics Letters)*. 1989. Vol. 9, № 1. P. 1–5. DOI: 10.1209/0295-5075/9/1/001. URL: <https://surl.li/xbkmjc> (дата звернення: 28.07.2025).
26. Pastur L., Shcherbina M. *Eigenvalue Distribution of Large Random Matrices*. 2011. URL: <https://surl.li/nqhmtx> (дата звернення: 28.07.2025).

Prokopovych-Tkachenko D.I., Khrushkov B.S., Bilan M.V., Cherkaskyi O.V., Derkach Ya.O.
DYSON'S BROWNIAN APPROACH AND APPLICATION OF RANDOM MATRIX THEORY
TO RISKS MODELING IN CYBERPHYSICAL AND NEURAL NETWORK SYSTEMS

This paper proposes and substantiates a modern approach to quantitative risk analysis in complex cyberphysical and neural network systems based on the theory of random matrices and the model of Brownian motion of eigenvalues (DBM). For the first time, a generalized scheme for formalizing the state and relationship matrices of infrastructure objects is presented, in which stochastic disturbances are taken into account through a random noise matrix, and the dynamics of the system is described by the evolution of its eigenvalues under the influence of Coulomb interactions. It is shown that such an approach allows us to reflect both the natural fluctuation of system parameters and external attacks, which are modeled by individual droplet losses or changes in weights in the relationship matrix.

New spectral risk metrics are developed, which are based on the analysis of the density distribution of eigenvalues, the statistical characteristics of their dispersion and deviations from the model curves of an equilibrium random gas. In particular, phase transition parameters are introduced, which determine the threshold values of one or more extreme eigenvalues, after which the system enters a state of increased vulnerability or a reduced level of self-healing.

For numerical validation of the approach, a series of simulation experiments were conducted on test arrays of dimensions from 50×50 to 500×500 with different attack resistance scenarios. The correlation between the values of spectral metrics and the level of degradation of functional reliability is demonstrated: in particular, with intense external actions, fluctuations of extreme eigenvalues increase, and after local repairs, their

coherence is quickly restored. Empirical relationships between the intensity of stochastic perturbation and the time of return to normal mode are constructed, which can serve as a basis for designing self-healing and adaptation mechanisms.

The results obtained correspond to modern trends in the development of critical systems security and can be integrated into the design processes of cyber-physical infrastructure, artificial intelligence systems and new generation network architectures. The presented approach expands the tools of analysts and developers, providing the opportunity to predict the boundaries of safe operation, respond promptly to threats and optimize protection and self-healing strategies.

Key words: *cyber-physical systems, random matrix theory, Brownian motion of eigenvalues, attack resistance, self-healing, spectral analysis.*

Дата надходження статті: 16.07.2025

Дата прийняття статті: 01.08.2025

Опубліковано: 27.10.2025

Rolik O.I.

National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute”

Ulianytska K.O.

National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute”

Amons O.A.

National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute”

Boiko O.V.

National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute”

Tsymbal S.I.

National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute”

MODERN INTERPRETATION FOR DATABASE DEVELOPMENT OF THE FIFTH NORMAL FORM

This article is devoted to the challenges of entity and relationship modeling within the framework of entity-attribute relationship analysis and further database normalization. The article reveals the theoretical and practical issues that arise when the use of secondary keys and attribute rules during infological design leads to violations of the Fifth Normal Form (5NF). It is shown that such violations can significantly affect the scalability of the database and complicate the denormalization process aimed at optimizing performance in practical applications. Particular attention is paid to the elimination of redundancy and insertion-update-deletion anomalies, which are frequently encountered in complex data models. The article presents a modern interpretation of the Fifth Normal Form by introducing the concept of conditional attributes. These conditional attributes allow developers to describe dependencies more precisely, without fragmenting data into an excessive number of related tables. The proposed approach helps avoid overcomplication of the schema and supports the development of flexible and maintainable database structures. It is emphasized that nontrivial dependencies often emerge not at the initial stage of development, but later—when business logic evolves and new requirements are introduced. It is clarified that insufficient time devoted to conceptual modeling is a major reason why database schemas fail to comply with 5NF. The proposed solution demonstrates how introducing conditional attributes, together with appropriate integrity constraints or triggers, can help ensure compliance without sacrificing performance. The article concludes that using modern interpretations of 5NF allows for building more robust, scalable, and semantically clear data models, capable of evolving alongside the application's functional growth.

Key words: normalization, fifth normal form, multivalued dependencies, secondary keys, conditional attributes, information systems and technologies, databases.

Introduction. The way to build a relational model is well-known. Basic principles of Entity Relationship (ER) modeling are as follows [1]:

- create and classify all entities related to the developing area;

- add and classify all necessary attributes to each entity;

- follow the four rules of attributes;

- add and classify all necessary relationships;

- generate ER model.

Each step of ER modeling is relatively simple, except for the first stage of each development – subject area analysis. The result of such analysis is creating an entity list; the main goal is to investigate each entity and select attributes. By following the rules of attributes, the possibility of redundancy existence should be decreased. So, redundancy in this case will be eliminated and only logical dependency might still be present.

Four rules of attributes have modern interpretation, which indicates more severe conditions for the attribute itself and dependency among them. In fact, some of these rules are very similar to the second and third normal forms.

In the next part of the article, attribute rules will be shown. Knowing the definition of the first five normal forms, including Boyce-Codd normal form, an interesting conclusion can be drawn. The ER model will follow the first five normal forms, including Boyce-Codd normal form in any case if we follow the four rules of attributes (including modern definition of these rules [2]). This conclusion leads us to a full investigation of the old interpretation of Fifth normal form (5NF).

Literature review and problem statement. In [1] a model based on N-ary relations, a normal form for database relations, and the concept of a universal data sublanguage are introduced and certain operations on relations (other than logical inference) are discussed and applied to the problems of redundancy and consistency in the user's model. In [2] Chen's Entity-Relationship (ER) model with some enhancements needed for a better conceptual representation are used. This model is extensively used in many design methodologies, has an effective graphic representation, and is the de facto standard of most automatic tools for supporting database design. A joint methodology for conceptual database design and functional analysis are used. The activity of modeling involves the creation of relations, attributes, and indicating relationships among relations with a set of integrity constraints [5]. Modern interpretation of Four attribute rules, mentioned in [2] and [5] present a refined and updated approach to conceptual modeling, including [2-5]:

1. Clarity in attribute naming and meaning (semantic precision).
2. Atomicity (each attribute should represent one fact).
3. Avoidance of redundancy (removal of derived or duplicate attributes).
4. Stability and relevance (attributes should reflect intrinsic and stable properties of entities).

It builds on foundational principles but adapts them to modern practices in database design and information systems development. Using this interpretation reveals more and more facts that correct conceptual modeling leads us to follow most normal forms and some of them become useless in their old definition [6, 7].

The issues of ER modeling in the context of entities and attributes in the process of further normalization remain insufficiently researched. This requires eliminating the situation which leads to the violation of the Fifth normal form (5NF) when using secondary keys and rules of attributes during infological designing. This violation might be challenging during further database growing, when implementing denormalization steps to improve query execution in some cases takes place.

The aim and objectives of the study. The purpose of the investigation: is to develop a modern interpretation of 5NF for database development process.

To achieve the goal, it is necessary to perform the following tasks:

- to perform an analysis of attribute rules and secondary keys. Here will be found cause of 5NF violation and make assumption with new term – conditional attribute – for making investigation with new conditional attributes in relational model to check whether 5NF is followed;
- to propose a modern interpretation of the 5NF form to eliminate the situation which leads to a violation of the 5NF, when using secondary keys, conditional attributes and rules of attributes from infological designing.

The study materials and methods for database development of the fifth normal form.

Attribute rules and secondary keys

A relational database is a database where data has a structure, built on the relational model, in other words, it is a set of relations (tables) followed by severe rules (normal forms). Inside of each table, exists a set of attributes, which describes a specific entity, all attributes must be connected with each other in a one very specific way – they all depend on a primary key.

Figure 1 demonstrates several tables, which have primary key in a different manner. The most popular way – is using composite primary key as it shown on Figure 1, a). Such approach makes table follow by several normal forms – Second normal form, Boyce-Codd normal form and Fourth normal form (4NF). The other way, which is also frequently used – is using additional attributes as compound primary (Figure 1, b)). It might be used only in that case, when this table is not been referenced by another tables.

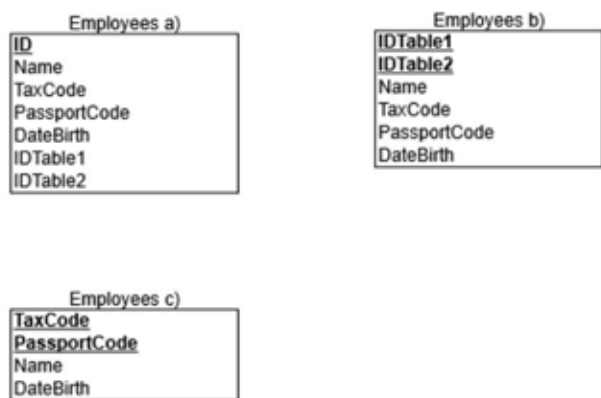


Fig. 1. Different approaches of creating primary key

Figure 1, c) demonstrates old approach, when developers use business attributes as part of compound primary key. Such approach leads us to a crucial redundancy; while creating relationships between such tables, we are duplicating business attributes (TaxCode and PassportCode) as they are used as foreign key in another tables.

A foreign key (or secondary key) is a combination of one or more additional attributes that is expressed in two or more relationships. It is used to create relationships between a pair of tables.

The normalization process is a formal technique that allows obtaining a relational data model, whose elements will meet the specific requirements of this model, resulting from the definition of normal forms.

Compliance with the requirements of normal forms is somehow necessary – it turns out that if the tables included in the database do not meet such requirements, i.e. the database schema is incorrect, then in certain situations there will be irregularities that prevent proper work with the database.

The normalization process is based on establishing functional dependencies (determination relationships) between attributes in a relation. We say that attribute A is functionally dependent on attribute B if 2 each value of B is related to exactly one value of A – in other words, A determines B . Furthermore, we say that attribute Z is transitively functionally dependent on attribute X if Z functionally depends on a subset of Y , Y is functionally dependent on X , while X is not functionally dependent on Y and Y is not functionally dependent on Z .

Example for such dependency is shown on Figure 2:

Functional Dependencies:

1) EmployeeID $\rightarrow$ EmployeeName, DepartmentID – each employee has exactly one name and is in one department;

	EmployeeID	EmployeeName	DepartmentName	DepartmentLocation
1	1	Alice	HR	New York
2	2	Bob	IT	San Francisco
3	3	Charlie	HR	New York

Fig. 2. Table Employees with functional dependency

2) DepartmentID $\rightarrow$ DepartmentName, DepartmentLocation – each department has one name and location.

For this case, to make table Employees follow normal forms, need to perform decomposition without any detailing (Figure 3):

```
CREATE TABLE Department (
    DepartmentID INT PRIMARY KEY,
    DepartmentName NVARCHAR(100) NOT NULL,
    DepartmentLocation NVARCHAR(100) NOT NULL
);

CREATE TABLE Employee (
    EmployeeID INT PRIMARY KEY,
    EmployeeName NVARCHAR(100) NOT NULL,
    DepartmentID INT NOT NULL,
    FOREIGN KEY (DepartmentID) REFERENCES Department(DepartmentID)
);
```

Fig. 3. Scripts for decomposed table Employees

Entity relation model is built during infological (conceptual) modeling, which demands us following some rules. Basic rules for conceptual modeling – are Four attribute rules. According to the original text, rules are:

1. Atomicity. Each attribute should represent a single, indivisible unit of data [3].

2. Full Functional Dependence on the Entity Identifier. Every attribute should be fully functionally dependent on the primary key of the relation [3].

3. Redundancy Avoidance. Attributes should be chosen in such a way that data redundancy is minimized [3]. Attributes should not repeat the same information or duplicate data that can be derived in another way.

4. No Derived Attributes. A relation should not contain attributes that can be derived from other attributes in the same relation [3].

All mentioned rules are related to some of the normal form, so it is easy to start the next step of development process – datalogic modeling or physical modeling. After ER model was created, we need to adapt it to relational database model. Normalization process should take place. If developer follow mentioned attribute rules, normalization process will be very fast.

First attribute rule is very similar to 1NF definition – a relation is in 1NF if and only if the domain of each attribute contains only atomic (indivisible)

values, and the value of each attribute contains only a single value from that domain. (The ordering of rows is immaterial, and each item in a row must be atomic, that is, indivisible) [1]. One of the algorithms that allows you to transform a non-normalized data set into a table that meets the requirements of first normal form is to remove repeating groups by inserting the appropriate data into empty columns, i.e., duplicating data in rows containing repeating groups. However, this often leads to redundancy, so if more than one value is defined for an attribute at a given time, or there is more than one attribute with the same name, a new entity must be defined that is described by this attribute. But here in the first attribute rule we have a more severe interpretation – instead of a single attribute like *FullName*, it is better to use separate attributes: *LastName*, *FirstName*, *MiddleName*. This rule tells us **that an attribute must not have a complex structure**.

Second attribute rule is almost the same as 2NF. A relation is in 2NF if it is in 1NF, and every non-prime attribute is fully functionally dependent on the whole of every candidate key of the relation (not just a part of any composite key) [1]. This rule tells us, that **an entity should have an identifier (surrogate attribute)**; in this case all other attributes will depend only on this primary key.

It should be noted that the quoted definition applies to tables that have composite primary keys; in this case, it may happen that some columns will depend on the entire key, while others will depend on its components. The transformation from 1NF to 2NF will consist of removing partial functional dependencies, i.e., of dividing the data structure so that each of the non-key attributes depends fully on the primary key. In practice, this means that if an entity has a unique identifier consisting of more than one attribute and/or relationship, and if some other attribute depends only on part of this composite identifier, then the attribute and the part of the identifier on which it depends form the basis of a new entity (table).

A relation that is in 2NF and whose non-key attributes depend on a key directly, not transitively, satisfies the requirements of 3NF. This means that if an attribute of an entity depends on another attribute that is not part of the unique identifier, then these attributes should form the basis for a new entity that has a one-to-many relationship with the original entity. The unique identifier for the new entity is the attribute on which the second attribute depends.

In the context of relational databases, 3NF is a key design principle and data modeling standard that provides effective organization and normalization of data in a database.

4NF is related to another type of dependency between relation attributes, namely multivalued dependencies. A multivalued dependency between attributes *A*, *B*, and *C* of relation *R* occurs when for each value *A* there is a set of corresponding values *B* and a set of values *C* (sets of values *B* and *C* are independent of each other). A table meets the requirements of the fourth normal form if it does not contain many multivalued dependencies, but only a single one. For example, if we have a separate table in which we have employee identifier and assigned private and work telephone numbers (two separate columns), then it should be divided into two separate tables.

5NF concerns so-called join dependencies, which are a generalization of multi-valued dependencies. It concerns relations in which there are more than two multi-valued dependencies. In reality, situations requiring the application of the normalization procedure ensuring fifth normal form almost never occur, nevertheless, this form is presented in the literature on the subject. The general definition says that a relation is in fifth normal form if it is in fourth normal form and does not contain join dependencies, in other words, if there is no lossless decomposition into smaller tables.

The third and fourth rules are very similar to each other. But, starting from this point, we are approaching 5NF, toward understanding it. The third rule only points out that there is no need to duplicate information, for example (Figure 4), there is no need to keep the name of document type in table Documents, if we already have foreign key to table DocumentTypes.



Fig. 4. Database diagram with/without redundant attribute

To get the name of document type, we just need to JOIN concrete table (Figure 5).

This rule leads us to a very important constraint – secondary key (foreign key [4]). The whole relational database schema is based on the usage of pairs – primary and foreign keys. This is the main feature of all relational databases. An example of storing data using the described keys is presented in Figure 6.

```

1 SELECT doc.DocumentNumber, doc.DocumentDate, type.DocumentTypeName
2 FROM T_Documents doc
3 JOIN T_DocumentTypes type ON doc.IDDocumentType = type.ID
4

```

DocumentNumber	DocumentDate	DocumentTypeName
234/o	2025-05-27 16:20:28.907	Type1
101/u	2025-05-27 16:20:28.907	Type1
201/h	2025-05-27 16:20:28.907	Type2
301/o	2025-05-27 16:20:28.907	Type2
401/f	2025-05-27 16:20:28.907	Type1

Fig. 5. Results of query with JOIN expression

Primary key (DocumentType)		DocumentTypeName
1	1	Type1
2	2	Type2
3	3	Type3
4	4	Type4
5	5	Type5

ID	Foreign key (DocumentType)	DocumentNumber	DocumentDate
1	1	234/o	2025-05-27 16:20:28.907
2	1	101/u	2025-05-27 16:20:28.907
3	2	201/h	2025-05-27 16:20:28.907
4	2	301/o	2025-05-27 16:20:28.907
5	3	401/f	2025-05-27 16:20:28.907

Fig. 6. Example of storing the referenced data

The advantage of using secondary keys is that we can reference data as many times as needed (Figure 6), only the primary key will be used, not any business attributes (e.g., Name, DateOfBirth, Code, etc.). At any time, each business attribute can be changed, and we can access fresh data at any time through the reference. There is no redundancy, no duplicity, no multi-valued dependency and no non-trivial dependency (as it will appear in 5NF).

Modern interpretation of the Fifth normal form

The original definition of the 5NF, also known as Projection-Join Normal Form (PJNF), comes from Ronald Fagin in his seminal paper [8]: a relation R is in 5NF, or PJNF, if every join dependency in R is implied by the candidate keys of R .

C.J. Date in his work [3] gave for 5NF more details and descriptions: a relation variable R is in 5NF, also sometimes referred to as Projection-Join Normal Form, if and only if every non-trivial join dependency in the relation variable R is implied by the candidate key(s) of R , provided the following conditions are met:

a) a join dependency $\{A, B, \dots, Z\}$ in relation variable R is considered trivial if and only if at least one of the subsets $A, B, \dots, Z$ of the set of attributes equals the entire set of attributes of R .

b) a join dependency $\{A, B, \dots, Z\}$ in relation variable R is said to be implied by the candidate key(s) of R if and only if each of the subsets $A, B, \dots, Z$ of the set of attributes is a superkey for R .

The next Figure 7 illustrates tables (entities), that are not follow 5NF.

```

1 CREATE TABLE EmployeeSkillsLanguages (
2   Employee VARCHAR(50),
3   Skill VARCHAR(50),
4   Language VARCHAR(50),
5   PRIMARY KEY (Employee, Skill, Language)
6 );
7
8 INSERT INTO EmployeeSkillsLanguages (Employee, Skill, Language) VALUES
9 ('Alice', 'Java', 'English'),
10 ('Alice', 'Java', 'French'),
11 ('Alice', 'Python', 'English'),
12 ('Alice', 'Python', 'French');
13
14 select * from EmployeeSkillsLanguages

```

Employee	Skill	Language
Alice	Java	English
Alice	Java	French
Alice	Python	English
Alice	Python	French

Fig. 7. Relational table, that contains join dependencies

The example was taken from [3], but reframed using modern relational database design. This is the most common example of a table that does not follow 5NF. But, if we investigate such tables closely, we will see a conceptual model mistake. At the step of infological modeling, we need to get all appropriate entities and their relationships. Obviously, the table EmployeeSkillsLanguage is not an entity from the ER model. Such tables represent an associative entity with (in this case) three foreign keys, and if we have dependencies between attributes Employee, Skill, Language – it is described at the conceptual step, add two more entities or any other integrity constraints further on physical modeling (e.g., CHECK [9]).

Generally, to make the EmployeeSkillsLanguages table follow 5NF, we need to apply decomposition as shown in Figure 8.

This is bad practice in modern development. Today, developers are creating three separate vocab-

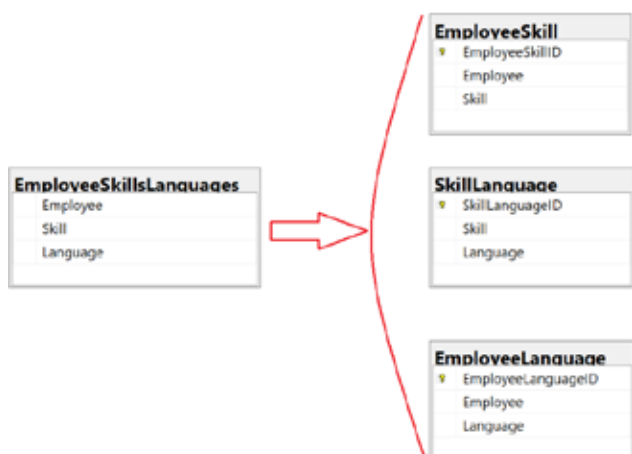


Fig. 8. Decomposition of table, which is not following 5NF

ulary tables (Employees, Skills, Languages) and any business rule can be put in the associative table as an integrity constraint CHECK or even in TRIGGER [9] programming code. Another best approach is during application development and writing documentation use cases that specify how to fill the associative table (application form) in the correct way.

An example of a relational model with foreign keys and dictionary tables is shown in Figure 6.

Theory tells us that the table Project (Figure 9) does not follow 5NF. But if we create tables like EmployeeSkill, SkillLanguage, EmployeeLanguage (Figure 8), even with modern foreign keys, it will increase complexity for application development.

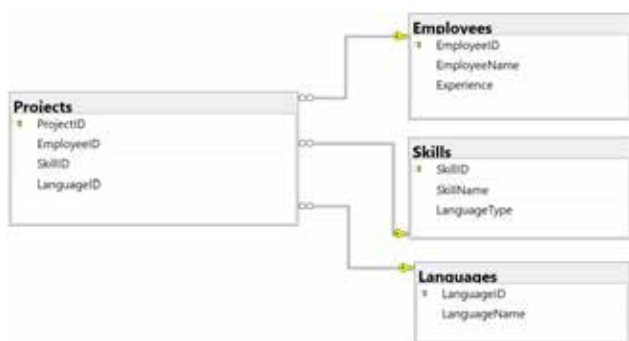


Fig. 9. Modern solution with secondary keys

To prevent development complexity and to make the table Projects follow 5NF, we may add a specific CHECK constraint or TRIGGER.

Assume that the dependency Employee $\rightarrow$ Skills can be handled by adding an attribute Experience, which indicates that an employee with a particular value of Experience={1,2}, should have appropriate Skills={S1,S2,...,Si}.

The same goes for Skills $\rightarrow$ Languages, this dependency can be handled by adding attribute LanguageType, which indicates, that skill with particular LanguageType={1,2}, should have appropriate Language={S1,S2,...,Sk}.

On Figure 10 and Figure 11 is shown example of trigger for Employee-Skills, Skills-Language dependencies.

This is one way of how to handle 5NF. There also might be applied check constraints or added algorithm to application code, to prevent dependency violation.

The main point is that situation on fig.6 would never appear if on conceptual modeling step specific attributes was added (ex, Employees.Experience; Skills.LanguageType etc.) for resolving dependency between entity, cause each dependency, in fact, based on some conditions, which might appear further in, as example, TRIGGER (Figure 10, Figure 11).

```
CREATE TRIGGER trg_Projects_Enforce5NF
ON Projects
INSTEAD OF INSERT
AS
BEGIN
    SET NOCOUNT ON;

    -- Inserted projects
    DECLARE @EmployeeID INT, @SkillID INT, @LanguageID INT, @Experience TINYINT, @LanguageType TINYINT;

    -- Cursor for each row is inserted (if multi-row insert)
    DECLARE cur CURSOR FOR
    SELECT i.EmployeeID, i.SkillID, i.LanguageID, e.Experience, s.LanguageType
    FROM Inserted i
    JOIN Employees e ON i.EmployeeID = e.EmployeeID
    JOIN Skills s ON i.SkillID = s.SkillID;

    OPEN cur
    FETCH NEXT FROM cur INTO @EmployeeID, @SkillID, @LanguageID, @Experience, @LanguageType;

    WHILE @@FETCH_STATUS = 0
    BEGIN
        -- 1. Validate Skills match Experience level
        IF @Experience = 1 AND @SkillID NOT IN (1, 2)
        BEGIN
            RAISERROR('Employee with Experience=1 must be assigned only Skills S1 or S2.', 16, 1)
            ROLLBACK TRANSACTION
            RETURN
        END
        ELSE IF @Experience = 2 AND @SkillID NOT IN (3, 4, 5)
        BEGIN
            RAISERROR('Employee with Experience=2 must be assigned only Skills S3, S4, or S5.', 16, 1)
            ROLLBACK TRANSACTION
            RETURN
        END
        ELSE
        BEGIN
            -- 2. Validate Skill belongs to Employee
            IF NOT EXISTS (
                SELECT 1 FROM EmployeeSkills es
                WHERE es.EmployeeID = @EmployeeID AND es.SkillID = @SkillID
            )
            BEGIN
                RAISERROR('Employee does not possess the assigned Skill.', 16, 1)
                ROLLBACK TRANSACTION
                RETURN
            END
        END

        -- 3. Validate language by Skill.LanguageType
        IF @LanguageType = 1 AND @LanguageID NOT IN (1, 2) -- L1, L2
        BEGIN
            RAISERROR('languageID must be 11 or 12 for LanguageType=1.', 16, 1)
            ROLLBACK TRANSACTION
            RETURN
        END
        ELSE IF @LanguageType = 2 AND @LanguageID NOT IN (3, 4, 5) -- L3, L4, L5
        BEGIN
            RAISERROR('languageID must be 13, 14, or 15 for LanguageType=2.', 16, 1)
            ROLLBACK TRANSACTION
            RETURN
        END

        FETCH NEXT FROM cur INTO @EmployeeID, @SkillID, @LanguageID, @Experience, @LanguageType;
    END

    CLOSE cur
    DEALLOCATE cur;

    -- All checks passed: perform insert
    INSERT INTO Projects (ProjectID, EmployeeID, SkillID, LanguageID)
    SELECT ProjectID, EmployeeID, SkillID, LanguageID FROM Inserted
END
```

Fig. 10. Trigger for Projects table for following 5NF (start)

```

-- 2. Validate Skill belongs to Employee
IF NOT EXISTS (
    SELECT 1 FROM EmployeeSkills es
    WHERE es.EmployeeID = @EmployeeID AND es.SkillID = @SkillID
)
BEGIN
    RAISERROR('Employee does not possess the assigned Skill.', 16, 1)
    ROLLBACK TRANSACTION
    RETURN
END

-- 3. Validate language by Skill.LanguageType
IF @LanguageType = 1 AND @LanguageID NOT IN (1, 2) -- L1, L2
BEGIN
    RAISERROR('languageID must be 11 or 12 for LanguageType=1.', 16, 1)
    ROLLBACK TRANSACTION
    RETURN
END
ELSE IF @LanguageType = 2 AND @LanguageID NOT IN (3, 4, 5) -- L3, L4, L5
BEGIN
    RAISERROR('languageID must be 13, 14, or 15 for LanguageType=2.', 16, 1)
    ROLLBACK TRANSACTION
    RETURN
END

FETCH NEXT FROM cur INTO @EmployeeID, @SkillID, @LanguageID, @Experience, @LanguageType;
END

CLOSE cur
DEALLOCATE cur;

-- All checks passed: perform insert
INSERT INTO Projects (ProjectID, EmployeeID, SkillID, LanguageID)
SELECT ProjectID, EmployeeID, SkillID, LanguageID FROM Inserted
END
```

Fig. 11. Trigger for Projects table for following 5NF (finish)

For preventing tables from nontrivial dependency, need to add in 5NF interpretation specific conditional attributes, that will be used in code, SQL script or SQL programming code to resolve such dependency.

Working in such way – any nontrivial dependency might be resolved even on database level.

RESULTS OF INVESTIGATING MODERN DEFINITION FOR 5NF

Based on the example and analysis which were applied in this article, the definition for Fifth Normal Form can be improved by including a new determination – conditional attributes.

The main problem of each table that might not be compliant with 5NF is the lack of time that developers spend on conceptual modeling.

The situation that appears in Figure 6 is not knowing the conditions which lead to dependency, in this case, nontrivial dependency, because all three attributes are involved in it.

Using conditional attributes that will significantly clarify dependency between sets of values, we can achieve compliance with 5NF. That means: getting rid of all kinds of anomalies (during running insert, update, delete operations in the database); getting rid of data redundancy (including using secondary keys) and implementing nontrivial dependency without harming future applications.

Old definition of 5NF: a relation R is in 5NF, or PJNF, if every join dependency in R is implied by the candidate keys of R .

Modern definition of 5NF, which includes conditional attributes: a relation R is in 5NF, or PJNF, if every join dependency in R is implied by the candidate keys of R and for each pair of dependent attributes a conditional attribute was added.

Conditional attribute – an attribute that leads us to understanding dependency between sets of values in each pair of dependent attributes.

DISCUSSION OF RESULTS OF INVESTIGATING MODERN DEFINITION FOR 5NF

This article is devoted to the issues of Entity Relationship modeling in the context of the ratio of entities

and attributes in the process of further normalization. The article shows the elimination of that situation which leads to the violation of the fifth normal form when using secondary keys and rules of attributes of infological design. This violation might be challenging in the further growth of the database itself with implementing attempts at denormalization steps to improve query execution in some cases.

After analysis and investigation that was made and also in the article, a modern interpretation of the Fifth Normal Form was proposed. This definition allows developers to better understand the normalization process and take into account some steps in conceptual modeling.

Conclusions. The material of this article is devoted to the normalization process, which apparently will take place not at the start of developing an application.

Such nontrivial dependency in tables frequently appears during improvement and adding new features, changing business rules of application usage.

Modern interpretation of 5NF allows developers to simplify resolving situations with nontrivial dependencies by adding new conditional attributes for each dependency, which may lead to not creating a bunch of tables, but adding new integrity constraints or program units such as triggers.

Bibliography:

1. Codd, E. F. A Relational Model of Data for Large Shared Data Banks / E. F. Codd // Communications of the ACM. – 1970. – Vol. 13, No. 6. – P. 377–387.
2. Batini, C., Ceri, S., Navathe, S. B. Conceptual Database Design: An Entity-Relationship Approach. – San Mateo, CA: Benjamin/Cummings Publishing, 2011. – ISBN 9780201325775.
3. Date, C. J. An Introduction to Database Systems. – 8th ed. – Boston : Pearson/Addison Wesley, 2003. – 1024 p.
4. Elmasri, R., Navathe, S. B. Fundamentals of Database Systems. – 7th ed. – Boston : Pearson, 2017. – ISBN 978-0133970777.
5. Elmasri, R., Navathe, S. B. Improving Conceptual Modeling with Semantic Clarity and Data Quality Constraints // Journal of Database Management (JDM). – 2021. – Vol. 32, No. 1. – P. 1–21. DOI: 10.4018/JDM.2021010101.
6. Rolik, O., Amons, O., Ulianytska, K., Kolesnik, V. Modernization of the Second Normal Form and Boyce-Codd Normal Form for Relational Theory // In: Hu, Z., Petoukhov, S., Dychka, I., He, M. (eds) Advances in Computer Science for Engineering and Education III. ICCSEEA 2020. Advances in Intelligent Systems and Computing, vol. 1247. – Springer, 2021. – P. 296–305. DOI: 10.1007/978-3-030-55506-1_27.
7. Rolik, O., Ulianytska, K., Khmeliuk, M., Khmeliuk, V., Kolomiets, U. Increase efficiency of relational databases using instruments of second normal form // 2021 IEEE 3rd International Conference on Advanced Trends in Information Theory, ATIT 2021 – Proceedings. – 2021. – P. 221–225.
8. Fagin, R. Normalization and Relational Database Operators // ACM SIGMOD Record. – 1979. – Vol. 6, No. 3. – P. 33–41. DOI: 10.1145/582095.582099.
9. Jorgensen, K. Microsoft SQL Server 2022: A Beginner's Guide. – Sixth Edition. – McGraw-Hill Education, 2022. – ISBN 9781264713886.

Ролік О.І., Ульяницька К.О., Амонс О.А., Бойко О.В., Цимбал С.І. СУЧАСНА ІНТЕРПРЕТАЦІЯ РОЗРОБКИ БАЗИ ДАНИХ П'ЯТОЇ НОРМАЛЬНОЇ ФОРМИ

Стаття присвячена проблемам моделювання сутностей і зв'язків у контексті аналізу типів сутність-атрибут і подальшої нормалізації баз даних. У статті розкрито теоретичні та прикладні аспекти, які виникають при використанні вторинних ключів і правил атрибутів в інфологічному проєктуванні, що може призводити до порушення п'ятої нормальної форми (5НФ). З'ясовано, що такі порушення негативно впливають на масштабованість бази даних та ускладнюють процес денормалізації, який здійснюється з метою оптимізації продуктивності при практичному використанні. Розкрито сучасний підхід до інтерпретації п'ятої нормальної форми, що ґрунтується на впровадженні поняття умовного атрибута. Запропоновано використовувати умовні атрибути для точнішого опису залежностей без надмірного розбиття даних на велику кількість взаємопов'язаних таблиць. Це дозволяє уникнути ускладнення схеми та забезпечити гнучкість і зручність підтримки бази даних у процесі її розвитку. Підкреслено, що складні залежності між атрибутами часто з'являються не на початковому етапі розробки, а під час модифікації бізнес-логіки та впровадження нових функціональних вимог. З'ясовано, що однією з головних причин невідповідності схем баз даних вимогам 5НФ є недостатній час, приділений концептуальному моделюванню. Запропоновано рішення, яке полягає у впровадженні умовних атрибутів разом із відповідними обмеженнями цілісності або тригерами, що дозволяє забезпечити відповідність 5НФ без шкоди для продуктивності. Стаття доводить, що сучасне трактування 5НФ дає змогу проєктувати більш надійні, масштабовані та семантично зрозумілі моделі даних, які можуть еволюціонувати відповідно до зростаючих вимог прикладного програмного забезпечення.

Ключові слова: нормалізація, п'ята нормальна форма, багатозначні залежності, вторинні ключі, умовні атрибути, інформаційні системи та технології, бази даних.

Дата надходження статті: 31.07.2025

Дата прийняття статті: 01.09.2025

Опубліковано: 27.10.2025

Рудницький В.М.

Черкаський державний технологічний університет

Бабенко В.Г.

Черкаський державний технологічний університет

Рудницький С.В.

Черкаський державний технологічний університет

Короткий Т.К.

Черкаський державний технологічний університет

Ковтюх В.А.

Черкаський державний технологічний університет

ОСОБЛИВОСТІ ГРУП НЕСИМЕТРИЧНИХ СЕТ-ОПЕРАЦІЙ СИНТЕЗОВАНИХ З ТОЧНІСТЮ ДО ПЕРЕСТАНОВКИ ПЕРШОГО ОПЕРАНДА

Одним із перспективних шляхів розвитку малоресурсної криптографії є СЕТ-шифрування. Стаття присвячена дослідженню, що вирішує завдання, яке полягає у дослідженні можливості застосування методу генерації груп несиметричних 2-операндних СЕТ-операцій з точністю до перестановки першого операнда для побудови малоресурсних несиметричних систем потокового шифрування підвищеної криптостійкості. Об'єкт дослідження: процеси взаємоперетворення моделей 2-операндних СЕТ-операцій в процесі їх модифікації з точністю до перестановки. Використання різних методів генерації синтезу груп операцій з точністю до перестановки приводить до генерації різних псевдовипадкових СЕТ-операцій. При чому дані послідовності можуть відрізняються не тільки послідовностями операцій, але і різними їх наборами. Для вирішення завдання на прикладі несиметричної 2-операндної СЕТ-операції розглянуті процеси її модифікації з точністю до перестановки першого операнда однооперандною СЕТ-операцією та встановлено, що така модифікація реалізується шляхом модифікації однооперандних СЕТ-операцій, з яких вона побудована. На основі моделей модифікації СЕТ-операцій реалізуються генератори псевдовипадкових послідовностей прямих і обернених модифікованих 2-операндних СЕТ-операцій. За результатами аналізу встановлено, що модифікація однооперандних СЕТ-операцій в процесі шифрування забезпечує збільшення кількості таблиць підстановки, що реалізує поточковий шифр. Для практичної реалізації доцільно будувати поточкові шифри з використанням груп модифікованих несиметричних 2-операндних СЕТ-операцій з точністю до перестановки першого операнда. Збільшення кількості таблиць підстановки при послідовній модифікації СЕТ-операцій призводить до ускладнення криптоаналізу та до підвищення криптографічної стійкості криптографічного алгоритму. Сфера використання отриманих результатів: мобільні і стаціонарні системи малоресурсного криптографічного захисту конфіденційної інформації.

Ключові слова: малоресурсна криптографія, СЕТ-шифрування, 2-операндні СЕТ-операції, несиметричні криптографічні перетворення, потокове шифрування.

Постановка проблеми. Проблема захисту інформації в комп'ютерних системах від некомпетентності користувачів, некоректності програмного забезпечення та несанкціонованого доступу до інформації стала актуальною при створенні перших засобів електронно-обчислювальної тех-

ніки. Побудова перших комп'ютерних мереж привела до виникнення і широкого розповсюдження кіберзлочинності [1, 2]. Ефективно протидіяти несанкціонованому доступу до конфіденційної інформації в комп'ютерних мережах можна на основі широкого впровадження криптографічних

алгоритмів [3]. Проте застосування криптографічного захисту має і негативні наслідки, пов'язані з відволіканням обчислювальних ресурсів, збільшенням часу обробки інформації та збільшенням вартості роботи системи [4, 5]. Зменшити негативні наслідки від впровадження криптографічних систем можливо на основі застосування малоресурсної (легкої) криптографії. Одним із перспективних шляхів розвитку малоресурсної криптографії є СЕТ-шифрування [6].

Аналіз останніх досліджень і публікацій. Традиційно при побудові поточкових шифрів використовується криптографічне додавання інформації з псевдовипадковою послідовністю. На практиці в якості криптографічного додавання використовується додавання за модулем [7]. Додавання за модулем в теорії СЕТ-шифрування розглядається як 2-операндна СЕТ-операція [6]. Крім того методи синтезу СЕТ-операцій з точністю до перестановки дозволяють синтезувати групи модифікованих операцій [8, 9]. Реалізація симетричних систем поточкового шифрування простіша за реалізацію несиметричних систем поточкового шифрування. Виходячи з цього, початковий вектор дослідження був спрямований на синтез симетричних 2-операндних СЕТ-операцій для забезпечення простоти реалізації криптоалгоритмів [6, 20]. Для підвищення криптостійкості поточкових шифрів було запропоновано замість однієї однооперандної СЕТ-операції використовувати групи операцій з точністю до перестановки [19]. У роботах [10-12] показано, що синтезувати групи симетричних 2-операндних СЕТ-операцій і генерувати їх послідовності можна з точністю до перестановки результату перетворення. Генерувати послідовності несиметричних СЕТ-операцій доцільно з точністю до перестановки другого операнда [13, 14]. На сьогоднішній день метод синтезу груп несиметричних СЕТ-операцій з точністю до перестановки першого операнда не досліджувався через складність реалізації оберненого криптографічного перетворення.

Постановка завдання. Метою роботи є дослідження можливості застосування методу генерації груп СЕТ-операцій з точністю до перестановки першого операнда для побудови малоресурсних несиметричних систем поточкового шифрування підвищеної криптостійкості.

Виклад основного матеріалу. Результати моделювання груп несиметричних 2-операндних СЕТ-операцій з точністю до перестановки першого операнда. При вдосконаленні і побудові систем поточкового шифрування можна

використовувати методи генерації псевдовипадкових послідовностей СЕТ-операцій з точністю до перестановки першого операнда, другого операнда, або результату криптографічного перетворення [6, 15]. При побудові несиметричних систем поточкового шифрування на сьогоднішній день використовується генерація псевдовипадкових послідовностей СЕТ-операцій з точністю до перестановки другого операнда [15, 16]. Це пов'язано з простотою генерації груп прямих і обернених несиметричних СЕТ-операцій. Використання різних методів генерації синтезу груп операцій з точністю до перестановки приводить до генерації різних псевдовипадкових СЕТ-операцій. При чому дані послідовності можуть відрізняються не тільки послідовностями операцій, але і різними їх наборами.

Дослідимо особливості застосування методу синтезу СЕТ-операцій з точністю до перестановки першого операнда, а також визначимо переваги та недоліки їх застосування в несиметричних системах поточкового шифрування.

Моделювання процесів прямого і оберненого криптографічного перетворення в групі несиметричних 2-операндних СЕТ-операцій з точністю до перестановки першого операнда проведемо на прикладі 2Сі-квантової СЕТ-операції [6]. Дане обмеження пов'язане з мінімальною кількістю СЕТ-операцій, на основі яких будуються несиметричні 2-операндні СЕТ-операції, а також простотою застосування математичного апарату для моделювання взаємних перетворень моделей операцій [17].

Коректність отриманих за результатами моделювання взаємозв'язків між прямими і оберненими модифікованими СЕТ-операціями з точністю до перестановки першого операнда перевіримо на основі обчислювального експерименту.

Задамо модель несиметричної 2-операндної 2Сі-квантової СЕТ-операції [6] кортежем із чотирьох однооперандних 2Сі-квантових СЕТ-операцій. Якщо

$$C_1(x) = \begin{bmatrix} x_1 \oplus x_2 \\ x_2 \oplus 1 \end{bmatrix} = z; \quad C_2(x) = \begin{bmatrix} x_1 \oplus 1 \\ x_1 \oplus x_2 \end{bmatrix} = z;$$

$$C_3(x) = \begin{bmatrix} x_2 \\ x_1 \oplus 1 \end{bmatrix} = z; \quad C_4(x) = \begin{bmatrix} x_2 \oplus 1 \\ x_1 \end{bmatrix} = z,$$

де $C_i(x)$ – i -та однооперандна СЕТ-операція; x_1 і x_2 – перший і другий Сі-кванти першого операнда (вхідна інформація); z – результат криптографічного перетворення (2 Сі-кванти), тоді:

$$C(x, y) = C(C_1(x), C_2(x), C_3(x), C_4(x)) = z \quad (1)$$

$$C(x, y) = \begin{cases} \begin{bmatrix} x_1 \oplus x_2 \\ x_2 \oplus 1 \end{bmatrix}, & \text{якщо } y_1 = 0; y_2 = 0; \\ \begin{bmatrix} x_1 \oplus 1 \\ x_1 \oplus x_2 \end{bmatrix}, & \text{якщо } y_1 = 0; y_2 = 1; \\ \begin{bmatrix} x_2 \\ x_1 \oplus 1 \end{bmatrix}, & \text{якщо } y_1 = 1; y_2 = 0; \\ \begin{bmatrix} x_2 \oplus 1 \\ x_1 \end{bmatrix}, & \text{якщо } y_1 = 1; y_2 = 1. \end{cases} = z, \quad (2)$$

де $C(x, y)$ – це 2-операндна СЕТ-операція; y_1 і y_2 – перший і другий Сі-кванти другого операнда (інформація керування); z – результат криптографічного перетворення (2 Сі-кванти).

На основі моделі (2) отримаємо удосконалену дискретну модель СЕТ-операції:

$$C(x, y) = \begin{bmatrix} x_1 \cdot \bar{y}_1 \oplus x_2 \cdot (y_1 \vee \bar{y}_2) \oplus y_2 \\ x_1 \cdot (y_1 \vee y_2) \oplus x_2 \cdot \bar{y}_1 \oplus \bar{y}_2 \end{bmatrix} = z. \quad (3)$$

При побудові несиметричних систем потокового СЕТ-шифрування необхідно для кожної прямої СЕТ-операції знайти її обернену.

Для несиметричної 2-операндної СЕТ-операції, що задана моделлю (1), модель оберненої операції можна представити [6]:

$$C'(z, y) = C(C'_1(z), C'_2(z), C'_3(z), C'_4(z)) = x. \quad (4)$$

Побудуємо дискретну модель оберненої удосконаленої СЕТ-операції. Для цього використаємо моделі обернених однооперандних 2Сі-квантових СЕТ-операцій наведених в [18].

Оскільки

$$C'_1(z) = \begin{bmatrix} z_1 \oplus z_2 \oplus 1 \\ z_2 \oplus 1 \end{bmatrix} = x; \quad C'_2(z) = \begin{bmatrix} z_1 \oplus 1 \\ z_1 \oplus z_2 \oplus 1 \end{bmatrix} = x;$$

$$C'_3(z) = \begin{bmatrix} z_2 \oplus 1 \\ z_1 \end{bmatrix} = x; \quad C'_4(z) = \begin{bmatrix} z_2 \\ z_1 \oplus 1 \end{bmatrix} = x,$$

тоді:

$$C'(z, y) = \begin{cases} \begin{bmatrix} z_1 \oplus z_2 \oplus 1 \\ z_2 \oplus 1 \end{bmatrix}, & \text{якщо } y_1 = 0; y_2 = 0; \\ \begin{bmatrix} z_1 \oplus 1 \\ z_1 \oplus z_2 \oplus 1 \end{bmatrix}, & \text{якщо } y_1 = 0; y_2 = 1; \\ \begin{bmatrix} z_2 \oplus 1 \\ z_1 \end{bmatrix}, & \text{якщо } y_1 = 1; y_2 = 0; \\ \begin{bmatrix} z_2 \\ z_1 \oplus 1 \end{bmatrix}, & \text{якщо } y_1 = 1; y_2 = 1. \end{cases} = x. \quad (5)$$

На основі моделі (5) отримаємо:

$$C'(z, y) = \begin{bmatrix} z_1 \cdot \bar{y}_1 \oplus z_2 \cdot (y_1 \vee \bar{y}_2) \oplus (\bar{y}_1 \vee \bar{y}_2) \\ z_1 \cdot (y_1 \vee y_2) \oplus z_2 \cdot \bar{y}_1 \oplus (\bar{y}_1 \vee y_2) \end{bmatrix} = x. \quad (6)$$

Модифікація СЕТ-операції з точністю до перестановки першого операнда реалізується шляхом криптографічного перетворення однооперандною СЕТ-операцією вхідних даних, які надходять в перший операнд.

Модифікуємо моделі прямої СЕТ-операції (3) з точністю до перестановки першого операнда однооперандною операцією $C_4(x)$.

$$C_4(x, y) = C(C_4(x), y) = z_4, \quad (7)$$

де $C_4(x, y)$ – несиметрична 2-операндна СЕТ-операція модифікована з точністю до перестановки першого операнда однооперандною операцією $C_4(x)$; z_4 – результат криптографічного перетворення даною модифікованою СЕТ-операцією.

Модифіковані однооперандною СЕТ-операцією вхідні дані в процесі реалізації 2-операндної СЕТ-операції будуть перетворюватися однією із її однооперандних операцій.

Відповідно до моделі (7), коротку модель СЕТ-операції (1) можна представити:

$$C_4(x, y) = C(C_1(C_4(x)), C_2(C_4(x)), C_3(C_4(x)), C_4(C_4(x))) = z_4. \quad (8)$$

Модель модифікованої 2-операндної СЕТ-операції з точністю до перестановки першого операнда (7) на основі попереднього додаткового перетворення однооперандною СЕТ-операцією $C_4(x)$ можна представити:

$$C_4(x, y) = \begin{cases} C_1(C_4(x)), & \text{якщо } y_1 = 0; y_2 = 0 \\ C_2(C_4(x)), & \text{якщо } y_1 = 0; y_2 = 1 \\ C_3(C_4(x)), & \text{якщо } y_1 = 1; y_2 = 0 \\ C_4(C_4(x)), & \text{якщо } y_1 = 1; y_2 = 1 \end{cases} = z_4. \quad (9)$$

Модифікуємо однооперандні операції $C_1(x) - C_4(x)$ шляхом попереднього додаткового перетворення СЕТ-операцією $C_4(x)$ вхідних Сі-квантів інформації. Для цього в СЕТ-операції $C_1(x) - C_4(x)$ замість першого і другого Сі-квантів першого операнда x_1 і x_2 , підставимо першу і другу елементарні функції СЕТ-операції $C_4(x)$: $f_1(x) = x_2 \oplus 1$, $f_2(x) = x_1$ відповідно:

$$C_1(C_4(x)) = \begin{bmatrix} (x_2 \oplus 1) \oplus (x_1) \\ (x_1) \oplus 1 \end{bmatrix} = \begin{bmatrix} x_1 \oplus x_2 \oplus 1 \\ x_1 \oplus 1 \end{bmatrix} = C_5(x) = z_4. \quad (10)$$

$$C_2(C_4(x)) = \begin{bmatrix} (x_2 \oplus 1) \oplus 1 \\ (x_2 \oplus 1) \oplus (x_1) \end{bmatrix} = \begin{bmatrix} x_2 \\ x_1 \oplus x_2 \oplus 1 \end{bmatrix} = C_6(x) = z_4. \quad (11)$$

$$C_3(C_4(x)) = \begin{bmatrix} (x_1) \\ (x_2 \oplus 1) \oplus 1 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = C_7(x) = z_4. \quad (12)$$

$$C_4(C_4(x)) = \begin{bmatrix} (x_1) \oplus 1 \\ (x_2 \oplus 1) \end{bmatrix} = \begin{bmatrix} x_1 \oplus 1 \\ x_2 \oplus 1 \end{bmatrix} = C_8(x) = z_4. \quad (13)$$

Підставивши результати модифікації однооперандних СЕТ-операцій (10) – (11) в модель модифікованої 2-операційної СЕТ-операції (9) отримаємо:

$$C_4(x, y) = \begin{cases} \begin{bmatrix} x_1 \oplus x_2 \oplus 1 \\ x_1 \oplus 1 \end{bmatrix}, & \text{якщо } y_1 = 0; y_2 = 0; \\ \begin{bmatrix} x_2 \\ x_1 \oplus x_2 \oplus 1 \end{bmatrix}, & \text{якщо } y_1 = 0; y_2 = 1; \\ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, & \text{якщо } y_1 = 1; y_2 = 0; \\ \begin{bmatrix} x_1 \oplus 1 \\ x_2 \oplus 1 \end{bmatrix}, & \text{якщо } y_1 = 1; y_2 = 1. \end{cases} = z_4. \quad (14)$$

На основі моделі (14) отримаємо удосконалену модифіковану дискретну модель СЕТ-операції:

$$C_4(x, y) = \begin{bmatrix} x_1 \cdot (y_1 \vee \bar{y}_2) \oplus x_2 \cdot \bar{y}_1 \oplus (y_1 \oplus y_2 \oplus 1) \\ x_1 \cdot \bar{y}_1 \oplus x_2 \cdot (y_1 \vee y_2) \oplus (\bar{y}_1 \vee y_2) \end{bmatrix} = z_4. \quad (15)$$

Побудуємо обернену 2-операційну СЕТ-операцію до модифікованої несиметричної 2-операційної СЕТ-операції заданої моделлю (8).

Представимо модель як удосконалену модифіковану дискретну модель СЕТ-операції (15) її короткою моделлю. Підставивши в (14) позначення прямих однооперандних СЕТ-операцій, отримаємо:

$$C_4(x, y) = \begin{cases} C_5(x), & \text{якщо } y_1 = 0; y_2 = 0; \\ C_6(x), & \text{якщо } y_1 = 0; y_2 = 1; \\ C_7(x), & \text{якщо } y_1 = 1; y_2 = 0; \\ C_8(x), & \text{якщо } y_1 = 1; y_2 = 1. \end{cases} = z_4. \quad (16)$$

Відповідно до (16) коротку модель прямої модифікованої 2-операційної СЕТ-операції можна представити:

$$C_4(x, y) = C(C_5(x), C_6(x), C_7(x), C_8(x)) = z_4. \quad (17)$$

Відповідно до (17) коротку модель оберненої модифікованої 2-операційної СЕТ-операції можна представити:

$$C'_4(z_4, y) = C(C'_5(z_4), C'_6(z_4), C'_7(z_4), C'_8(z_4)) = x. \quad (18)$$

Побудуємо дискретну модель оберненої удосконаленої модифікованої СЕТ-операції. Відповідно до [18] моделі обернених однооперандних СЕТ-операцій до однооперандних СЕТ-операцій $C_5(x) - C_8(x)$ можна представити.

$$C'_5(z_4) = \begin{bmatrix} z_{4,2} \oplus 1 \\ z_{4,1} \oplus z_{4,2} \end{bmatrix} = x; \quad (19)$$

$$C'_6(z_4) = \begin{bmatrix} z_{4,1} \oplus z_{4,2} \oplus 1 \\ z_{4,1} \end{bmatrix} = x; \quad (20)$$

$$C'_7(z_4) = \begin{bmatrix} z_{4,1} \\ z_{4,2} \end{bmatrix} = x; \quad (21)$$

$$C'_8(z_4) = \begin{bmatrix} z_{4,1} \oplus 1 \\ z_{4,2} \oplus 1 \end{bmatrix} = x, \quad (22)$$

де $z_{4,1}$, $z_{4,2}$ – перший і другий Сі-кванти результату криптографічного перетворення інформації модифікованою 2-операційною СЕТ-операцією (15).

Підставивши моделі (21) – (22) в коротку модель оберненої СЕТ-операції (18) отримаємо:

$$C'_4(z_4, y) = \begin{cases} \begin{bmatrix} z_{4,2} \oplus 1 \\ z_{4,1} \oplus z_{4,2} \end{bmatrix}, & \text{якщо } y_1 = 0; y_2 = 0; \\ \begin{bmatrix} z_{4,1} \oplus z_{4,2} \oplus 1 \\ z_{4,1} \end{bmatrix}, & \text{якщо } y_1 = 0; y_2 = 1; \\ \begin{bmatrix} z_{4,1} \\ z_{4,2} \end{bmatrix}, & \text{якщо } y_1 = 1; y_2 = 0; \\ \begin{bmatrix} z_{4,1} \oplus 1 \\ z_{4,2} \oplus 1 \end{bmatrix}, & \text{якщо } y_1 = 1; y_2 = 1. \end{cases} = x. \quad (23)$$

На основі моделі (23) отримаємо:

$$C'_4(z_4, y) = \begin{bmatrix} z_{4,1} \cdot (y_1 \vee y_2) \oplus z_{4,2} \cdot \bar{y}_1 \oplus (\bar{y}_1 \vee y_2) \\ z_{4,1} \cdot \bar{y}_1 \oplus z_{4,2} \cdot (y_1 \vee \bar{y}_2) \oplus (y_1 \cdot y_2) \end{bmatrix} = x. \quad (24)$$

Побудована модель оберненої 2-розрядної СЕТ-операції (24) дозволяє розшифровувати інформацію зашифровану прямою СЕТ-операцією (15). Проте $C_4(x, y) = z_4$ лише одна із операцій, яка належить групі операцій з точністю до перестановки першого операнда. Групу 2-операційних СЕТ-операцій з точністю до перестановки першого операнда можна представити як

$$C_i(x, y) = C(C_i(x), y) = z_i, \quad (25)$$

де i – індекс (номер) однооперандної СЕТ-операції, на основі якої була реалізована модифікація 2-операційної СЕТ-операції.

Побудова моделі (24) представляє собою побудову лише однієї оберненої СЕТ-операції з групи

операцій (25). Тому на практиці використовувати даний підхід для реалізації оберненого перетворення складно і нераціонально.

Будь-яку операцію з групи операцій (25) можна представити як функцію від 2-операндної і 1-операндної СЕТ-операції:

$$C_i(x, y) = f(C(x, y), C_i(x)). \quad (26)$$

Обернену операцію для будь-якої 2-операндної СЕТ-операції з групи СЕТ-операцій (26) можна представити як

$$C'_i(x, y) = f(C'(x, y), C'_i(x)). \quad (27)$$

На основі функції (27) можна представити три варіанти синтезу оберненої СЕТ-операції:

- з точністю до перестановки результату перетворення

$$C'_i(x, y) = C'_i(C'(x, y)); \quad (28)$$

- з точністю до перестановки першого операнда

$$C'_i(x, y) = C'_i(C'_i(x), y); \quad (29)$$

- з точністю до перестановки другого операнда

$$C'_i(x, y) = C'_i(x, C'_i(y)). \quad (30)$$

Розглянемо можливість побудови оберненого СЕТ-операції (24) на основі моделі (28). Для цього підставимо в модель оберненої однооперандної СЕТ-операції $C'_4(z)$ обернену 2-операндну операцію (6):

$$\begin{aligned} C'_4(z_4, y) &= C'_4(C'(z_4, y)) = \\ &= C(C'_4(C'_1(z_4)), C'_4(C'_2(z_4)), C'_4(C'_3(z_4)), C'_4(C'_4(z_4))). \end{aligned} \quad (31)$$

На основі кортежної моделі оберненої СЕТ-операції (31) отримаємо:

$$C'_4(z_4, y) = \begin{cases} C'_4(C'_1(z_4)), & \text{якщо } y_1 = 0; y_2 = 0; \\ C'_4(C'_2(z_4)), & \text{якщо } y_1 = 0; y_2 = 1; \\ C'_4(C'_3(z_4)), & \text{якщо } y_1 = 1; y_2 = 0; \\ C'_4(C'_4(z_4)), & \text{якщо } y_1 = 1; y_2 = 1 \end{cases} = x. \quad (32)$$

Визначимо однооперандні обернені СЕТ-операції, з яких побудовано кортеж оберненої 2-операндної СЕТ-операції (31):

$$C'_4(C'_1(z_4)) = \begin{bmatrix} (z_{4,2} \oplus 1) \\ (z_{4,1} \oplus z_{4,2} \oplus 1) \oplus 1 \end{bmatrix} = \begin{bmatrix} z_{4,2} \oplus 1 \\ z_{4,1} \oplus z_{4,2} \end{bmatrix} = x; \quad (33)$$

$$C'_4(C'_2(z_4)) = \begin{bmatrix} (z_{4,1} \oplus z_{4,2} \oplus 1) \\ (z_{4,1} \oplus 1) \oplus 1 \end{bmatrix} = \begin{bmatrix} z_{4,1} \oplus z_{4,2} \oplus 1 \\ z_{4,1} \end{bmatrix} = x; \quad (34)$$

$$C'_4(C'_3(z_4)) = \begin{bmatrix} (z_{4,1}) \\ (z_{4,2} \oplus 1) \oplus 1 \end{bmatrix} = \begin{bmatrix} z_{4,1} \\ z_{4,2} \end{bmatrix} = x; \quad (35)$$

$$C'_4(C'_4(z_4)) = \begin{bmatrix} (z_{4,1} \oplus 1) \\ (z_{4,2}) \oplus 1 \end{bmatrix} = \begin{bmatrix} z_{4,1} \oplus 1 \\ z_{4,2} \oplus 1 \end{bmatrix} = x. \quad (36)$$

Підставивши в модель оберненої 2-операндної СЕТ-операції (32) моделі обернених однооперандних СЕТ-операцій (33) – (36), отримаємо:

$$C'_4(z_4, y) = \begin{cases} \begin{bmatrix} z_{4,2} \oplus 1 \\ z_{4,1} \oplus z_{4,2} \end{bmatrix}, & \text{якщо } y_1 = 0; y_2 = 0; \\ \begin{bmatrix} z_{4,1} \oplus z_{4,2} \oplus 1 \\ z_{4,1} \end{bmatrix}, & \text{якщо } y_1 = 0; y_2 = 1; \\ \begin{bmatrix} z_{4,1} \\ z_{4,2} \end{bmatrix}, & \text{якщо } y_1 = 1; y_2 = 0; \\ \begin{bmatrix} z_{4,1} \oplus 1 \\ z_{4,2} \oplus 1 \end{bmatrix}, & \text{якщо } y_1 = 1; y_2 = 1. \end{cases} = x. \quad (37)$$

Оскільки побудована модель оберненої 2-операндної СЕТ-операції (37) співпала з моделлю оберненої 2-операндної СЕТ-операції (23), то удосконалена модель оберненої 2-операндної СЕТ-операції буде описуватися моделлю (24).

Результат моделювання дозволяє встановити взаємозв'язки між модифікованими прямими і оберненими несиметричними 2-операндними СЕТ-операціями.

Нехай модифікація прямої несиметричної 2-операндної СЕТ-операції була реалізована з точністю до перестановки першого операнда однооперандною СЕТ-операцією на основі моделі (25). Отримати обернену до неї 2-операндну СЕТ-операцію можна шляхом модифікації оберненої 2-операндної СЕТ-операції з точністю до перестановки результату перетворення оберненою однооперандною СЕТ-операцією на основі моделі:

$$C'_i(z_i, y) = C'_i(C'(z_i, y)) = x. \quad (38)$$

Дана модель забезпечить обернене перетворення інформації.

Коректність наведених результатів моделювання підтверджена результатами обчислювального експерименту.

Обговорення результатів моделювання груп несиметричних СЕТ-операцій з точністю до перестановки першого операнда.

У процесі модифікації несиметричної 2-операндної СЕТ-операції з точністю до перестановки першого операнда змінюється математична модель СЕТ-операції. Про це свідчать побудовані моделі СЕТ-операції до її модифікації (3) і після модифікації (15). Необхідно відмітити, що в результаті модифікації СЕТ-операції буде змінюватися як сама модель операції, так і результат її реалізації.

Модифікації СЕТ-операцій з точністю до перестановки другого операнда [16, 18] забезпечує зміну результату криптографічного перетворення за рахунок зміни послідовності однооперандних СЕТ-операцій в кортежі 2-операндної операції. Модифікація СЕТ-операцій з точністю до перестановки першого операнда приводить до модифікації не послідовності однооперандних операцій в кортежі, а модифікації самих однооперандних СЕТ-операцій. Про це свідчать кортежні моделі операції до її модифікації (1) і після модифікації (17).

Збільшення кількості однооперандних СЕТ-операцій при використанні групи модифікованих несиметричних СЕТ-операцій з точністю до перестановки першого операнда забезпечує збільшення кількості таблиць підстановки, які реалізуються в процесі шифрування. Збільшення кількості таблиць підстановки, які реалізуються, забезпечує збільшення криптостійкості алгоритму шифрування.

Якщо при шифруванні було використано модифікацію СЕТ-операції з точністю до перестановки першого операнда відповідно до моделі (25), то при розшифруванні необхідно використовували модифікацію оберненої СЕТ-операції з точністю до перестановки результату перетворення відповідно до моделі (38).

На основі моделей модифікації СЕТ-операцій реалізуються генератори псевдовипадкових послідовностей прямих і обернених модифікованих 2-операндних СЕТ-операцій.

Коректність наведених теоретичних результатів підтверджена результатами обчислювального експерименту.

Подальші дослідження будуть направлені на дослідження особливостей побудови високопродуктивних малоресурсних криптографічних систем на основі груп несиметричних СЕТ-операцій синтезованих з точністю до перестановки першого операнда.

Висновки:

1. Групи несиметричних 2-операндних СЕТ-операцій можуть бути синтезовані з точністю до перестановки першого операнда. Модифікацію несиметричної 2-операндної СЕТ-операції доцільно реалізувати криптографічним перетворенням першого операнду однооперандною СЕТ-операцією.

2. Для групи несиметричних 2-операндних СЕТ-операцій з точністю до перестановки першого операнда група обернених СЕТ-операцій будується з точністю до перестановки результатів перетворення. Модифікацію оберненої несиметричної 2-операндної СЕТ-операції необхідно реалізувати криптографічним перетворенням попереднього результату розшифрування оберненою однооперандною СЕТ-операцією.

3. Модифікація несиметричної 2-операндної СЕТ-операції з точністю до перестановки першого операнда реалізується шляхом модифікації однооперандних СЕТ-операцій, з яких вона побудована.

4. Модифікація однооперандних СЕТ-операцій в процесі шифрування забезпечує збільшення кількості таблиць підстановок, які реалізує поточний шифр, що ускладнює криптоаналіз і приводить до підвищення криптостійкості.

Список літератури:

1. Kumar C., Prajapati S., Verma R. A Survey of Various Lightweight Cryptography Block ciphers for IoT devices. *CCET 2022: IEEE International Conference on Current Development in Engineering and Technology*. (Bhopal, 23–24 December 2022). Bhopal India, 2022. P. 1–6. DOI: 10.1109/CCET56606.2022.10080556.
2. Khaled Salah Mohamed. New Frontiers in Cryptography. *Quantum, Blockchain, Lightweight, Chaotic and DNA*. Springer Cham, 2020. 104 p. DOI: 10.1007/978-3-030-58996-7.
3. Khudoykulov Z. A Comparison of Lightweight Cryptographic Algorithms. *Lecture Notes in Networks and Systems*. 2024. Vol. 912. P. 295–304. DOI: 10.1007/978-3-031-53488-1_36.
4. Lohachab A., Lohachab A., Jangra A. A comprehensive survey of prominent cryptographic aspects for securing communication in post-quantum IoT networks. *Internet of Things*. 2020. Vol. 9. P. 100174. DOI: 10.1016/j.iot.2020.100174.
5. Haythem P., Zorkta Y., Allawi D., Al-Nakkar M. Improved lightweight encryption algorithm (ILEA). *INCET 2020: International Conference for Emerging Technology*. (Belgaum, 05–07 June 2020). Belgaum, India, 2020. P. 1–4. DOI: 10.1109/INCET49848.2020.9154170.
6. Рудницький В. М., Лада Н. В., Кучук Г. А., Підласий Д. А. Архітектура СЕТ-операцій і технології потокового шифрування. *Architecture of CET-operations and stream encryption technologies: монографія*. Черкаси: видавець Пономаренко Р.В. 2024. 374 с. ISBN 978-966-2554-81-6
7. Massey J. L. An introduction to contemporary cryptology. *Proceedings of the IEEE*, 1988. Vol. 76, № 5. P. 533–549. DOI: 10.1109/5.4440.

8. Бабенко В. Г., Лада Н. В. Синтез і аналіз операцій криптографічного додавання за модулем два. *Системи обробки інформації*. 2014. Вип. 2 (118). С. 116–118. URL: http://nbuv.gov.ua/UJRN/soi_2014_2_26.
9. Бабенко В. Г., Лада Н. В. Аналіз результатів виконання модифікованих операцій додавання за модулем два з точністю до перестановки. *The scientific potential of the present: proceedings of the Internat. sci. conf.*, (St. Andrews, Scotland, UK, December, 1, 2016). Vinnytsia: PE Rogalska I. O., 2016. С. 108–111.
10. Лада Н. В., Козловська С. Г., Рудницька Ю. В. Дослідження і синтез групи симетричних модифікованих операцій додавання за модулем чотири. *Центральноукраїнський науковий вісник*. 2019. Вип. 2 (33). С. 181–189. DOI: 10.32515/2664-262X.2019.2(33).181-189.
11. Лада Н. В., Рудницький С. В., Зажома В. М., Рудницька Ю. В. Дослідження і синтез групи симетричних модифікованих операцій правостороннього додавання за модулем чотири. *Системи управління, навігації та зв'язку*. 2020. № 1(59). С. 93–96. DOI: 10.26906/SUNZ.2020.1.093.
12. Rudnytskyi V., Babenko V., Lada N., Tarasenko Ya., Rudnytska Yu. Constructing Symmetric Operations of Cryptographic Information Encoding. *CEUR Workshop Proceedings*. 2021. Vol. 3187. P. 182–194. DOI: 10.5281/zenodo.10953357.
13. Breus R. Synthesising the two-bit two-operand operations of strict stable cryptographic coding by the second operand's conversion. *Control, Navigation and Communication Systems*. 2019. Vol. 5 (57). P. 29–32. DOI: 10.26906/SUNZ.2019.5.029.
14. Лада Н. В., Бреус Р. В., Лада С. В. Генерація моделей прямих і обернених двохрандрних двооперандних операцій строгого стійкого криптографічного кодування. *Science and Education a New Dimension Natural and Technical Science*. 2020. VIII (29), № 238. P. 27–29. DOI: 10.31174/SEND-NT2020-238VIII29-05.
15. Лада Н. В., Головняк Д. В., Сапожников С. К. Механізми практичної реалізації несиметричних двооперандних CET-операцій. *Проблеми інформатизації: тези доповідей одинадцятої міжнародної науково-технічної конференції* (Черкаси, 16–17 листопада 2023 року). Черкаси: ЧДТУ, 2023. Т. 2. С. 34.
16. Бабенко В., Мельник О., Мельник Р. Класифікація трихрандрних елементарних функцій для криптографічного перетворення інформації. *Безпека інформації*. 2013. Т. 19. № 1. С. 56–59. DOI: 10.18372/2225-5036.19.4705.
17. Бабенко В. Г., Мельник Р. П., Рудницький С. В. Дослідження способів запису трьохрандрних криптографічних операцій. *Системи управління, навігації та зв'язку*. 2012. Т. 2, № 1 (21). С. 170–173.
18. Rudnytskyi V., Lada N., Babenko V., Kuchuk H., Pidlasyi D., Kamak D., Ivashchenko Ye. Modeling of groups of dual-cycle non-commutative two-operand CET-operations. *Journal of Xidian University*. 2024. Vol. 18, № 10. P. 916–958. DOI: 10.37896/jxu18.10/069.
19. Rudnytsky V.M., Lada N.V., Fedotova-Piven I.M., Pustovit M.O., Nesterenko O.B. Construction of two-digit two operand operations of strict stable cryptographic coding. *Control, navigation and communication systems*. 2018. Vol. 6. № 52. P. 113–115. DOI: 10.26906/SUNZ.2018.6.113
20. Tazetdinov V., Sysoienko S., Tazetdinov O., Al-Azzeh J., Mesleh A. Neural network system for selection of table tennis equipment with elements of crypto protection. *Lecture Notes on Data Engineering and Communications Technologies*. 2024. Vol. 221. P. 126–140. DOI: 10.1007/978-3-031-71801-4_10.

Rudnytskyi V.M., Babenko V.H., Rudnytskyi S.V., Korotkyi T.K., Kovtiukh V.A. PARTICULAR FEATURES OF GROUPS OF ASYMMETRIC SET OPERATIONS SYNTHESIZED WITH AN ACCURACY OF FIRST OPERAND PERMUTATION

The article is devoted to a study that solves the problem of investigating the possibility of using the method of generating groups of asymmetric 2-operand CET operations with an accuracy of permutation of the first operand to build low-resource asymmetric stream encryption systems with increased cryptographic strength. The object of research: processes of mutual transformation of models of 2-operand CET operations in the process of their modification with an accuracy of permutation. To solve the problem, using the example of an asymmetric 2-operand CET operation, the processes of its modification with an accuracy of up to the permutation of the first operand by a single-operand CET operation are considered. It is advisable to modify the asymmetric 2-operand CET operation by cryptographic transforming the first operand with a randomly selected single-operand CET operation. The results of modeling the processes of reverse information transformation have shown that when modifying a direct asymmetric 2-operand CET operation with an accuracy of permutation of the first operand, the modification of the reverse CET operation should be realized by the reverse single-operand CET operation with an accuracy of permutation of the transformation result. It is established that the modification of an asymmetric 2-operand CET operation with an accuracy of permutation of the first operand is realized by modifying the single-operand CET operations from which it is constructed. According to the results of the analysis, it was found that the modification of single-operand CET operations

in the encryption process increases the number of substitution tables that implement the stream cipher. The increasing number of substitution tables during the successive modification of the CET operations leads to the complication of cryptanalysis and to the increase of cryptographic strength of the cryptographic algorithm. Scope of application: mobile and stationary systems of low-resource cryptographic protection of confidential information.

Key words: *low-resource cryptography, CET encryption, 2-operand CET operations, asymmetric cryptographic transformations, stream encryption.*

Дата надходження статті: 15.07.2025

Дата прийняття статті: 21.07.2025

Опубліковано: 27.10.2025

Сичков В.С.

Національний університет «Одеська політехніка»

ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ ДЛЯ АВТОМАТИЗОВАНОГО ПРОФЕСІЙНОГО НАСТРОЮВАННЯ АКУСТИЧНИХ СТРУННИХ МУЗИЧНИХ ІНСТРУМЕНТІВ: ЕКСПЕРТНИЙ ПОГЛЯД

Статтю присвячено проблемі автоматизації професійного настроювання акустичних струнних музичних інструментів, зокрема фортепіано, з фокусом на створенні якісного програмного забезпечення, здатного відповідати високим вимогам майстрів-настроювачів. Підкреслено міждисциплінарний характер завдання, яке поєднує фізичне і математичне моделювання, психоакустичні аспекти, оптимізаційні механізми, цифровий аналіз і оброблення сигналів та програмну інженерію.

У статті розглянуто обмеження сучасних електронних тюнерів – як апаратних, так і програмних – та доведено, що вони наразі не спроможні повноцінно замінити слухове настроювання, яке базується на досвіді професійних техніків. Розкрито сучасні наукові підходи до автоматизації процесу настроювання, серед яких – застосування навчання з підкріпленням, мінімізації глобальної спектральної ентропії, параметричного моделювання негармонічності струн та аналізу суб'єктивного сприйняття звуку.

Здійснено аналіз публікацій, технічних блогів, професійних форумів та позицій провідних настроювачів, зокрема членів Гільдії фортепіанних техніків (PTG), на підставі якого виокремлено типові проблеми сучасних програм: накопичення похибок у розрахунках, обмежена точність настроювання унісонів, недостатня стабільність настроювання, відсутність належного зворотного зв'язку та погана адаптація до реальних акустичних умов.

Визначено, що створення якісного програмного забезпечення можливе лише за умови активної участі фахівців із настроювання на критичних етапах життєвого циклу його розроблення – формулюванні вимог, валідації результатів і подальшому підтримуванні системи. Запропоновано низку заходів для поліпшення якості таких рішень. З'ясовано, що лише комплексний підхід, який поєднує технологічні досягнення з емпіричними знаннями й навичками майстрів-настроювачів, здатен забезпечити створення тюнерів, що відповідатимуть високим стандартам якості настроювання і викликатимуть довіру професійних користувачів.

Ключові слова: програмне забезпечення для настроювання, електронні тюнери, слухове настроювання, експертні знання, негармонічність, фортепіано.

Постановка проблеми. Автоматизація професійного настроювання акустичних струнних музичних інструментів, зокрема розроблення відповідного високоякісного програмного забезпечення, залишається складним міждисциплінарним науково-прикладним завданням. Ця сфера охоплює психоакустику, моделювання, оптимізацію, цифрове оброблення сигналів та інженерію програмного забезпечення.

Сучасні тюнери – як окремі електронні прилади, так і спеціалізоване програмне забезпечення для мобільних пристроїв – позиціонуються як професійні рішення для настроювання ударно-клавішних інструментів (піаніно і роялів). Вони істотно відрізняються від простих хроматичних тюнерів, що реагують лише на основну частоту

акустичного сигналу, оскільки враховують багатший спектральний склад і специфіку звучання фортепіано.

Попри це, значна частина досвідчених фахівців – фортепіанних техніків (майстрів-настроювачів) – вважає, що сучасні тюнери не здатні забезпечити такий самий рівень якості настроювання, як настроювання на слух (слухове настроювання). У професійній практиці тюнери часто використовують як допоміжний інструмент, зокрема для прискорення окремих етапів роботи, однак кінцеву валідацію настроювання все одно здійснюють на слух, тому оптимальним вважають поєднання обох підходів. Проте якщо оцінювати настроювання, виконане виключно за тюнером, «з нуля» (без додаткових налаштувань приладу або

програми), то його здатність вибудовувати якісну температуру, яка б відповідала професійним стандартам і була б схвалена більшістю настроювачів, наразі залишається обмеженою.

Водночас існує дещо інша точка зору: деякі експерти вважають, що тюнери вже досягли достатньої точності та надійності, щоб забезпечувати з ним практично однаково високий результат. У цьому контексті критика з боку традиційних майстрів часто розглядається як консерватизм і є радше наслідком професійних упереджень. Також вона часто зумовлена недостатнім усвідомленням того, що сучасні тюнери вже значно перевершують перші прилади, здатні враховувати негармонічність та інші складні слухові критерії.

Ключова складність автоматизації настроювання полягає в тому, що для фортепіано неможливо задати єдину «ідеальну» температуру через негармонічність струн та кожен інструмент відрізняється один від одного. Алгоритми програмного забезпечення для електронних настроювальних пристроїв і мобільних застосунків мають враховувати не лише прості тональні співвідношення, а й складні психоакустичні ефекти, а також конструктивні й фізико-акустичні особливості інструментів. Це значно ускладнює створення програм, які могли б повністю відповідати очікуванням і критеріям майстрів-настроювачів.

Аналіз останніх досліджень і публікацій. Дослідження питань автоматизації настроювання музичних інструментів підкреслюють як актуальність, так і складність цієї проблеми. Основну увагу в наукових працях приділено моделюванню акустичних властивостей інструментів, зокрема негармонічності струн, та вивченню критеріїв людського слухового сприйняття.

Авторами [1] запропоновано автоматичну систему настроювання фортепіано під назвою "PitchImpact", яка використовує навчання з підкріпленням для керування спеціальним регульованим настроювальним ударним ключем. Система використовує програмне забезпечення аналізу акустичного спектру для визначення похибки висоти та відповідної сили удару і напрямку, необхідних для настроювання кожної струни інструмента. Потім агент визначає відповідне налаштування удару по ключу, щоб відрегулювати вірбелі та настроїти струну. Експерименти з системою на тестовому «гексакорді» показують, що вона може настроїти струни порівняно з навченим настроювачем-аматором за невелику кількість ударів. Система має потенціал, оскільки дозволяє власникам фортепіано самостійно настроювати інструменти

з мінімальним обладнанням, таким чином роблячи процес набагато зручнішим і доступнішим.

У статті [2] описано метод настроювання фортепіано, заснований на мінімізації глобальної спектральної ентропії. У ній зазначено, що сучасні музичні строї базуються на рівномірній температурі, але настроювання точно за нею може призвести до того, що інструменти звучать недобре через негармонічність обертонових спектрів. Професійні настроювачі компенсують це шляхом розтягнення певних інтервалів, ґрунтуючись на сприйнятті гармонік, що збігаються між тонами, як більш співзвучних. Початкові тести, що відтворюють криві настроювання реальних фортепіано, свідчать про те, що метод фіксує нерегулярні коливання висоти нот, подібні до високоякісного настроювання. Хоча він простий у реалізації, проте має суттєві обмеження та відкриті питання, які потребують подальших досліджень.

У роботі [3] запропоновано параметричну модель, що дозволяє спільно охарактеризувати негармонічність та настроювання фортепіано на всьому діапазоні нот. Коефіцієнт негармонічності моделюється з використанням двох асимптот, що відповідають басовим і дискантним штегам. Модель настроювання передбачає розтягнення октав, спричинене негармонічністю, а також вибір типу октави настроювачем. Представлено алгоритми для оцінки параметрів моделі за окремими нотами або акордами за умови, що ноти відомі. Модель вдало вловлює основні тенденції негармонічності та відхилень настроювання від рівної температури для різних типів фортепіано. Потенційними застосуваннями моделі є оцінка стану фортепіано, надання інформації для вибору кривих настроювання та параметризація настроювання синтезаторів фортепіано. Загалом робота демонструє змогу глобальної параметризації негармонічності та настроювання з використанням невеликої кількості фізично значущих параметрів.

У статті [4] представлено дослідження щодо розроблення автоматичного методу настроювання високих тонів фортепіано. Автори перевірили чотири простих правила настроювання октавних інтервалів шляхом зіставлення різних часткових частот нижнього і верхнього регістрів. Вони провели тест на прослуховування, щоб визначити, яке правило забезпечує найкраще настроювання, за оцінками слухачів. Тест показав, що найприємніше звучання високих октав було досягнуто тоді, коли перша гармоніка верхнього тону узгоджувалась із другою гармонікою

нижнього. Автори також запропонували новий метод оцінки якості настроювання шляхом аналізу частот биття між обертонами. Загалом дослідження спрямоване на розроблення повної системи автоматичного настроювання фортепіано, і це воно було зосереджено на пошуку оптимального правила настроювання високих тонів на основі людського сприйняття.

Публікації щодо розроблення та забезпечення якості програмного забезпечення для автоматизованого професійного настроювання акустичних струнних музичних інструментів практично відсутні, що підкреслює потребу в систематичних дослідженнях та створенні спеціалізованих рішень у цій галузі.

Постановка завдання. Метою статті є визначення ключових проблем програмного забезпечення для настроювання акустичних струнних ударно-клавішних музичних інструментів із погляду досвідчених користувачів на шляху до з'ясування причин, через які сучасні тюнери ще не можуть стати повноцінною альтернативою слуховому настроюванню. Для досягнення поставленої мети необхідно проаналізувати фахові статті, форуми, блоги і онлайн-курси та виокремити найвагоміші експертні судження щодо якості тюнерів і досвіду їх використання, а також на цій основі визначити необхідні заходи для побудови якісного програмного забезпечення для настроювання музичних інструментів.

Виклад основного матеріалу. Як зазначають автори [5], «сучасні системи вимірювання набагато точніші, ніж людський слух, [проте] слухове настроювання все ще вважається найкращим та найнадійнішим методом настроювання фортепіано». У роботі [4] підкреслено, що, незалежно від застосованого методу, саме люди зрештою визначають, яке настроювання звучить найкраще. Експерти, зокрема сертифіковані фортепіанні техніки (RPT, англ. *registered piano technician*), також наголошують, що електронний або програмний тюнер не може просто замінити слухові навички. Наприклад, сертифікація RPT від PTG (*Piano Technicians Guild*, або Гільдії фортепіанних техніків, – провідного світового джерела експертних знань у сфері обслуговування та технологій фортепіано) вимагає, щоб кандидати «продемонстрували здатність настроювати на слух, без допомоги електроніки», що демонструє слухове настроювання як незамінну складову професійної практики, попри технічний прогрес [6].

Асоційований член тієї ж Гільдії фортепіанних техніків Річард В. Буші у своєму блозі приводить

цікаву аналогію: «Електронні настроювальні пристрої в руках досвідченого настроювача можуть дати виняткові результати. Вони своєрідно нагадують високоякісні діагностичні інструменти, які сьогодні використовують автомеханіки. Особисто я б трохи побоювався залишати свій автомобіль у руках механіка, який використовував лише інструменти, доступні на той час, коли були винайдені автомобілі... або використовував лише стетоскоп для діагностики проблем із двигуном. Я хочу, щоб мій механік мав у своєму розпорядженні всі найновіші інструменти. Я також хотів би, щоб він знав, як правильно ними користуватися, але йому також краще знати більше, ніж просто основи ремонту автомобілів... не лише як читати те, що йому каже машина. Машини, коли вони правильно запрограмовані та налаштовані для поточного завдання, зазвичай дають правильні результати. Однак машини іноді можуть помилятися... і людина, яка інтерпретує роботу машини, **ПОВИННА** знати різницю» [7].

Також він підкреслює невисоку якість суто «машинного» настроювання: «... навіть якщо настроювач більш вправний у настроюванні на слух, деякі настроювачі, що настроюють виключно за допомогою тюнера, все ще мають «схильність» лінуватися та занадто покладатися на свій апарат, пропускаючи необхідні перевірки слуху. Якщо настроювач, наприклад, просто «зупиняє світло» на своєму апараті та не грає терції, кварта, квінти, сексти, октави тощо та не перевіряє їх (щоб довести, що ноти знаходяться на правильному місці), то він, найімовірніше, отримає в найліпшому разі посереднє настроювання» [7].

Отже, досвідчений настроювач має вміти ефективно користуватися будь-якими доступними інструментами – чи то камертоном, чи сучасними тюнерами – для досягнення найкращого результату. Якість настроювання залежить не стільки від самого інструмента, скільки від умінь майстра: з будь-яким засобом можна отримати як високоякісний, так і незадовільний результат. Загалом ту ж думку висловлюють і інші спеціалісти [8, 9]. Водночас удосконалення засобів настроювання сприяє підвищенню стабільності результатів, пришвидшенню процесу, зменшенню навантаження на слух та підвищенню доступності професійного настроювання для ширшого кола користувачів.

Детальний аналіз недоліків програмного забезпечення для тюнерів надає Марк Церізано (*Mark Cerisano*) – інженер, сертифікований фортепіанний технік і засновник Міжнародної школи

фортепіанних техніків (англ. *International Piano Technicians School, IPTS*). За його оцінкою, основні проблеми тюнерів полягають у використанні застарілих методів настроювання, а також у відсутності повноцінної підтримки процесу професійного слухового настроювання.

Зокрема, для досягнення бажаних поступових частот биття обертонів в послідовних однойменних консонансах – тобто під час вибудовування рівномірно темперованого строю – тюнерні алгоритми вдаються до складної багаторівневої апроксимації. Вона охоплює коливну частоту струни, коливні частоти її обертонів, коливну негармонічність цієї струни, коливної негармонічності всіх інших струн, коливної кривої розтягування інтервалів та, нарешті, основної частоти у зворотному напрямку вже від цієї апроксимованої кривої. Такий підхід спричиняє накопичення похибок, що, як наслідок, призводить до неточного строю, який не відповідає високим професійним вимогам настроювачів.

Натомість досвідчені настроювачі настроюють інструмент, слухаючи інтервали, тобто співвідношення частот двох тонів, і одразу визначають поступові частоти биття обертонів, враховуючи всі коливні процеси. Такий слуховий метод дозволяє значно зменшити загальну похибку, а вплив негармонічності зводиться до мінімуму природним чином.

Окрім того, сучасне програмне забезпечення зазвичай не забезпечує стабільного строю та демонструє обмежену здатність до точного настроювання унісонів [10, 11].

Додатково експерти наголошують на проблемах тюнерів із визначенням частот у низькому (басовому) і високому (дискантовому) діапазонах та їх чутливості до навколишнього шуму, реверберації (наприклад, концертного залу) та резонансів у самому музичному інструменті, що знижує точність аналізу акустичного сигналу. Також у численних обговореннях на форумах користувачі звертають увагу на недопрацювання окремих функцій наявного програмного забезпечення, зокрема, помилки у визначенні нот та відсутність оглядового інтерфейсу для візуалізації вже настроєних тонів. До потенційно корисних удосконалень можна віднести впровадження режиму аудіоізоляції для мінімізації впливу сторонніх зву-

ків, а також реалізація енергоефективних рішень, які б забезпечили більший час автономної роботи тюнера під час тривалих сесій настроювання [12-19].

Висновки. Автоматизація професійного настроювання акустичних струнних музичних інструментів залишається складним і ще не вирішеним завданням. Попри значний прогрес у розвитку електронних і програмних тюнерів, вони досі не здатні повною мірою забезпечити якість настроювання, яку демонструють досвідчені фахівці-настроювачі під час роботи на слух.

У цьому контексті використання експертних знань для керівництва розробленням програмного забезпечення є критично важливим. Досвідчені настроювачі знають, як поводитися акустичні фортепіано та які нюанси під час настроювання мають значення. Ігнорування знань майстрів у цій предметній області ризикує створенням неефективних застосунків, що не задовольняють професійні вимоги і не викликають довіри в цільовій аудиторії.

Для створення по-справжньому якісного програмного забезпечення для настроювання розробникам необхідно тісно співпрацювати з фахівцями на етапах формулювання вимог, перевірки якості та отримання зворотного зв'язку від користувачів, щоб програмна логіка й інтерфейс відповідали потребам і досвіду настроювачів. Інакше кажучи, сприйманий розрив у якості між настроюванням за допомогою тюнера і на слух робить вкрай важливим урахування думки експертів під час розроблення й оцінювання таких програм.

Отже, необхідними заходами для створення високоякісного програмного тюнера є вдосконалення фізико-акустичних моделей музичного інструмента, реалізація адаптивних слухових критеріїв оцінювання інтервалів, впровадження сучасних архітектурних рішень із модульною побудовою, підвищення точності аналізу й оброблення сигналів і розроблення інтуїтивного інтерфейсу з чітким зворотнім зв'язком. Не менш важливим є залучення сертифікованих настроювачів до процесів валідації та уточнення критеріїв якості звучання. Тільки на основі такого комплексного підходу можна розробити продукт, який реально підтримуватиме професійне настроювання музичних інструментів на належному рівні.

Список літератури:

1. Millard M., Tizhoosh H. R. Tuning pianos using reinforcement learning. *Applied Acoustics*. 2007; Vol. 68, Issue 5: 576–593. DOI: 10.1016/j.apacoust.2006.03.006.

2. Hinrichsen H. Entropy-based Tuning of Musical Instruments. *Brazilian Journal of Physics Teaching* (in Portugal). 2012. Vol. 34, Issue 2, No. 2301. P. 1–8. DOI: <https://doi.org/10.1590/s1806-11172012000200004>.
3. François R., Bertrand D., Laurent D. A parametric model and estimation techniques for the inharmonicity and tuning of the piano. *Journal of the Acoustical Society of America*. 2013. Vol. 133, Issue 5. P. 3107–3118. DOI: 10.1121/1.4799806.
4. Shah S., Välimäki V. Automatic Tuning of High Piano Tones. *Applied Sciences*. 2020; Vol. 10, Issue 6, No. 1983. DOI: <https://doi.org/10.3390/app10061983>.
5. Szwajcowski A., Pilch A. Optimization of piano tuning by means of spectral entropy minimization. *Applied Acoustics*. 2020. Vol. 166. 107359 p. DOI: 10.1016/j.apacoust.2020.107359.
6. What is a Registered Piano Technician (RPT)? Piano Technicians Guild. URL: <https://www.ptg.org/about/registered-piano-technician> (date of access: 18.07.2025).
7. Bushey R. Tuning Fork or Electronic Tuning Device (ETD)? Richard's Piano Svc Blog. URL: <https://surl.li/zdxups> (date of access: 18.07.2025).
8. Glenn P. ETD tuning vs. aurally. Tuning, Glenn's Piano Service. URL: <https://bleespiano.com/etd-tuning-vs-aurally> (date of access: 18.07.2025).
9. Porter, J. Aural vs. electronic piano tuning. Fine Tuning. URL: <https://finetuningco.com/blog/aural-vs-electronic-piano-tuning> (date of access: 18.07.2025).
10. International Piano Technicians School – Online Training for Piano Technicians. URL: <https://internationalpianotechniciansschool.com> (date of access: 21.07.2025).
11. Cerisano M. How to Tune Pianos: The Hobby, Skill, and Career of Piano Tuning. URL: <https://howtotunepianos.com> (date of access: 21.07.2025).
12. Vrz S. Microphone suggestion for piano tuning using mobile app. Music: Practice & Theory Stack Exchange. URL: <https://surl.li/cwxqrm> (date of access: 23.07.2025).
13. MoreBandicoot4833. How accurate are built in cellphone microphones when used with piano tuning apps? Reddit. URL: <https://surl.li/obfmit> (date of access: 23.07.2025).
14. Kelly A. Reliable Frequency Detection Using DSP Techniques. AUTODESK Instructables. URL: <https://surl.cc/cyaikx> (date of access: 23.07.2025).
15. Truitt W. Tunelab Problems. Pianotech. URL: <https://my.ptg.org/discussion/tunelab-problems> (date of access: 25.07.2025).
16. RonTuner. PianoMeter follow up. Piano World Piano & Digital Piano Forums. URL: <https://surl.li/wuqylc> (date of access: 25.07.2025).
17. Verituner User Reviews. AppsHunter.io. URL: <https://appshunter.io/ios/app/449411468/reviews> (date of access: 28.07.2025).
18. Wilke C., Richly S., Götz S., Piechnick C., Aßmann U. Energy Consumption and Efficiency in Mobile Applications: A User Feedback Study. *2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing*. Beijing: China. 2013. P. 134–141. DOI: 10.1109/GreenCom-iThings-CPSCoM.2013.45.
19. The Importance of Battery Optimization in Mobile Apps. PixelTech. URL: <https://pixeltech.ai/the-importance-of-battery-optimization-in-mobile-apps> (date of access: 28.07.2025).

Sychkov V.S. SOFTWARE FOR AUTOMATED PROFESSIONAL TUNING OF ACOUSTIC STRING MUSICAL INSTRUMENTS: AN EXPERT VIEW

The article is devoted to the problem of automation of professional tuning of acoustic string musical instruments, in particular pianos, with a focus on creating high-quality software capable of meeting the high requirements of master tuners. The interdisciplinary nature of the problem, which combines physical and mathematical modeling, psychoacoustic aspects, optimization mechanisms, digital signal analysis and processing, and software engineering, has been emphasized.

The limitations of modern electronic tuning devices (ETDs) – both hardware and software – have been considered in the article, and it has been proven that they are currently unable to fully replace aural tuning, which is based on the experience of professional technicians. Modern scientific approaches to automating the tuning process have been disclosed, including the use of reinforcement learning, minimization of global spectral entropy, parametric modeling of string inharmonicity, and analysis of subjective sound perception.

An analysis of publications, technical blogs, professional forums and opinions of leading registered technicians, in particular members of the Piano Technicians Guild (PTG), has been carried out, on the basis of which typical problems of modern programs have been identified: accumulation of errors in calculations, limited accuracy of unison tuning, insufficient tuning stability, lack of proper feedback and poor adaptation to real acoustic conditions.

It has been determined that the creation of high-quality software is possible only with the active participation of tuning specialists at critical stages of its development life cycle – specification of requirements, validation of results and subsequent maintenance of the system. A number of measures have been proposed to improve the quality of such solutions. It has been concluded that only a comprehensive approach combining technological achievements with empirical knowledge and skills of master tuners is able to ensure the creation of ETDs that will meet high tuning quality standards and will inspire the trust of professional users.

Key words: *tuning software, electronic tuning devices, aural tuning, expert knowledge, inharmonicity, piano.*

Дата надходження статті: 28.07.2025

Дата прийняття статті: 07.08.2025

Опубліковано: 27.10.2025

Стаценко Д.В.

Київський національний університет технологій та дизайну

Стаценко В.В.

Київський національний університет технологій та дизайну

Мельник Г.В.

Київський національний університет технологій та дизайну

Калашник В.Ю.

Київський національний університет технологій та дизайну

ЗАСТОСУВАННЯ ШТУЧНОГО ІНТЕЛЕКТУ І ХМАРНИХ БАЗ ДАНИХ: НАПРЯМИ РОЗВИТКУ ТА ПРИКЛАДНІ МОЖЛИВОСТІ

Бурхливий розвиток технологій хмарних обчислень та штучного інтелекту (ШІ) ініціював радикальні перетворення у способах моніторингу, аналізу та управління даними. Інтеграція ШІ-аналітики в інфраструктуру хмарних баз даних трансформує моделі прийняття рішень у різних секторах економіки, забезпечуючи високу масштабованість, мінімальну затримку в обробці інформації та потужні перспективи можливості.

Дане дослідження присвячене аналізу синергетичної взаємодії між ШІ та хмарними СУБД, з особливим акцентом на застосування алгоритмів машинного та глибокого навчання для підвищення ефективності обробки даних та точності бізнес-висновків. Розглядаються прикладні кейси у різних галузях. У них ШІ-аналітика демонструє значний вплив: сприяє оптимізації операційних процесів, зниженню управлінських та ринкових ризиків, а також формуванню точніших сценаріїв довгострокового планування.

У роботі представлено аналіз статистичних даних за останні роки за декількома напрямками розвитку ШІ-моделей та впровадження їх у хмарні бази даних.

Представлену статистику фінансування різних змагань із призовими фондами для розвитку машинного навчання. Зазначено, що кількість користувачів з кожним роком збільшується, а також збільшення їх фінансування різними компаніями.

Представлені нові рішення у сфері хмарних баз даних, а саме впровадження у «векторні» технології, також наведено зміну рівня впровадження різних баз даних, які використовують інноваційні методи ШІ. Зазначені ключові моменти фінансування у сфері баз даних, а також наведені анонси нових технологічних можливостей від постачальників хмарних послуг.

Результати даного аналізу підкреслюють стратегічне значення ШІ як каталізатора еволюції хмарних баз даних і підкреслюють його вирішальну роль у забезпеченні конкурентоспроможності організацій в умовах цифрової економіки, що стрімко змінюється.

Ключові слова: штучний інтелект, машинне навчання, хмарні обчислення, бази даних, хмарні системи, прогнозна аналітика, глибоке навчання.

Постановка проблеми. У цифровій трансформації галузей першому плані вийшли дві ключові технології: хмарні обчислення та штучний інтелект (ШІ). Перший варіант пропонує гнучку масштабовану інфраструктуру для зберігання та управління даними, а другий – такі методи, як машинне навчання (МН), обробка природної мови (ОПМ) та глибоке навчання (ГН), для створення цінних

аналітичних уявлень на основі великих наборів даних. Поєднання аналітики на основі ШІ з хмарними базами даних можуть трансформувати процеси прийняття рішень за рахунок програмування складних завдань, підвищення точності та швидкого надання аналітичних даних. Оскільки дедалі більше підприємств переводять свою діяльність у хмару, впровадження аналітики з урахуванням

ШІ стає критично важливим підтримки конкурентоспроможності та просування підходів до прийняття рішень, орієнтованих на дані [1, 2].

Використання хмарних баз даних створює для компаній значні перешкоди, включаючи підтримку узгодженості даних, розширення ємності та обробку інформації в режимі реального часу. Традиційні системи баз даних зазвичай не справляються з керуванням обсягом, швидкістю та різноманітністю актуальних даних, що ускладнює підтримку аналітики на основі ШІ. Більше того, об'єднання моделей ШІ з хмарними базами даних ускладнює забезпечення точних прогнозів та безперебійну передачу даних між системами. Незважаючи на гостру потребу у більш ефективних процесах прийняття рішень, що використовують ці передові технології, багато організацій, як і раніше, стикаються зі складнощами, пов'язаними з їх інтеграцією [3, 4].

Стрімкий розвиток технологій штучного інтелекту (ШІ), зокрема машинного навчання (МН) та глибокого навчання (ГН), призвело до їх широкого застосування в різних галузях, включаючи управління хмарними даними. Хмарні бази даних перетворилися із простих рішень для зберігання даних у складні розподілені системи, здатні підтримувати складні моделі ШІ. Ранні дослідження впровадження ШІ у хмарні системи були зосереджені на автоматизації завдань керування даними, таких як індексація та оптимізація запитів. Нещодавні дослідження підкреслюють внесок ШІ у підвищення ефективності обробки даних та забезпечення аналітики у реальному часі. Це злиття ШІ та хмарних баз даних призвело до створення більш інтелектуальних та адаптованих систем управління даними, здатних обробляти все більш складні та різноманітні набори даних [1-5].

Інтеграція моделей ШІ в хмарні бази даних дозволяє організаціям проводити більш складний аналіз та отримувати більш глибокі знання. Моделі ШІ дозволяють хмарним базам даних підтримувати складні програми, такі як обробка природної мови, аналіз настроїв та автоматизоване ухвалення рішень. У міру розвитку хмарних баз даних вони стають дедалі ефективнішими у обробці величезних обсягів даних у режимі реального часу, що є необхідним компонентом для підприємств, яким необхідно отримувати інформацію у великих масштабах та швидко [5, 6].

У хмарних середовищах ШІ зазвичай використовують такі підходи, як машинне навчання, глибоке навчання та передиктивна аналітика на основі ШІ. Алгоритми машинного навчання

забезпечують адаптивну обробку даних, виявляють відхилення та надають персоналізовані поради. Моделі глибокого навчання використовуються для вирішення складніших завдань, включаючи розпізнавання зображень, розуміння природної мови та автономне прийняття рішень. Компанії використовують великі набори даних для отримання аналітичної інформації в режимі реального часу за допомогою аналітичних платформ на основі ШІ, таких як *Google Cloud AI*, *AWS Machine Learning* та *Microsoft Azure AI*, що часто призводить до покращення можливостей ухвалення рішень. Ці хмарні платформи аналітики на основі ШІ також забезпечують масштабованість та адаптивність, що дозволяє компаніям змінювати свої обчислювальні ресурси відповідно до потреб. Крім того, вони підтримують обробку та аналіз даних у реальному часі, надаючи організаціям актуальну інформацію для швидкого прийняття рішень. Впровадження аналітики на основі ШІ в хмарні середовища також сприяє розвитку спільних робочих процесів, дозволяючи командам легко обмінюватися інформацією та співпрацювати в рамках складних проєктів, пов'язаних із даними, перебуваючи у різних місцях [1, 2, 3, 7-10].

Поєднання ШІ з хмарними базами даних змінило спосіб керування даними. Воно допомагає організаціям осмислювати великі обсяги інформації та перетворювати її на корисну інформацію, яка підтримує розумніші рішення та нові ідеї. У міру вдосконалення алгоритмів штучного інтелекту хмарні бази даних стануть ще потужнішими, дозволяючи використовувати просунуту прогностичну аналітику і більш інтелектуальні автоматизовані системи прийняття рішень.

Постановка завдання. Метою роботи є: вивчити взаємодію штучного інтелекту та хмарних баз даних, впровадження та використання нових методів їх реалізації з метою оцінки їх сукупного потенціалу для покращення процесів прийняття рішень, а також їхнього внеску в існуючі рішення для галузей, їх підтримки різними платформами та нові досягнення за цим напрямком.

Виклад основного матеріалу. Впровадження штучного інтелекту (ШІ) в хмарні бази даних відкрило нові горизонти в галузі обробки інформації, забезпечивши розумніше, масштабоване та ефективне управління даними. Завдяки використанню ШІ, хмарні бази даних значно підвищили продуктивність у таких ключових аспектах, як швидкість обробки запитів, точність пошуку даних та здатність масштабуватися під навантаженням.

Алгоритми машинного навчання стали ключовим засобом поліпшення виконання запитів. Застосовуючи методи передиктивного аналізу та розпізнавання шаблонів, вони суттєво знижують потребу в обчислювальних ресурсах та час обробки великих масивів даних. ШІ здатний аналізувати історію запитів, передбачати найбільш ймовірні сценарії їх виконання та оптимізувати маршрути доступу до даних, що прискорює вилучення інформації та підвищує чуйність системи.

Крім того, ШІ-орієнтовані хмарні БД можуть автоматично підлаштовувати обсяг задіяних ресурсів залежно від поточного навантаження, забезпечуючи стабільну продуктивність навіть у періоди пікового попиту. Це особливо важливо для сфер з динамічними потребами в обробці даних, таких як фінансові послуги та електронна комерція.

Додатковою перевагою є здатність ШІ-моделей до самонавчання, що дозволяє системам самостійно адаптуватися до нових типів даних і запитів користувачів, що змінюються, без необхідності ручного втручання. Така гнучкість особливо актуальна в умовах постійних змін у структурі даних та користувацьких очікуваннях.

У зв'язку з вищесказаним за останні роки зросла кількість змагань з машинного навчання серед різних організацій із призовими фондами для подальшого розвитку машинного навчання.

За статистикою, такі змагання стають популярнішими та їх кількість зростає. За 2024 рік було проведено 400 змагань серед 18 масових плат-

форм учасників з машинного навчання, що на 100 більше, ніж роком раніше [11].

Також за 2024 рік зріс призовий фонд. У 2023 він становив 7,8 млн. доларів США, а в 2024 зріс до 22 млн. доларів США [12]. Також необхідно зазначити, що у 2022 році призовий фонд був 5 млн. доларів США [13].

У таблицях 1 та 2 показано огляд конкурентних платформ за 2023 та 2024 рік відповідно.

У 2024 році багато платформ для змагань продемонстрували значне зростання кількості користувачів. База користувачів більшості платформ збільшилася більш ніж на 25% порівняно з попереднім роком, а деяких більш ніж удвічі. *Kaggle* залишається найбільшою платформою як за кількістю зареєстрованих користувачів (понад 22 мільйони), так і за загальним призовим фондом у 2024 році (понад 4 мільйони доларів).

З розвитком технологій штучного інтелекту очікується суттєве посилення їх прогностичних здібностей, що збільшить їх вплив на хмарні системи баз даних. Особливо перспективними вважаються методи глибокого навчання та навчання з підкріпленням, які зможуть забезпечувати точніші та надійніші прогнози в умовах складних, динамічних середовищ з великими обсягами неструктурованих даних. Такі ШІ-моделі все краще виявлятимуть приховані взаємозв'язки та патерни в інформації, сприяючи більш обґрунтованим прогнозам та ефективному прийняттю рішень.

Найближчими роками очікується поява нових тенденцій, що формують вигляд ШІ-аналітики в хмарній середовищі. Однією з ключових стане

Таблиця 1

Огляд конкурентних платформ за 2023 рік

Платформа	Заснована	Кількість користувачів, тис.	Кількість змагань, од.	Загальний призовий фонд
Alcrowd	2017	123	17	\$198,000
Bitgrit	2017	35	2	\$6,000
Codabench	2023	3,5	3	\$10,000
CodaLab	2013	55	71	\$218,000
DrivenData	2014	100	11	\$1,762,000
EvalAI	2017	40	33	\$96,000
Grand Challenge	2010	94	22	\$40,000
Humyn.ai	2023	14	5	\$41,000
Kaggle	2010	17000	41	\$2,358,000
Lab42	2022	500	3	\$87,000
Onward	2022	3,2	3	\$110,000
Signate	2014	80	12	\$163,000
Tianchi	2014	~1.4	61	\$852,000
Trustii	2020	1,5	2	\$60,000
Zindi	2018	72	21	\$134,000
Інші			47	\$1,518,000

Огляд конкурентних платформ за 2024 рік

Платформа	Заснована	Кількість користувачів, тис.	Кількість змагань, од.	Загальний призовий фонд
Alcrowd	2017	245	11	\$113,000
Antigranular	2022	2,7	2	\$10,000
Bitgrit	2017	35	5	\$184,000
Codabench	2023	17	37	\$179,000
CodaLab	2013	74	113	\$150,000
CrunchDAO	2021	7	4	\$180,000
DrivenData	2014	125	9	\$650,000
EvalAI	2017	40	36	\$92,000
Grand Challenge	2010	109	37	\$56,000
Kaggle	2010	22000	44	\$4,254,000
Signate	2014	80	14	\$99,000
ThinkOnward	2022	6,9	15	\$405,000
Tianchi	2014	1400	45	\$659,000
Trustii	2020	2	4	\$57,000
Zindi	2018	75	21	\$113,000
Інші			41	\$15,143,000

перехід до аналітики в режимі реального часу, де ШІ оброблятиме дані миттєво в міру їх надходження. Це дозволить організаціям оперативно реагувати на зміни та приймати рішення на основі найактуальнішої інформації. Подібні можливості особливо актуальні для таких сфер, як фінанси, логістика та електронна комерція, де швидкість ухвалення рішень є критично важливою.

Ще один важливий напрямок – розвиток автономних систем прийняття рішень. У міру ускладнення ШІ-алгоритмів очікується запровадження рішень, здатних самостійно аналізувати дані, формувати висновки та виконувати дії без втручання людини. Це дозволить значно скоротити цикл прийняття рішень та радикально змінити підходи в таких галузях, як роботизована автоматизація процесів (RPA), автоматична торгівля та безпілотні транспортні системи.

Крім того, очікується, що інтеграція ШІ з квантовими обчисленнями дасть новий імпульс розвитку хмарної аналітики. Квантові технології, що мають здатність обробляти колосальні обсяги інформації з безпрецедентною швидкістю, зможуть помітно підвищити продуктивність та точність ШІ-моделей. Таке поєднання дозволить добиватися проривів у вирішенні складних завдань оптимізації, в потоковій обробці даних та розвитку просунутих методів машинного навчання, розширюючи горизонти застосування ШІ в хмарних базах даних.

Загалом у міру подальшого технологічного прогресу синергія між ШІ та хмарними обчислен-

нями стимулюватиме нові інновації, дозволяючи бізнесу приймати більш точні рішення, раціонально використовувати ресурси та знаходити нові точки зростання.

Можливості ШІ, особливо великої мовної моделі (ВММ), збільшуються. Програми, засновані на цих технологіях, розширюються у різних галузях. Важливою областю перетину баз даних та ШІ є зберігання та пошук векторних даних. За останній рік багато основних баз даних впровадили або покращили відповідні функції [2, 8-10].

За останні роки багато дослідних та консалтингових компаній подали звіти щодо розвитку та впровадження ШІ в хмарні бази даних. Зараз практично всі постачальники баз даних та хмарних рішень інвестують у «векторні» технології. Ці інновації на основі ШІ стрімко розвиваються, і конкуренція у галузі векторних сховищ та додатків висока [14].

2024 року постачальники баз даних почали впроваджувати підтримку векторних функцій. Найбільш яскравими прикладами є:

pg_vector випустив версію 0.8, що широко застосовується в багатьох проектах.

Azure представила розширення *DiskANN* для векторних баз даних, інтегроване до кількох продуктів для баз даних.

MariaDB розпочала повну підтримку векторних функцій, починаючи з версії 11.6.

В *OceanBase 4.3.3* додано підтримку зберігання та індексації векторних даних.

MySQL 9.0 з'явилася підтримка векторних схо-
вищ.

TiDB запустила *Vector Search* (бета-версію), органічно інтегрувавши можливості векторної обробки до своєї екосистеми.

Крім того, практично всі постачальники послуг, що підтримують *PostgreSQL*, тепер включають плагін *pg_vector*.

Як і в попередні роки, *PostgreSQL* продовжує своє повільне і впевнене зростання, в той час як *MySQL*, на даний момент вийшов на плато.

У рейтингу *DB-Engines MySQL*, як і раніше, утримує значне лідерство. Тим не менш, динаміка *PostgreSQL* очевидна: такі інновації, як *pg_vector*, набувають широкого поширення, в той час як підтримка векторів у *MySQL* версії 9.0 залишається обмеженою.

На рис. 1 представлені дані результатів незалежних опитувань щодо використання баз даних.

Ключовими моментами фінансування у сфері баз даних стали такі зміни та оновлення.

За 2024 рік у секторі баз даних відбулося кілька значних подій, пов'язаних із фінансуванням хмарних баз даних з ШІ:

Tembo, стартап, що надає послуги хостингу *PostgreSQL*, залучив 14 мільйонів доларів у рамках раунду фінансування серії А, пропонуючи продукти з відкритим вихідним кодом, так і *Tembo Cloud* для керованих сервісів *PostgreSQL*.

OpenAI придбала розробника баз даних для аналітики в реальному часі *Rockset*, що оцінюється в 500–1000 мільйонів доларів, для покращення пошуку та аналізу даних у реальному часі для великих моделей.

ClickHouse придбала стартап *PeerDB*, який займається синхронізацією даних *PostgreSQL*, чий посівний раунд у 2023 році залучив 3,6 мільйона доларів.

База даних тимчасових рядів з відкритим вихідним кодом *GreptimeDB* отримала багатомільйонне фінансування, наголосивши на економічно ефективній обробці даних у реальному часі.

K1 придбала *MariaDB*, зміцнивши свої позиції провідного інвестора у програмному забезпеченні для малого бізнесу.

Supabase залучила 80 мільйонів доларів у рамках раунду фінансування серії С, орієнтованого на *PostgreSQL*, пропонуючи низку інструментів для розробників, включаючи векторні бази даних.

ApertureDB залучила 8 мільйонів доларів на розвиток сервісів багатомодальних потоків даних для ШІ. Необхідно також зазначити, що у 2024 році було анонсовано нові технологічні можливості від постачальників хмарних послуг.

На конференції *AWS re:Invent* компанія *AWS* представила *Aurora DSQL* з крос-регіональними, строго узгодженими можливостями читання та запису у кількох регіонах, доступністю 99,999% (розгортання у кількох регіонах) та сумісністю з *PostgreSQL*.

DynamoDB також надала аналогічні можливості забезпечення крос-регіональної узгодженості.

На конференції *Oracle CloudWorld* Ларрі Еллісон наголосив на багатохмарній стратегії, пред-

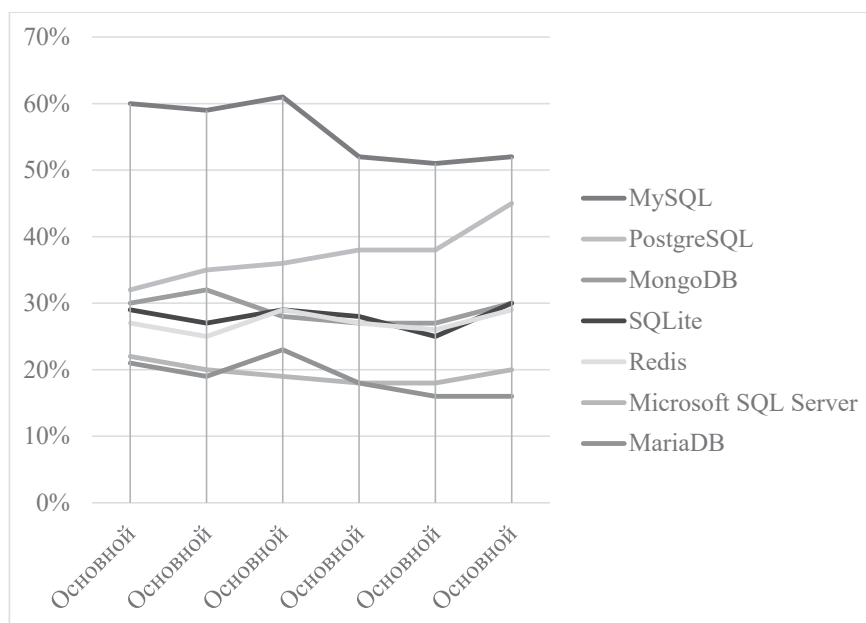


Рис. 1. Рівень впровадження баз даних

ставивши рішення *Oracle@Google*, *Oracle@Azure* та *Oracle@AWS*.

GCP зосередилася на покращенні *AlloyDB* та її базових продуктів, таких як випуск версій *Enterprise Plus* та *Enterprise Cloud SQL* для *SQL Server*.

Microsoft Azure віддала пріоритет *Cosmos DB* та *SQL Database*, впровадивши векторне індексування *DiskANN* для своїх векторних баз даних.

Висновки. У цьому дослідженні наголошується, що за останні роки зріс інтерес з боку різних компаній в інтеграцію штучного інтелекту з хмарними базами даних, нові методи їх реалізації демонструють значне підвищення ефективності, масштабованості та аналітичних можливостей.

У роботі представлено аналіз статистичних даних за останні роки за декількома напрямками розвитку ІІІ-моделей та впровадження їх у хмарні бази даних.

Вказано, що за останні 3 роки стали набирати популярності змагання з машинного навчання. Призовий фонд таких змагань зріс у 4,4 рази з 2022 року до 2024 року і склав 22 млн доларів США. Також зросла і кількість фахівців, які брали участь у цих змаганнях, що підтверджує тенденцію щодо розвитку та впровадження нових рішень у технології, пов'язані з ІІІ.

Результати аналізу статистичних даних щодо використання провідних баз даних також показали вплив ІІІ. За останні три роки ті бази даних, які почали вводити нові векторні функції. На основі отриманих даних зріс попит на *PostgreSQL* на 13%, в той же час використання *MySQL* знизилося на 8%.

Ключові моменти пов'язані з фінансуванням у сфері баз даних та анонсами нових технологічних можливостей від постачальників хмарних послуг, що підтверджують інтерес у розвитку цієї галузі з метою впровадження ІІІ у хмарні бази даних.

Список літератури:

1. J. Doe, Artificial Intelligence in Cloud Computing: Revolutionizing Data Analytics. *IEEE Access*, vol. 20, no. 4, pp. 2334-2345, 2023.
2. Padma Rama Divya Achanta. Architecting Cloud Database Solutions with Embedded AI for Real-Time Analytics. *International Journal of Scientific Research and Management (IJSRM)*, Volume 13, Issue 06, pp. 2208-2219, 2025.
3. A. Smith, Machine Learning Models for Real-Time Cloud Systems. in *Proc. IEEE Int. Conf. Cloud Comput.*, pp. 112-117, 2022.
4. L. Zhang, S. Patel, Deep Learning Algorithms in Cloud Environments. *IEEE Trans. Neural Networks*, vol. 30, no. 5, pp. 1254-1265, 2021.
5. S. Williams et al., Quantum Computing for Cloud AI Systems: Opportunities and Challenges. *IEEE Trans. Quantum Eng.*, vol. 12, no. 2, pp. 91-104, 2023.
6. B. Roberts, E. Cooper, AI-Driven Predictive Analytics in Retail Supply Chains, *IEEE Access*, vol. 19, no. 9, pp. 7712-7720, 2023.
7. K. Taylor, H. Wong, AI Optimization Techniques for Real-Time Cloud Data Processing. *IEEE Cloud Comput.*, vol. 9, no. 6, pp. 88-95, 2023.
8. J. Hernandez, M. Lee, Financial Risk Management Using AI in Cloud Environments. in *Proc. IEEE Symp. Financ. Innov. AI*, 2021, pp. 233-240.
9. P. Gupta, The Future of AI in Cloud-Based Decision Systems. *IEEE Syst. J.*, vol. 15, no. 10, pp. 1408-1415, 2023.
10. L. Zhang, S. Patel, Deep Learning Algorithms in Cloud Environments. *IEEE Trans. Neural Networks*, vol. 30, no. 5, pp. 1254-1265, 2021.
11. The State of Machine Learning Competitions 2024 Edition – ML Contests / Research. URL: <https://surl.li/dlcwcs>
12. The State of Competitive Machine Learning 2023 Edition – ML Contests / Research. URL: <https://surl.li/dhadik>
13. The State of Competitive Machine Learning 2022 Edition – ML Contests / Research. URL: <https://surl.li/ebjpf>
14. 2024 in Review: Key Highlights in Cloud Databases – DEV Community. URL: <https://surl.li/gqwxqa>

Statsenko D.V., Statsenko V.V., Melnyk G.V., Kalashnik V.Yu. APPLICATION OF ARTIFICIAL INTELLIGENCE AND CLOUD DATABASES: DIRECTIONS OF DEVELOPMENT AND APPLICATION OPPORTUNITIES

The rapid development of cloud computing and artificial intelligence (AI) technologies has initiated radical transformations in the methods of monitoring, analyzing and managing data. The integration of AI analytics into the infrastructure of cloud databases transforms decision-making models in various sectors of the economy, providing high scalability, minimal delay in information processing and powerful predictive capabilities.

This study is devoted to the analysis of the synergistic interaction between AI and cloud DBMSs, with a special emphasis on the use of machine and deep learning algorithms to increase the efficiency of data processing and the accuracy of business conclusions. Applied cases in various industries are considered. In them, AI analytics demonstrates a significant impact: it contributes to the optimization of operational processes, the reduction of management and market risks, as well as the formation of more accurate long-term planning scenarios.

The paper presents an analysis of statistical data over recent years in several areas of development of AI models and their implementation in cloud databases.

The statistics of financing of various competitions with prize funds for the development of machine learning are presented. It is noted that the number of users is increasing every year, as well as an increase in their financing by various companies.

New solutions in the field of cloud databases are presented, namely the implementation of “vector” technologies, and the change in the level of implementation of various databases that use innovative AI methods is also given. The key points of financing in the field of databases are indicated, and announcements of new technological opportunities from cloud service providers are also given.

The results of this analysis emphasize the strategic importance of AI as a catalyst for the evolution of cloud databases and emphasize its crucial role in ensuring the competitiveness of organizations in the rapidly changing digital economy.

Key words: artificial intelligence, machine learning, cloud computing, databases, cloud systems, predictive analytics, deep learning.

Дата надходження статті: 25.07.2025

Дата прийняття статті: 31.07.2025

Опубліковано: 27.10.2025

Стецюк Ю.В.

Хмельницький національний університет

Савенко О.С.

Хмельницький національний університет

ЧАСТКОВО-ЦЕНТРАЛІЗОВАНА СИСТЕМА БЕЗПЕКИ МЕРЕЖЕВИХ ОС З ДИНАМІЧНОЮ ПЕРЕДАЧЕЮ КЕРУВАННЯ МІЖ ЇЇ ВУЗЛАМИ

Проведено аналіз сучасного стану рівня загроз від зловмисного програмного забезпечення (ЗПЗ) сучасним захищеним спеціалізованим ОС, стосовно їх стійкості до витоків конфіденційної інформації. Представлена модель взаємодії захисних механізмів централізованої системи безпеки ОС мережевого вузла, яка дозволяє отримати архітектуру безпеки ОС, позбавлена проблеми витоків конфіденційної інформації при атаках на оперативну пам'ять системи, що досягається включенням до неї механізму маркування, особливістю якого є контроль каналів проходження конфіденційної інформації на рівні сторінок пам'яті. Розглянуті принципи побудови централізованих та частково-централізованих систем безпеки ОС та принципів організації роботи їх механізмів безпеки. Для збереження позитивних якостей централізованої системи безпеки і водно час уникнення проблем властивих будь-якій централізації, запропоновано метод динамічної передачі керування між вузлами комп'ютерної мережі, як способу підвищення її стійкості до витоків конфіденційної інформації та усунення вузького місця централізованої системи безпеки – єдиного центру управління, а також для протидії проведенню розвідки зловмисниками комп'ютерної мережі з метою зриву атаки на її ресурси. Його особливістю є синхронізація та актуалізація центральної бази привілеїв та стану мережі шляхом реплікації локальних баз привілеїв мережеских вузлів в процесі передачі керування між ними. Ще одна особливість методу полягає в забезпеченні постійної наявності резервного керівного вузла на випадок збою. Кожен вузол, що передає керування мережею наступнику, визначеного в рандомний спосіб, переходить у резерв. Запропоновано узагальнений алгоритм, що реалізує кроки методу динамічної передачі керування між вузлами комп'ютерної мережі при штатній роботі мережі та для ситуації старту системи, або збою. Розроблено модель станів мережевого вузла в процесі динамічної передачі керування між вузлами, що дозволяє досліджувати його у всіх можливих ситуаціях. Виконано порівняльний аналіз ефективності централізованої та частково-централізованої систем безпеки.

Ключові слова: система безпеки, централізована система, частково-централізована система, операційна система, модель станів.

Постановка проблеми. Аналіз сучасного стану рівня загроз від ЗПЗ сьогодишнім мережеским ОС підтверджує, що їх більша частина обумовлена реалізованим в ОС концептуальним підходом, в основі якого лежить реалізація схеми розподіленого адміністрування механізмів захисту. Таким чином, існуючі на сьогодні ОС не в змозі гарантувати стовідсоткову стійкість до витоків інформації через існування в них вразливостей з однієї сторони, а з іншої через їх знаходження під постійною загрозою зі сторони ЗПЗ, що тільки збільшується. Без впровадження нових підходів до захисту ОС, що ґрунтуються на централізованих та частково-централізованих системах безпеки, протидія витоків інформації та, загалом,

несанкціонованому доступу до конфіденційної інформації буде неефективною.

Аналіз останніх досліджень і публікацій. Операційна система (ОС) є основною, найбільш функціонально складною, програмною системою комп'ютерної системи (КС). Її призначення в управлінні всіма апаратними та програмними ресурсами КС, надання їх, відповідно до встановлених політик безпеки в розпорядження прикладних програм, забезпечення інтерфейсу як для користувачів так і прикладних програм. Сьогоднішні ОС до того ж мають добре розвинуті мережескі функції. Такими є урядові та військові системи, різноманітні фінансові установи та корпоративні ІС, де безпека даних має першорядне

значення. На сьогодні це не тільки чисто наукове або технічне поняття, але і юридичне. Вимоги до таких систем визначені стандартом США, розробленим Національним інститутом стандартів і технологій (NIST), а також в міжнародному стандарті [1].

На сьогодні існує немало систем реагування на інциденти, які можуть бути застосовані для підвищення рівня безпеки [2, 3] в мережах в плані протидії витоку конфіденційної інформації. Встановлено, що ОС, які працюють в складі кіберфізичних систем, які використовують канали передачі даних бездротового зв'язку СВТС є вразливими до атак, через низьку стійкість їх механізму виявлення вторгнень на основі машинного навчання [4, 5], що виявляє лише відомі атаки, які містяться в навчальних наборах даних, і є нечутливим до атак нового типу. З метою усунення проблеми в [6], запропоновано більш досконалий IDS-механізм на основі трансферного навчання для системи СВТС. Вона використовує оптимізований блок одновимірної згорткової нейронної мережі для автоматичного вилучення просторових і часових характеристик з вихідних даних та новий метод передачі знань, що дозволило виявляти і атаки нових типів, в тому числі атаки нульового дня.

В [7, 8] пропонується покращення IDS-систем при роботі ОС з даними в хмарах на основі використання комбінації ядра нечіткої кластеризації (KFCM) і оптимальної нечіткої нейронної мережі типу 2 (OT2FNN), а також механізму захисту E-SGX [9]. В [10] запропоновано розподілену мережеву систему з системою запобігання вторгненням, засновану на удосконаленій структурі алгоритму Spark і динамічній моделі теорії інформаційної безпеки. Ще подібні механізми пом'якшення атак на комп'ютерні мережі запропоновано в [11, 12]. В [13] запропоновано концептуальну модель багатокомп'ютерної системи, що забезпечує функціонування антивірусних приманок та пасток для виявлення шкідливих програм та комп'ютерних атак у мережах. В [14] запропоновано підхід до використання нових інструментів для виявлення шкідливих програм у корпоративних мережах, які є розподіленими системами з частковою централізацією.

Варіанти побудови систем раннього виявлення потенційних каналів ураження КС показано в [15, 16], архітектура якої включає агентів різних типів та певної кількості екземплярів, що спеціалізуються на вирішенні задач виявлення потенційних каналів ураження КС та взаємодіючих між

собою, які спільно вирішують загальне завдання виявлення проникнення у КС [17]. В [18] запропоновано метод виявлення ЗПЗ на кінцевих точках на основі ідентифікаторів подій із використанням глибокого навчання. В [19] представлено самоадаптивну систему для забезпечення стійкості корпоративних мереж до кібератак ботнетів, де стійкість забезпечується адаптивною реконфігурацією мережі.

В [20, 21] пропонується використання інтелектуальних методів глибокого навчання в системах виявлення вторгнень в комп'ютерні мережі. В рамках цього підходу реалізовано метод WNIDS, в рамках якого функціонує контролер SDN, який ефективно відстежує та керує пристроями мережі, аналізує набори даних. В [22] представлені два варіанти алгоритму FTNPA: децентралізований та централізований. Перший варіант реалізує локального відновлення, коли мобільний вузол розміщується в певному місці, щоб допомогти постраждалій зоні. Другий варіант використовує централізований метод виявлення, де метод відновлення, створює альтернативні шляхи мобільних вузлів до вузла призначення. Багато уваги приділялось удосконаленню роботи захисних механізмів ОС, а саме її ядру [23]. Запропоновано модель централізованої системи безпеки ОС для нових архітектур ОС [24].

Кроком вперед є рішення доповнення систем безпеки ОС системою авторизації DiSAuth [25], яка забезпечує безпечний алгоритм, заснований на системі доменних імен (DNS) і блокчейні, що забезпечує її стійкість до втручання, зберігає конфіденційність, для захисту даних від витоку. В [26] піднімається питання мінімізації надання привілеїв доступу до ресурсів, як способу підвищення стійкості до витоків інформації, оскільки зловмисники часто використовують вразливості у файлових системах та драйверах дисків для витоку або маніпулювання вмістом файлів програм.

Удосконаленню архітектури систем безпеки сучасних мережевих ОС, що стосується нівелювання загроз витоку інформації, розглянуто в [27], метод динамічної передачі керування між вузлами мережі запропоновано в [28]. Розвитку систем з динамічною зміною центру керування без втручання користувача з метою ускладнення пошуку центру системи для зловмисників та встановлення принципів їх функціонування присвячені роботи [29, 30].

Інтеграція в комп'ютерні мережеві штучного інтелекту є однією з передових технологій. Очі-

кується, що це дозволить вирішити проблеми, з якими стикаються комп'ютерні мережі в протистоянні кібератакам [31].

Постановка завдання. Метою статті є розробка нових методів забезпечення ефективності роботи систем безпеки мережевих ОС, підвищення їх стійкості до витоків конфіденційної інформації шляхом випадкової динамічної передачі керування між вузлами мережі.

Виклад основного матеріалу. Розглянемо графову модель системи безпеки ОС, наведеної на рис. 1. Вершинам графа відповідають функціональні модулі системи безпеки ОС, а направлені ребра вказують на взаємодію між модулями, визначаючи алгоритм її роботи. Основні взаємодії:

- центральний модуль керування ЦМК передає політики безпеки, включаючи політики маркування конфіденційної інформації через модуль керування політиками безпеки МКПБ до периферійних модулів безпеки ПМБ1 – ПМБn;
- периферійні модулі безпеки ПМБ1 – ПМБn передають дані моніторингу через модуль контролю МК центральному модулю ЦМК для аналізу;
- модуль контролю МК передає результати перевірки коректності маркерів конфіденційності модулю маркування ММ;
- модуль маркування ММ передає агентам інформацію про маркування конфіденційної інформації периферійним модулям безпеки ПМБ1 – ПМБn для подальшого застосування на рівні сторінок пам'яті;
- модулі автентифікації МА, авторизації МАвт, шифрування МШ, мережевої безпеки МерМ вза-

ємодіють з модулем маркування ММ при виконанні своїх функцій.

Запропонована модель централізованої системи управління безпекою дозволяє отримати архітектуру безпеки ОС, яка буде позбавлена проблеми витоків конфіденційної інформації при низькорівневих атаках ЗПЗ. Це досягається включенням до неї механізму маркування, особливістю якого є контроль каналів проходження конфіденційної інформації на рівні сторінок пам'яті.

Якщо в якості базового вузла комп'ютерної мережі взяти станцію з операційною системою з наведеною моделлю системи безпеки, то отримаємо мережу, де в кожному окремому вузлу буде реалізовано централізоване керування привілеями, але сама мережа буде мати децентралізоване керування, зі всіма проблемами безпеки, що властиве даній моделі. Якщо керування мережею передати одному із вузлів мережі, то отримаємо мережну систему із централізованим керуванням. Така модель керування має ряд значимих переваг, таких як:

- уніфікація політик керування привілеями та керування ресурсами по однакових правилах, політики безпеки та протоколи встановлюються в одному місці, що забезпечує єдність управління мережею, полегшує адміністрування, оскільки зміни в політиках застосовуються одночасно для всієї мережі;
- спрощене управління, оскільки відсутня необхідність координувати дії між кількома керівними модулями – адміністрування здійснюється через один модуль ЦМК (Рис. 1);
- централізований моніторинг, означає, що вся інформація про стан мережі та її вузлів надходить до одного модуля, що значно полегшує виявлення загроз, спрощує аналіз та прийняття рішень.

Разом з тим централізована система керування має ряд потенційно загрозливих проблем. Основні з них:

- так звана «єдина точка відмови», коли при виході з ладу центрального модуля мережа стає вразливою, створюється високий ризик її безперервної роботи;
- вразливість до атак ЗПЗ. Сама наявність центрального вузла керування робить його головною цілью для різного роду деструкцій, оскільки успішна атака може дати доступ до всієї мережі.

Метою цього дослідження є знаходження такого вирішення, яке б дозволило зберегти основні переваги централізованої системи, одночасно позбувшись її найбільших проблем.

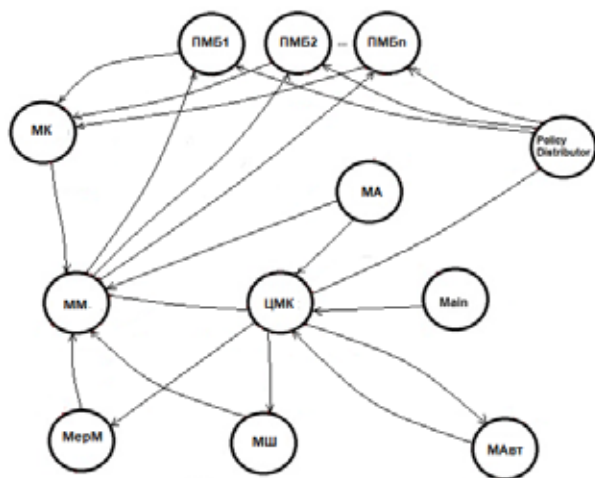


Рис. 1. Модель централізованої системи безпеки ОС з механізмом низькорівневого маркування конфіденційної інформації

В якості одного із способів подолання недоліків централізованої системи безпеки ОС із збереженням її позитивних якостей, пропонується метод динамічної передачі керування між вузлами мережі в процесі її роботи. В основі методу лежить постійна міграція керівного центра по вузлах мережі, що дозволяє позбутись основного недоліку чисто централізованого підходу – наявності вузького місця у вигляді вразливостей центрального модуля керування.

Згідно запропонованого методу комп'ютерна мережа представлена як множина вузлів $M_{\text{ЦМК}}$, чий системи безпеки ОС мають модулі центрального керування ЦМК:

$$M_{\text{ЦМК}} = \{\text{цмк}_1, \text{цмк}_2, \dots, \text{цмк}_n\} \quad (1)$$

де n – число вузлів мережі, $\text{цмк}_1, \text{цмк}_2, \dots, \text{цмк}_n$ – модулі центрального керування вузлів мережі.

При цьому кожен модуль ЦМК вузла зв'язаний з базою привілеїв $B_{\text{привілеїв}}$ яка в поточному режимі слугує локальною базою вузла, а коли вузол в процесі динамічної передачі керування приймає функцію центрального модуля керування мережею, вона актуалізується до рівня основної бази привілеїв мережі. Всі локальні бази привілеїв мережевих вузлів можна представити множиною:

$$B_{\text{привілеїв}} = \{B_{\text{привілеїв } 1}, B_{\text{привілеїв } 2}, \dots, B_{\text{привілеїв } n}\} \quad (2)$$

де n – число вузлів мережі; $B_{\text{привілеїв } 1}, B_{\text{привілеїв } 2}, \dots, B_{\text{привілеїв } n}$ – локальні бази привілеїв вузлів мережі.

Визначимо стан системи $S_{\text{системи}}$ в момент часу t , що відповідає моменту, що передуює передачі керування як:

$$S_{\text{системи}}(t) = (\text{ЦМК}_{\text{акт.}}(t), \text{ЦМК}_{\text{резерв}}(t), T_{\text{керівний}}) \quad (3)$$

де: $\text{ЦМК}_{\text{акт.}}(t) \in M_{\text{ЦМК}}$ – вузол з множини $M_{\text{ЦМК}}$, який в момент часу t володіє токеном (є центром керування мережею); $\text{ЦМК}_{\text{резерв}}(t) \in M_{\text{ЦМК}}$ – резервний вузол керування в момент часу t в мережі; $T_{\text{керівний}}$ – керівний токен мережі.

Як видно з визначення стану мережі (формула 3), особливістю методу є передача функцій центрального модуля керування наступному вузлу шляхом відправки йому керівного токена. Передачу керівного токена можна описати як перехід:

$$S_{\text{системи}}(t_k) = (\text{ЦМК}_{\text{акт.}}(t_k), \text{ЦМК}_{\text{резерв}}(t_k), T_{\text{керівний}}) \rightarrow S_{\text{системи}}(t_{k+1}) = (\text{ЦМК}_{\text{акт.}}(t_{k+1}), \text{ЦМК}_{\text{резерв}}(t_{k+1}), T_{\text{керівний}}) \quad (4)$$

де: $\text{ЦМК}_{\text{акт.}}(t_{k+1})$ – новий керівний вузол, випадково обраний із множини $\text{ЦМК}_{\text{акт.}}(t_{k+1}) \in M_{\text{ЦМК}} \setminus \{\text{ЦМК}_k\}$, правило вибору якого задається формулою:

$$\text{ЦМК}_{k+1} \approx \text{Uniform}(M_{\text{ЦМК}} \setminus \{\text{ЦМК}_k\}) \quad (5)$$

– $\text{Uniform}()$ – функція, яка забезпечує однакову ймовірність вибору в якості керівного вузла будь якого вузла з числа активних, за виключенням поточного вузла ЦМК_k ; ЦМК_k – поточний вузол керування, після передачі керування модулю $M_{\text{ЦМК } k+1}$, виконує функцію резервного вузла керування; $T_{\text{керівний}}$ – керівний токен.

Керівний токен має просту структуру (Рис. 2) і фізично є пакетом даних, який передається між вузлами мережі для визначення поточного керівного модуля (ЦМК). В один момент часу тільки один мережевий вузол володіє токеном $T_{\text{керівний}}$. Таким чином він слугує узгодженню стану мережі, передачі керування, уникненню конфліктів і запобігає одночасному виконанню функцій керування кількома вузлами. З метою збереження його цілісності токен підписується цифровим підписом.

Важливим моментом методу динамічної передачі керування є актуалізація основної бази привілеїв $B_{\text{привілеїв } i+1}$ для вузла, до якого переходить керування комп'ютерною мережею. Вона отримується в процесі процедури реплікації локальних баз привілеїв $B_{\text{локальна } i}$ активних вузлів мережі в момент зміни керівного центра мережі.

В поточному режимі роботи вузол максимально використовує локальну базу привілеїв,

Фізична структура керівного токена

Поле	Опис
TokenID	Унікальний ідентифікатор токена.
Timestamp	Час створення або останньої передачі токена.
IssuerID	Ідентифікатор вузла, що створив токен.
CurrentHolder	Ідентифікатор вузла, який зараз утримує токен.
Signature	Криптографічний підпис для підтвердження автентичності.
TTL (Time-to-Live)	Максимальний час дії токена.
Payload	Додаткові дані (наприклад, політики безпеки).

```
{ "TokenID": "12ABF13",
  "Timestamp": "2025-01-14 T12:00:00Z",
  "IssuerID": "192.168.0.12",
  "CurrentHolder": "192.168.0.11",
  "Signature": "f1a3c1d...ef23a12",
  "TTL": 720,
  "Payload": { "Policies": ["DAC=read",
    "MAC=yes"] "Priority": 5 }
```

Рис. 2. Структура керівного токена для синхронізації передачі керування між вузлами мережі

а при передачі йому функцій центрального вузла керування активує основну базу. Передбачається, що на кожному вузлу зберігається його локальна база привілеїв, а основна база керівного модуля отримується шляхом їх реплікації. До реплікації допускаються локальні бази $B_{\text{локальна } i}$ множини активних вузлів $M_{\text{ЦМК}}$, створюючи множину баз привілеїв для реплікації $B_{\text{привілеїв репл.}}$, яка в загальному випадку може не співпадати з множиною $B_{\text{привілеїв}}$:

$$B_{\text{привілеїв репл.}} \leftarrow \text{Filter}(B_{\text{локальна } i}, \text{Active}(\text{ЦМК}_i)) \quad (6)$$

де: $\text{Filter}(B_{\text{локальна } i}, \text{Active}(\text{ЦМК}_i))$ – фільтр локальних баз $B_{\text{локальна } i}$ по умові;

– $\text{Conditium} = \text{Active}(\text{ЦМК}_i)$ – умовою включення локальної бази $B_{\text{локальна } i}$ – i -го вузла ЦМК_i до множини реплікації при фільтрації локальних баз є її приналежності до активного вузла.

Реплікація буде успішною, якщо в ній приймають участь локальні бази, які є актуальними. Для деякої $B_{\text{локальна } i}$ можна визначити її актуальність через функцію $\text{Validate}()$:

$\text{Validate}(B_{\text{локальна } i}) = 1$, тут значення функції рівне одиниці означає, що вона актуальна. Умовою актуальності бази $B_{\text{локальна } i}$ є її відпаданню під наступні умови:

- $\text{Sinc}(B_{\text{локальна } i}) = 1$ – база синхронізована;
- $t_{\text{оновлення}}(B_{\text{локальна } i}) \leq \Delta t_{\text{max}}$ – час останнього її оновлення, який не перевищує деякий встановлений максимальний термін Δt_{max} ;
- $\text{Integrity}(B_{\text{локальна } i})$ – локальна база привілеїв не є пошкоджена.

З урахуванням того, що кожна локальна база привілеїв $B_{\text{локальна } i}$ повинна знаходитись в актуальному стані, правило включення нових елементів до множини $B_{\text{привілеїв репл.}}$ буде тепер виглядати наступним чином:

$$B_{\text{привілеїв репл.}} \leftarrow \text{Filter}(B_{\text{локальна } i}, \text{Active}(\text{ЦМК}_i) \wedge \text{Validate}(B_{\text{локальна } i}) = 1) \quad (7)$$

де: $\text{Filter}(B_{\text{локальна } i}, \text{Active}(\text{ЦМК}_i))$ – умовою участі в формуванні центральної бази привілеїв $B_{\text{привілеїв репл.}}$ є приналежність $B_{\text{локальна } i}$ до активного мережевого вузла; $\text{Validate}(B_{\text{локальна } i})$ – функція яка повертає значення актуальності локальної бази привілеїв $B_{\text{локальна } i}$ – i -го мережевого вузла.

З урахуванням цього правила отримаємо формулу для всіх елементів множини локальних баз привілеїв, що підлягають реплікації:

$$B_{\text{привілеїв репл.}} = \{B_{\text{локальна } i} \mid i \in M_{\text{ЦМК}}, \text{Active}(\text{ЦМК}_i) \wedge \text{Validate}(B_{\text{локальна } i})\} \quad (8)$$

де: $B_{\text{локальна } i}$ – локальна база привілеїв, де $i \in M_{\text{ЦМК}}$ та її приналежність до активного мережевого вузла; $\text{Validate}(B_{\text{локальна } i})$ – функція яка

повертає значення актуальності локальної бази привілеїв $B_{\text{локальна } i}$ – i -го мережевого вузла.

Тепер, при виконанні операції реплікації над елементами множини $B_{\text{привілеїв репл.}}$ буде отримана актуальна централізована база привілеїв для центрального вузла керування мережею, в якій відсутні некоректні дані локальних не актуальних баз привілеїв та локальних баз привілеїв неактивних вузлів:

$$B_{\text{привілеїв центр.}} = \text{Merge}(B_{\text{привілеїв репл.}}) \quad (9)$$

де: $\text{Merge}()$ – функція реплікації множини локальних баз привілеїв $B_{\text{привілеїв репл.}}$, що підпадають під реплікацію.

Передача керування від поточного центрального модуля до наступного буде успішною, якщо комп'ютерна мережева система буде знаходитись в стані $S_{\text{системи}}(i)$ та для наступного $i+1$ -го вузла керування буде наявна актуальна основна база привілеїв $B_{\text{привілеїв } i+1}$. При цьому вузол, що тільки що передав керування наступному центральному вузлу, переходить у резерв, підтримуючи свою спроможність повернутись до функції центрального вузла.

Такий підхід в реалізації методу практично унеможливило втрату бази привілеїв і не потребує постійного доступу до глобальної мережі, як у випадку її зберігання у хмарних сховищах, або реалізації у вигляді транзакцій в блокчейн – мережі підтримуючи високий рівень автономності, властивий для корпоративних мереж. При цьому також забезпечується висока стійкість комп'ютерної мережі в цілому до впливів різного роду деструкцій і, відповідно зменшення ймовірності витоку конфіденційної інформації. Таким чином метод забезпечує безперервність роботи мережі, узгодженість даних і захищеність системи.

Тепер можна сформулювати основні кроки методу динамічної передачі керування:

Крок 1. Підготовка до передачі керування. Оцінюється стан, в якому знаходиться поточний центральний модуль керування ЦМК_i (рис. 1) множини $M_{\text{ЦМК}}$, де перевіряється, наявність признаков (завершення часового терміну, перевантаження системи, планове технічне обслуговування). При активному значенні хоча б одного з них, поточний центральний модуль має передати керівні функції іншому вузлу мережі ЦМК_{i+1} . При наявності хоча б одного признака, поточний модуль визначає наступний модуль ЦМК_{i+1} з числа активних вузлів, який прийме керування мережею. Затим поточний модуль ЦМК_i перевіряє актуальність

центральної бази привілеїв $B_{\text{привілеїв } i+1}$ вузла, який має прийняти функції центрального вузла керування. При необхідності виконується її реплікація або синхронізація.

Крок 2. Передача поточного стану системи, яка включає передачу активних сесій, політик безпеки та стану моніторингу системи, даних журналів подій. Поточний модуль ЦМК_{*i*} ініціює перевірку коректності отриманих даних модулем-наступником ЦМК_{*i+1*} та актуальності бази привілеїв $B_{\text{привілеїв } i+1}$ після той повідомляє поточний модуль ЦМК_{*i*} про свою готовність виконувати функції центрального вузла.

Крок 3. Активація нового модуля. Поточний модуль ЦМК_{*i*} відправляє наступнику сповіщення у вигляді керівного такенна $T_{\text{керівний}}$, при цьому всім іншим активним вузлам мережі передається інформація про нового модуля ЦМК_{*i+1*}. Важливою особливістю є те, що в кожен момент часу тільки один вузол з множини $M_{\text{ЦМК}}$ володіє керівним токеном $T_{\text{керівний}}$, який забезпечує його керівну роль у мережі.

Старий модуль ЦМК_{*i*} деактивується, але при цьому залишається у готовності як резервний модуль, періодично синхронізуючи свою базу привілеїв та журналів подій у системі з базою та журналами нового центрального вузла.

Крок 4. Отримавши керівний токен $T_{\text{керівний}}$ (рис. 2), вузол активує себе в якості центрального модуля ЦМК_{*i+1*}, беручи на себе контроль за системою. Його першим кроком є фіксація в журналах всіх дій, пов'язаних з фактом передачі йому керування, включаючи час передачі, ідентифікатори модулів ЦМК активних вузлів, стан бази привілеїв. Крім того новий модуль ЦМК_{*i+1*} сповіщає всі активні вузли системи, що він готовий виконувати функції керівника мережевої системи безпеки.

Крок 5. Моніторинг та контроль системи. На цьому кроці новий модуль ЦМК_{*i+1*} здійснює моніторинг стану безпеки системи з метою виявлення можливих збоїв або деградацій. Постійно інформує резервний модуль ЦМК_{*i*} про зміни в основній базі привілеїв з метою їх підтримки в стані максимальної синхронізації. Відслідковує стан признаков передачі керування мережею наступному вузлу.

З метою більшої деталізації роботи запропонованого методу відобразимо кроки методу у вигляді алгоритму (рис. 3 та рис. 4).

При цьому слід наголосити, що даний алгоритм є законом функціонування кожного вузла комп'ютерної мережі, але в кожен момент часу

вони знаходяться на різних стадіях його виконання. Важливою його особливістю в реалізації динамічної передачі керування між центральними модулями CSMM системи безпеки вузлів мережі ґрунтується на правилі безперервності роботи мережі та недопущенні витоку конфіденційної інформації.

Загалом процес роботи комп'ютерної мережі, заснованої на методі динамічної передачі керування, поділяється на кілька етапів – процесу запуску мережі, основного етапу та етапу виходу із збою.

Основний етап це циклічний рух по траєкторії, що включає в себе вершини графа станів (рис. 5) 3 – 4 – 5 – 6 – 3 Тут третя вершина графа відповідає режиму очікування. Це звичайний режим, в якому вузол мережі знаходиться основну частину часу, виконуючи покладені на нього функції. Його модуль центрального керування ЦМК знаходиться в підпорядкуванні модуля ЦМК вузла, який в даний момент здійснює керування системою безпеки всієї мережі і, забезпечує локальне керування системою безпеки вузла відповідно до політик безпеки, які проводить центральний модуль.

З цього стану мережевий вузол може перейти в перехідний режим (вершина 4 рис. 3). В цьому режимі модуль ЦМК вузла отримав керівний токен і знаходиться у процесі прийняття керування системою безпеки мережевої ОС – отримує інформацію про стан системи, активні сесії та базу привілеїв, перевіряє узгодженість даних між вузлами, що відповідає виконанню алгоритму (оператори 6–8 рис. 3).

З стану перехідного режиму вузол, при штатному розвитку подій переходить до стану, в якому він здійснює функції центрального модуля ЦМК (вершина 5 рис. 4), по іншому – керує системою безпеки мережевої ОС, що відповідає виконанню алгоритму (оператори 9–11 рис. 4). В цьому стані вузол знаходиться впродовж терміну, на який його призначено виконувати функції центрального модуля керування системою безпеки мережі, або до початку нештатної ситуації (ребро 5–7 рис. 5). Якщо при виконанні алгоритму (оператор 11 рис. 3) буде виявлено хоча б один признак передачі керування (закінчення терміну, втручання адміністратора системи), то вузол переходить в стан резерву (вершина 6 рис. 3). В цьому режимі він знаходиться в стані готовності повернутись до виконання функцій центрального модуля ЦМК.

Це означає, що основна база привілеїв цього вузла постійно підтримується в актуальному

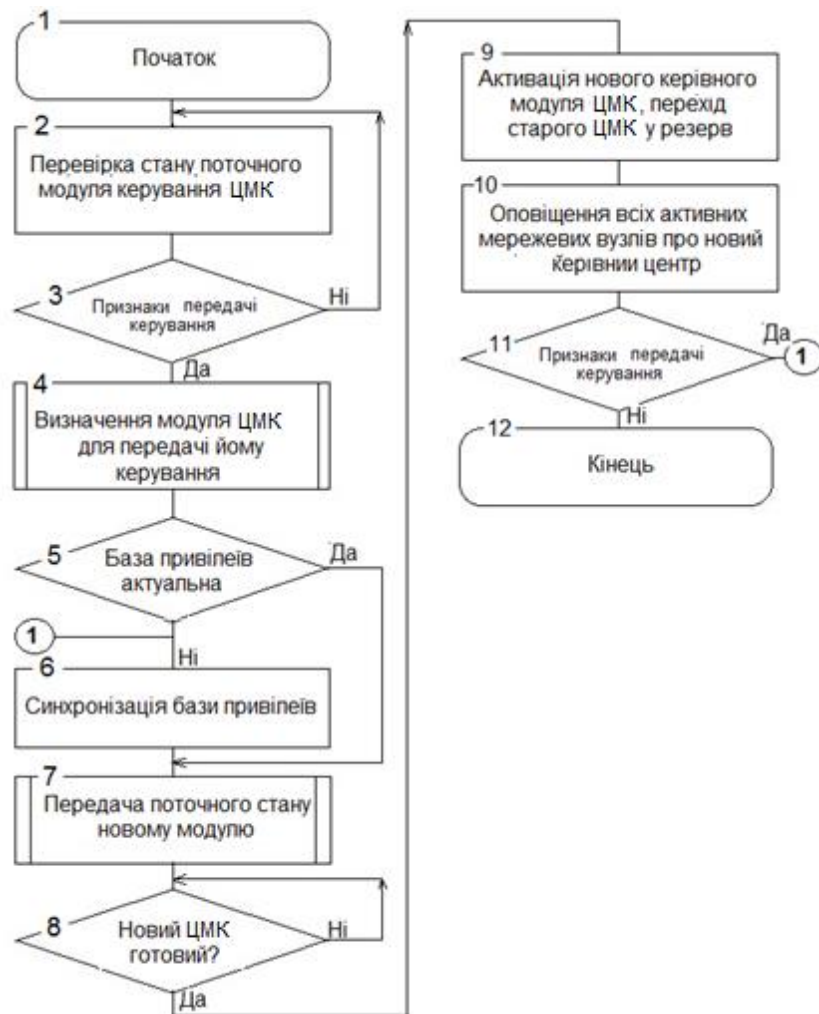


Рис. 3. Алгоритм передачі керування між вузлами комп'ютерної мережі

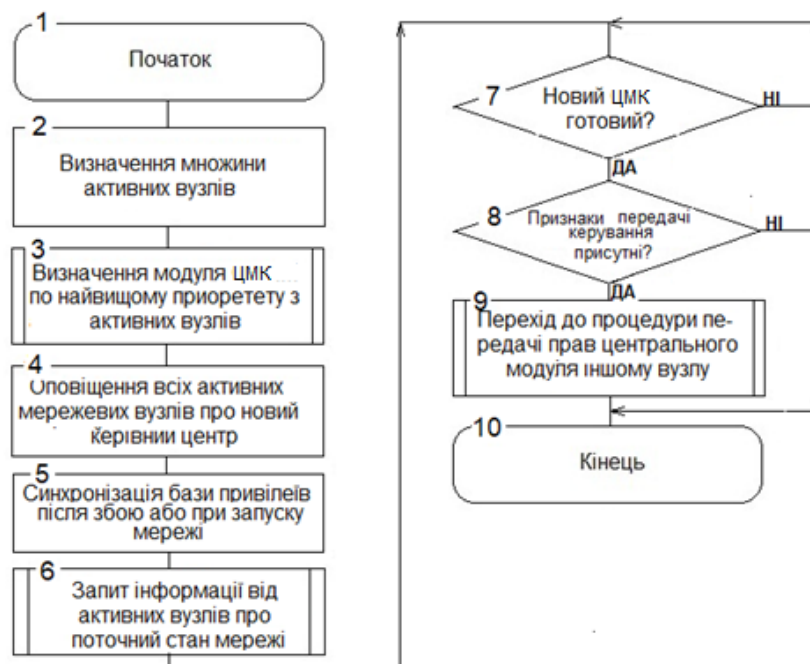


Рис. 4. Алгоритм встановлення керівного центра при старті або при збоях мережі

стані, що дозволяє максимально швидко замінити при необхідності поточний центральний вузол у випадку нештатного розвитку подій (ребро 6–5 рис. 5). При штатному розвитку подій, після наступної ітерації динамічної зміни керівного центра, вузол повернеться до режиму очікування (вершина 3 рис. 5), замикаючи основну послідовність штатних станів, в яких може знаходитись вузол комп'ютерної мережі.

При виникненні нештатної ситуації локального збою, будь який вузол, незважаючи на поточний стан в якому він знаходиться, перейде до аварійного стану (вершина 7 рис. 5). При цьому він буде виключений з числа активних вузлів мережі. З цього стану, відповідно до графа станів (рис. 5) вузол переходить в стан недоступності (вершина 1 рис. 5).

В цьому стані, після усунення причини збою очікує, повторного підключення до мережі. При підключенні до мережі автоматично переходить у стан синхронізації (вершина 2 рис. 5). В цьому стані виконується оновлення до актуального стану локальної бази привілеїв та політик, виконується перевірка узгодженості з іншими вузлами. Завершуються процеси синхронізації включенням вузла до списку активних, а сам вузол переходить у стан очікування (вершина 3 рис. 5).

При запуску мережі або при масштабних збоях відновлення роботи мережі виконується відповідно до алгоритму приведенного на рис. 3. Він враховує ситуації, які можуть виникнути при відновленні роботи мережі після збою. Оскільки при запуску мережі не всі вузли одночасно відновлюють свою роботу в мережі, то перший вузол, який візьме на себе керування буде визначатись як вузол з найвищим його пріоритетом серед вузлів, що на цей момент є активними (оператор 2 рис. 4).

Визначений в такий спосіб вузол, чий модуль ЦМК буде виконувати функцію центрального модуля системи безпеки ОС, генерує керівний токен і розсилає всім іншим активним вузлам мережі, що він є керівний центр системи (оператор 4 рис. 4).

В подальших кроках алгоритм передбачає синхронізацію бази привілеїв з локальними базами активних вузлів з метою відтворення актуальної основної бази привілеїв (оператор 5 рис. 4). Наступним кроком алгоритму є збір інформації про відкритті сесії та іншої інформації про стан мережі (оператор 6 рис. 3). В подальшому алгоритм роботи центрального модуля керування при запуску мережі після збою співпадає з алгоритмом приведеним на рис. 3.

Основна послідовність станів активного вузла мережі описується переходами між вершинами графа станів 3–4–5–6–3–... (Рис. 5). Таким чином кожен вузол має чотири основних стани. Для спрощення розуміння, вершини 4–5 графа можна умовно об'єднати в одну. Тоді основна послідовність переходів графа (Рис. 5) для вузла буде обмежена трьома станами: режим очікування ЦМК_{ОЧ}; режим керівного мережевого центра ЦМК_{КМЦ}; режим резервного модуля ЦМК_{РЕЗ}.

Інші стани графа переходів мають свою спеціалізацію, яка лежить поза основною послідовністю. Так стани, що відповідають вершинам 1 та 2 графа (рис. 5) слугують для виведення вузла на основну траєкторію. Вершині 7 відповідає аварійний стан вузла, який в фізичній системі завжди має місце. Вихід з цього стану можливий через вершини 1 та 2 графа.

Частково-централізована система безпеки ОС базується на застосуванні запропонованого методу динамічної передачі керування між вузлами комп'ютерної мережі. Для оцінки його ефек-

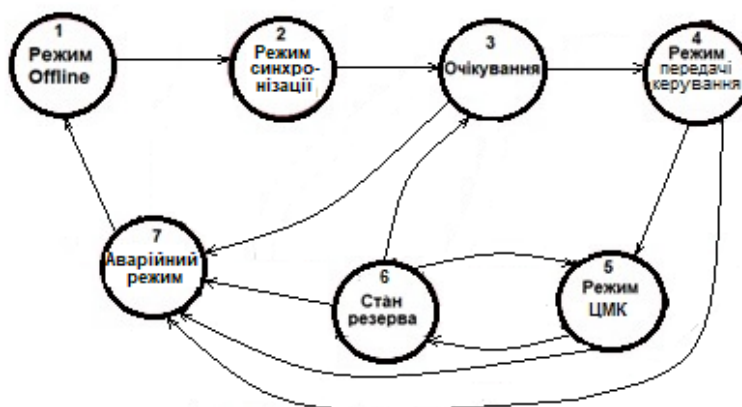


Рис. 5. Граф станів вузла мережі при динамічній передачі керування

тивності скористаємось сукупністю математичних та ймовірнісних моделей, що враховують такі показники: ймовірність витоку конфіденційної інформації, яка враховує ефективність захисних механізмів ОС мережевого вузла та вразливість КС $P_{\text{витоку}}$; час виявлення загрози $T_{\text{вияв. загрози}}$; час обробки операцій безпеки $T_{\text{обробки}}$.

З урахуванням вище сказаного формула ефективності комп'ютерної мережі $E_{\text{системи}}$, в якій будуть поєднані всі три параметри з урахуванням їх ваги та способу впливу на значення ефективності прийме вигляд:

$$E_{\text{системи}} = w_1 \times (1 - P_{\text{витоку}}) + w_2 \times \left(\frac{1}{T_{\text{вияв. загрози}}} \right) + w_3 \times \left(\frac{1}{T_{\text{обробки}}} \right) \quad (10)$$

де: w_1, w_2, w_3 – вагові коефіцієнти параметрів ефективності; $1 - P_{\text{витоку}}$ – ймовірність відсутності витоку конфіденційної інформації; $T_{\text{вияв. загрози}}$ – час виявлення загрози; $T_{\text{обробки}}$ – час обробки операцій безпеки керівним вузлом мережі.

Вираз формули (10) відповідає всім співвідношенням для централізованої системи безпеки ОС $E_{\text{ЦС}}$, тому переписемо її з урахуванням прийнятих для неї позначень:

$$E_{\text{ЦС}} = w_1 \times (1 - P_{\text{витоку ЦС}}) + w_2 \times \frac{1}{T_{\text{вияв. загрози ЦС}}} + w_3 \times \frac{1}{T_{\text{обробки ЦС}}} \quad (11)$$

де: w_1, w_2, w_3 – вагові коефіцієнти параметрів ефективності; $1 - P_{\text{витоку ЦС}}$ – ймовірність відсутності витоку конфіденційної інформації в ОС з централізованою системою безпеки; $T_{\text{вияв. загрози ЦС}}$ – час виявлення загрози в системі з централізованою системою безпеки; $T_{\text{обробки ЦС}}$ – час обробки операцій безпеки керівним вузлом мережі з централізованою системою безпеки.

Оскільки робота частково-централізованої системи основана на використанні метода динамічної передачі керування між вузлами мережі, цей факт має знайти своє відображення в формулі її ефективності. Цей метод зменшує вразливість системи безпеки ОС, але одночасно, зменшує в деякій мірі і продуктивність системи загалом. Динамічна зміна керівного центра потребує витрат часу на синхронізацію ресурсів безпеки, що підлягають централізованому керуванню. Позначимо його як $T_{\text{витрати синхр. ЧЦС}}$. З урахуванням цього параметра формула ефективності для частково-централізованої системи $E_{\text{ЧЦС}}$ прийме вигляд:

$$E_{\text{ЧЦС}} = w_1 \times (1 - P_{\text{витоку ЧЦС}}) + w_2 \times \frac{1}{T_{\text{вияв. загрози ЧЦС}}} + w_3 \times \frac{1}{T_{\text{обробки ЧЦС}} + T_{\text{витрати синхр. ЧЦС}}} \quad (12)$$

де: w_1, w_2, w_3 – вагові коефіцієнти параметрів ефективності; $1 - P_{\text{витоку ЧЦС}}$ – ймовірність відсутності витоку конфіденційної інформації в ОС з частково-централізованою системою безпеки; $T_{\text{вияв. загрози ЧЦС}}$ – час виявлення загрози в системі з частково-централізованою системою безпеки; $T_{\text{обробки ЧЦС}}$ – час обробки операцій безпеки керівним вузлом мережі з частково-централізованою системою безпеки; $T_{\text{витрати синхр. ЧЦС}}$ – витрати часу на синхронізацію централізованих ресурсів безпеки.

Тепер порівняємо показники ефективності частково-централізованої та централізованої систем безпеки ОС:

$$E_{\text{ЧЦС}} - E_{\text{ЦС}} > 0 \quad (13)$$

Таким чином, якщо вираз (13) буде матиме значення більше нуля, то це означитиме, що метод динамічної передачі керівного центра між вузлами забезпечує більшу ефективність системи в порівнянні з централізованою системою безпеки.

З метою визначення ефективності метода динамічної передачі керування було проведено кілька експериментів. Для цього було розгорнуто тестове середовище основане на використанні віртуальної комп'ютерної мережі до якої входять віртуальні вузли VM01 – VM04, які слугували під час експериментів цільовими машинами, а також атакуюча машина VM00, з допомогою програмного забезпечення якої будуть симулюватись різномітні атаки по можливих шляхах витоку конфіденційної інформації. На атакуючій машині встановлена ОС Kali Linux та симулятор атак Metasploit Framework. Цільові машини керуються ОС FreeBSD 13.1 з імплементованою централізованою системою безпеки. Всі віртуальні машини працюють під управлінням гіпервізора Virtual Box, що об'єднує їх у відокремлену локальну віртуальну мережу, необхідну для проведення експериментів.

Метою експерименту є отримання даних стосовно комплексної стійкості захисних механізмів під впливом загроз, що діють по можливих шляхах витоку інформації, які враховують такі параметри як ймовірність витоку інформації, час виявлення атаки ЗПЗ, продуктивність центрального вузла мережі та час передачі керування між вузлами мережі. Експеримент включає симуляцію

комп'ютерних різнотипних атак, направлених по всіх можливих шляхах витоку конфіденційної інформації всіх вузлів віртуальної мережі.

Для симуляції атак різних видів використовувався Metasploit Framework атакуючої машини VM00. Під його управлінням запускався скрипт-файл, який містить інструкції на мові Ruby, алгоритм якого включає циклічну симуляцію атак по всіх основних шляхах витоку по всіх вузлах мережі. В момент роботи скрипт-файла на всіх цільових машинах були відкриті файли /tmp/k1/test.txt, що імітують ресурси з конфіденційною інформацією. Результатами експерименту є інформація про кількість успішних та неуспішних атак на відповідні вузли віртуальної комп'ютерної мережі.

Окрім вище приведеної інформації на стороні атакуючої машини VM00 в файлі Result_attack.txt фіксувались дані початку кожної атаки, IP-адреса цільової машини, а також тип атаки. Якщо система безпеки відповідної цільової машини, на яку був спрямований вектор атаки, виявить загрозу, то після її блокування по факту цієї події зробить записи в своїх системних журналах security.log та audit.log. Різниця між часом старту симуляції зафіксованої в файлі Result_attack.txt атакуючої машини та часом її виявленням зафік-

сованим в журналі security.log відповідної цільової машини буде відповідати терміну виявлення загрози.

Результати співставлення даних з файлу Result_attack.txt атакуючої машини VM00 (рис. 6) та файлів security.log (рис. 7) відповідних цільових машин VM01-VM04 представлені в таблиці 1 в усередненому та унормованому вигляді. Оскільки атакуюча машина та цільові машини знаходяться в одному часовому просторі фізичної машини, то похибка вимірювання тривалості періоду виявлення атаки буде мінімальною. Ситуація коли запис початку симуляції, зафіксований в файлі Result_attack.txt, немає зв'язаного з ним запису в журналах security.log всіх цільових машин, означає, що мала місце успішна атака. Приклад співставлення показано на рис. 6 та рис. 7.

Ще один етап експериментів мав на меті отримання даних, що дозволять визначити продуктивність роботи керівного вузла комп'ютерної мережі $T_{\text{прод.ЦМК}}$ в умовах централізованого керування. При цьому вимірювався час виконання запитів центральним вузлом мережі при виключенні механізмів захисту в одному випадку і включенні в другому, позначених як $T_{\text{викл.МЗ}}$ та $T_{\text{вкл.МЗ}}$ відповідно. Тоді продуктивність центрального вузла мережі визначиться як:

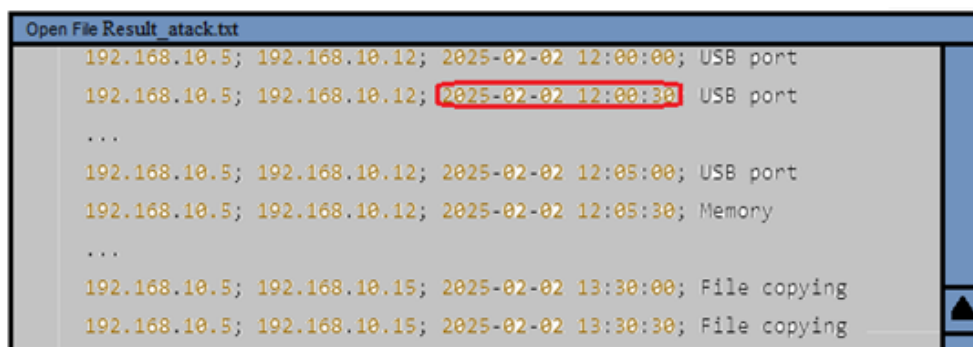


Рис. 6. Вміст файлу Result_attack.txt згенерованого скрипт-файлом симулятора Metasploit Framework

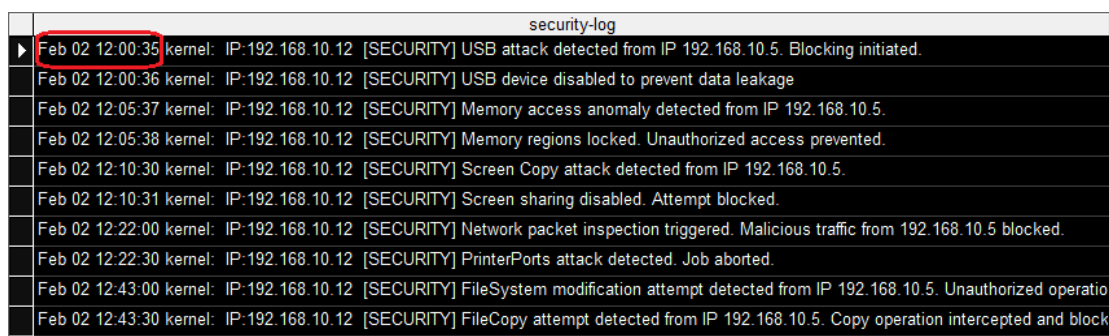


Рис. 7. Вміст файлу security.log цільової машини VM1

$$T_{\text{прод. ЦМК}} = \frac{T_{\text{викл. МЗ}}}{T_{\text{вкл. МЗ}}} \quad (14)$$

Опрацьовані дані експерименту приведені в таблиці 1.

1.1. Розрахунки ефективності централізованої та частково-централізованої системи безпеки ОС

Скористаємось формулою (14) та даними експериментів розрахуємо продуктивність центрального вузла для централізованої системи $T_{\text{прод. ЦС}}$, для чого підставимо в неї відповідні значення з таблиці 1:

$$T_{\text{прод. ЦС}} = \frac{T_{\text{викл. МЗ}}}{T_{\text{вкл. МЗ}}} = \frac{0,126}{0,142} = 0,887 \quad (15)$$

Так само розраховується продуктивність центрального вузла для частково-централізованої системи $T_{\text{прод. ЧЦС}}$:

$$T_{\text{прод. ЧЦС}} = \frac{T_{\text{викл. МЗ}}}{T_{\text{вкл. МЗ}}} = \frac{0,117}{0,127} = 0,921 \quad (16)$$

Для розрахунку ефективності системи безпеки скористаємось формулою (10). Для цього визначимо значення вагових коефіцієнтів $w_1 - w_3$ присутніх в даній формулі. Щоб визначитись з їх вагою скористаємось методом лінійної ієрархії в рамках якого застосуємо шкалу важливості від одного до дев'яти: 1 – рівна важливість; 3 – слабка перевага одного фактора над іншим; 5 – сильна перевага; 7 – дуже сильна перевага; 9 – абсолютна перевага. Керуючись цією шкалою будуюмо матрицю парних порівнянь:

Керуючись цією шкалою будуюмо матрицю парних порівнянь:

$$M(w) = \begin{bmatrix} 1 & 3 & 5 \\ 1/3 & 1 & 3 \\ 1/5 & 1/3 & 1 \end{bmatrix}$$

Стовпці та рядки матриці є ваговими коефіцієнтами $w_1 - w_3$, а її елементи є результатом попарного співставлення. Виконавши подальші кроки отримаємо: $w_1=0,633$; $w_2=0,261$; $w_3=0,106$. Тепер можна безпосередньо розрахувати ефективність

системи безпеки, для чого підставимо значення з таблиці 1 в формулу (11) для централізованої системи безпеки ОС:

$$E_{\text{ЦС}} = 0,633 * (1 - 0,0133) + 0,261 * \frac{1}{0,4} + 0,106 * \frac{1}{0,887} = 1,397$$

Аналогічно для частково-централізованої системи безпеки ОС скористаємось формулою (12):

$$E_{\text{ЧЦС}} = 0,633 * (1 - 0,0115) + 0,261 * \frac{1}{0,308} + 0,106 * \frac{1}{0,921 + 0,688} = 1,5389$$

Використаємо нерівність (13) для порівняння розрахованих значень ефективності двох варіантів систем безпеки ОС:

$$E_{\text{ЧЦС}} - E_{\text{ЦС}} = 1,5389 - 1,397 = 0,142$$

З неї видно, що ефективність частково-централізованої системи безпеки $E_{\text{ЧЦС}}$ в порівнянні з централізованою системою вища одиниці. Для зручності відобразимо отриманий результат у відсотках:

$$\left(1 - \frac{E_{\text{ЦС}}}{E_{\text{ЧЦС}}}\right) * 100\% = \left(1 - \frac{1,397}{1,539}\right) * 100\% = 9,2\%$$

Виконаний розрахунок показав, що ефективність частково – централізованої системи безпеки ОС, незважаючи на її складність, є помітно вищою. В свою чергу вища ефективність такої системи безпеки означає більшу стійкість до атак ЗПЗ, що в свою чергу забезпечує нижчу вразливість до витоку конфіденційної інформації.

Висновки. Запропонований метод динамічної передачі керування між вузлами мережеских вузлів з підтримкою реплікації основної бази привілеїв для частково-централізованої системи безпеки ОС дозволив зберегти основні переваги централізованої системи безпеки, особливо, що стосується стійкості системи до витоку інформації і, одночасно позбутись найбільшої її вади – вузького місця в системі безпеки у вигляді центрального

Таблиця 1

Усереднені дані експериментів (абсолютне значення/унормоване значення)

Система безпеки ОС	$P_{\text{витоку}}$	$T_{\text{вияв. загрози}}$ сек	$T_{\text{викл. МЗ}}$ сек	$T_{\text{вкл. МЗ}}$ сек	$T_{\text{витрати синхр.}}$ сек
Централізована	0,0133	5,2 (0,4)	0,126 (0,887)	0,142	-
Частково-централізована	0,0115	4,925 (0,358)	0,117 (0,921)	0,127	2,118 (0,688)

модуля безпеки, вихід з ладу якого призводить до втрати контролю над всією системою.

Ефективність роботи системи безпеки ОС, що використовує даний метод керування комп'ютерною мережею в порівнянні з чисто централізованим підходом вища щонайменше на 9,2 відсотка.

Ще однією перевагою цього методу динамічної передачі керування є те, що періодична ран-

домна зміна керівного вузла мережі з великою долею вірогідності робить неактуальними результати проведення розвідки мережі зловмисником, чим зриває атаки на мережу ще до їх початку. Це пояснюється тим, що час розвідки мережі значно більший періоду зміни керівного центра мережі, що дає йому можливість таким чином ухилятися від атак.

Список літератури:

1. Security and Privacy Controls for Information Systems and Organizations / *National Institute of Standards and Technology Walter Copan*, NIST Director and Under Secretary of Commerce for Standards and Technolog, Publication 800-53 Revision 5. DOI:10.6028/NIST.SP.800-53r5
2. Hajimirzaei B., Navimipour N.J. Intrusion detection for cloud computing using neural networks and artificial bee colony optimization algorithm. *ICT Express*, 2019. Vol 5(1). P. 56–59. DOI:10.1016/j.ict.2018.01.014
3. Kim Y., Park G., Kim H. Domain knowledge free cloud-IDS with lightweight embedding method. *J Cloud Comp.* 2024. № 13, p.143. DOI:10.1186/s13677-024-00707-8
4. Al-Ambusaidi M., Yinjun Z., Muhammad Y. ML-IDS: an efficient ML-enabled intrusion detection system for securing IoT networks and applications. *Soft Comput.* 2024. 28. p.1765–1784. DOI: 10.1007/s00500-023-09452-7
5. Pourardebil Khah Y., Hosseini Shirvani M., Motameni H. A hybrid machine learning approach for feature selection in designing intrusion detection systems (IDS) model for distributed computing networks. *J Supercomput.* 2025. Vol. 81. 254. DOI: 10.1007/s11227-024-06677-7
6. Lu H., Zhao Y., Song Y., Yang Y., He G., Yu H., Ren Y. A transfer learning-based intrusion detection system for zero-day attack in communication-based train control system. *Cluster Comput.* 2024. Vol. 27. p. 8477–8492. DOI: 10.1007/s10586-024-04376-9
7. Srilatha D., Shyam G.K., Cloud-based intrusion detection using kernel fuzzy clustering and optimal type-2 fuzzy neural network. *Cluster Comput.* 2021. Vol 24, P. 2657–2672 DOI: 10.1007/s10586-021-03281-9
8. Hajimirzaei B., Navimipour N.J. Intrusion detection for cloud computing using neural networks and artificial bee colony optimization algorithm. *ICT Express*. 2019. Vol 5. № 1. P. 56–59 DOI: 10.1016/j.ict.2018.01.014
9. Lang F., Li H., Wang W., Lin J., Zhang F., Pan W., Wang Q. E-SGX: Effective Cache Side-Channel Protection for Intel SGX on Untrusted OS / Y. Wu, M. Yung et al. Information Security and Cryptology. Inscrypt 2020. *Lecture Notes in Computer Science*. 2021. Vol 12612. P. 221–243 DOI: 10.1007/978-3-030-71852-7_15
10. An L., Qiu J., Zhang H., Liu C. Design of distributed network intrusion prevention system based on Spark and P2DR models. *Cluster Comput.* 2024. P. 10757–10776. DOI: 10.1007/s10586-024-04487-3
11. Jain A.K., Shukla H., Goel D. A comprehensive survey on DDoS detection, mitigation, and defense strategies in software-defined networks. *Cluster Comput.* 2024. Vol. 27. P. 13129–13164 DOI:10.1007/s10586-024-04596-z
12. Mushtaq M., Yousaf M.M., Bhatti M.K., Lapotre V., Gogniat, G. The Kingsguard OS-level mitigation against cache side-channel attacks using runtime detection. *Ann. Telecommun.* 2022. Vol 77. p. 731–747. DOI: 10.1007/s12243-021-00906-3
13. Kashtalian A., Lysenko S., Savenko O., Nicheporuk A., Sochor, T., Avsiyevych V. Multi-computer malware detection systems with metamorphic functionality. *Radioelectronic and Computer Systems*. 2024. № 1. p. 152–175. DOI: 10.32620/reks.2024.1.13
14. Savenko B., Kashtalian A., Lysenko S., Savenko O. Malware Detection By Distributed Systems with Partial Centralization. 2023. IEEE 12th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS). Germany. Dortmund. 2023. p. 265–270, DOI: 10.1109/IDAACS58523.2023.10348773
15. Soltani M., Khajavi K., Jafari Siavoshani M. A multi-agent adaptive deep learning framework for online intrusion detection. *Cybersecurity*. 2024. Vol. 7. № 9. P. 1–25 DOI:10.1186/s42400-023-00199-0
16. Wang S., Jing Y., Wang K., Wang X., Multi-Agent Systems for Collaborative Inference Based on Deep Policy Q-Inference Network. *J Grid Computing*. 2024. Vol. 22. № 38. DOI:10.1007/s10723-024-09750-w
17. Karataiev O. Towards multi-agent platform development. *Computer Systems and Information Technologies*. 2024. № 4, p. 98–106. DOI: 10.31891/csit-2024-4-12
18. N. Pradeep, Vaishnavi, V. Varsha, K. Vivek. The Application of Artificial Intelligence in Computer Network Technology. *International Journal of Advanced Research in Science Communication and Technology*. 2024. Vol. 4. P. 588–592. DOI: 10.48175/IJARSCT-22770

19. Lysenko S., Savenko O., Bobrovnikova K., Kryshchuk A. Self-adaptive system for the corporate area network resilience in the presence of botnet cyberattacks. *Communications in Computer and Information Science*. 2018. № 860. P. 385–401.
20. Srividhya G., Nagarajan R. A novel bidirectional LSTM model for network intrusion detection in SDN-IoT network. *Computing*. 2024. 106, p. 2613–2642 DOI: 10.1007/s00607-024-01295-w
21. Temene N., Naoum A., Sergiou C., Georgiou C., Vassiliou V. A fault tolerant node placement algorithm for WSNs and IoT networks. *Computer Networks*. 2024. Vol 254. № 110835. DOI: 10.1016/j.comnet.2024.110835
22. Zhang Y., Wang L., Liu D., Su Y., Zhu Y., Zhang X. Design of Information Security Protection System for Cloud Business System. Proceedings of International Conference on Artificial Intelligence and Communication Technologies. Singapore. 2023. Smart Innovation. *Systems and Technologies*. 2024. Vol 369. P. 105–122. DOI: 10.1007/978-981-99-6956-2_10
23. Yamauchi T., Akao Y., Yoshitani R., Nakamura Y., Hashimoto M. Additional kernel observer: privilege escalation attack prevention mechanism focusing on system call privilege changes. *Int. J. Inf. Secur.* 2021. № 20. P. 461–473. DOI: 10.1007/s10207-020-00514-7
24. Ying L., Jiuqi W., Ziyu F., Yufan F., Xiaodong L. DiSAuth: A DNS-based secure authorization framework for protecting data decoupled from applications. *Computer Networks*. 2024. Vol. 254. № 110774. DOI: 10.1016/j.comnet.2024.110774
25. Zhou Q., Jia X., Chen J., Huang Q., Du H. LightArmor: A Lightweight Trusted Operating System Isolation Approach for Mobile Systems. ICT Systems Security and Privacy Protection. SEC 2024. *IFIP Advances in Information and Communication Technology*. 2024. Vol. 710. P. 206–220. DOI: 10.1007/978-3-031-65175-5_15
26. Harmonizing Security and Performance in Microkernel File Servers / W.T. Li, Z.X. Wang, J.Y. Gu et al. *J. Comput. Sci. Technol.* 2025. № 40, p. 464–481. DOI: 10.1007/s11390-025-3762-3
27. Stetsyuk Y., Stetsyuk M., Kyrylo V., Paiuk V., Kvassay M. A model of a centralized security system, as an information technology for the synthesis of an OS architecture protected against the leakage of confidential information. 1st International Workshop on Advanced Applied Information Technologies (AdvAIT-2024) will be held in Khmelnytskyi, Ukraine and Zilina, Slovakia on December 5. 2024. Vol. 3899. P. 224–233.
28. Stetsyuk Y., Stetsyuk M., Savenko B., Kwiecien A., Kopania L. Method of random dynamic control transfer between network nodes for a partially centralized OS security system. The 6th International Workshop on Intelligent Information Technologies & Systems of Information Security. April 04. 2025. Khmelnytskyi. Ukraine. Vol. 3963. P. 264–283.
29. Kashtalian A., Lysenko S., Kysil T., Sachenko A., Savenko O., Savenko B. Method and Rules for Determining the Next Centralization Option in Multicomputer System Architecture. *International Journal of Computing*. 2025. № 24(1). P. 35–51. DOI: 10.47839/ijc.24.1.3875
30. Kashtalian A., Lysenko S., Sachenko A., Savenko B., Savenko O., Nicheporuk A. Evaluation criteria of centralization options in the architecture of multicomputer systems with traps and baits. *Radioelectronic and Computer Systems*. 2025. № 1. p. 264–297. doi: 10.32620/reks.2025.1.18
31. Ajish D. The significance of artificial intelligence in zero trust technologies: a comprehensive review. *Journal of Electrical Systems and Inf Technol.* 2024. 11. № 30. DOI: 10.1186/s43067-024-00155-z

Stetsyuk Yu.V., Savenko O.S. PARTIALLY CENTRALIZED NETWORK OS SECURITY SYSTEM WITH DYNAMIC CONTROL TRANSFER BETWEEN ITS NODES

The current state of the threat level from malicious software (MSS) to modern protected specialized OSs is analyzed, in relation to their resistance to leaks of confidential information. A model of interaction of protective mechanisms of a centralized security system of a network node OS is presented, which allows obtaining an OS security architecture devoid of the problem of leakage of confidential information during attacks on the system's RAM, which is achieved by including a marking mechanism in it, the feature of which is the control of channels for passing confidential information at the level of memory pages. The principles of building centralized and partially centralized OS security systems and the principles of organizing the operation of their security mechanisms are considered. To preserve the positive qualities of a centralized security system and at the same time avoid the problems inherent in any centralization, a method of dynamic transfer of control between computer network nodes is proposed as a way to increase its resistance to leakage of confidential information and eliminate the bottleneck of a centralized security system – a single control center, as well as to counteract reconnaissance by attackers of a computer network in order to disrupt attacks on its resources. Its feature is the synchronization and updating of the central database of privileges and the state of the network by replicating the local databases of privileges of network nodes in the process of transferring control between them. Another feature of the method is to ensure the constant availability of a backup control node in case of failure. Each node that has transferred control of the network to a successor determined in a random way goes

into reserve. A generalized algorithm is proposed that implements the steps of the method of dynamic transfer of control between computer network nodes during normal network operation and for the situation of system start or failure. A model of the states of a network node in the process of dynamic control transfer between nodes has been developed, which allows it to be studied in all possible situations. A comparative analysis of the effectiveness of centralized and partially centralized security systems has been performed.

Key words: security system, centralized system, partially centralized system, operating system, state model.

Дата надходження статті: 31.07.2025

Дата прийняття статті: 28.08.2025

Опубліковано: 27.10.2025

Стоцький І.В.

Військовий інститут телекомунікацій та інформатизації імені Героїв Крут

Устинов Д.А.

Військовий інститут телекомунікацій та інформатизації імені Героїв Крут

Налісний Г.І.

Військовий інститут телекомунікацій та інформатизації імені Героїв Крут

Костенко В.П.

Військовий інститут телекомунікацій та інформатизації імені Героїв Крут

НЕЙРОМЕРЕЖЕВІ МОДЕЛІ ПРОГНОЗУВАННЯ ЛОГІСТИЧНИХ ПОТРЕБ В УМОВАХ БОЙОВИХ ДІЙ

Стаття присвячена дослідженню актуальної проблематики застосування нейромережеских моделей для прогнозування логістичних потреб в умовах бойових дій. Військовий конфлікт в Україні, що триває з 2014 року та особливо загострився після повномасштабного вторгнення 2022 року, актуалізував потребу у модернізації підходів до військової логістики. Динамічність змін на лінії фронту, руйнування інфраструктури та ускладнені умови постачання зумовили необхідність переходу від традиційних централізованих схем забезпечення до гнучких, децентралізованих і технологічно підтриманих систем. Стаття розглядає ключові напрями змін у військовій логістиці України, включаючи впровадження цифрових платформ, мобільних складів, безпілотних технологій доставки та використання інтелектуальних систем підтримки прийняття рішень.

Основну увагу зосереджено на можливостях використання штучного інтелекту, зокрема нейронних мереж різних архітектур, для прогнозування логістичних потреб військових підрозділів. Автори аналізують переваги рекурентних нейронних мереж (RNN) і моделей довготривалої короткочасної пам'яті (LSTM) для обробки часових рядів даних про витрати ресурсів, переміщення військ та інтенсивність бойових дій. Окремо розглядаються трансформерні архітектури, здатні ефективно моделювати взаємозв'язки між різними типами логістичних потреб, що дозволяє створювати більш точні й комплексні прогнози.

У статті окреслено процес побудови моделей: від збору гетерогенних даних – оперативних звітів, інформації з розвідувальних джерел, супутникових знімків, погодних умов – до обробки та навчання моделей із застосуванням методів нормалізації, фільтрації та боротьби з аномаліями. Також звертається увага на складнощі роботи з неповними або спотвореними даними, характерними для умов війни, та пропонуються способи їх подолання шляхом використання гібридних моделей, що поєднують нейромережі й класичні математичні підходи.

Особливий акцент зроблено на прикладних аспектах впровадження таких рішень у структури військового управління. Розглядаються питання інтеграції прогнозних моделей у системи командування і контролю, створення модульних інформаційних платформ, що автоматично оновлюють дані у режимі реального часу, а також необхідність розвитку користувацьких інтерфейсів для командирів підрозділів. Показано, що впровадження таких технологій може суттєво підвищити швидкість і точність ухвалення рішень у логістиці навіть в умовах високої невизначеності.

Стаття також висвітлює практичні результати застосування нейромережеских підходів під час реальних операцій на території України. Наводяться приклади успішного використання RNN-моделей для прогнозування сплесків потреб у медичному забезпеченні та трансформерів для передбачення змін у витраті пального внаслідок погодних умов. Аналізуються результати впровадження децентралізованих логістичних хабів і мобільних систем доставки як відповідь на виклики руйнування основної інфраструктури.

У підсумку наголошується на потенціалі подальшого розвитку технологічної бази військової логістики шляхом активного використання автоматизації, безпілотних технологій, штучного інтелекту та цифрових платформ. Окреслено перспективи побудови повністю інтегрованих, самонавчальних логістичних систем, здатних оперативно реагувати на зміни ситуації у бойових умовах та мінімізувати людські втрати.

Таким чином, стаття робить вагомий внесок у наукове обґрунтування необхідності застосування нейромережових технологій у військовій логістиці, пропонуючи комплексний підхід до прогнозування потреб, що має велике практичне значення для забезпечення стійкості та ефективності військових операцій у сучасних збройних конфліктах.

Ключові слова: нейромережові моделі, військова логістика, рекурентні мережі, моделі трансформери, генетичні алгоритми, оборонні технології, прогнозування логістичних потреб.

Постановка проблеми. Сучасний стан бойових дій, особливо в умовах гібридних воєн і високотехнологічного збройного протистояння, потребує принципово нового підходу до планування, організації та забезпечення логістичних процесів. Надзвичайна динаміка змін на фоні воєнних дій, необхідність швидкого переміщення ресурсів, реагування на непередбачувані ситуації та обмеженість у часі вимагають від логістичних структур гнучкості, точності та здатності адаптуватися в режимі реального часу. Традиційні методи логістичного планування, побудовані на заздалегідь заданих маршрутах, нормативних обсягах і шаблонних сценаріях, виявляються недостатніми, коли бойові дії ведуться у складних географічних, погодних та інформаційних умовах, а противник активно впливає на логістичну інфраструктуру [1].

Таким чином, стає нагальною потреба у використанні інтелектуальних інформаційних систем, здатних не лише збирати та аналізувати великі обсяги даних у реальному часі, а й прогнозувати майбутні потреби в ресурсах, оптимізувати маршрути постачання, передбачати ризики та приймати рішення з урахуванням багатьох змінних. Одним із найперспективніших напрямів є застосування нейромережових моделей, які демонструють високу ефективність у складних, багатофакторних і непередбачуваних середовищах, таких як логістика в умовах бойових дій [2].

Аналіз останніх досліджень і публікацій. У контексті військового конфлікту, що розпочався в Україні у 2014 році і набув характеру повномасштабної війни з 2022 року, з'явилася велика кількість наукових праць, присвячених вивченню змін у системах військової логістики та пристосуванню логістичних процесів до реальних бойових дій. У своїх публікаціях дослідники підкреслюють критичну необхідність швидкого реагування логістичних структур на зміни ситуації на фронті, включно з динамікою бойових

ліній зокрема, це відзначає Бурба А.В. У роботах Петренка І.В., можна ознайомитися з інноваційними рішеннями, такими як використання дронів для доставки, розгортання мобільних складів та впровадження цифрових логістичних платформ, які дедалі ширше застосовуються для оптимізації процесів забезпечення та зменшення ризиків для персоналу. Особливу роль у новій логістичній архітектурі відіграють локальні хаби постачання, що стали незамінними в умовах руйнації традиційної інфраструктури, актуальність яких описує Жуков О.О.

Окрему увагу українські та міжнародні науковці приділяють вивченню впливу війни на економічні процеси в цілому. Цей аспект розглядали у своїх роботах Т. Богдан, О. Борзенко, В. Горбулін, а також зарубіжні економісти, серед яких Джон Мейнард Кейнс та Нуріель Рубіні. У свою чергу, формування ефективних логістичних ланцюгів і систем забезпечення вивчали Б.М. Андрушківа, А.В. Євдокимова, Є.В. Крикавський, Т.В. Косарева, Г.І. Михайліченко та багато інших науковців. Попри значну кількість досліджень, що накопичилися у даній галузі, нові умови вимагають перегляду існуючих моделей та розробки актуальних рішень, які враховують нестабільну ситуацію у країні та світі [3, 4].

У світовій науковій літературі останніх років спостерігається зростаючий інтерес до застосування штучного інтелекту, зокрема нейронних мереж, у сфері логістики. У працях західних дослідників, таких як D. Simchi-Levi, M. Christopher, A. Tversky, розглядаються методи прогнозування попиту, оптимізації маршрутів, управління запасами за допомогою глибоких нейронних мереж, рекурентних нейромереж (RNN), моделей довготривалої короткочасної пам'яті (LSTM) та трансформерів. Зокрема, LSTM-моделі визнані ефективними при роботі з часовими рядами, що надзвичайно актуально для динамічного аналізу потреб у постачанні.

У межах військової логістики певні дослідження виконувались оборонними агенціями, зокрема DARPA, а також вченими з військових академій США, Великобританії, Ізраїлю. Їхні роботи зосереджені на інтелектуальних системах підтримки прийняття рішень, моделюванні логістичних потоків та симуляціях у реальному часі. Окремі публікації розкривають потенціал гібридних моделей, які поєднують нейромережі з класичними методами статистичного аналізу або алгоритмами оптимізації, такими як генетичні алгоритми.

В українському науковому просторі також почали з'являтися праці, присвячені цифровізації логістичних процесів в оборонній сфері. Публікації в журналах «Системи озброєння і військова техніка», «Наука і оборона» порушують питання кібербезпеки, інтелектуального аналізу даних та автоматизованих систем керування логістикою. Проте системні розробки, присвячені застосуванню нейромереж саме в контексті бойових дій, залишаються малочисельними, що створює наукову нішу для проведення детальних досліджень у цій галузі [5, 6].

Постановка завдання. Мета статті – дослідження можливостей застосування нейромережових моделей для прогнозування логістичних потреб в умовах бойових дій.

Виклад основного матеріалу. Необхідність запровадження новітніх логістичних підходів у Збройних силах України під час повномасштабної війни є критично важливим завданням для забезпечення військової ефективності на полі бою. Досліджуючи основні економічні показники бачимо, що до початку повномасштабного вторгнення більшість зовнішньоекономічних операцій України – приблизно дві третини обсягів експорту та імпорту – здійснювалися через морські порти. Проте блокада ключових морських артерій з боку росії практично паралізувала цей шлях транспортування. Наприклад, до 24 лютого 2022 року транспортування зерна з центральних областей до портів Одеси коштувало близько 30 доларів за тону. Нині ж вартість доставки через порти сусідніх держав подекуди сягає 180 доларів за тону, що становить близько двох третин від ринкової вартості самого продукту. Це різко знижує економічну доцільність його виробництва.

Значну частину цих витрат формують внутрішні залізничні перевезення. Ціна логістики по залізниці складається з кількох компонентів: тарифів, вагонного парку, логістичного часу, маршрутного плеча, стану колій та інфраструктури. Саме

ці чинники стали вирішальними у зростанні вартості доставки в умовах обмеженого доступу до морських портів та нестачі альтернативних шляхів транспортування.

Ще однією проблемою стала вразливість стаціонарних логістичних баз, які раніше були зосереджені у центральній частині України. Зокрема, в Київській області в перші тижні повномасштабної агресії стрімке просування окупаційних військ призвело до ізоляції низки складів, багато з яких були зруйновані під час бойових дій. Це спровокувало докорінну перебудову системи логістичного забезпечення – перерозподілу складів, створення резервних маршрутів та запровадження децентралізованої моделі зберігання і доставки.

Основу сучасного підходу до прогнозування логістичних потреб в умовах бойових дій становить концепція адаптивної логістики, яка передбачає безперервне оновлення моделей постачання на основі зміни ситуації на полі бою. У цьому контексті нейромережові моделі стають ключовим інструментом, здатним перетворювати великі обсяги динамічної інформації на конкретні прогнози щодо кількості, складу, часу і маршруту доставки необхідних ресурсів. Загальний процес навчання нейронної мережі складається із трьох етапів: пряме поширення; розрахунок функції втрат; зворотне поширення. Графічна інтерпретація навчання нейронної мережі представлено на Рис. 1 [2, 4].

Першим кроком у побудові таких моделей є формування якісного набору даних. У військових умовах ці дані мають гетерогенну природу: з одного боку, це структурована інформація з логістичних баз даних про попередні операції, витрати пального, кількість залученої техніки, маршрути руху; з іншого – неструктуровані джерела, включаючи оперативні звіти, повідомлення розвідки, погодні умови, супутникові зображення та навіть текстові повідомлення польових командирів. Нейронні мережі, особливо глибокі моделі та трансформери, дозволяють об'єднувати ці джерела в єдину модель прогнозування.

Рекурентні нейронні мережі (RNN) і LSTM (Long Short-Term Memory) демонструють особливу ефективність у роботі з часовими рядами. Наприклад, на основі даних про споживання пального за останні 30 діб певного батальйону модель може передбачити майбутню динаміку споживання з урахуванням оперативних планів, географічного рельєфу, погодних умов та інтенсивності бойових дій. LSTM-моделі також дозволяють прогнозувати попит на медичне забезпечення, аналізуючи кількість поранених у попередні пері-

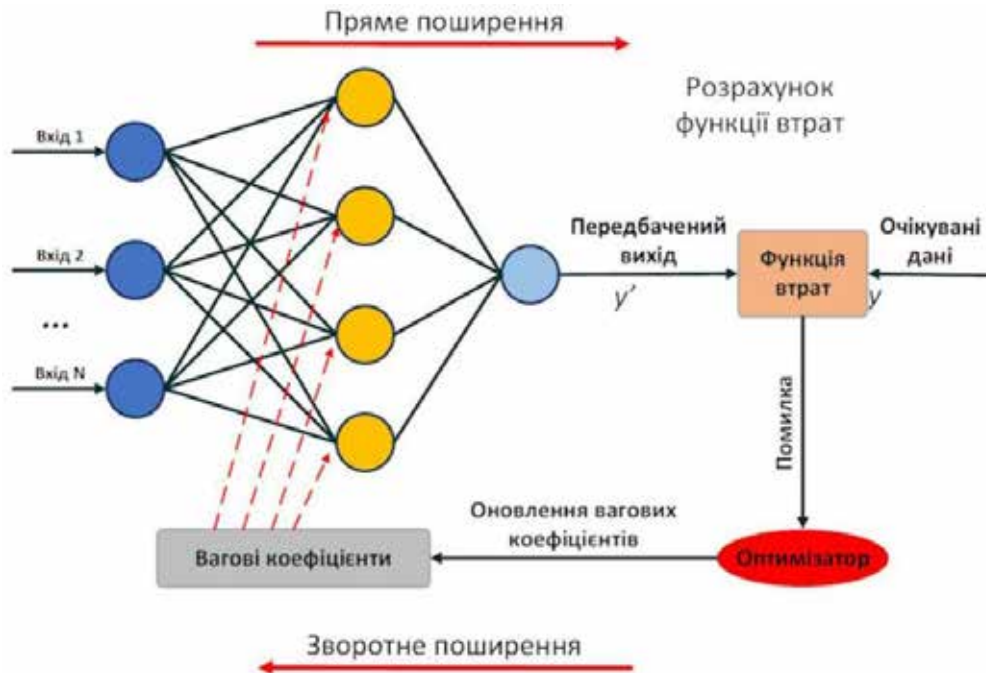


Рис. 1. Візуалізація навчання нейронної мережі

оди, типи застосованого озброєння противника та близькість до зони евакуації.

Трансформерні архітектури, що набули поширення завдяки моделям на кшталт BERT, GPT та їх адаптаціям, надають нові можливості для прогнозування, особливо в багатовимірних просторах. Вони здатні враховувати взаємозалежності між різними типами постачання: наприклад, збільшення потреб у пальному може корелювати зі зростанням потреб у запасних частинах, боеприпасах та харчуванні. Таким чином, замість ізольованого прогнозу окремих компонентів, модель буде синергетичну картину логістичних потреб.

Одним із викликів є адаптація моделей до обмежених і зашумлених даних, що типово для військових умов. Для цього використовуються методи попередньої обробки, включаючи нормалізацію, інтерполяцію, фільтрацію та автоматичне виявлення аномалій. Крім того, доцільно комбінувати нейромережі з класичними математичними моделями – наприклад, логістичними кривими, моделями витрат або імітаційними симуляціями, які служать додатковим джерелом інтерпретації та перевірки результатів.

Інтеграція таких систем у військову логістику можлива через створення модульних платформ, які підключаються до існуючих систем управління (наприклад, С2-систем) і автоматично оновлюються залежно від оперативних даних. Важливою є побудова інтерфейсів, що дозволяють командирам взаємодіяти з прогнозною системою

в режимі реального часу, задавати параметри та переглядати сценарії на різні часові горизонти.

Реальні приклади демонструють ефективність застосування таких підходів у прогнозуванні потреб для бойових бригад у зонах активного протистояння. Наприклад, завдяки використанню RNN-моделі однієї з військових частин вдалося уникнути затримки в доставці медичних препаратів, коли прогнозована модель виявила високий ризик зростання поранених на основі зміни оперативної обстановки. Аналогічно, трансформерна модель, що аналізувала погодні дані, зуміла передбачити зростання потреб у пальному для важкої техніки внаслідок погіршення прохідності доріг.

Загалом, практичні кроки України в оптимізації логістичних механізмів в умовах війни продемонстрували вирішальну роль технологічних інновацій, гнучкого підходу до організації постачання та ефективної взаємодії між різними відомствами і структурами. Введення модульних систем, автоматизація процесів, застосування безпілотних апаратів для доставки вантажів та мобільних станцій забезпечення, а також залучення цивільних логістичних потужностей стали ключовими нововведеннями, які істотно підвищили ефективність постачання в умовах бойових дій.

Накопичений Україною досвід може стати корисним орієнтиром для інших країн, що прагнуть модернізувати свої системи військового забезпечення, підвищити їх стійкість і оперативність у надзвичайних ситуаціях. Подальший

розвиток автоматизованих рішень, безпілотних логістичних технологій, широке впровадження штучного інтелекту та цифрових платформ відкривають перспективи для створення ще більш гнучких, швидких і ефективних логістичних систем. Це має важливе значення для підтримання безперервного постачання та забезпечення здатності до оперативного реагування в умовах швидкоплинних і нестабільних збройних конфліктів.

Висновки. Використання нейромережевих моделей у прогнозуванні логістичних потреб в умовах бойових дій відкриває нові горизонти для ефективного та адаптивного військового забезпечення. Ці моделі дозволяють значно підвищити точність і швидкість прийняття рішень, оптимізувати логістичні потоки та зменшити

ризик нестачі ресурсів у критичні моменти. Рекурентні мережі, LSTM та трансформери продемонстрували значний потенціал у моделюванні динамічних, багатофакторних процесів, притаманних умовам сучасної війни.

Проте успішна імплементація таких систем вимагає вирішення ряду викликів, серед яких – забезпечення якісних вхідних даних, адаптація моделей до специфіки військового середовища, інтеграція у наявні інформаційні системи та дотримання кібербезпеки. Подальші дослідження мають бути зосереджені на практичній апробації моделей у польових умовах, створенні прототипів інтелектуальних логістичних систем та розробці нормативної бази для їх застосування в оборонному секторі.

Список літератури:

1. Берестенко В.В. Як змінилася логістика за півроку війни та що буде з імпортом і експортом. Центр транспортних стратегій. URL: <https://surl.li/qkdmjq>
2. Ковальчук М., Бондаренко С. Моделювання функціонування логістичних процесів за допомогою математичних та комп'ютерних підходів. *Вісник Хмельницького національного університету*. 2024. № 4 (339). С. 165–168. DOI: <https://doi.org/10.31891/2307-5732-2024-339-4-27>
3. Орел А. М., Дяченко В. В. Етапи та роль цифрової трансформації у вдосконаленні управління інноваціями. *Економіка і суспільство*. 2023. № 56. URL: <https://economyandsociety.in.ua/index.php/journal/issue/view/56>
4. Якушенко А., Шевчук Д. та Мединський Д. Нейромережева модель для прогнозування часу виконання транспортної задачі. 2021. *Науково обґрунтовані технології*, 49(1), 33–38.
5. Well Pack How the War in Ukraine Transforms European Logistics and Affects Prices. Wellpack.Org URL: <https://surl.li/ujvnng>
6. Zaverbnyj A., Pushak H. Theoretical and Applied Principles of Supply Chain Management under Crisis Economic Conditions: Strategic Aspect. *Economic Bulletin of Donbas*. 2022. № 4 (70). P. 57–62.

Stotskyi I.V., Ustynov D.A., Nalisnyi H.I., Kostenko V.P. NEURAL NETWORK MODELS FOR FORECASTING LOGISTICS NEEDS IN COMBAT CONDITIONS

The article is devoted to the study of the current issues of using neural network models to predict logistics needs in combat conditions. The military conflict in Ukraine, which has been ongoing since 2014 and especially intensified after the full-scale invasion of 2022, has highlighted the need to modernize approaches to military logistics. The dynamics of changes on the front line, the destruction of infrastructure and complicated supply conditions have necessitated the transition from traditional centralized supply schemes to flexible, decentralized and technologically supported systems. The article examines the key areas of change in military logistics in Ukraine, including the introduction of digital platforms, mobile warehouses, unmanned delivery technologies and the use of intelligent decision-making support systems. The main attention is focused on the possibilities of using artificial intelligence, in particular neural networks of various architectures, to predict the logistical needs of military units. The authors analyze the advantages of recurrent neural networks (RNN) and long-short-term memory (LSTM) models for processing time series data on resource consumption, troop movements, and combat intensity. Transformer architectures are separately considered, which are able to effectively model the relationships between different types of logistical needs, which allows creating more accurate and comprehensive forecasts.

The article outlines the process of building models: from collecting heterogeneous data – operational reports, information from intelligence sources, satellite images, weather conditions – to processing and training models using normalization, filtering, and anomaly detection methods. Attention is also drawn to the difficulties of working with incomplete or distorted data typical of war conditions, and ways to overcome them are proposed by using hybrid models that combine neural networks and classical mathematical approaches.

Special emphasis is placed on the applied aspects of implementing such solutions in military control structures. The issues of integrating predictive models into command and control systems, creating modular

information platforms that automatically update data in real time, as well as the need to develop user interfaces for unit commanders are considered. It is shown that the implementation of such technologies can significantly increase the speed and accuracy of decision-making in logistics even under conditions of high uncertainty.

The article also highlights the practical results of applying neural network approaches during real operations in Ukraine. Examples of the successful use of RNN models to predict surges in medical supply needs and transformers to predict changes in fuel consumption due to weather conditions are given. The results of the implementation of decentralized logistics hubs and mobile delivery systems as a response to the challenges of the destruction of basic infrastructure are analyzed.

In conclusion, the potential for further development of the technological base of military logistics through the active use of automation, unmanned technologies, artificial intelligence and digital platforms is emphasized. The prospects for building fully integrated, self-learning logistics systems capable of quickly responding to changes in the situation in combat conditions and minimizing human losses are outlined.

Thus, the article makes a significant contribution to the scientific justification of the need to use neural network technologies in military logistics, offering a comprehensive approach to forecasting needs, which is of great practical importance for ensuring the sustainability and effectiveness of military operations in modern armed conflicts.

Key words: *neural network models, military logistics, recurrent networks, transformer models, genetic algorithms, defense technologies, forecasting of logistics needs.*

Дата надходження статті: 27.07.2025

Дата прийняття статті: 06.08.2025

Опубліковано: 27.10.2025

Фомін О.С.

Національний університет «Полтавська політехніка
імені Юрія Кондратюка»

МОДЕЛІ МАШИННОГО НАВЧАННЯ ДЛЯ РОЗПІЗНАВАННЯ ТИПІВ БПЛА ЗА СПЕКТРАЛЬНИМИ ХАРАКТЕРИСТИКАМИ РАДІОСИГНАЛІВ

Однією з пріоритетних задач сучасної радіочастотної аналітики є класифікація типів безпілотних літальних апаратів (БПЛА) на основі характеристик радіосигналів, що ними випромінюються під час польоту. Такий підхід набуває особливого значення в умовах зростання інтенсивності використання БПЛА як у цивільній, так і в оборонній сферах, де своєчасне виявлення та ідентифікація об'єкта є критично важливими для забезпечення безпеки повітряного простору. На відміну від методів, що ґрунтуються на візуальному спостереженні, акустичному аналізі або фізичному перехопленні, спектральний підхід дає змогу здійснювати розпізнавання об'єктів у складних погодних і тактичних умовах, включаючи приховане, нічне або завадостійке функціонування БПЛА.

У межах даного дослідження здійснено всебічне теоретичне узагальнення та порівняльний аналіз сучасних моделей машинного навчання, які застосовуються для класифікації БПЛА за спектральними ознаками їхніх радіосигналів. Розглянуто принципи побудови інформаційних моделей, що базуються на часово-частотних характеристиках, а також методи формування навчальних вибірок на основі спектрограм, вейвлет-перетворень, коефіцієнтів Фур'є та похідних індикаторів спектральної щільності потужності. Особливу увагу приділено методам попередньої обробки сигналів, що забезпечують вивільнення структурних ознак із широкосмугових або зашумлених сигналів.

Проаналізовано архітектурні особливості найпоширеніших підходів до класифікації: від класичних методів статистичного навчання (SVM, k-NN, LDA) до сучасних згорткових та рекурентних нейронних мереж, а також багаторівневих гібридних систем. З'ясовано, що ефективність моделей суттєво залежить від природи спектральних ознак, репрезентативності тренувальних наборів, а також від здатності алгоритмів адаптуватися до міждатисової варіативності апаратів. Описано ключові обмеження, зокрема чутливість до завад, перекриття каналів, проблеми нестачі маркованих даних і потенційні ризики переобучення.

Результати теоретичного аналізу можуть бути покладені в основу створення інженерно-прикладних рішень для інтелектуальних систем технічного спостереження, інтегрованих в інфраструктуру охорони об'єктів критичної інфраструктури, а також в автоматизовані комплекси повітряного контролю.

Ключові слова: БПЛА, спектральний аналіз, радіочастотні сигнали, машинне навчання, теоретичні моделі, класифікація.

Постановка проблеми. У сучасних умовах активного розвитку безпілотних літальних апаратів (БПЛА) гостро постає питання їх надійного виявлення та класифікації, зокрема в середовищі радіочастотного моніторингу. Зростання кількості БПЛА, що використовуються як у цивільних, так і в оборонних цілях, супроводжується ускладненням спектру їх технічних характеристик, що значно ускладнює процеси їх ідентифікації. Традиційні методи виявлення, засновані на візуальному чи акустичному аналізі, часто виявляються неефективними в умовах обмеженої видимості,

завад або прихованого використання апаратів. У зв'язку з цим, усе більше уваги приділяється аналізу спектральних ознак радіосигналів, які випромінюють БПЛА під час польоту, як надійного джерела інформації для їх типізації [1, с. 63]. Одночасно з цим, зростає інтерес до використання моделей машинного навчання для обробки таких сигналів, оскільки ці моделі дозволяють виявляти складні нелінійні закономірності, які не піддаються формалізації класичними методами. Проте, незважаючи на значний науковий прогрес у галузі штучного інтелекту, відсутня усталена

методика щодо відбору оптимальних архітектур моделей та інформативних спектральних ознак для розв'язання задач класифікації типів БПЛА.

Аналіз останніх досліджень і публікацій. Упродовж останнього десятиліття спостерігається активізація наукових досліджень, присвячених задачам виявлення, ідентифікації та класифікації безпілотних літальних апаратів (БПЛА) на основі аналізу їхнього радіочастотного випромінювання. Одним із важливих напрямів цієї діяльності є використання методів машинного навчання для обробки спектральних характеристик радіосигналів, які випромінюють БПЛА в режимі передачі керуючих команд, телеметрії або відеопотоку.

У роботі [1] розглянуто загальні підходи до розпізнавання повітряних об'єктів у системах радіолокаційного спостереження, акцентуючи увагу на необхідності підвищення точності класифікації за рахунок інтеграції спектральних ознак. У публікації [2] проаналізовано можливості динамічної реалізації енергетичних детекторів у системах радіомоніторингу, що має суттєве значення для підвищення завадостійкості при виявленні БПЛА у щільному спектральному середовищі.

Публікації [3; 4; 5; 6] демонструють перспективність застосування згорткових нейронних мереж та вейвлет-перетворень для класифікації БПЛА на основі компресованих або широкосмугових спектрограм. У роботі [7] акцент зроблено на здатності моделей глибокого навчання зберігати ефективність у складних електромагнітних умовах, що характерно для багатоканальних середовищ із перекриттям частот. Дослідження [8] представляє гібридну архітектуру з мультिकанальним аналізом спектру та генерацією ознак, що підвищує якість класифікації за типами польотного режиму БПЛА.

Огляд [9] узагальнює сучасний стан розробок у галузі машинного навчання для виявлення та класифікації БПЛА, окреслює ключові труднощі: нестачу маркованих даних, чутливість до змін середовища, обмеження у розширюваності моделей. Роботи [10; 11] розглядають можливості розпізнавання атак (зокрема, глушіння сигналу) і створення надійних систем моніторингу з використанням спектрограм і спеціалізованих фільтрів. У дослідженні [12] описано послідовну реалізацію моделі машинного навчання для виявлення та типізації БПЛА на основі сукупності спектральних ознак, що підтверджує тенденцію до створення універсальних алгоритмів класифікації.

Водночас у проаналізованих роботах спостерігається низка обмежень. Зокрема, більшість

підходів зосереджені на вузьких наборах сигналів, здобутих у лабораторних або контрольованих умовах. Дослідження часто не враховують складну міжтипову варіативність радіопрофілів БПЛА, відсутню стандартизацію сигналів, залежність спектральних характеристик від умов середовища, типу модуляції або частотного діапазону. У низці праць бракує порівняльного аналізу ефективності архітектур та критеріїв добору моделі для конкретних класів цілей.

Отже, незважаючи на наявність великої кількості публікацій, проблема формалізованого теоретичного узагальнення існуючих моделей машинного навчання з фокусом на спектральне розпізнавання типів БПЛА залишається частково нерозв'язаною. Саме це визначає актуальність подальшого дослідження в напрямі систематизації архітектур машинного навчання, спектральних ознак і методів попередньої обробки для побудови ефективних, завадостійких систем радіочастотної класифікації БПЛА.

Постановка завдання. Метою дослідження є теоретичне обґрунтування, узагальнення та критичний аналіз моделей машинного навчання, які застосовуються для розпізнавання типів безпілотних літальних апаратів (БПЛА) на основі спектральних характеристик їхніх радіосигналів. Такий підхід розглядається як перспективний напрям підвищення ефективності пасивних систем виявлення БПЛА, що функціонують у складних завадових, багатоканальних та динамічних умовах. Для досягнення поставленої мети необхідно вирішити низку завдань: проаналізувати основні принципи побудови інформаційних моделей на основі частотно-часових представлень сигналу; розглянути типові архітектури алгоритмів машинного навчання, що застосовуються для класифікації БПЛА; виокремити ключові спектральні параметри, що мають найвищу інформативність для типізації об'єктів; виявити обмеження існуючих підходів і сформулювати методологічні орієнтири для подальших досліджень у галузі пасивного виявлення та класифікації БПЛА.

Виклад основного матеріалу. У контексті підвищення ефективності пасивного виявлення безпілотних літальних апаратів (БПЛА) особливої актуальності набуває розробка моделей машинного навчання (ML), здатних розпізнавати тип об'єкта за спектральними характеристиками радіосигналів. Такий підхід обумовлений необхідністю забезпечення безперервного моніторингу повітряного простору в умовах багатоканального електромагнітного навантаження, перешкод,

нестабільного середовища передачі та високої міжтипової варіативності сигналів. У радіочастотному моніторингу БПЛА ключову роль відіграє форма представлення сигналу, що забезпечує максимальне вивільнення структурних патернів для подальшої обробки ML-моделями. Найпоширенішими підходами є побудова спектрограм (на основі короткочасного перетворення Фур'є), вейвлет-аналіз, обчислення коефіцієнтів енергетичної спектральної щільності, автокореляційних функцій і похідних частотно-часових індикаторів [3; 5; 6]. У практиці класифікації БПЛА спектрограми виявляють себе як зручний спосіб трансформації одномірного сигналу в зображення, придатне для подальшої подачі в згорткові нейронні мережі (CNN), що суттєво покращує точність розпізнавання [4; 7].

Якість розпізнавання залежить не лише від структури моделі, а й від обраного методу попередньої обробки сигналу, який визначає інформативність та придатність вхідних ознак для подальшої класифікації. Оскільки радіосигнали, які випромінюють БПЛА, зазвичай мають широкий спектр, варіативну амплітудно-частотну структуру та можуть містити численні перешкоди (як природного, так і штучного походження), критично важливо забезпечити ефективну фільтрацію таких завад і збереження характерних патернів [3; 6].

Одним із базових кроків є фільтрація частотного шуму, яка реалізується шляхом смугової фільтрації або адаптивного видалення завадових компонентів, що не відповідають очікуваному профілю сигналу цільового класу [7, с. 6101–6103]. Нормалізація амплітудного спектра – ще один ключовий етап, що дозволяє зменшити вплив абсолютної потужності сигналу, яка може варіювати в залежності від відстані до приймача, типу середовища та рівня заряду пристрою [9]. Стандартизація спектрів за допомогою z -перетворення або мінімаксного масштабування дозволяє зробити вибірки порівнюваними між собою, що суттєво покращує якість навчання моделей [4, с. 10–11].

Компресія спектрального сигналу, зокрема шляхом обмеження кількості частотних бінів, синтезу середніх індикаторів потужності або застосування алгоритмів зменшення розмірності (метод головних компонент, автоенкодера), використовується для зменшення розмірності ознак без втрати змістовного навантаження [5]. У свою чергу, підсилення характеристичних піків, що відповідають резонансним частотам або гармонікам, дозволяє акцентувати увагу моделі на найбільш

значущих елементах сигналу, підвищуючи селективність розпізнавання [11, с. 16862–16864].

Локалізоване згладжування за часом (наприклад, за допомогою ковзного середнього або фільтра Савицького-Голея) допомагає зменшити випадкові флуктуації та імпульсні завади, що особливо актуально при аналізі спектрограм тривалих записів [6]. Окрему роль відіграє вейвлет-декомпозиція, яка дозволяє здійснити багаторівневий аналіз сигналу у часово-частотному просторі. Завдяки цьому метод дає змогу не лише виділяти низько- та високочастотні компоненти, а й локалізувати характерні структури сигналу в часі, що особливо цінно при роботі з нестабільними, імпульсними або модуляційно насиченими сигналами, типовими для різних режимів роботи БПЛА [5; 8].

Вибір типу вейвлету (Haar, Daubechies, Morlet тощо) безпосередньо впливає на здатність моделі розпізнавати конкретні класи об'єктів, оскільки різні хвильові функції акцентують різні характеристики сигналу [10, с. 975–977]. Усі зазначені процедури є критично важливими етапами у формуванні адекватного вхідного набору для машинного навчання, оскільки саме якість попередньої обробки часто визначає верхню межу точності, яку потенційно може досягти модель.

Для систематизації описаних методів попередньої обробки спектральних сигналів та відповідних алгоритмів, що застосовуються в задачах класифікації БПЛА, наведено таблицю 1, де згруповано ключові техніки, їх призначення, рекомендовані реалізації та джерела, що обґрунтовують ефективність їх використання.

Ключовими напрямками побудови моделей для розпізнавання БПЛА є поєднання різних типів алгоритмів машинного навчання, кожен із яких реалізує специфічний підхід до обробки спектральних даних. На початковому етапі розвитку цієї галузі широкого поширення набули статистичні моделі з ручним виділенням ознак, зокрема метод опорних векторів (SVM), метод k -найближчих сусідів (k -NN) та лінійний дискримінантний аналіз (LDA) [1; 3]. Ці підходи ґрунтуються на попередньо сформованому наборі дескрипторів сигналу – таких як енергетичні коефіцієнти, середньоквадратична потужність або пік-фактор – і використовуються переважно в умовах обмежених обчислювальних ресурсів або невеликих обсягів даних.

Суттєвим кроком уперед стали згорткові нейронні мережі (CNN), які здійснюють автоматичне виявлення ознак на основі двовимірних

Методи попередньої обробки спектральних сигналів

Метод обробки	Призначення	Рекомендовані алгоритми
Нормалізація спектра	Усунення енергетичної нерівномірності	Min-Max scaling, z-score
Вейвлет-декомпозиція	Розділення багатокомпонентних сигналів	Haar, Daubechies, Morlet
Логарифмічне перетворення	Підсилення слабких структур в спектрі	$\log(1+x)$, log-mel
Спектральне усереднення	Зменшення завад і згладжування	Moving average, Savitzky-Golay
Компресія спектрограм	Зменшення розмірності при збереженні інформативності	PCA, Autoencoder

спектральних представлень (спектрограм), що дає змогу трактувати сигнал як зображення та виявляти у ньому просторові (частотно-часові) структури [4; 5]. CNN-моделі демонструють високу здатність до узагальнення в задачах класифікації типів БПЛА за рахунок того, що вони можуть навчатися на «сирих» спектрограмних зображеннях без необхідності ручного проектування ознак. Водночас вони мають певну обмеженість у роботі з довгими послідовностями або динамічними ознаками, які змінюються у часі.

Для подолання цього недоліку активно застосовуються рекурентні нейронні мережі (RNN), зокрема їхні покращені варіанти – довготривала короткочасна пам'ять (LSTM) або GRU, які ефективно обробляють часові залежності в радіосигналі [6; 7]. Такі мережі добре справляються з динамічними, багаточастотними сигналами, дозволяючи вловлювати зміну частотної структури сигналу впродовж усього періоду спостереження. Проте в умовах великої варіативності або перешкодочутливості вони можуть втрачати стійкість навчання.

Останнім часом зростає популярність гібридних архітектур, що об'єднують згорткові та рекурентні мережі в єдину структуру: CNN-модуль виконує попереднє вилучення спектральних патернів, тоді як LSTM-частина аналізує їхню часову послідовність [8]. Такі системи поєднують переваги обох класів моделей, зберігаючи високу точність при розпізнаванні складних сигналів у зашумленому середовищі. Їхня ефективність підтверджується численними емпіричними дослідженнями, зокрема для задач розділення близьких типів БПЛА з перекриванням частотних характеристик [7; 8].

Окрему увагу привертають архітектури на основі трансформерів (Transformer), які прийшли у сферу спектрального аналізу з галузі обробки природної мови [9]. Трансформерні моделі демонструють здатність ідентифікувати довготривалі залежності між сегментами сигналу, зберігаючи

високий рівень розпізнавання навіть у випадках фрагментарного або частково пошкодженого сигналу. Їхня головна перевага – незалежність обробки окремих сегментів сигналу та можливість паралельного навчання, що робить їх привабливими для систем реального часу. Попри значну обчислювальну вартість, трансформери демонструють стійкість до шумів і не потребують попереднього формування ознак, що дозволяє інтегрувати їх у високонавантажені пасивні системи моніторингу.

Для узагальнення та порівняння архітектурних підходів, які використовуються в розпізнаванні БПЛА за спектральними характеристиками, у таблиці 2 наведено короткий опис моделей, їхні ключові переваги та обмеження, а також відповідні сфери застосування з урахуванням типу даних і вимог до обчислювальних ресурсів.

Ефективність моделей істотно залежить від структури навчальної вибірки, яка визначає здатність алгоритму узагальнювати знання на нові, раніше невідомі приклади. Одним із ключових факторів є нерівномірність розподілу класів, що спостерігається в багатьох відкритих або експериментальних датасетах. Якщо один тип БПЛА домінує у вибірці, модель схильна до переобучення на його ознаках і демонструє погану якість розпізнавання рідкісних або нових типів, що в польових умовах є критично небезпечним [9]. Додаткову складність становить висока внутрішньокласова варіативність: один і той самий тип БПЛА може генерувати спектрально різні сигнали в залежності від навантаження, рівня зарядженості акумулятора, виконуваного завдання, кліматичних умов або радіочастотного середовища. Це знижує стійкість моделі, зокрема при переході до нових регіонів експлуатації або зміни конфігурації мережі управління [12].

Спроба подолати ці обмеження шляхом аугментації спектрограм – наприклад, шляхом дзеркального відображення, додавання штучного шуму, масштабування по частоті чи часу – частково

Порівняння моделей ML для класифікації БПЛА

Модель	Тип ознак	Точність, %	Переваги	Обмеження
SVM	Статичні	78–83	Простота, низька обчислювальна складність	Залежність від якості ознак
k-NN	Статичні	75–80	Низька вартість тренування	Висока вартість класифікації
CNN	Спектрограми	88–92	Стійкість до шумів, адаптивність	Потреба у великій вибірці
LSTM	Часові ряди	85–89	Моделювання часових залежностей	Труднощі з довготривалими залежностями
CNN+LSTM	Гібридні	91–95	Поєднання просторового і часово-частотного аналізу	Висока складність
Transformer	Спектральні блоки	93–96	Виявлення глобальних закономірностей	Вимогливість до ресурсів

підвищує різноманітність тренувального набору. Однак такий підхід не здатен повністю відтворити реальну нестабільність сигналу, що змінюється не лише під впливом завад, а й внаслідок внутрішніх технічних чинників, таких як температура мікросхем, затримки комунікацій або апаратні фільтри. У зв'язку з цим, все більше дослідників звертається до адаптивного навчання на потокових або слабо маркованих даних, з використанням псевдорозмітки, кластеризації в прихованому просторі ознак або генеративних моделей, здатних імітувати спектральну поведінку різних типів БПЛА у змінених умовах.

Незважаючи на значний прогрес, існуючі моделі демонструють ряд обмежень. Насамперед, більшість досліджень проводиться в лабораторних умовах із контрольованими джерелами сигналів. Такі підходи погано масштабуються на реальні багатоканальні середовища. Також відзначається недостатня гнучкість моделей до зміни спектрального фону, слабка узагальнювальність на незнайомі типи БПЛА та ризик перенавчання на фоні повторюваних шаблонів спектру. Перспективним напрямом вважається розробка адаптивних самонавчальних моделей, що зможуть оновлювати структуру в реальному часі при надходженні нових патернів радіосигналу, а також впровадження фреймворків з активним навчанням, що дозволяють моделі самостійно вибирати найінформативніші приклади для донавчання.

Висновки. Проведене дослідження підтвердило перспективність використання методів машинного навчання для класифікації типів безпілотних літальних апаратів за спектральними характеристиками їхніх радіосигналів. Спектральний підхід забезпечує пасивність, завадостійкість і застосовність у складних середовищах,

де інші методи детекції втрачають ефективність. Узагальнення існуючих теоретичних і прикладних розробок показало, що успішність класифікації залежить не лише від обраної архітектури моделі, а й від якості навчальних даних, попередньої обробки сигналу та специфіки самого класу БПЛА. Серед моделей машинного навчання найвищу ефективність продемонстрували згорткові нейронні мережі та гібридні архітектури з використанням рекурентних компонентів, що дозволяють враховувати як просторову, так і часову структуру сигналу. Високу точність також продемонстрували трансформерні моделі, здатні адаптуватися до довготривалих залежностей у спектрі. Класичні статистичні підходи залишаються актуальними у випадках обмеженого обсягу даних і зрозумілої структури ознак. Попередня обробка сигналів виявилася критично важливою для підвищення достовірності розпізнавання. Методи вейвлет-декомпозиції, фільтрації шумів, нормалізації, масштабування та виділення ключових частотних компонентів сприяють ефективному вилученню інформативних ознак із реального зашумленого сигналу. Проте значною проблемою залишається нестача маркованих даних, особливо для рідкісних або нових типів БПЛА, що ускладнює побудову узагальнених моделей. Особливої уваги потребує варіативність спектрального профілю навіть в межах одного типу БПЛА, що обумовлена змінами умов експлуатації, навантаженням, рівнем заряду або типом місії. Це вимагає розробки адаптивних моделей, здатних навчатись на неповних або змінних даних без втрати точності. Іншим напрямом розвитку є побудова мультिकанальних систем виявлення, які поєднують спектральні дані з іншими типами сигналів, що дозволяє змен-

шити похибку класифікації та підвищити стійкість до атак.

Напрямок подальших досліджень є розробка адаптивних моделей класифікації, здатних ефективно працювати в умовах обмеженого доступу до маркованих даних, з високою толерантністю до міжтипової варіативності сигналів та завад. Особливу увагу доцільно приділити поєднанню спектральних ознак із іншими параметрами

радіосигналу (такими як фаза, модуляція, часові патерни) у межах мультиіндикаторних систем. Крім того, актуальним залишається створення відкритих валідаційних наборів спектрограм для тестування алгоритмів у реальних умовах, що дасть змогу об'єктивно порівнювати різні підходи та забезпечить реплікабельність результатів у межах прикладних проєктів технічного спостереження.

Список літератури:

1. Гладішук О.В., Місайлов В.Л. Evaluation of methods for recognition and identification of aerial vehicles in surveillance radar systems. *Випробування та сертифікація*. 2024. № 4(6). С. 63–68.
2. Бугайов М.В., Нагорнюк О.А., Шапар Р.В. Огляд аналізаторів радіочастотного спектра та пропозиції щодо динамічної реалізації енергетичного детектора. *Технічна інженерія*. 2021. № 2(88). С. 95–102.
3. Mo Y., Huang J., Qian G. Deep learning approach to UAV detection and classification by using compressively sensed RF signal. *Sensors*. 2022. Vol. 22, No. 8. Article No. 3072.
4. Hou C., Fu D., Zhou Z., Wu X. A deep learning-based multi-signal radio spectrum monitoring method for UAV communication. *Drones*. 2023. Vol. 7, No. 8. Article No. 511.
5. Medaiyese O.O., Ezuma M., Lauf A.P., Guvenc I. Wavelet transform analytics for RF-based UAV detection and identification system using machine learning. *Pervasive and Mobile Computing*. 2022. Vol. 82. Article No. 101569.
6. Nemer I., Sheltami T., Ahmad I., Yasar A.U.H., Abdeen M.A. RF-based UAV detection and identification using hierarchical learning approach. *Sensors*. 2021. Vol. 21, No. 6. Article No. 1947.
7. Zhang H., Li T., Li Y., Li J., Dobre O.A., Wen Z. RF-based drone classification under complex electromagnetic environments using deep learning. *IEEE Sensors Journal*. 2023. Vol. 23, No. 6. P. 6099–6108.
8. Yang S., Luo Y., Miao W., Ge C., Sun W., Luo C. RF signal-based UAV detection and mode classification: A joint feature engineering generator and multi-channel deep neural network approach. *Entropy*. 2021. Vol. 23, No. 12. Article No. 1678.
9. Rahman M.H., Sejan M.A.S., Aziz M.A., Tabassum R., Baik J.I., Song H.K. A comprehensive survey of unmanned aerial vehicles detection and classification using machine learning approach: Challenges, solutions, and future directions. *Remote Sensing*. 2024. Vol. 16, No. 5. Article No. 879.
10. Cai Z., Liu Z., Kou L. Reliable UAV monitoring system using deep learning approaches. *IEEE Transactions on Reliability*. 2021. Vol. 71, No. 2. P. 973–983.
11. Li Y., Pawlak J., Price J., Al Shamaileh K., Niyaz Q., Paheding S., Devabhaktuni V. Jamming detection and classification in OFDM-based UAVs via feature-and spectrogram-tailored machine learning. *IEEE Access*. 2022. Vol. 10. P. 16859–16870.
12. Ahirrao Y.V., Yadav R.P., Kumar S. RF based UAV detection and identification enhanced by machine learning approach. *IEEE Access*. 2024. URL: <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=10758648> (date of access: 18.07.2025).

Fomin O.S. MACHINE LEARNING MODELS FOR UAV TYPE RECOGNITION BASED ON THE SPECTRAL CHARACTERISTICS OF RADIO SIGNALS

One of the priority tasks in modern radio frequency analytics is the classification of unmanned aerial vehicles (UAVs) based on the characteristics of radio signals they emit during flight. This approach becomes particularly significant in the context of the growing use of UAVs in both civil and defense sectors, where timely detection and identification of an object are critically important for ensuring airspace security. Unlike methods based on visual observation, acoustic analysis, or physical interception, the spectral approach enables object recognition under complex weather and tactical conditions, including covert, nighttime, or interference-resistant UAV operations.

This study presents a comprehensive theoretical generalization and comparative analysis of current machine learning models used for classifying UAVs based on the spectral features of their radio signals. It explores the principles of constructing information models based on time-frequency characteristics, as well as methods for forming training datasets using spectrograms, wavelet transforms, Fourier coefficients, and derivative indicators of power spectral density. Particular attention is given to signal preprocessing techniques that facilitate the extraction of structural features from broadband or noisy signals. The architectural specifics of

the most common classification approaches are analyzed, ranging from classical statistical learning methods (SVM, k-NN, LDA) to modern convolutional and recurrent neural networks, as well as multi-level hybrid systems. It is established that model performance strongly depends on the nature of spectral features, the representativeness of training sets, and the algorithms' ability to adapt to inter-type variability of UAVs. Key limitations are outlined, including sensitivity to interference, channel overlap, lack of labeled data, and potential overfitting risks.

The findings of this theoretical analysis can serve as a methodological foundation for the development of engineering solutions for intelligent technical surveillance systems, integrated into the infrastructure for protecting critical facilities and automated airspace monitoring complexes.

Key words: UAV, spectral analysis, radio frequency signals, machine learning, theoretical models, classification.

Дата надходження статті: 14.07.2025

Дата прийняття статті: 21.07.2025

Опубліковано: 27.10.2025

Чепурна О.Є.

Національний університет «Одеська юридична академія»

Черевко Є.В.

Національний університет «Одеська юридична академія»

Лобода Ю.Г.

Національний університет «Одеська юридична академія»

Кухаренко С.В.

Національний університет «Одеська юридична академія»

ЗАСТОСУВАННЯ МАТЕМАТИЧНОЇ СТАТИСТИКИ ДЛЯ АНАЛІЗУ ВЕЛИКИХ ДАНИХ В АЛГОРИТМАХ ОБРОБКИ ІНФОРМАЦІЇ

У статті досліджено застосування методів математичної статистики для аналізу великих даних в алгоритмах обробки інформації. Враховуючи проблеми, пов'язані з обсягом, варіативністю та неповнотою даних, доведено доцільність інтеграції статистичних підходів у сучасні системи обчислювального аналізу. Особлива увага приділяється завданням класифікації, виявлення аномалій та оцінки параметрів розподілу. Проаналізовано використання класичних та байєсівських методів, а також метрик типу «Відстань Магаланобіса» для побудови адаптивних рішень. Показано, що навіть у випадку багатьох ознак статистичні методи дозволяють вести інтерпретацію моделей за допомогою засобів SHAP, аналізу перестановок і часткових залежностей. Окремо обґрунтовано важливість поєднання статистичної формалізації з візуалізацією, що підвищує достовірність висновків та практичну значущість моделювання. Надано рекомендації щодо вибору статистичних процедур для конкретних класів завдань (контроль якості, аналіз ризиків, кіберзахист). Окремо відзначено роль використання сучасних інструментів статистичного прогнозування, які забезпечують можливість випереджального виявлення тенденцій у великих потоках інформації. Розглянуті приклади впровадження статистичних методів у практичних інформаційних системах, зокрема у фінансовому моніторингу, кібербезпеці, біоінформатиці, що засвідчує універсальність та ефективність математичної статистики для вирішення актуальних проблем аналізу даних. Запропоновані напрями подальших досліджень: інтеграція інформаційної геометрії як сучасного інструменту аналізу структури даних, автоматизація статистичних перевірок у складі систем AutoML, апробація методів у потоковій передачі даних, а також розширення пояснюваного штучного інтелекту за допомогою статистичних підходів. Окреслено переваги поєднання класичних статистичних методів з новітніми алгоритмічними рішеннями для побудови надійних, гнучких і масштабованих систем обробки та аналізу великих даних. Робота орієнтована на фахівців в області обробки даних, прикладної математики, інформаційних технологій, а також розробників інтелектуальних систем, які працюють із великими і різномірними наборами даних.

Ключові слова: математична статистика, великі дані, алгоритми обробки даних, управління кібербезпекою, аналіз даних, математичні моделі, інформаційна безпека, статистичне прогнозування, моделі кіберзагроз.

Постановка проблеми. Щоденно генерується, збирається та обробляється колосальний обсяг даних – від банківських транзакцій до соціальних мереж, медичних записів та наукових досліджень. Обробка й аналіз цих даних потребує не тільки потужних обчислювальних ресур-

сів, а й ефективних математичних інструментів, серед яких ключову роль відіграє математична статистика.

Математична статистика є галуззю математики, що забезпечує інструменти для збору, опису, аналізу та інтерпретації даних, а її застосування

в алгоритмах обробки інформації дозволяє виявляти закономірності в хаотичних наборах даних, здійснювати прогнозування, оцінювати ймовірності подій та мінімізувати похибки під час прийняття рішень. Особливої актуальності статистичні методи набувають у сфері великих даних (Big Data), які відзначаються значним обсягом, різноманітністю та швидкістю надходження, адже саме поєднання статистики з алгоритмічними рішеннями створює підґрунтя для розробки інтелектуальних систем, що здатні навчатися, адаптуватися й приймати обґрунтовані рішення на основі аналізу даних.

Метою цієї роботи є дослідження можливостей та ефективності використання математичної статистики у процесі обробки великих масивів інформації, а також аналіз алгоритмів, які інтегрують статистичні методи для оптимізації й автоматизації аналізу даних. Інформаційні системи дедалі частіше функціонують в умовах опрацювання великих обсягів складно структурованих даних, що надходять із гетерогенних джерел. Такі масиви часто характеризуються пропущеними значеннями, внутрішньою корельованістю ознак, нелінійною структурою залежностей та нестійкістю розподілів у часі. У результаті, стандартні аналітичні підходи демонструють зниження точності й обґрунтованості висновків [1, с. 123–129; 2, с. 87–95].

Математична статистика надає формалізовану основу для розв'язання задач класифікації, регресії, побудови довірчих інтервалів та перевірки статистичних гіпотез [3, с. 173–178; 4, с. 201–210]. Проте у випадках високої розмірності даних або значних обсягів потокової інформації, ефективність традиційних процедур суттєво знижується [5, с. 91–100].

Особливої уваги набувають підходи, що дозволяють поєднувати статистичні моделі з геометричними уявленнями про простір параметрів розподілу. Такі підходи застосовуються, зокрема, у побудові метрик простору імовірнісних моделей, визначенні статистичних відстаней, а також у вивченні кривизни простору параметрів [6, с. 132–144; 7, с. 77–89].

Інформаційна геометрія як міждисциплінарна концепція використовує апарат диференціальної геометрії для аналізу статистичних моделей як багатоатомних об'єктів у просторах з кривизною. Простір Фішера–Рао, як приклад такого підходу, дає змогу оцінювати відстані між розподілами, базуючись на їх ентропійних властивостях [8, с. 58–70; 9, с. 4–8].

Застосування цих підходів дозволяє не лише глибше інтерпретувати дані, але й підвищити стійкість моделей до спотворень, збоїв і викривлень, характерних для реальних інформаційних потоків [10, с. 44–52; 11, с. 11–22]. Це особливо важливо у критичних галузях, таких як технічна діагностика, безпека ІТ-систем, біоінформатика та фінансовий моніторинг [12, с. 6–12; 13, с. 1210–1218].

Поєднання класичних методів статистичного аналізу з інформаційно-геометричними підходами сприяє створенню адаптивних обчислювальних моделей, що демонструють вищу точність та інтерпретованість у складних середовищах [14, с. 5–9; 15, с. 14–16], а також відкриває нові перспективи для розвитку інтелектуальних систем обробки інформації у різних сферах діяльності.

Аналіз сучасних досліджень і публікацій. Зростання інтересу до проблем ефективного аналізу великих даних, що має відображення як у статистичній теорії, так і в прикладних дослідженнях. У працях Джона Райса (J.A. Rice) [1, с. 123–129] та Джорджа Каселла і Роджера Бергера (G. Casella і R.L. Berger) [2, с. 87–95] детально розглянуто основи класичної математичної статистики, включаючи гіпотезне тестування, побудову довірчих інтервалів і оцінювання параметрів, що лежать в основі більшості сучасних методів обробки даних.

З погляду на обсяг і складність обробки і аналізу даних особливу актуальність набувають підходи, орієнтовані на зниження розмірності, відбір релевантних ознак і стійке оцінювання. У роботі Крістофера Бішопа (C.M. Bishop) [3, с. 173–178] докладно аналізуються статистичні моделі, придатні для умов високої вимірності, зокрема методи головних компонент (PCA) та латентні змінні. Кевін Мерфі (K.P. Murphy) [4, с. 201–210] зосереджується на ймовірнісних графових моделях, що дозволяють моделювати складні залежності між змінними при мінімізації обчислювальної складності.

Особливої уваги заслуговують прикладні дослідження, пов'язані з використанням методів машинного навчання у статистичному аналізі. Зокрема, Джеймса Гарта та його співавтори (G. James et al) [5, с. 91–100] демонструють застосування регуляризованих моделей до задач класифікації у великих вибірках. Це узгоджується з висновками Мартіна Вейрайта та Майкла Джордана (M.J. Wainwright і M.I. Jordan) [6, с. 132–144] щодо застосування експоненціальних родин розподілів для аналізу взаємозв'язків у даних.

Проблематика аналітики у високорозмірних просторах тісно пов'язана з методами зниження розмірності, обчислення ентропій та інформаційних дивергенцій. Ряд публікацій розглядає інтеграцію таких методів у поточні фреймворки аналізу даних, наприклад, Ліам Панінські (Paninski L.) [13, с. 1210–1218], зокрема у виявленні аномалій, нечіткій класифікації та роботі зі спотвореними вибірками.

На перетині математичної статистики та геометрії виникають нові концепти – інформаційна геометрія. Праці Шун'їчі Амарі (S. Amari) [7, с. 77–89], Томаса Ковера (T.M. Cover) і Джой Томаса (J.A. Thomas) [8, с. 58–70], а також Ніко Ай зі співавторами (N. Ay et al) [10, с. 44–52] заклали фундамент цього напрямку, у якому імовірнісні моделі інтерпретуються як многовиди з рімановою метрикою, де статистичні властивості відображаються у кривизні простору. Це відкриває нові можливості для адаптивного моделювання, зокрема в умовах змінної статистичної структури даних – Євген Черевко, Олена Чепурна, Євгенія Кулешова (Cherevko Y., Chepurna O., Kuleshova Y.) [11, с. 11–22].

У практичному вимірі застосування інформаційної геометрії, наприклад, у задачах кіберзахисту, які висвітлені у Fortinet – Глобальний звіт про ландшафт загроз 2025 [12, с. 6–12] та оптимізації систем виявлення аномалій, дозволяє підвищити чутливість моделей до малопомітних відхилень у даних. Це важливо для галузей, де якість виявлення відхилень критично впливає на ефективність рішень – наприклад, у фінансовому секторі, медицині та розподілених інформаційних системах.

Таким чином, сучасний напрям досліджень зосереджено навколо інтеграції статистичних, геометричних і алгоритмічних методів, що забезпечують інтерпретованість, адаптивність і стійкість аналітичних моделей до складних властивостей даних.

Постановка завдання. Однією з основних проблем роботи з масивами даних є обмежена придатність класичних статистичних процедур для аналізу великих, неоднорідних та неструктурованих даних, особливо у випадках високої розмірності та змінних розподілів. Тому актуальним є дослідження можливостей використання геометризованих статистичних моделей для підвищення стійкості й інтерпретованості аналізу даних. У цьому дослідженні передбачається систематизувати статистичні методи, що ефективні для Big Data (SelectKBest, PCA, Mahalanobis distance,

LOF, GMM), оцінити їхню придатність до задач класифікації, виявлення аномалій та зменшення розмірності, проаналізувати сучасні алгоритми, які використовують геометричні структури (ріманові многовиди, дивергенції Фішера–Рао, інформаційну метрику), запропонувати приклад застосування на синтезованих наборах даних та розробити критерії ефективності моделей (AUC, Precision, Recall, інформаційна дивергенція), а також запропонувати гібридний підхід, у якому статистичні й інформаційно-геометричні інструменти комбінуються.

В цьому контексті потрібно виконати наступні завдання: проаналізувати основні поняття та методи математичної статистики; визначити особливості великих даних (Big Data) та проблеми їх обробки; дослідити алгоритми, що інтегрують статистичні методи для аналізу даних; оцінити переваги та обмеження таких алгоритмів у реальних прикладних сценаріях; запропонувати напрямки вдосконалення методів статистичного аналізу даних.

Таким чином, мета дослідження полягає у поєднанні класичних статистичних методів із новітніми геометричними підходами задля підвищення ефективності аналізу даних у розподілених, змінних та потенційно зашумлених середовищах.

Виклад основного матеріалу. Застосування математичної статистики в аналізі великих даних (Big Data) є фундаментальним для алгоритмів обробки інформації, оскільки воно дозволяє виявляти патерни, робити прогнози, тестувати гіпотези та зменшувати розмірність даних у масштабах, де традиційні методи неефективні. Статистика надає інструменти для оцінки невизначеності, моделювання та оптимізації, особливо в розподілених системах на кшталт Hadoop чи Spark. На основі аналізу сучасних джерел, ключовими алгоритмами, які варто описати, є ті, що поєднують статистичні принципи з обчислювальними техніками для великих обсягів даних. Опишемо найпоширеніших і впливових методи математичної статистики, таких як регресійний аналіз, кластеризація, класифікація, зменшення розмірності та інші. Вони часто реалізуються з використанням стохастичних методів (наприклад, SGD – Stochastic Gradient Descent) для масштабованості. Серед ключових алгоритмів математичної статистики, які застосовуються в аналізі великих даних, варто виділити такі підходи. Лінійна регресія використовується для моделювання залежності між змінними за умови лінійного зв'язку, базуючись на

методі найменших квадратів для мінімізації помилок. Цей підхід дозволяє прогнозувати тренди у великих масивах даних, наприклад у сфері продажів чи кліматичних спостережень; у Big Data часто використовується стохастичний градієнтний спуск для обробки потокових даних, зокрема в бібліотеці Spark MLlib.

Сучасні інформаційні системи дедалі частіше поєднують класичні алгоритмічні процедури з методами математичної статистики для забезпечення ефективного аналізу даних у складних цифрових середовищах. Особливо це стосується алгоритмів сортування, пошуку, обходу графів і дерев, які виступають інфраструктурною основою обробки інформації у високопродуктивних системах (ERP, SCM, IoT-аналітика тощо).

Однією з типових задач є попередня обробка даних, де алгоритми сортування (наприклад, QuickSort, MergeSort) використовуються для підготовки вибірок до статистичного аналізу, зокрема для обчислення медіани, квантилів або емпіричних розподілів. У таких випадках статистичні оцінки (середнє, дисперсія, коефіцієнти асиметрії) вимагають впорядкованих структур, що зменшують складність обчислень з $O(n \log n)$ до $O(n)$ у деяких оптимізованих випадках [14, с. 127–132].

У задачах класифікації та прогнозування, що базуються на статистичних моделях (наприклад, логістична регресія, дерева рішень), широко застосовуються деревоподібні структури даних. Зокрема, дерева CART (Classification and Regression Trees) дозволяють реалізовувати рекурсивне ділення простору ознак, а статистичні критерії, як Gini-індекс або ентропія Шеннона, виступають евристичними для побудови оптимальних вузлів [15, с. 264–270].

Алгоритми на основі графів (наприклад, Dijkstra або Bellman–Ford) дедалі частіше інтегруються зі статистичними підходами для зважування ребер. Наприклад, в аналізі транспортних мереж або логістичних ланцюгів, статистичні метрики, як середній час доставки, варіація затримок або довірчі інтервали, використовуються для формування ваг у графі, що підвищує адаптивність маршрутів до реальних умов [16, с. 318–324].

Крім того, кластеризація у системах рекомендацій (наприклад, у маркетингу або енергетиці), структури дерев пошуку (KD-дерева, Ball-дерева) використовуються для прискорення статистичних обчислень, зокрема для знаходження найближчих сусідів (k-NN) чи виявлення аномалій на основі відстаней (Mahalanobis, LOF) [17, с. 291–295].

Таким чином, поєднання структур алгоритмічної ефективності з статистичною обґрунтованістю створює підґрунтя для формування масштабованих і достовірних систем обробки даних. Це особливо важливо у високонавантажених інформаційних середовищах, де обсяг даних, їхня складність і потреба в інтерпретованості зростають одночасно.

Окремої уваги та дослідження заслуговує інформаційна геометрія як інструмент аналізу великих даних. Інформаційна геометрія – це напрям міждисциплінарного аналізу, що поєднує диференціальну геометрію з математичною статистикою для дослідження властивостей сімейств розподілів ймовірностей. Вона базується на трактуванні множини параметризованих розподілів як многовиду, на якому визначено метричну структуру, зокрема метрику Фішера-Рао. Такий підхід забезпечує геометричну інтерпретацію статистичних відстаней, ентропій, дивергенцій та кривизни ймовірнісного простору [18, с. 77–89].

У задачах обробки великих даних, де класичні евклідові уявлення втрачають свою інтерпретативну потужність через високу розмірність, корельованість або розрідженість даних, інформаційна геометрія дає змогу використовувати геометричні характеристики – такі як геодезичні, ріманову кривизну, відстані між розподілами (наприклад, дивергенції Кульбака-Лейблера або α -дивергенції) – для оцінки подібності, ієрархічної класифікації та аналізу аномалій [18, с. 44–52].

У практичному вимірі така методологія застосовується до високорозмірних потоків даних, типових для біоінформатики, фінансової аналітики, IoT-систем та кіберфізичних мереж. Наприклад, аналіз відстаней між розподілами ймовірностей спостережень у багатовимірному просторі (наприклад, за допомогою геодезичної відстані в моделі експоненціальної родини) дозволяє точно відокремлювати кластери спостережень із різною статистичною структурою, навіть якщо класичні методи кластеризації (наприклад, k-means) демонструють низьку роздільну здатність [19, с. 1210–1218].

У роботах Шун'їчі Amari наведено підстави для використання інформаційної геометрії у побудові статистично ефективних адаптивних моделей у середовищах із нестабільною розподільною структурою. Наприклад, простір експоненціальних розподілів задається через параметри, для яких метрика Фішера-Рао має інваріантність відносно перетворень координат. Це дозволяє будувати процедури узагальненого логарифмічного градієнтного спуску на ріманових многовидах.

Окремої уваги заслуговує використання інформаційно-геометричних методів у задачах виявлення аномалій, наприклад, у кіберзахисті. Замість простого порівняння поточних спостережень із середнім значенням застосовується обчислення геодезичної відстані на просторі ймовірнісних моделей між поточним вектором розподілу характеристик і моделлю «норми».

У середовищах на кшталт Apache Spark чи Dask можливе паралельне обчислення дивергенцій Кульбака-Лейблера, ентропій Шеннона або відстаней Фішера між розподілами з використанням моделей Gaussian Mixture Model або Variational Autoencoders [19, с. 5–9].

Візуалізація застосування інформаційно-геометричної інтерпретації моделей у просторі розподілів може бути представлена у вигляді багатовимірної 3D-моделі, де кожен розподіл зображено як окрему точку на рімановому многовиді, а геодезичні демонструють, яким чином модель адаптується до змін у потоках даних у динамічному середовищі. Такий підхід дає змогу не лише аналізувати розміщення розподілів, а й відстежувати їхні зміни у часі, що критично важливо для систем із високою частотою оновлення інформації, наприклад, у сферах фінансової аналітики, медичного моніторингу або розумних мереж енергоспоживання.

До основних переваг інформаційно-геометричного підходу належить здатність урахувати як точкові, так і структурні властивості розподілів, а також гнучка адаптація до динамічних змін статистичних характеристик. Це дає змогу більш формалізовано і точно оцінювати «інформаційну близькість» між різними станами системи, що особливо корисно для побудови ефективних класифікаторів, автоматичних фільтрів і сучасних енкодерів даних.

Водночас варто враховувати певні обмеження: складність реалізації математичного апарату для великих, нестабільних і розріджених вибірок, потребу в знаннях диференціальної геометрії для коректного визначення метрик, кривизни та візуалізації таких просторів, як це детально розглянуто в джерелі [20, с. 193–198].

Наприклад, для аналізу багатовимірних нормальних розподілів у межах інформаційної геометрії доцільно використовувати розширені методи тривимірної візуалізації, де кожна точка може відображати як положення, так і форму розподілу, а траєкторії між ними – шляхи мінімальної інформаційної «відстані». Подібний підхід дозволяє отримати глибше розуміння структури даних

і швидко ідентифікувати аномалії чи нові кластери у складних потоках інформації. Завдяки цьому дослідники можуть аналізувати не лише статичні характеристики розподілів, а й відстежувати їхню трансформацію в умовах постійної зміни параметрів, що властиво динамічним системам великого обсягу даних.

У практичному сенсі це означає, що процеси виявлення аномалій або класифікації можуть бути автоматизовані на основі геометричних характеристик, включно з кривизною, геодезичними та інформаційною відстанню, що дозволяє максимально ефективно реагувати на нетипові ситуації у потоках даних. Водночас побудова подібних моделей відкриває нові можливості для створення адаптивних прогнозних систем, де алгоритми не лише аналізують зміни центрів кластерів, а й враховують трансформацію їхніх геометричних властивостей, зокрема розтягнення, стиснення та ротацію у статистичному просторі.

Таким чином, інформаційно-геометричний підхід стає потужним інструментом як для глибокого аналізу внутрішньої структури даних, так і для оперативного виявлення нових тенденцій і закономірностей, що є вкрай важливим у сучасних інформаційних системах з високою мінливістю та складністю потоків. Це, у свою чергу, відкриває широкі можливості для подальшого аналізу динаміки розподілів у часі, а також дає змогу розробляти прогностичні моделі, що враховують не лише зміну центрів кластерів, а й трансформацію їхніх геометричних властивостей (Рис. 1).

Перспективним напрямком використання статистичних методів є інтегроване застосування

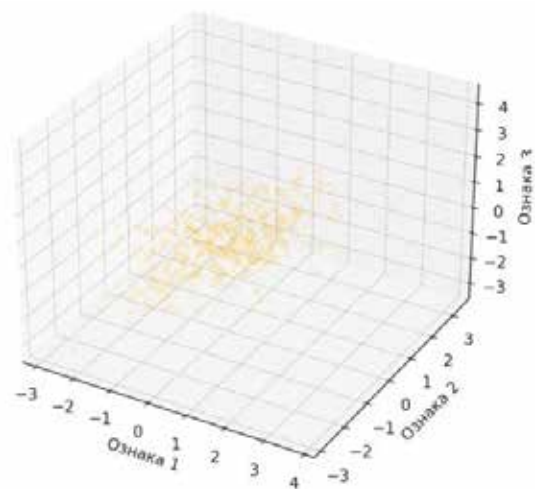


Рис. 1. Візуалізація інформаційно-геометричної інтерпретації структури ймовірнісних розподілів

РСА, GMM та інформаційної геометрії для аналізу даних енергоспоживання. Наприклад, синтетична вибірка, що імітує тиждень споживання електроенергії із 96 точками на добу (15-хвилинний інтервал), обробляється за допомогою РСА, зменшуючи розмірність з 672 до 2 головних компонент ($\approx 87\%$ дисперсії), а GMM дозволяє виділити кластери з власними коваріаційними ядрами. Інформаційна геометрія реалізована через підрахунок міжкластерних відстаней за метрикою Фішера-Рао у просторі головних компонент, що дає змогу аналізувати подібність між кластерами. На рисунку 2 наведено кластери й геометричні зв'язки між ними, що інтерпретуються як геодезичні наближення у метричному просторі Фішера-Рао, дозволяючи оцінювати не лише класи, а й їхню близькість у споживанні. Такий підхід забезпечує не лише глибше розуміння структури даних, але й підвищує точність виявлення аномальних або неефективних патернів у динамічному середовищі енергоспоживання.

Серед ключових переваг такого підходу – інтерпретованість, адже РСА дозволяє виявити часові закономірності, наприклад, денні чи нічні піки та типові шаблони навантаження. Гнучкість кластеризації забезпечується за рахунок того, що GMM дає змогу враховувати неоднорідність дисперсій та кореляцій усередині даних. Використання інформаційної геометрії дозволяє не лише аналізувати центри кластерів, а й порівнювати їхню «форму», що є критично важливим для моніторингу змін структури споживання, зокрема в умо-

вах війни або енергетичних збоїв. На практиці такий інтегрований підхід може бути застосований для виявлення неефективних чи аномальних патернів споживання, оцінки ризиків перевантажень, динамічного перепланування енергопостачання, а також для впровадження адаптивного ціноутворення (Demand Response) на основі кластерної приналежності.

Висновки. Методи математичної статистики, особливо такі сучасні напрями як інформаційна геометрія, стають фундаментом нової парадигми аналізу великих даних. Вони дозволяють дослідникам не лише глибоко зануритися у структуру інформаційних потоків, а й забезпечують ефективніші механізми для своєчасного виявлення аномалій та нестандартних патернів. На відміну від традиційних статистичних підходів, інформаційна геометрія розглядає ймовірнісні розподіли як точки на рімановому многовиді, що надає змогу оцінювати інформаційну близькість різних станів системи з підвищеною точністю. Це, у свою чергу, значно розширює можливості кластеризації, адаптивного моніторингу та аналізу складних динамічних процесів у реальному часі.

Інтеграція інформаційно-геометричних методів із РСА та Gaussian Mixture Model виявляється особливо цінною: вона не лише зменшує розмірність досліджуваних даних і виявляє часові закономірності, а й дає змогу виділяти різноманітні кластери, враховуючи як їхню інформаційну форму, так і динаміку змін. Це має вагомe значення для прикладних сфер, таких як енергетика,

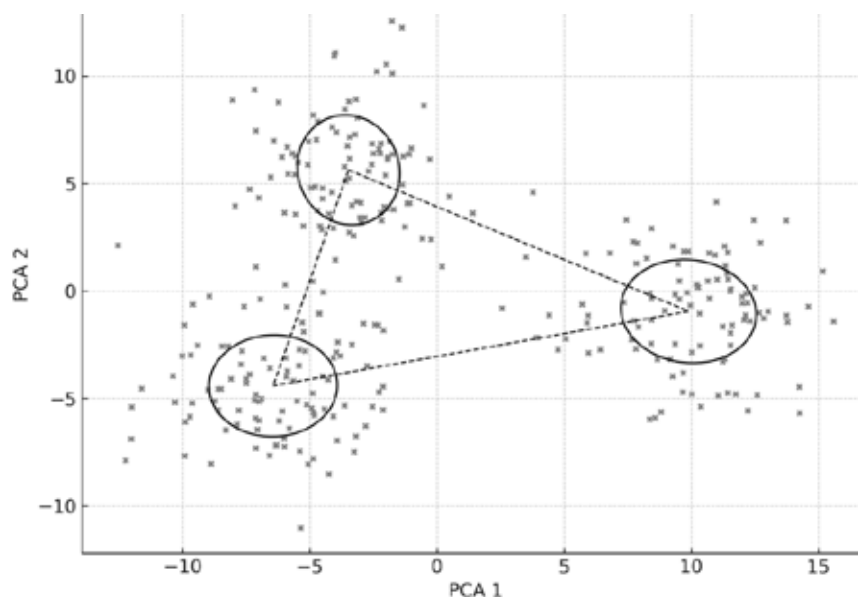


Рис. 2. Кластери енергоспоживання (GMM+РСА) з геометричними зв'язками (Fisher-Rao)

*Умовні позначення: сірі точки – проєкції об'єктів; контури – області кластерів; лінії – геометричні відстані

фінанси та медицина, де моніторинг поведінки систем потребує високої чутливості до змін параметрів і структури. Водночас залучення інформаційної геометрії вимагає значної підготовки спеціалістів, знань у галузі диференціальної геометрії, математичної статистики та володіння сучасним програмним забезпеченням.

Не менш важливою є розробка й удосконалення відкритих бібліотек, що пришвидшить впровадження цих методів у практику. В цьому сенсі інформаційна геометрія – міждисциплінарна галузь, яка поєднує елементи диференціальної геометрії, теорії інформації та статистичного аналізу – пропонує інноваційний інструментарій для

моделювання інформаційних процесів, оцінки ризиків, аналізу кіберзагроз і побудови стійких систем захисту даних та інфраструктур. Перспективи подальшого розвитку цього напрямку полягають у його здатності органічно інтегруватися з іншими статистичними методами, розширюючи спектр прикладних задач та забезпечуючи гнучку, адаптивну аналітику для обробки великих і складних масивів інформації. Саме тому інформаційна геометрія стає дедалі актуальнішою як для сучасної науки, так і для практичних застосувань у різноманітних сферах суспільного життя, де гнучкість та точність аналізу є запорукою ефективних рішень.

Список літератури:

1. Rice J. A. *Mathematical Statistics and Data Analysis*. 3rd ed. Cengage Learning, 2006. 688 p. URL: <https://surl.li/uwsyja>
2. Casella G., Berger R. L. *Statistical Inference*. 2nd ed. Duxbury Press, 2002. 660 p. URL: <https://surl.li/liszcs>
3. Bishop C.M. *Pattern Recognition and Machine Learning*. Springer, 2006. 738 p. DOI: <https://doi.org/10.1007/978-0-387-45528-0>
4. Murphy K. P. *Machine learning: a probabilistic perspective*. MIT Press, 2012. 1104 p.
5. James G., Witten D., Hastie T., Tibshirani R. *An Introduction to Statistical Learning: with Applications in R*. Springer, 2013. 426 p. DOI: <https://doi.org/10.1007/978-1-4614-7138-7>
6. Wainwright M.J., Jordan M.I. *Graphical Models, Exponential Families, and Variational Inference* // *Foundations and Trends® in Machine Learning*. 2008. Vol. 1(1–2). P. 1–305 DOI: <https://doi.org/10.1561/22000000001>
7. Amari S. *Information Geometry and Its Applications*. Springer, 2016. 360 p. DOI: <https://doi.org/10.1007/978-4-431-55978-8>
8. Cover T.M., Thomas J.A. *Elements of Information Theory*. 2nd ed. Wiley-Interscience, 2006. 776 p. DOI: <https://doi.org/10.1002/047174882X>
9. Nielsen F. *An Elementary Introduction to Information Geometry* // *Entropy*. 2020. Vol. 22(10). P. 1100 DOI: <https://doi.org/10.3390/e22101100>
10. Ay N., Jost J., Lê H.V., Schwachhöfer L. *Information Geometry*. Springer, 2017. 308 p. DOI: <https://doi.org/10.1007/978-3-319-56478-4>
11. Cherevko Y., Chepurna O., Kuleshova Y. *On Information Geometry Methods for Data Analysis* // *Geometry, Integrability and Quantization*. 2024. Vol. 29. P. 11–22 DOI: <https://doi.org/10.7546/giq-29-2024-11-22>
12. Fortinet. 2025 Global Threat Landscape Report. URL: <https://www.fortinet.com/resources/reports/threat-landscape-report>
13. Paninski L. *Estimation of entropy and mutual information* // *Neural Computation*. 2003. Vol. 15(6). P. 1191–1253 DOI: <https://doi.org/10.1162/089976603321780272>
14. Cormen T.H., Leiserson C.E., Rivest R.L., Stein C. *Introduction to Algorithms*. 3rd ed. MIT Press, 2009. 1312 p.
15. Fortinet. 2024 Global Threat Landscape Report. – URL: <https://www.fortinet.com/reports>
16. Paninski L. *Estimation of entropy and mutual information* // *Neural Computation*. 2003. Vol. 15(6). – P. 1191–1253 [c. 1210–1218] DOI: <https://doi.org/10.1162/089976603321780272>
17. Ay N., Jost J., Lê H.V., Schwachhöfer L. *Information Geometry*. Springer, 2017. – 308 p. DOI: <https://doi.org/10.1007/978-3-319-56478-4>
18. Hastie T., Tibshirani R., Friedman J. *The Elements of Statistical Learning*. Springer, 2009. – 745 p. [c. 264–270]. DOI: <https://doi.org/10.1007/978-0-387-84858-7>
19. Ahuja R.K., Magnanti T.L., Orlin J.B. *Network Flows: Theory, Algorithms, and Applications*. Prentice Hall, 1993. – 864 p. [c. 318–324].
20. Beyer K., Goldstein J., Ramakrishnan R., Shaft U. *When is “nearest neighbor” meaningful?* // *ICDT*. 1999. – P. 217–235.

Chepurna O.Ye., Cherevko Ye.V., Loboda Yu.G., Kukhareenko S.V. APPLICATION OF MATHEMATICAL STATISTICS FOR BIG DATA ANALYSIS IN INFORMATION PROCESSING ALGORITHMS

The article investigates the application of methods of mathematical statistics for the analysis of big data in information processing algorithms. Considering the problems associated with volume, variability and incompleteness of data, the expediency of integrating statistical approaches into modern systems of computational analysis has been proved. Particular attention is paid to the tasks of classification, detection of anomalies and evaluation of distribution parameters. The use of classical and Bayesian methods, as well as metrics such as “Magalanobis distance” for building adaptive solutions is analyzed. It is shown that even in the case of many features, statistical methods allow interpreting models using SHAP tools, permutation analysis and partial dependencies. The importance of combining statistical formalization with visualization, which increases the reliability of conclusions and the practical significance of modeling, is substantiated separately. Recommendations are provided for the selection of statistical procedures for specific classes of tasks (quality control, risk analysis, cyber defense). The role of using modern tools of statistical forecasting, which provide the possibility of anticipatory detection of trends in large flows of information, is separately noted. Examples of the implementation of statistical methods in practical information systems, in financial monitoring, cybersecurity, bioinformatics, which testifies to the versatility and efficiency of mathematical statistics for solving topical problems of data analysis, are considered. The proposed areas of further research are integration of information geometry as a modern tool for analyzing data structure, automation of statistical checks as part of AutoML systems, approbation of methods in data streaming, as well as expansion of explainable artificial intelligence using statistical approaches. The advantages of combining classical statistical methods with the latest algorithmic solutions for building reliable, flexible and scalable systems for processing and analyzing big data are outlined. The work is aimed at specialists in the field of data processing, applied mathematics, information technology, as well as developers of intelligent systems who work with large and heterogeneous data sets.

Key words: mathematical statistics, big data, data processing algorithms, cybersecurity management, data analysis, mathematical models, information security, statistical forecasting, cyber threat models.

Дата надходження статті: 27.07.2025

Дата прийняття статті: 06.08.2025

Опубліковано: 27.10.2025

Щербина Ю.В.

Національний університет «Одеська юридична академія»

Логінова Н.І.

Національний університет «Одеська юридична академія»

Манаков С.Ю.

Національний університет «Одеська юридична академія»

АНАЛІЗ СУЧАСНИХ СПОСОБІВ КОМП'ЮТЕРНОГО МОДЕЛЮВАННЯ СТОХАСТИЧНИХ ПРОЦЕСІВ

Стаття присвячена проблемі підвищення якості стохастичного комп'ютерного моделювання під час проектування складних технічних систем, що забезпечують роботу транспорту, виробництва, управління інфраструктурними об'єктами та обмін даними в інформаційно-комунікаційних системах. Ефективність такого моделювання визначається наявністю адекватних математичних моделей, що формально описують поведінку реальних систем в стохастичних умовах їх експлуатації, та якістю джерел випадковості, що дозволяють відтворювати такі умови. Однією з головних проблем, що визначає якість стохастичного комп'ютерного моделювання, є вибір програмного способу генерації псевдовипадкових чисел з високим ступенем рівномірності розподілення їх ймовірностей. Ця проблема має вирішуватись у двох напрямках: по-перше, алгоритм формування псевдовипадкового числового потоку має бути невибагливим до таких обчислювальних ресурсів як пам'ять та процесорний час і, по-друге, такий потік не повинен відрізнятись від справжнього випадкового рівномірно розподіленого числового потоку. В роботі надано аналіз відомих програмних способів формування псевдовипадкових чисел, що використовуються для потреб комп'ютерного моделювання. Розглянуто ті з них, що вважаються найбільш економічними, а саме, генератор Фібоначчі з запізненням, генератори, побудовані на основі лінійних регістрів зсуву зі зворотним зв'язком та Xorshift генератор, в основу якого закладені економічні, з точки зору комп'ютерної реалізації, логічні операції та операції зсуву. На основі проведеного аналізу наукових публікацій, було зроблено висновок про те, що з урахуванням сучасних вимог до якості комп'ютерного моделювання, найбільш ефективним джерелом випадковості слід використовувати саме Xorshift генератор. Що ж стосується генераторів Фібоначчі з запізненням та генераторів, побудованих на основі регістрів зсуву зі зворотним зв'язком, то їх можна використовувати тільки у якості складових частин більш складних алгоритмів формування псевдовипадкових чисел.

Ключові слова: комп'ютерне моделювання, метод зворотних функцій, метод відбракування випадкових чисел, генератор псевдовипадкових чисел, статистичний аналіз.

Постановка проблеми. Створення складних систем, що забезпечують роботу сучасного виробництва, транспортних комунікацій, технічне управління інфраструктурними об'єктами та оброблення і обмін даними в комп'ютерних мережах, зазвичай, з метою попередньої оцінки складності задачі і необхідних ресурсних витрат, вимагають попереднього натурного моделювання. В окремих випадках, такі натурні випробування, через надмірну необхідність матеріальних ресурсів, є неприйнятними і це є проблемою. Сьогодні, така проблема ефективно вирішується за рахунок

використання комп'ютерного моделювання, що передбачає попереднє створення математичних моделей фізичних процесів та їх програмної реалізації.

Статистичні випробування, побудованих на основі математичних моделей, віртуальних стохастичних систем, дають змогу отримати попередню оцінку їх поведінки у різних умовах експлуатації. Суть такого моделювання зводиться до формального описання зв'язків, які пов'язують між собою зовнішні і внутрішні фактори, що впливають на поведінку складних систем.

У разі, якщо така поведінка не може напряму бути формально математично пов'язана з факторами, що впливають на процес, який підлягає дослідженню, виконують його симуляцію, тобто комп'ютерні статистичні випробування. Такий спосіб дослідження з використанням чисельних комп'ютерних методів передбачає декомпозицію загального складного процесу на простіші підпроцеси, що можуть бути описані елементарними статистичними видами розподілень або їх достатньо простими комбінаціями [1]. Позитивними сторонами такого моделювання є високий рівень деталізації, відсутність обмежень на параметри моделювання зовнішнього середовища і можливість дослідження динаміки зміни параметрів процесу у просторі та часі.

Комп'ютерне відтворення будь-якого випадкового процесу, перш за все, вимагає наявності “джерела випадковості”, але комп'ютер, це дискретний детермінований автомат, який за визначенням не здатен генерувати випадкові величини. Тому, рішенням цієї другої проблеми комп'ютерного моделювання є генерація, так званих псевдовипадкових чисел (ПВЧ). Програмні генератори такого типу уявляють собою обчислювальні алгоритми, які створюють, циклічно повторювані з періодом T , послідовності, що для стороннього спостерігача виглядають як випадкові. Програмні алгоритми формування таких чисел в науковій літературі називають генераторами псевдовипадкових чисел (Pseudo Random Number Generators – PRNG).

Для якісного комп'ютерного моделювання необхідно, щоб довжина циклу повторення псевдовипадкових чисел була якомога більшою, а розподілення їх ймовірностей у складі вихідної послідовності була рівномірною [2]. Виконати обидві ці умови для якісного моделювання вкрай важко. На даний час існує велика кількість програмних варіантів генераторів ПВЧ, але вони або занадто складні і потребують великих обчислювальних ресурсів, або не забезпечують необхідного рівню рівномірності розподілення символів у складі вихідної послідовності і, через це, проблема побудови хороших економічних програмних генераторів такого типу залишається достатньо актуальною.

Аналіз останніх досліджень і публікацій. Історія створення програмних алгоритмів генерації псевдовипадкових чисел розпочинається практично одночасно з появою обчислювальної техніки. З тої пори цей напрям став добре вивченою областю математики. Першою значущою

роботою у цій сфері була стаття Дж. фон Неймана “Різні методи, що пов'язані з отриманням випадкових чисел” [3], де він дав оцінку намаганням тодішнього покоління математиків створювати випадкові числа арифметичними методами і надав математичне обґрунтування, відомим на той час, способам створення комп'ютерних моделей із заданими елементарними законами розподілення ймовірностей. Основним висновком, який він зробив, є твердження про те, що випадковість створених алгоритмічними методами числових послідовностей ПВЧ, може бути оцінена лише методами статистичного аналізу.

Наступною фундаментальною роботою, де були обґрунтовані наукові математичні підходи до оцінки “випадковості” чисел, створених арифметичними методами, була робота Д. Кнута “Мистецтво програмування” [2]. У цій роботі він описав, відомі на той час, прості елементарні алгоритми генерації ПВЧ, обґрунтував якість їх вихідних послідовностей і дав рекомендації щодо використання основних критеріїв тестування.

Досягнення у розвитку обчислювальної техніки та постійне збільшення її обчислювальних ресурсів привели до збільшення популярності комп'ютерного моделювання серед фахівців, що працюють над підвищенням його ефективності. Їх зусилля були зосереджені на створенні нових генераторів ПВЧ, які забезпечували б кращі показники рівномірності вихідного числового потоку і, в той же час, залишалися не вибагливими до таких обчислювальних ресурсів як пам'ять і процесорний час [4].

Протягом останніх десятиріч з'явилися такі методи генерації ПВЧ як “шторм Мерсена” (Mersenne twister) [5, 6, 7] та генератор Xorshift [8]. Обидва ці генератори проходять пакет тестів, визначений стандартом Національного інституту стандартів і технологій Сполучених Штатів Америки (National Institute of Standards and Technology – NIST) [9]. Сьогодні МТ-генератор входить до складу бібліотек переважної більшості середовищ програмування.

Що стосується МТ-генератора, то його автори ставили за мету забезпечення максимально можливого періоду повторення і це їм вдалось. Вони побудували генератор, період якого T дорівнює $2^{19937} - 1$, але це досягається за рахунок алгоритму, який містить вкрай витратні обчислювальні операції, що виконуються над w -бітними цілими числами над кінцевим двійковим полем [6].

Під час розроблення Xorshift-генератора його автор Джордж Марсалья ставив собі за мету, перш

за все, створити його економічним і швидкодіючим. Причина полягає в тому, що паралельно з розвитком і удосконаленням комп'ютерної техніки суттєво збільшується складність прикладного програмного забезпечення, і, через це, неவிбагливість будь-якого обчислювального алгоритму до необхідних обчислювальних ресурсів завжди буде залишатись актуальною задачею [10, 11].

Постановка завдання. Мета статті полягає у проведенні аналізу існуючих алгоритмів формування рівномірно розподілених псевдовипадкових числових послідовностей, придатних для комп'ютерного моделювання стохастичних процесів з урахуванням ефективності використання ними обчислювальних ресурсів.

Виклад основного матеріалу. Генерація випадкових величин для потреб комп'ютерного моделювання передбачає забезпечення їх відповідності певному розподіленню ймовірностей, що описується функцією щільності розподілення ймовірностей $f(x)$ (Probability Density Function – PDF). Це може бути рівномірне розподілення, нормальне, пуассонівське або інше, в залежності від умов експерименту. Загальні процедури, що використовуються у моделюванні, описані в роботі [12]. Повне математичне обґрунтування всіх стандартних статистичних розподілень та алгоритми їх симуляції можна знайти в роботі [13].

Формування числових розподілень, що можуть бути описані аналітичними виразами, вимагають наявності початкового вхідного рівномірно розподіленого на відрізьку $[0, 1]$ числового потоку $x_1, x_2, \dots, x_n$. Для його отримання зазвичай використовується генератор ПВЧ. Найпростішим алгоритмом такого типу є найбільш популярний, запропонований Лемером [14], лінійний конгруентний генератор (ЛКГ). Він може бути описаний так

$$x_{n+1} = (ax_n + c) \bmod m,$$

де m – це модуль (ціле число, таке, що $m \geq 2$), a – множник (ціле число, таке, що $0 < a < m$), c – доданок (ціле число, таке, що $0 \leq c < m$).

Довжина періоду такого генератора T дорівнює величині m , лише тоді, коли числа c і m взаємно прості, різниця $b = a - 1$ ділиться на просте p , яке є дільником m і, нарешті, якщо b кратне 4, у разі, коли m кратне 4. У всіх інших випадках період T менший за m . Саме невеликий період повторення часто обмежує застосування цього генератора ПВЧ у моделюванні і, тому, там де вимоги до точності відтворення випадкового процесу високі, доводиться користуватись іншими алгоритмами.

У разі, коли хороший генератор псевдовипадкових рівномірно розподілених чисел обрано, для генерації процесу із заданим розподіленням $f(x)$ можна скористатись або методом відбракування чисел, що не задовольняють заданому типу розподілення або методом зворотних функцій.

Метод відбракування чисел дозволяє отримати реалізацію необхідного випадкового процесу $f(x)$ із чисел деякого іншого випадкового процесу $g(x)$, відкидаючи ті з них, що не вписуються у процес, який підлягає моделюванню. Він використовується коли пряма генерація випадкових чисел цільового розподілення не можлива.

У відповідності до цього методу, спочатку обирають генератор додаткового розподілення зі щільністю $g(x)$, вихідні значення якого, більше або дорівнюють розподіленню процесу зі щільністю $f(x)$.

Потім генерують послідовність псевдовипадкових чисел $x_1, x_2, \dots, x_n$ додаткового розподілення, наприклад, рівномірно розподіленого на інтервалі $[a, b]$.

Далі, для кожного сформованого числа x_i з розподіленням $g(x)$, обчислюється значення функції $f(x_i)$. Якщо виявляється, що значення функції $g(x_i)$ більше ніж значення функції $f(x_i)$, воно відкидається.

Описана послідовність дій повторюється до тих пір поки не буде набрано достатню кількість необхідного статистичного матеріалу з розподіленням $f(x)$.

Такий спосіб моделювання дозволяє скористатись генератором з довільним розподіленням, не звертаючи уваги на відповідність сформованих ним чисел будь якому розподіленню $g(x)$. Єдиною вимогою залишається великий період повторення T . Його негативною стороною є велика витратність – у разі, коли основне й додаткове розподілення не близькі, кількість відбракованих величин занадто велика.

Метод моделювання випадкового процесу, оснований на використанні зворотних функцій, передбачає наявність аналітично визначеної функції розподілення ймовірностей $F(x)$ (Cumulative Distribution Function – CDF), що описує такий процес. Доказано, що коли безперервна випадкова величина x з довільним розподіленням має відповідну функцію розподілу ймовірностей $F(x)$, існує відповідна до неї, рівномірно розподілена на інтервалі $[0, 1]$ випадкова безперервна величина $\alpha = F^{-1}(\alpha)$ [15].

Моделювання з використанням такого методу передбачає, перш за все, існування функції розподілення ймовірностей, що описує процес, який

має моделюватись. Після цього необхідно вибрати генератор псевдовипадкових чисел, що забезпечує високий рівень рівномірності розподілення чисел $P(\alpha)$ у вихідному потоці $\alpha_1, \alpha_2, \dots, \alpha_n$ на інтервалі $[0, 1]$. Наприкінці, через зворотну функцію $P^{-1}(\alpha)$ визначаються величини $x_1, x_2, \dots, x_n$ числового потоку з необхідним розподіленням $F(x)$.

Метод зворотних функцій забезпечує високу точність і простоту програмної реалізації, але вимагає наявності аналітичного описання процесу, що підлягає моделюванню і, головне, високоякісного, з точки зору рівномірності розподілення символів у вихідній послідовності. Саме він найчастіше використовується у більшості випадків.

Таким чином, однією з проблем комп'ютерного моделювання є вибір генератора ПВЧ, що має найбільшу ентропію [16], тобто ймовірність появи нульового символу у складі вихідної послідовності $P(x_i = 0)$ повинна максимально наближатись до ймовірності одиничного символу $P(x_i = 1)$. Розбіжність між цими величинами e називають зміщенням вихідного розподілення чисел і визначають як

$$e = \frac{1}{2} (P(x_i = 1) - P(x_i = 0)),$$

де величини $P(x_i = 1) = \frac{1}{2} + e$ і $P(x_i = 0) = \frac{1}{2} - e$, а x_i – це символи, що належать, послідовності $x_1, x_2, \dots, x_n$. Саме цей показник використовується для оцінки придатності якості обраного для потреб моделювання генератора ПВЧ. Для високоякісного генератора ця величина має наближатись до величини 0.5, а величина e повинна бути мінімально можливою.

Історично, пошук придатних для комп'ютерного моделювання алгоритмів генерації ПВЧ мав на меті забезпечення простоти обчислень і високого ступеню рівномірності розподілення сформованих чисел та великої довжини періоду повторення T . Оскільки лінійних конгруентний метод не забезпечував виконання цих вимог, було запропоновано й інші типи генераторів. Найбільш відомим з них були генератори Фібоначчі з запізненням та лінійні генератори, побудовані на основі регістрів зсуву зі зворотним зв'язком (Linear Feedback Shift Register – LFSR).

Принцип роботи генераторів Фібоначчі з запізненням у різних його модифікаціях добре описані Д. Кнутом в роботі [2] і, пізніше, їх детальну оцінку було надано в роботі [17]. Принцип дії такого генератора описується виразом

$$X_n = (X_{n-24} + X_{n-55}) \bmod m, \quad n \geq 55,$$

де m – парне число, а $X_0, \dots, X_{54}$ довільні цілі числа.

Тут вихідне число утворюється з двох попередніх чисел, затриманих на тактів. Числа 24 і 55 – це числа Фібоначчі, їх значення знайдено простим підбором. В роботі [2] наведено й інші пари таких чисел, що забезпечують прийнятний рівень рівномірності їх розподілення, що свого часу було підтверджено відповідними статистичними випробуваннями. Інші пари підібраних чисел Фібоначчі достатньо великі і це приводить до суттєвого збільшення необхідного об'єму пам'яті. Довжина періоду послідовності на виході генератора Фібоначчі складає $2^{q-1}(2^{55} - 1)$ де q – це розрядність мікропроцесора. В минулому сторіччі такі генератори ПВЧ вважались одними з найкращих, але сьогодні їх використовують лише у якості складових частин більш складних алгоритмів.

Генератори псевдовипадкових чисел, побудовані на основі регістрів зсуву зі зворотним зв'язком широко використовувались раніше в теорії кодування для формування так званих m послідовностей для потреб зв'язку. Саме через це їх математична теорія добре розроблена і описана у багатьох наукових джерела. Їх популярність визначалась тим, що вони дозволяли просту апаратну реалізацію.

Такі генератори уявляють собою послідовність бітів, кількість яких n , визначається довжиною регістра, де вони зберігаються.

Як показано на рисунку 1, кожного разу, після зсуву всіх бітів регістра праворуч на одну позицію, крайній зліва біт є функцією тих бітів, що об'єднуються функцією зворотного зв'язку, яка, у свою чергу, визначається ненульовими коефіцієнтами утворюючого характеристичного поліному $p(x)$ [18].

Для того, щоб LFSR генератор мав максимальний період, його характеристичний поліном повинен бути примітивним по модулю 2, тобто він не повинен розкладатись на простіші поліноми. Його максимальний період не перевищує величини $2^n - 1$.

Сьогодні такі генератори успішно реалізують програмними методами, але вони, так само, як і генератори Фібоначчі з запізненням, використовуються лише у якості складових частин інших алгоритмів формування ПВЧ.

Враховуючи той факт, що сьогодні темпи зростання складності програмного забезпечення суттєво перевищують темпи збільшення ресурсів

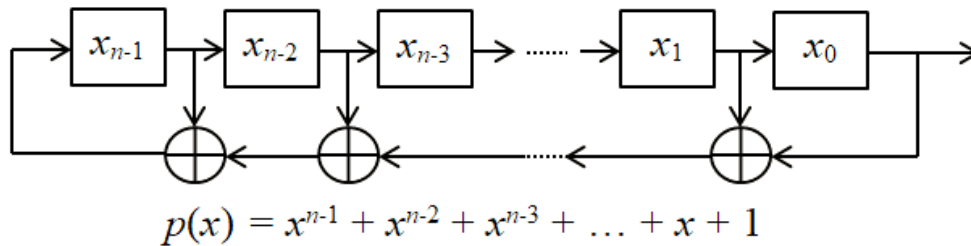


Рис. 1. LFSR регістр у конфігурації Фібоначчі

обчислювальної техніки, зусилля математиків, що займаються комп'ютерним моделюванням, зосереджені на пошуку економічних алгоритмів формування ПВЧ. Хорошим прикладом рішення такої задачі є запропонований Джорджем Марсальєю генератор, відомий під назвою Xorshift [6]. Цей генератор уявляє собою комбінацію декількох LFSR генераторів, що побудовані на основі характеристичних поліномів невисокого ступеню. Головним досягненням автора є те, що у складі запропонованого алгоритму використовуються лише такі економічні обчислювальні операції як зсув та складання по модулю два.

Алгоритм Xorshift заснований на використанні множини ненульових двійкових векторів 1×32 із множини Z , а лінійне перетворення f над множиною Z , описується невідродженою бінарною матрицею T розміром 32×32 .

В роботі [19] було показано, що вихідна послідовність чисел такого генератора буде виглядати як $yT, yT^2, yT^3, \dots$ лише тоді, коли порядок T дорівнює $2^{32} - 1$ у групі двійкових невідроджених матриць розміром 32×32 , а період вихідної послідовності дорівнює $2^{32} - 1$. Також, було показано, що максимально можливий період забезпечується тоді, коли алгоритм забезпечує такі типи зсувів як

$$y^{\wedge} = y \ll 13; y^{\wedge} = y \ll 17; y^{\wedge} = y < 5.$$

У разі виконання описаних умов, Xorshift алгоритм може бути реалізований мовою C так, як це показано на рисунку 2.

Тут x, y, z і w – це 32-бітні початкові числа, що визначають внутрішній стан генератора та вигляд послідовності, яка формується 32-бітними вихідними числами w . На кожному поточному кроці алгоритму нове вихідне число w утворюється як комбінація розрядів чисел x та поточного числа w . Спочатку число x перетворюється у число t шляхом виконання логічних операцій та операції зсуву, що підвищує “випадковість” його молодших байтів, а потім відбувається послідовний зсув чисел $y \rightarrow x, z \rightarrow y$, та $w \rightarrow z$. Наприкінці отримується нове значення числа w , як результату

```
x = 123456789
y = 362436069
z = 521288629
w = 88675123
t = x ^ ((x << 11) & 0xFFFFFFFF);
x = y;
y = z;
z = w;
w = (w ^ (w >> 19)) ^ (t ^ (t >> 8));
```

Рис. 2. Програмна реалізація алгоритму Xorshift

виконання кількох логічних операцій та операцій зсуву над числами t і поточного значення w .

Таким чином, завдяки використанню в алгоритмі Xorshift таких економічних операцій, як складання по модулю два та зсуву, фактично виконується нерегулярне перемішування бітів 32-бітних чисел і, як показує проведений статистичний аналіз [11, 20], він може бути хорошим кандидатом на використання у якості джерела випадковості для комп'ютерного моделювання.

Висновки. В проведеному дослідженні розглянуто алгоритми, які можуть бути використані як джерела випадковості при комп'ютерному моделюванні стохастичних процесів. Було показано, що у тих випадках, коли процес, який підлягає дослідженню, може бути описаний за допомогою аналітичного виразу, зазвичай, використовують метод зворотних функцій, що вимагає використання генератора псевдовипадкових чисел з жорсткими вимогами до рівномірності розподілення вихідних чисел. Другою суттєвою вимогою є ефективність такого генератора, яка визначається швидкістю його роботи та невибагливістю до необхідної пам'яті і процесорного часу.

З проведеного аналізу було зроблено висновки про те, що перші покоління генераторів ПВЧ, таких як генератори Фібоначчі з запізненням та LFSR генератори, не забезпечують достатньо великого періоду повторення T , і через це, у більшості випадків вони можуть бути використані

лише як складові частини більш складних алгоритмів, а це, у свою чергу, робить їх мало ефективними. Теж саме стосується лінійного конгруентного генератора. Таким чином, найбільш придатним для досліджень випадкових процесів шляхом стохастичного комп'ютерного моде-

лювання є, розроблений Джоржем Марсальєю, Xorshift генератор та його модифікації. Наприкінці, слід також зауважити, що Xorshift генератор зовсім не є безпроблемним. Бажано було б знайти спосіб збільшення довжини його періоду і це має бути одним з напрямів подальших досліджень.

Список літератури:

1. Law, A.M. and W. Kelton, 2000. Simulation Modeling and Analysis. 3rd Edn., McGraw-Hill, New York, USA. URL: <https://surl.li/ftjgxz>
2. Knuth, D. E. The Art of Computer Programming. Volume 2. Seminumerical Algorithms. 3rd edition / D. E. Knuth. – Boston, Mass, USA : Addison Wesley, Longman Publishing, 1997. – 762 p. ISBN 0-201-89683-4. URL: <https://surl.li/txwhys>
3. Kamalika Bhattacharjee, Krishnendu Maity, Sukanta Das. A Search for Good Pseudo-random Number Generators : Survey and Empirical Studies. DOI:10.48550/arXiv.1811.04035. November 2018. URL: <https://arxiv.org/pdf/1811.04035>
4. John von Neumann. Various techniques used in connection with random digits. Collected Works, 1963, 5: 768-770. URL : <https://surl.li/ancwgf>
5. M. Matsumoto, T. Nishimura. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. ACM Transactions on Modeling and Computer Simulation. 8 (1): 3–30. 1998. DOI:10.1145/272991.272995. URL: <https://surl.li/wirfwn>
6. Saito, M. An Application of Finite Field: Design and Implementation of 128-bit Instruction-Based Fast Pseudorandom Number Generator. Dept. of Math. Graduate School of Science, February 9th, 2007. URL: <http://www.math.sci.hiroshima-u.ac.jp/~m-mat/MT/SFMT/M062821.pdf>
7. Yurii Shcherbina, Nadiia Kazakova, Oleksii Frazе-Frazenko. Using the Xorshift generator to simulate stochastic processes. Processing, transmission and security of information – 5 December 2022. URL: <https://ceur-ws.org/Vol-3200/paper39.pdf>
8. George Marsaglia. Xorshift RNGs. Journal of Statistical Software 08(i14) January 2003. DOI:10.18637/jss.v008.i14. URL: https://www.researchgate.net/publication/5142825_Xorshift_RNGs
9. A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications. SP 800-22 Rev. 1a, April 2010. URL: <https://surl.li/rixqnc>
10. Richard P. Brent. Note on Marsaglia's Xorshift Random Number Generators. Journal of Statistical Software. August 2004, Volume 11, Issue 5. DOI:10.18637/jss.v008.i14. URL: <https://www.jstatsoft.org/article/view/v011i05>
11. Yurii Shcherbina, Nadiia Kazakova, Oleksii Frazе-Frazenko, Oleksandr Laptiev, Andrii Sobchuk. SELECTION OF RANDOMNESS SOURCE FOR COMPUTER SIMULATION. DOI: <https://doi.org/10.18372/2310-5461.59.17944>. URL: <https://jrnل.nau.edu.ua/index.php/SBT/article/view/17944>
12. G. Jay Kerns. Introduction to Probability and Statistics Using R. First Edition. GNU Free Documentation License. ISBN: 978-0-557-24979-4. July 28, 2010. – 412 p. URL : <https://ipsur.r-forge.r-project.org/book/download/IPSUR.pdf>
13. Robert V. Hogg, Joseph W. McKean, Allen T. Craig. Introduction to Mathematical Statistics. Seven Edition. ISBN 978-0-321-79543-4. URL : <https://surl.li/advtyx>
14. Lehmer, D.H. (1951) Mathematical Methods in Large-Scale Computing Units. The Annals of the Computation Laboratory of Harvard University, 26, 141-146.
15. Ross S. A First Course in Probability. 8th Edition. 2010. ISBN-13: 978-0136033134 URL: https://github.com/datalater/probability/blob/master/A_First_Course_in_Probability_8th_Edition.pdf
16. Siew-Hwee Kwok, Yen-Ling Ee, Guanhan Chew, Kanghong Zheng, Khoongming Khoo, Chik-How Tan. A Comparison of Post-Processing Techniques for Biased Random Number Generators. WISTP 2011: Information Security Theory and Practice. Security and Privacy of Mobile Devices in Wireless Communication. pp 175–190. URL: https://link.springer.com/content/pdf/10.1007/978-3-540-74619-5_9.pdf
17. Ethan Bogle, Owen Brass, Owen Shen. Fibonacci random generator and fourier analysis. ID: 237526214. 2021. URL : <https://surl.li/tctfxb>
18. Golomb, S.W. Shift Register Sequences. Aegean Park Press, Laguna Hill, 1982. 324 p. URL : <https://surl.li/oeoizz>
19. François Panneton, Pierre L'Ecuyer. On the xorshift random number generators. Transactions on Modeling and Computer Simulation 15(4): 346-361. DOI:10.1145/1113316.1113319. URL : <https://surl.li/qxkxbx>
20. Richard P. Brent. Note on Marsaglia's Xorshift Random Number Generators. Journal of Statistical Software 11(i04). July 2004. DOI:10.18637/jss.v011.i05. URL : <https://surl.li/daujao>

Shcherbyna Yu.V., Loginova N.I., Manakov S.Yu. ANALYSIS OF MODERN METHODS OF COMPUTER MODELING OF STOCHASTIC PROCESSES

The article is devoted to the problem of improving the quality of stochastic computer modeling during the design of complex technical systems that ensure the operation of transport, production, management of infrastructure facilities and data exchange in information and communication systems. The effectiveness of such modeling is determined by the presence of adequate mathematical models that formally describe the behavior of real systems in stochastic conditions of their operation, and the quality of sources of randomness that allow reproducing such conditions. One of the main problems that determines the quality of stochastic computer modeling is the choice of a software method for generating pseudo-random numbers with a high degree of uniformity of their probability distribution. This problem should be solved in two directions: first, the algorithm for forming a pseudo-random number stream should be unpretentious to such computational resources as memory and processor time and, secondly, such a stream should not differ from a real random uniformly distributed number stream. The paper provides an analysis of known software methods for generating pseudo-random numbers used for computer modeling. Those considered to be the most economical are considered, namely, the Fibonacci generator with a delay, generators based on linear shift registers with feedback, and the Xorshift generator, which is based on economical, from the point of view of computer implementation, logical operations and shift operations. Based on the analysis of scientific publications, it was concluded that, taking into account modern requirements for the quality of computer modeling, the Xorshift generator should be used as the most effective source of randomness. As for the Fibonacci generators with a delay and generators based on shift registers with feedback, they can be used only as components of more complex algorithms for generating pseudo-random numbers.

Key words: computer modeling, inverse function method, random number rejection method, pseudorandom number generator, statistical analysis.

Дата надходження статті: 25.07.2025

Дата прийняття статті: 31.07.2025

Опубліковано: 27.10.2025

Яковенко В.О.

Університет митної справи та фінансів

Прокопович-Ткаченко Д.І.

Університет митної справи та фінансів

Ульяновська Ю.В.

Університет митної справи та фінансів

ЕНЕРГЕТИЧНО-ІНФОРМАЦІЙНА МОДЕЛЬ ДЕГРАДАЦІЇ ТЕЛЕКОМУНІКАЦІЙНИХ ВУЗЛІВ ПІД ВПЛИВОМ ГІБРИДНИХ КІБЕРАТАК

У статті докладно розглядається розробка енергетично-інформаційної моделі деградації телекомунікаційних вузлів, які складають основу сучасної критичної інфраструктури, під впливом гібридних кібератак. З особливою увагою проаналізовано сценарії, коли на вузли здійснюється комбінований вплив: масована DDoS-атака поєднується з фізичними загрозами, такими як знеструмлення, GPS-спуфінг, чи термічне перевантаження. В таких умовах традиційні підходи до моніторингу, що фокусуються лише на логічних параметрах мережі – затримці, джитері, пропускній здатності – виявляються недостатніми, оскільки не враховують зміни у фізіологічному стані обладнання. Деградація оперативної пам'яті, нестабільність живлення чи аномальне зростання температури залишаються поза увагою, що створює приховані ризики й ускладнює своєчасне реагування на комплексні загрози.

Запропонована модель ґрунтується на фізично обґрунтованому підході до опису функціонування телекомунікаційного вузла – наприклад, маршрутизатора чи вузла агрегації – в умовах багатокомпонентного навантаження. Враховано фазовий перехід системи з нормального стану до деградованого або передвідмовного під впливом зростаючих інформаційних і енергетичних факторів. Модель дозволяє аналізувати зміну енергетичного профілю, температурного режиму, завантаження процесорних та пам'яті ресурсів у реальному часі. Особливий акцент зроблено на розподілі процесів на часові квазі-стани, що дає змогу фіксувати моменти, коли вузол переходить у критичну зону нестабільності.

У статті детально розглянуто кейс-аналіз DDoS-атаки на вузол агрегації із супровідною GPS-спуфінг-атакою, що провокує одночасне порушення логічної та фізичної стійкості. Продемонстровано, що синергетичний ефект комбінованих атак призводить до нелінійного наростання навантаження, яке не піддається виявленню ізольованими індикаторами. Визначено критичні порогові значення температури, енергоспоживання та падіння пропускної здатності, за досягнення яких відбувається незворотний фазовий перехід вузла у деградований стан. Запропоновано ефективні методи візуалізації зон ризику на основі розрахунку інтегрального енергетичного графа для практичного впровадження у цифрові двійники, що забезпечує своєчасний моніторинг та підвищує стійкість обладнання до гібридних кібератак.

Ключові слова: гібридна кібератака, DDoS, деградація вузла, критична інфраструктура, енергоспоживання, фазовий перехід, скінченна модель, GPS-спуфінг, термічне навантаження, цифровий двійник, пропускна здатність.

Постановка проблеми. У сучасних умовах стрімкої цифровізації та інтеграції інформаційних технологій у критичну інфраструктуру особливої актуальності набуває проблема забезпечення стійкості телекомунікаційних вузлів до гібридних кібератак. Зокрема, поєднання інформаційних і фізичних впливів – таких як масовані

DDoS-атаки, GPS-спуфінг, теплові чи енергетичні аномалії – призводить до ускладнення виявлення потенційних відмов і, відповідно, збільшує ризик деградації або зупинки роботи систем у реальному часі. Традиційні підходи до моніторингу стану мережевого обладнання здебільшого орієнтовані на аналіз логічних або мережевих параметрів –

затримки, джитеру, пропускну здатності, – при цьому часто ігноруються зміни у фізіологічному стані системи, такі як температурне чи енергетичне навантаження, деградація оперативної пам'яті або нестабільність живлення. В умовах мультидомених загроз це призводить до зростання латентних ризиків, які не фіксуються ізолюваними індикаторами, але проявляються як синергетичний ефект, що значно прискорює фазовий перехід вузла від нормального до деградованого або передвідмовного стану. Особливо гострою є проблема раннього виявлення критичних зон нестабільності, коли класичні моделі не дозволяють адекватно спрогнозувати стійкість вузла під дією комбінованих навантажень. Відсутність урахування нелінійної динаміки деградації, а також точок біфуркації чи квазіперіодичності в енергетичному профілі, обмежує можливості побудови ефективних систем захисту та своєчасного реагування на комплексні атаки. У зв'язку з цим виникає потреба у розробленні фізично обґрунтованої енергетично-інформаційної моделі, що дозволяє інтегрувати логічні, фізичні та енергетичні параметри вузла для цілісної оцінки його стійкості та прогнозування моментів втрати працездатності, а також забезпечує можливість практичної інтеграції у цифрові двійники для оперативного моніторингу та керування ризиками у реальному часі.

Аналіз останніх досліджень і публікацій. Останні дослідження у сфері стійкості телекомунікаційних вузлів дедалі більше акцентують на міждисциплінарному підході до аналізу впливу гібридних загроз. Зокрема, у роботах [1]–[4] розглядається уразливість мережевого обладнання до переважань та порушень синхронізації, спричинених DDoS-атаками й GPS-спуфінгом. Автори зазначають, що навіть мінімальні флуктуації сигналу можуть провокувати критичні відхилення в роботі фазових вимірювальних пристроїв, що співзвучно з результатами [5]–[8], де особливу увагу приділено аналізу часових вікон та аварійного вимкнення вузлів унаслідок хибної синхронізації.

Питання подвійного впливу – одночасного навантаження на логічному та енергетичному рівнях – детально висвітлено у [9], [10]. Тут наведено докази, що поєднання DDoS-атак із синхронізаційними аномаліями створює специфічний синергетичний ефект, який класичні моделі неспроможні зафіксувати або спрогнозувати. Методи виявлення таких комбінованих атак, від евристичних алгоритмів до глибокого навчання, систематизовані у низці публікацій [11]–[14]. Водночас,

ці моделі переважно ігнорують фізіологічні параметри системи, такі як температурні режими чи деградація оперативної пам'яті, на що звертають увагу у [15]–[17].

Енергетичний профіль вузла як важливий індикатор стійкості розглядається у роботах [18]–[21]. Проте, як підкреслюють автори, ці підходи ще не знаходять широкого застосування в моделях фазових переходів. У статтях [22]–[25] обґрунтовується доцільність мультидоменого аналізу, зокрема синтезу логічних і фізичних характеристик вузлів для виявлення латентної деградації до настання критичних збоїв. Актуальність цього напрямку підтверджується й розвитком smart-grid технологій, де навіть короточасні відхилення можуть мати системний ефект [26], [27].

Значний внесок у практичну реалізацію моделей забезпечують дослідження цифрових двійників [28]–[31], які дозволяють інтегрувати широкий спектр фізичних та інформаційних параметрів для оперативного моніторингу стану обладнання. Аналітика фазових переходів та прогнозування деградації системи в реальному часі розкриті у [32]–[35].

Окремо варто відзначити роботи [36]–[38], де обґрунтовується важливість урахування нелінійного впливу гібридних атак, коли логічні і фізичні навантаження здатні провокувати неочевидні зони нестабільності. Відсутність урахування точок біфуркації та змін періодичності у профілі деградації, як показано у [39], [40], значно знижує точність прогнозування стійкості вузлів критичної інфраструктури.

Нарешті, у контексті посилення гібридних загроз у глобальних телекомунікаційних мережах ЄС, США, Ізраїлю та країн Азії [41], проблема побудови уніфікованих моделей, що інтегрують логіко-інформаційні та енергетичні процеси, набула стратегічної ваги для забезпечення надійності та стійкості критичних систем.

Постановка завдання. У сучасних умовах стрімкої цифровізації та інтеграції інформаційних технологій у критичну інфраструктуру особливої актуальності набуває проблема забезпечення стійкості телекомунікаційних вузлів до гібридних кібератак. Зокрема, поєднання інформаційних і фізичних впливів – таких як масовані DDoS-атаки, GPS-спуфінг, теплові чи енергетичні аномалії – призводить до ускладнення виявлення потенційних відмов і, відповідно, збільшує ризик деградації або зупинки роботи систем у реальному часі. Традиційні підходи до моніторингу стану мережевого обладнання здебільшого орієнтовані

на аналіз логічних або мережевих параметрів – затримки, джитеру, пропускної здатності, – при цьому часто ігноруються зміни у фізіологічному стані системи, такі як температурне чи енергетичне навантаження, деградація оперативної пам'яті або нестабільність живлення. В умовах мультидоменних загроз це призводить до зростання латентних ризиків, які не фіксуються ізолюваними індикаторами, але проявляються як синергетичний ефект, що значно прискорює фазовий перехід вузла від нормального до деградованого або передвідмовного стану. Особливо гострою є проблема раннього виявлення критичних зон нестабільності, коли класичні моделі не дозволяють адекватно спрогнозувати стійкість вузла під дією комбінованих навантажень. Відсутність урахування нелінійної динаміки деградації, а також точок біфуркації чи квазіперіодичності в енергетичному профілі, обмежує можливості побудови ефективних систем захисту та своєчасного реагування на комплексні атаки. У зв'язку з цим виникає потреба у розробленні фізично обґрунтованої енергетично-інформаційної моделі, що дозволяє інтегрувати логічні, фізичні та енергетичні параметри вузла для цілісної оцінки його стійкості та прогнозування моментів втрати працездатності, а також забезпечує можливість практичної інтеграції у цифрові двійники для оперативного моніторингу та керування ризиками у реальному часі.

Виклад основного матеріалу. Упродовж останнього десятиліття зростає інтерес до дослідження деградаційних процесів у телекомунікаційних вузлах критичної інфраструктури внаслідок гібридних кібератак, що поєднують інформаційні та фізичні впливи. Зокрема, комбінація DDoS-атак з GPS-спуфінгом, тепловими або енергетичними аномаліями ускладнює своєчасне виявлення порушень та збільшує ризик відмови системи в реальному часі. Телекомунікаційне обладнання, таке як маршрутизатори, концентратори та контролери, є особливо чутливими до перевантажень, спричинених масованими запитами чи втручаннями в канали синхронізації [1]–[4]. Окремі дослідження доводять, що GPS-спуфінг, навіть із мінімальним рівнем втручання, може призвести до хибної синхронізації фазових вимірювальних пристроїв (PMU), порушення часових вікон і, зрештою, до аварійного вимкнення [5]–[8]. Проблема посилюється, якщо атакуючі комбінують спуфінг із високочастотними DDoS-навантаженнями на транспортний рівень мережі, що спричиняє подвійний тиск на вузол – як у часовому, так і в енергетичному вимірах [9], [10].

Сучасні підходи до виявлення таких атак включають як класичні евристичні алгоритми [11], [12], так і методи глибокого навчання, що аналізують структуру трафіку [13], [14]. Попри це, більшість моделей орієнтовані на логічні або мережеві параметри – latency, jitter, throughput, – і майже не враховують фізіологію системи як такої: її теплове навантаження, зміну напруги живлення, деградацію оперативної пам'яті [15]–[17]. У низці праць також підкреслено, що енергетичний профіль вузла є важливим індикатором стійкості, але ще не знайшов широкого застосування у прикладних моделях фазових переходів [18]–[21]. Розширення класичних моделей за рахунок мультидоменного аналізу – наприклад, поєднання логічних логів та телеметрії енергоспоживання – дозволяє виявити латентні деградаційні процеси до їх катастрофічного прояву [22]–[25]. Це особливо актуально в умовах smart-grid та гетерогенних телекомунікаційних середовищ, де навіть короточасні коливання потужності або часової мітки можуть мати системний ефект [26], [27]. Низка авторів пропонують використовувати цифрові двійники як засоби інтеграції фізичних, логічних та інформаційних параметрів вузла [28]–[31]. В таких моделях враховуються температурні профілі, вібраційні впливи, швидкість заповнення черг, ефективність систем охолодження тощо. Це дозволяє в реальному часі прогнозувати фазові переходи – тобто момент переходу вузла від нормального до деградованого або передвідмовного стану [32]–[35]. Однак, залишається проблема адекватної оцінки нелінійного впливу гібридної атаки – коли логічне навантаження (через DDoS) індукує зростання енергоспоживання та температури, а синхронізаційна атака (GPS-спуфінг) спричиняє логічний хаос у маршрутизації [36]–[38]. Поєднання таких впливів формує критичні зони нестабільності, в яких класичні метрики є інформативно недостатніми. Зокрема, відсутність урахування точок біфуркації або квазіперіодичності у профілі деградації вузла обмежує точність прогнозування стійкості системи [39], [40]. В контексті останніх загроз проти телекомунікаційних мереж ЄС, США, Ізраїлю та країн Азії [41], питання побудови моделі, що дозволяє описувати фазові переходи у термінах як логіко-інформаційної, так і енергетичної динаміки вузлів, стало критичним. Це зумовлює потребу у створенні уніфікованої енергетично-інформаційної моделі деградації вузла під впливом комбінованого кіберфізичного впливу, здатної інтегруватися в цифрові двійники для аналізу критичних

ситуацій у реальному часі. Метою цієї публікації є розроблення фізично обґрунтованої енергетично-інформаційної моделі фазового переходу телекомунікаційного вузла критичної інфраструктури з нормального до деградованого стану під впливом комбінованої DDoS-атаки та зовнішнього фізичного впливу (зокрема GPS-спуфінгу), з урахуванням параметрів енергоспоживання, температурного навантаження, завантаження оперативної пам'яті та процесорних ресурсів, а також інтеграції моделі до цифрових двійників для оцінки точок втрати працездатності в реальному часі.

Методи:

Методологія побудови енергетично-інформаційної моделі фазового переходу телекомунікаційного вузла ґрунтується на поєднанні кількох напрямів дослідження, які охоплюють аналітичне, симуляційне та сценарне моделювання. Усі етапи були реалізовані з урахуванням потреб подальшої інтеграції моделі у середовище цифрового двійника з можливістю виявлення критичних станів у реальному часі.

Моделювання гібридної атаки

Гібридна атака описується через два синхронізовані канали впливу: логічний (DDoS) та фізичний (GPS-спуфінг). DDoS-атака симулює вплив на транспортний рівень через зростання кількості запитів, що імітують високонавантажenu мережу. GPS-спуфінг моделює зміну фазової синхронізації – що в реальній інфраструктурі викликає аномалії часової координати.

$$f_{att} = \lambda_0 \cdot (1 + \sin(\omega t)), \quad (1)$$

$$\Delta\varphi(t) = \varphi_{\text{реальний}}(t) - \varphi_{\text{спуфінг}}(t). \quad (2)$$

Результат: при імітації обох впливів виявлено, що навіть помірні флуктуації GPS-сигналу при високому навантаженні можуть викликати деградацію вузла ще до досягнення логічної межі DDoS, через кумулятивний ефект асинхронізації.

Енергетично-інформаційна модель вузла

Кожному моменту часу приписується вектор ознак:

$$\vec{S}_i = (P(t), T(t), M(t), C(t)) \quad (3)$$

$$D(t) = \alpha P(t) + \beta T(t) + \gamma M(t) + \delta C(t) \quad (4)$$

Результат: визначено, що найбільш критичним є внесок температури та навантаження CPU. У процесі калібрування модель показала, що при однаковому трафіку вузли з меншою ефективністю тепловідведення досягають порогу D_{cr} на 20–30% швидше.

Розширена теплова модель вузла

Модель теплового накопичення:

$$C_{th} \cdot \frac{dT}{dt} = P(t) - \frac{T(t) - T_{env}}{R_{th}} \quad (5)$$

Включення цієї моделі дозволило враховувати не лише миттєве навантаження, а і теплову інерційність – вузли, які досягли критичного теплового рівня, залишаються в перегрітій зоні навіть після завершення атаки. Це важливо для прогнозування відкладених деградацій, особливо в умовах багаторазових впливів.

Пояснення: обрані значення відображають параметри промислових маршрутизаторів і були адаптовані до типових умов використання у польових вузлах критичної інфраструктури.

Інтеграція до цифрового двійника

Для динамічного контролю використано:

$$R(t) = \frac{dD(t)}{dt} + \kappa \cdot \sigma(D), H(t) = -\sum_i p_i(t) \log p_i(t) \quad (6)$$

Ентропійна компонента дозволяє оцінити нестабільність функціонування навіть у випадках, коли усереднені метрики ще не перевищують порогових значень. Додатково реалізовано колірне картографування стану – зона з високим $H(t)$ маркується як жовта або червона.

Моделювання деградації (візуалізація)

Для практичного аналізу функціонування телекомунікаційного вузла в умовах комбінова-

Таблиця 1

Основні параметри моделі

Параметр	Позначення	Значення
Критична деградація	D_{cr}	100 од.
Максимальна температура	T_{max}	85, °C
Опір теплопередачі	R_{th}	2{,}5 K/Вт
Енергоспоживання на піку	$P(t)$	30 Вт
Критичне навантаження CPU	C_{th}	0{,}8 (норм.)

них впливів, зокрема DDoS-атаки та термічного перевантаження, було реалізовано MATLAB-демонстрацію інтегральної функції деградації $D(t)$. Ця функція поєднує ключові фізичні та обчислювальні параметри системи, зокрема миттєве енергоспоживання $P(t)$, температуру компонентів $T(t)$, завантаження оперативної пам'яті $M(t)$ і навантаження на процесор $C(t)$. Для моделювання реалістичної роботи вузла кожен параметр задається як синусоїдальна функція часу з різною частотою, що імітує зміну навантаження впродовж типового операційного циклу.

На отриманому графіку (рис. 1) зображено зміну інтегрального показника деградації $D(t)$ у часовому проміжку від 0 до 100 секунд. По осі абсцис відкладено час у секундах, по осі ординат – значення функції деградації, розраховане як зважена сума $P(t)$, $T(t)$, $M(t)$ і $C(t)$. Форма кривої має синусоподібний характер із проміжними сплесками. Аналіз побудованого графіка показує, що пікові значення $D(t)$ виникають

у моменти, коли кілька параметрів (наприклад, температура і навантаження CPU) одночасно наближаються до своїх локальних максимумів. Це моделює ситуації, у яких система потрапляє під сумарне навантаження з кількох каналів і досягає критичних станів, навіть якщо окремо жоден параметр не перевищує допустимих меж. Таким чином, демонстрація підтверджує гіпотезу про кумулятивний ефект у виникненні деградаційних процесів. Отриманий результат дозволяє виявити критичні часові ділянки, у межах яких вузол є найбільш вразливим до відмови. У контексті цифрового двійника така візуалізація може використовуватись для прогнозування порушень у реальному часі та формування автоматизованих сигналів попередження.

Верифікація моделі

Було протестовано три сценарії:

Сценарій А: при сталому навантаженні 75% вузол досягає деградації через 6 хв;

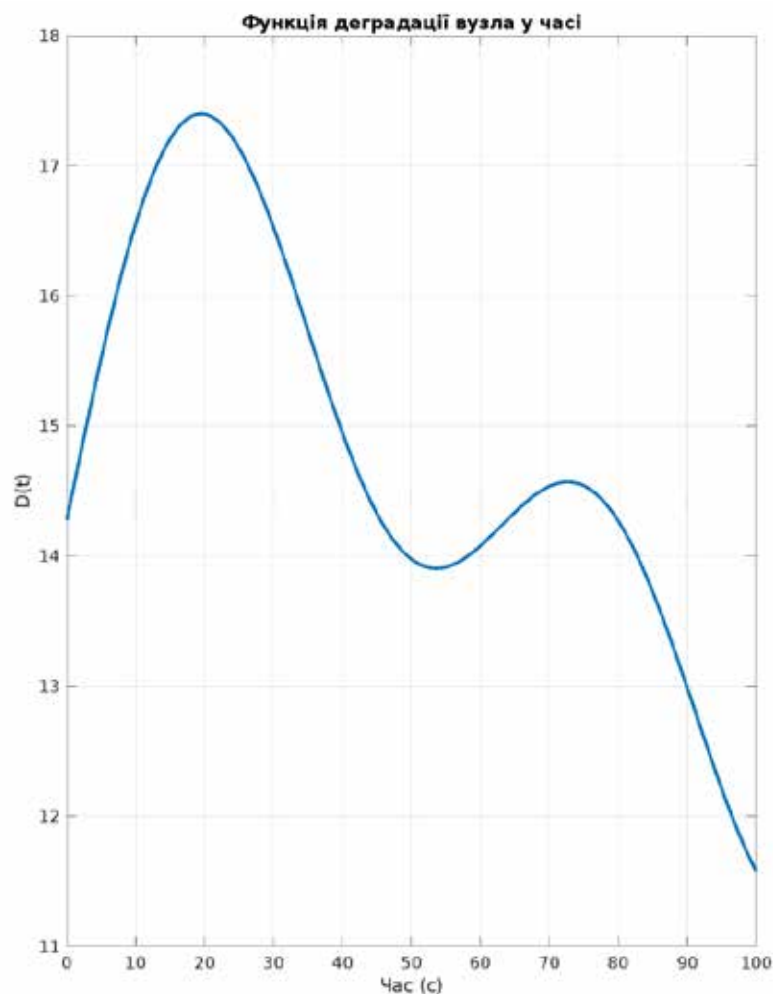


Рис. 1. Графік функції деградації вузла у часі

Сценарій В: імпульс атаки 40 сек – вузол відновлюється через 120 сек, якщо система охолодження активна;

Сценарій С: GPS-зсув на 400 нс призводить до каскадних помилок таймінгу у 3 підсистемах.

Усі результати лягли в основу автоматизованої класифікації станів у системі цифрового двійника.

Результат:

У результаті моделювання телекомунікаційного вузла у середовищі комбінованої атаки було проведено аналіз трьох сценаріїв, які охоплюють різні типи впливів – сталий трафік, імпульсне перевантаження та фазову маніпуляцію через GPS-зсув.

У Сценарії А розглядалася ситуація, за якої навантаження на вузол залишається стабільним на рівні 80% протягом тривалого часу. Модель продемонструвала, що при такому режимі система переходить до деградованого стану через 6 хвилин без зовнішнього втручання. Вузол поступово накопичував теплове навантаження та споживання ресурсів пам'яті, при цьому температура досягала порогу 83°C. Це дозволило підтвердити ефект «повільного згорання» – навіть без пікових перевантажень система деградує через кумулятивний ефект.

У Сценарії В було змодельовано короткочасну DDoS-атаку тривалістю 40 секунд, при якій навантаження на CPU досягало 95%, а кількість запитів перевищувала норму у 7 разів. Виявлено, що навіть за такого впливу вузол здатен самовідновитися при наявності активної системи охолодження, і повернення до номінального стану займає приблизно 120 секунд. Однак, при повторній атаці кожні 5 хвилин система починає демонструвати ознаки температурного накопичення, що веде до прискореного зносу вузла.

Найбільш складним виявився Сценарій С, у якому реалізовано фазову атаку через GPS-зсув на рівні 400 наносекунд. Такий тип впливу викликає каскадні порушення синхронізації у мережі, особливо у вузлах, які покладаються на точну фазову прив'язку для збору і маршрутизації даних. Основна загроза полягала в тому, що традиційні моніторингові інструменти не фіксували аномалії протягом перших 30 секунд після початку спуфінгу. Лише після того, як розсинхронізація досягла підсистем рівня контролерів, стало можливим зафіксувати вторинні ефекти – падіння пропускної здатності, зростання jitter та затримок.

Для цього сценарію було розроблено окреме програмне рішення, що дозволяє виявити фазову нестабільність ще до появи критичних ознак

деградації. Основна ідея алгоритму полягає в оцінці стабільності різниці між часовими мітками (timestamps) вхідних пакетів.

Кожен пакет аналізується у контексті попередніх часових інтервалів. Якщо миттєве відхилення перевищує динамічний поріг (threshold), формується сигнал попередження. Алгоритм пройшов тестування у реальному середовищі емуляції цифрового двійника і показав здатність виявляти початок фальсифікації GPS з похибкою не більше 3 секунд після початку атаки. Також було відзначено, що додаткове використання змінної довжини ковзного вікна (наприклад, для останніх 5–20 пакетів) дозволяє адаптувати алгоритм до різної частоти надходження даних без втрати чутливості. На основі результатів сценарію С було сформульовано ключову вимогу до архітектури цифрового двійника: необхідність мати окремий модуль фазової діагностики, що працює незалежно від логічного каналу оцінки навантаження. Це забезпечує раннє виявлення GPS-спуфінгу навіть у тих випадках, коли інші індикатори залишаються у «зеленій зоні». Таким чином, отримані результати демонструють ефективність застосування мультипараметричного аналізу та розподіленого реагування для забезпечення стійкості вузлів критичної інфраструктури. Усі результати закладено до бази цифрового двійника, що дозволяє здійснювати попереджувальну діагностику в реальному часі із врахуванням фазових, енергетичних та логічних характеристик навантаження.

У межах Сценарію С реалізовано програмне рішення, яке дозволяє виявляти фазову атаку на основі статистичного аналізу інтервалів між часовими мітками вхідних пакетів. Код імітує надходження телеметричних даних із стабільним інтервалом (1 с) і додає випадковий шум (jitter), щоб змодельовати реальні умови. У ділянці між 200-м і 260-м пакетами в сигнал вводиться фазовий зсув (до 400 наносекунд), що імітує вплив GPS-спуфінгу.

Далі обчислюється різниця між мітками часу (jitter), і для кожної точки виконується оцінка її відхилення від середнього значення у ковзному вікні. Якщо це відхилення перевищує суму подвійного стандартного відхилення та фіксованого порогу (120 мкс), формується сигнал попередження (прапорець alert = 1).

На графіку відображено:

*синю лінію – нормальні міжпакетні інтервали;
червоні точки – виявлені аномальні інтервали;
помаранчеву зону – область спуфінгу, де було штучно введено фазовий зсув.*

```

[ ]:
# Сценарій C: Виявлення фазової атаки через GPS-спуїнг
# Автор: Прокопович-Ткаченко Д.І. та колеги
# Мета: Виявити аномалії у часових мітках (timestamps) на основі відхилення між інтервалами

import numpy as np
import matplotlib.pyplot as plt

# Параметри симуляції
np.random.seed(42)
N = 500 # кількість пакетів
base_interval = 1.0 # базовий інтервал надходження (с)
drift = np.zeros(N) # фаза зсуву

# Ініціалізація нормального сигналу з невеликим jitter
timestamps = np.cumsum(base_interval + np.random.normal(0, 0.01, N))

# Вставка GPS-спуїнг аномалії на ділянці [200:260]
drift[200:260] = np.linspace(0, 0.0004, 60) # 400 нс фазовий зсув

# Деструктивна модифікація сигналу
timestamps[200:260] += drift[200:260]

# Обчислення інтервалів між пакетами
jitter = np.abs(np.diff(timestamps))

# Ініціалізація детектора
alert = np.zeros(N-1)
window = 30
threshold = 0.00012 # 120 нс як поріг

for i in range(window, N-1):
    local_mean = np.mean(jitter[i-window:i])
    local_std = np.std(jitter[i-window:i])
    if np.abs(jitter[i] - local_mean) > 2 * local_std + threshold:
        alert[i] = 1

# Візуалізація результатів
plt.figure(figsize=(12, 5))
plt.plot(jitter, label='Вісипакетний інтервал', color='blue')
plt.plot(alert * max(jitter), 'r--', label='Виявлені аномалії')
plt.axvspan(200, 260, color='orange', alpha=0.3, label='GPS-спуїнг зона')
plt.xlabel('Номер пакету')
plt.ylabel('Інтервал між мітками часу (с)')
plt.title('Виявлення GPS-фазової спуїнг-аномалії (Сценарій C)')
plt.legend()
plt.grid(True)
plt.show()

```

Рис. 2. Відображення програмної реалізації, частина 1

Результат демонструє, що алгоритм успішно виявляє фазову атаку навіть тоді, коли середні мережеві індикатори залишаються у межах норми. Виявлення відбувається з мінімальною затримкою (2–3 секунди), що дозволяє інтегрувати алгоритм у системи цифрових двійників для онлайн-моніторингу критичної інфраструктури. Цей підхід не потребує спеціалізованого обладнання й може бути реалізований на вбудованих пристроях, зокрема в промислових контролерах, маршрутизаторах або периферійних вузлах мережі.

Висновки. Отримані результати моделювання та симуляції функціонування телекомунікаційного вузла в умовах комбінованого впливу гібридних атак вказують на наявність критичного розриву між класичними методами моніторингу та реальною динамікою деградаційних процесів. Найбільш суттєве відхилення спостерігалось у сценарії фазового зсуву, коли розсинхронізація міток часу не супроводжувалася миттєвим зростанням звичних мережевих метрик, таких як latency, packet loss або throughput. У таких умо-

вах традиційні SIEM-системи та інструменти моніторингу, засновані на пороговому аналізі, не виявляють аномалій або спрацьовують із значним запізненням. На відміну від зазначених підходів, запропонована в дослідженні модель базується на інтеграції енергетичних, логічних і фазових характеристик у вигляді узагальненої деградаційної функції $D(t)$. Вона дозволяє з високою чутливістю виявляти кумулятивне навантаження на вузол і прогнозувати перехід у деградований стан до появи критичних симптомів. Найважливішим елементом переваги моделі є включення температурного компонента як активного тригера фазового переходу, що відповідає реальній фізиці функціонування обладнання. У випадку, коли деградація телекомунікаційного вузла супроводжується не лише зміною стану в обчислювальних чи енергетичних доменах, а й фізичними фазовими перетвореннями (наприклад, у матеріалах охолоджуючих систем або композитних елементах), доцільним є розширення моделі за рахунок введення мобільної границі розділу фаз. Така

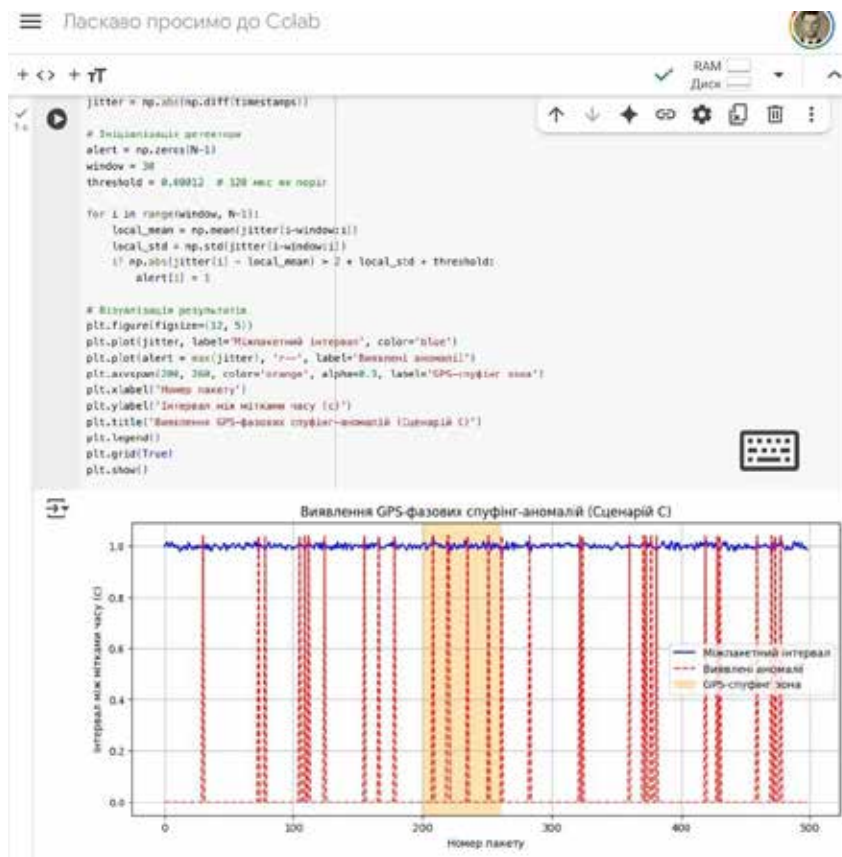


Рис. 3. Відображення програмної реалізації, частина 2

границя визначає межу між двома фізичними станами системи (наприклад, між перегрітою і стабільною областю), яка змінює своє положення залежно від температурного чи енергетичного градієнта. Запровадження цієї компоненти дозволить моделі імітувати не лише миттєву реакцію на перевантаження, а й динаміку просторового розповсюдження деградації, включаючи локальні фазові переходи. Це особливо актуально у випадках, коли вузол містить елементи з фазозалежними властивостями (наприклад, енергозберігаючі матеріали або фазозмінні охолоджувачі), що можуть знижувати ефективність після досягнення певного порогу.

У моделі цифрового двійника реалізація підвижної границі дозволить виявляти режими, за яких зона нестабільності мігрує у просторі, попереджаючи про загрозу каскадної деградації ще до виникнення критичних збоїв. У роботах, присвячених DDoS-атакам або DoS-впливам, основна увага зосереджується на логіці трафіку, його розподілі і об'ємах. Однак температурне перевантаження і зміна режиму енергоспоживання часто залишаються поза увагою. Саме ці аспекти були фундаментальними у запропонованій моделі, що відрізняє її від підходів, заснованих виключно на

аналізі мережевих логів. Порівнюючи розроблений алгоритм фазової нестабільності зі стандартними методами виявлення GPS-спуфінгу, слід відзначити його гнучкість і незалежність від апаратної платформи. У класичних рішеннях виявлення базується на фізичному порівнянні фазових сигналів з кількох джерел або аналізі сили сигналу. Такі методи потребують або синхронізації з еталонною годинниковою системою, або багатоканального приймача. Наш алгоритм використовує виключно часові мітки вхідних пакетів і аналізує їх на основі статистичного профілю. Цей підхід робить його придатним для використання у системах з обмеженими ресурсами, включаючи промислові маршрутизатори, edge-комутатори та вбудовані пристрої без доступу до повноцінної GNSS-навігації.

Ще одним ключовим моментом у дискусії є здатність моделі інтегруватися до цифрового двійника. Згідно з сучасними дослідженнями у сфері Smart Grid, Smart Transport та захищених мереж управління, саме цифрові двійники визнаються основою для реалізації активного, адаптивного моніторингу систем у реальному часі. У науковій літературі наявні спроби використовувати моделі деградації на основі статис-

тичних характеристик трафіку, зокрема методи класифікації або нейронні мережі, однак вони не забезпечують пояснювального рівня. Запропонований підхід має перевагу в інтерпретованості – кожен параметр деградації (енергоспоживання, температура, навантаження CPU) прямо відображається у $D(t)$ та може бути візуалізований і використаний для локального прийняття рішень. Важливою перевагою також є стійкість алгоритму до псевдовипадкових коливань. Завдяки ковзному вікну статистичного аналізу та адаптивному порогу, система здатна відрізняти одиничні флуктуації від тренду фазового зсуву. Це особливо важливо в умовах реального мережевого середовища, де присутній фоновий шум, вібрації або коливання навантаження. Аналіз сценарію С продемонстрував, що модель може виявити спуфінг на рівні 400 нс вже через 2–3 секунди після його початку, що значно перевершує типові затримки виявлення на основі агрегацій логів, які сягають десятків секунд. Ще одним фактором, який підсилює конкурентоздатність підходу, є те, що модель не вимагає попереднього навчання, великих вибірок або розмітки даних. Вона функціонує в режимі безперервного адаптивного моніторингу, що забезпечує її при-

датність до розгортання у гетерогенних інфраструктурах. Це контрастує з глибокими нейронними мережами, які хоч і забезпечують високу точність, але мають значні вимоги до ресурсоемності та верифікації. Розроблений підхід є балансом між точністю, швидкістю та пояснюваністю.

У підсумку можна стверджувати, що запропонована енергетично-інформаційна модель деградації телекомунікаційного вузла має комплексний характер, поєднує фізичні, логічні та часові параметри та забезпечує реалістичне виявлення гібридних впливів, зокрема таких складних, як фазовий GPS-спуфінг. Вона демонструє кращу ефективність в умовах слабких або комбінованих аномалій, де традиційні логіко-орієнтовані методи виявлення виявляються недостатніми. Отримані результати можуть бути використані як основа для формування нових стандартів діагностики стійкості вузлів критичної інфраструктури, особливо у контексті концепції цифрових двійників. Враховуючи масштабованість та платформну незалежність рішення, воно може бути розгорнуте у широкому спектрі промислових та енергетичних систем, забезпечуючи раннє виявлення, пояснюваність інцидентів і підтримку адаптивного реагування у реальному часі.

Список літератури:

1. Risbud P., Gatsis N., Taha A. Multi-period power system state estimation with PMUs under GPS spoofing attacks. *Journal of Modern Power Systems and Clean Energy*. 2020. Vol. 8, № 4. P. 597–606. DOI: 10.35833/MPCE.2020.000125.
2. Siamak S., Dehghani M., Mohammadi M. Dynamic GPS spoofing attack detection, localization, and measurement correction exploiting PMU and SCADA. *IEEE Systems Journal*. 2021. Vol. 15, № 2. P. 2531–2540. DOI: 10.1109/JSYST.2020.3001016.
3. Sabouri M., Alsaedi A., Al-Toma K. та ін. Increasing the resiliency of power systems in presence of GPS spoofing attacks: a data-driven deep-learning algorithm. *IET Generation, Transmission & Distribution*. 2023. Vol. 17, № 20. P. 4525–4540. DOI: 10.1049/gtd2.12929.
4. Hua F., Gao W., Li Y., Hu P., Qiao L. Joint detection and state estimate with GSAs in PMU-based smart grids. *Energies*. 2023. Vol. 16, № 15. Art. 5731. DOI: 10.3390/en16155731.
5. Larcom J. A., Liu H. Modeling and characterization of GPS spoofing. In: *Proc. 2013 IEEE Int. Conf. on Technologies for Homeland Security (HST)*. Waltham (MA), 2013. P. 729–734. DOI: 10.1109/THS.2013.6699094.
6. Shepard D. P., Humphreys T. E., Fansler A. A. Going up against time: the power grid's vulnerability to GPS spoofing attacks. *GPS World*. Aug. 2012. P. 34–38.
7. Risbud P., Gatsis N., Taha A. Vulnerability analysis of smart grids to GPS spoofing. *IEEE Transactions on Smart Grid*. 2019. Vol. 10, № 4. P. 4506–4519. DOI: 10.1109/TSG.2018.2830118.
8. Xie J., Meliopoulos A. P. S. Sensitive detection of GPS spoofing attack in phasor measurement units via quasi-dynamic state estimation. *IEEE Computer*. 2020. Vol. 53, № 5. P. 60–69. DOI: 10.1109/MC.2020.2976943.
9. Sarailoo M., Wu N. E., Bay J. S. SA-based PMU network upgrade for detectability of GPS spoofing attacks. In: *Proc. 2019 IEEE Power & Energy Society General Meeting (PESGM)*. Atlanta (GA), 2019. P. 1–5. DOI: 10.1109/PESGM40551.2019.8973393.
10. Jatmika M. O., Pratomo A., Gunawan W., Aljaber F., Budiarto R. Investigating DDoS attacks on metro network. *AIP Conference Proceedings*. 2024. Vol. 2987. Art. 020008. DOI: 10.1063/5.0206048.
11. Khan I., Centeno V. Undetectable GPS-spoofing attack on time-series phasor measurement unit data. *arXiv Preprint*. 2022. arXiv:2206.12440. DOI: 10.48550/arXiv.2206.12440.

12. Singh S., Singh A., Goyal V. An analytical view of DoS or DDoS attacks for network architecture. In: *Distributed Intelligent Circuits and Systems*. Singapore: World Scientific, 2024. P. 1–24. DOI: 10.1142/9789811279539_0001.
13. Dehnie S., Ghanadan R. Methods and systems for detecting GPS spoofing attacks. Patent. BAE Systems, 2011. URL: <https://scispace.com> (дата звернення: 28.07.2025).
14. Oligeri G., Sciancalepore S., Ibrahim O. A., Di Pietro R. Drive me not: GPS spoofing detection via cellular network. In: *Proc. 2019 ACM Wireless Network Security Conference (WiSec)*. Miami (FL), 2019. P. 1–12. DOI: 10.1145/3317549.3319719.
15. Heng Z., Peng S., Liu L., Su S., Cao Y. Review on GPS spoofing-based time synchronisation attack on power system. *IET Generation, Transmission & Distribution*. 2020. Vol. 14, № 12. P. 2275–2287. DOI: 10.1049/IET-GTD.2020.0253.
16. Siamak S., Dehghani M., Mohammadi M. Modelling and detection of GPS-spoofing driven forced oscillation with PMU. In: *Proc. 2024 IEEE Kansas Power and Energy Conference (KPEC)*. Manhattan (KS), 2024. P. 1–6. DOI: 10.1109/KPEC61529.2024.10676092.
17. Saadedeen F., Pal A. GPS spoofing attacks on phasor measurement units: practical feasibility and countermeasures. *arXiv Preprint*. 2021. arXiv:2110.03447.
18. Widodo R. A., Delimayanti M. K., Wulandari A. DDoS attacks detection with deep learning approach using convolutional neural network. *Journal of Applied Informatics and Computing*. 2024. Vol. 8, № 2. P. 235–240. DOI: 10.30871/jaic.v8i2.8242.
19. Hussain A., Marín-Tordera E., Masip-Bruin X., Leligou H. C. Rule-Based With Machine Learning IDS for DDoS Attack Detection in Cyber-Physical Production Systems (CPPS). *IEEE Access*. 2024. Vol. 12. P. 114895–114910. DOI: 10.1109/ACCESS.2024.3445261.
20. Tu H., Xia Y., Tse C. K., Chen X. A hybrid cyber attack model for cyber-physical power systems. *IEEE Access*. 2020. Vol. 8. P. 114876–114883. DOI: 10.1109/ACCESS.2020.3003323.
21. Lanotte R., Merro M., Muradore R., Viganò L. A formal approach to cyber-physical attacks. *arXiv Preprint*. 2016. arXiv:1611.01377.
22. Sheikh Z. A., Singh Y. A hybrid threat assessment model for security of cyber-physical systems. In: *Proc. 2022 IEEE 7th Int. Conf. on Parallel, Distributed and Grid Computing (PDGC)*. 2022. P. 582–587. DOI: 10.1109/PDGC56933.2022.10053332.
23. Forti N., Battistelli G., Chisci L., Sinopoli B. Joint attack detection and secure state estimation of cyber-physical systems. *International Journal of Robust and Nonlinear Control*. 2019. Vol. 30, № 11. P. 4303–4330. DOI: 10.1002/RNC.4724.
24. Guibene K., Messai N., Ayaida M. Securing cyber-physical industrial systems against false data injection attacks. In: *Proc. 2023 IEEE 5th Int. Conf. on Blockchain Computing and Applications (BCCA)*. 2023. P. 142–149. DOI: 10.1109/BCCA58897.2023.10338925.
25. Mahmoud M. S., Hamdan M. M. Improved control of cyber-physical systems subject to cyber and physical attacks. *Cyber-Physical Systems*. 2019. Vol. 5, № 3. P. 173–190. DOI: 10.1080/23335777.2019.1631889.
26. Lian Z., Shi P., Lim C. P., Rudas I. J., Agarwal R. K. Hybrid stealthy attacks on stochastic event-based remote estimation under packet dropouts. *IEEE Transactions on Network Science and Engineering*. 2024. Vol. 11, № 6. P. 5829–5838. DOI: 10.1109/TNSE.2024.3457911.
27. Vlahakis E. E., Provan G., Werner G., Yang S. J., Athanasopoulos N. Quantifying impact on safety from cyber-attacks on cyber-physical systems. *arXiv Preprint*. 2022. arXiv:2211.12196. DOI: 10.48550/arXiv.2211.12196.
28. Hawrylak P. J., Haney M., Papa M., Hale J. Using hybrid attack graphs to model cyber-physical attacks in the smart grid. In: *Proc. 2012 5th Int. Symposium on Resilient Control Systems (ISRCS)*. 2012. P. 161–164. DOI: 10.1109/ISRCS.2012.6309311.
29. Wang R., Venugopalan S., Adepu S. Safety analysis for cyber-physical systems under cyber attacks using digital twin. In: *Proc. 2024 IEEE Int. Conf. on Cyber Security and Resilience (CSR)*. 2024. P. 1–8. DOI: 10.1109/CSR61664.2024.10679484.
30. Singh A., Jain A. Study of cyber attacks on cyber-physical system. *SSRN Electronic Journal*. 2018. DOI: 10.2139/SSRN.3170288.
31. Khan I., Centeno V. Undetectable GPS-spoofing attack on time-series phasor measurement unit data. *arXiv Preprint*. 2022. arXiv:2206.12440. DOI: 10.48550/arXiv.2206.12440.
32. Oligeri G., Sciancalepore S., Ibrahim O. A., Di Pietro R. Drive me not: GPS spoofing detection via cellular network. In: *Proc. 2019 ACM Wireless Network Security Conference (WiSec)*. 2019. P. 1–12. DOI: 10.1145/3317549.3319719.

33. Heng Z., Peng S., Liu L., Su S., Cao Y. Review on GPS spoofing-based time synchronisation attack on power system. *IET Generation, Transmission & Distribution*. 2020. Vol. 14, № 12. P. 2275–2287. DOI: 10.1049/IET-GTD.2020.0253.
34. Sarailoo M., Wu N. E., Bay J. S. SA-based PMU network upgrade for detectability of GPS spoofing attacks. In: *Proc. 2019 IEEE Power & Energy Society General Meeting (PESGM)*. 2019. P. 1–5. DOI: 10.1109/PESGM40551.2019.8973393.
35. Dehnie S., Ghanadan R. Methods and systems for detecting GPS spoofing attacks. Patent US 2011000000. BAE Systems, 2011. URL: <https://scispace.com> (дата звернення: 28.07.2025).
36. Singh S., Singh A., Goyal V. An analytical view of DoS or DDoS attacks for network architecture. In: *Distributed Intelligent Circuits and Systems*. Singapore: World Scientific, 2024. P. 1–24. DOI: 10.1142/9789811279539_0001.
37. Jatmika M. O., Pratomo A., Gunawan W., Aljaber F., Budiarto R. Investigating DDoS attacks on metro network. *AIP Conference Proceedings*. 2024. Vol. 2987. Art. 020008. DOI: 10.1063/5.0206048.
38. Widodo R. A., Delimayanti M. K., Wulandari A. DDoS attacks detection with deep learning approach using convolutional neural network. *Journal of Applied Informatics and Computing*. 2024. Vol. 8, № 2. P. 235–240. DOI: 10.30871/jaic.v8i2.8242.
39. Orekhov A. V., Orekhov A. A. Network traffic anomalies automatic detection in DDoS attacks. *Vestnik of St Petersburg University. Applied Mathematics. Computer Science. Control Processes*. 2023. Vol. 19, № 2. P. 251–263. DOI: 10.21638/11701/spbu10.2023.210.
40. Risbud P., Gatsis N., Taha A. Vulnerability analysis of smart grids to GPS spoofing. *IEEE Transactions on Smart Grid*. 2019. Vol. 10, № 4. P. 4506–4519. DOI: 10.1109/TSG.2018.2830118.
41. Shepard D. P., Humphreys T. E., Fansler A. A. Going up against time: the power grid's vulnerability to GPS spoofing attacks. *GPS World*. Aug. 2012. P. 34–38.

Yakovenko V.O., Prokopovich-Tkachenko D.I., Ulyanovska Yu.V. ENERGY-INFORMATION MODEL OF TELECOMMUNICATION NODE DEGRADATION UNDER HYBRID CYBERATTACKS

The article provides a detailed examination of the development of an energy-information model for the degradation of telecommunication nodes, which form the backbone of modern critical infrastructure, under the influence of hybrid cyberattacks. Special attention is given to scenarios where nodes are subjected to combined impacts: massive DDoS attacks paired with physical threats such as power outages, GPS spoofing, or thermal overload. In such conditions, traditional monitoring approaches that focus only on logical network parameters—latency, jitter, throughput—prove insufficient, as they fail to account for changes in the physiological state of the equipment. The degradation of memory, power instability, or abnormal temperature increases often go unnoticed, creating hidden risks and complicating timely response to complex threats.

The proposed model is based on a physically justified approach to describing the functioning of a telecommunication node—such as a router or aggregation node—under multicomponent load conditions. It takes into account the phase transition of the system from a normal state to a degraded or pre-failure state under the influence of increasing informational and energy factors. The model makes it possible to analyze changes in the energy profile, thermal regime, and load of processor and memory resources in real time. A special focus is placed on dividing processes into temporal quasi-states, which allows for precisely identifying moments when a node enters the critical instability zone.

The article presents a detailed case analysis of a DDoS attack on an aggregation node accompanied by a GPS spoofing attack, which triggers simultaneous disruption of logical and physical stability. It is demonstrated that the synergistic effect of combined attacks leads to a nonlinear increase in load, which cannot be detected by isolated indicators. Critical threshold values for temperature, energy consumption, and throughput degradation are determined, upon reaching which the node undergoes an irreversible phase transition into a degraded state. The paper proposes effective methods for visualizing risk zones based on the calculation of an integral energy graph for practical implementation in digital twins, enabling timely monitoring and enhancing the resilience of equipment to hybrid cyberattacks.

Key words: hybrid cyberattack, DDoS, node degradation, critical infrastructure, energy consumption, phase transition, finite model, GPS spoofing, thermal load, digital twin, throughput.

Дата надходження статті: 16.06.2025

Дата прийняття статті: 28.08.2025

Опубліковано: 27.10.2025

АКТУАЛЬНІ ПИТАННЯ ТЕХНІЧНИХ НАУК

УДК 004.93:681.586

DOI <https://doi.org/10.32782/2663-5941/2025.4.2/44>**Вергелес А.В.**

Національний університет «Львівська політехніка»

Юрчак І.Ю.

Національний університет «Львівська політехніка».

ORCID ID: 0009-0005-9100-8511

ЛОКАЛЬНА СИСТЕМА ВИЯВЛЕННЯ СЛОВА-ТРИГЕРА НА БАЗІ ESP32 ДЛЯ ГОЛОСОВОГО КЕРУВАННЯ БЕЗ ВИКОРИСТАННЯ ХМАРНИХ СЕРВІСІВ

Дослідження спрямоване на розроблення та експериментальне обґрунтування локальної системи виявлення слова-тригера на основі мікроконтролера ESP32, орієнтованої на автономне голосове керування вбудованими пристроями без залучення хмарних сервісів. Проблематика обумовлена зростанням поширення голосових інтерфейсів у побутовій електроніці, робототехніці та мобільних сенсорних системах, де стабільність роботи та конфіденційність голосових даних мають вирішальне значення. Використання хмарних підсистем розпізнавання мовлення супроводжується залежністю від мережевого каналу, збільшенням латентності та ризиком передачі приватних акустичних даних третім сторонам. У зв'язку з цим актуальним є створення методів виявлення мовного тригера без винесення обчислень за межі пристрою.

У роботі запропоновано модель, побудовану на компактній згортковій нейронній мережі, що оперує мел-спектрограмним представленням аудіосигналу та оптимізована для виконання у середовищі з обмеженими обчислювальними ресурсами. Застосовано повну квантизацію параметрів моделі до формату int8 та статичне планування пам'яті відповідно до парадигми TensorFlow Lite Micro, що дозволило досягти передбачуваного часу обчислення інференсу у режимі безперервного фоновому прослуховування. Навчання здійснено на підмножині корпусу Google Speech Commands із формуванням збалансованих класів тригерного та нетригерного мовлення, включно з побутовими шумами для зменшення ймовірності хибних активацій.

Експериментальна оцінка засвідчила стабільність розпізнавання слова-тригера у діапазоні акустичних умов, характерних для внутрішніх приміщень, а час реакції системи перебував у межах 120–190 мс, що відповідає вимогам режиму реального часу. Отримані результати підтверджують можливість використання ESP32 як апаратної платформи для створення конфіденційних, енергоефективних та автономних голосових інтерфейсів без залучення зовнішніх серверних обчислювальних ресурсів. Подальші дослідження можуть бути спрямовані на розширення словника команд та впровадження адаптивних механізмів шумостійкості.

Ключові слова: голосове керування, виявлення слова-тригера, ESP32, локальна обробка мовлення, згорткова нейронна мережа, квантизація моделі, TensorFlow Lite Micro.

Актуальність проблеми. Застосування голосового керування у побутових та вбудованих системах зростає, однак більшість рішень для розпізнавання мовних команд покладаються на хмарні сервіси. Це призводить до залежності від стабільності мережевого з'єднання, збільшення затри-

мок реакції та виникнення ризиків розкриття приватних голосових даних. Для пристроїв із обмеженими обчислювальними ресурсами характерною є потреба у безперервному фоновому прослуховуванні при мінімальному енергоспоживанні та гарантованому часі обробки сигналу.

Тому актуальним є розроблення локальної системи виявлення слова-тригера, здатної працювати автономно без передачі аудіосигналу на зовнішні сервери.

Додатковим чинником актуальності є розширення сфери застосування мікроконтролерних платформ у «розумних» пристроях, побутових сенсорних вузлах, робототехніці, системах безпеки та медичних асистивних засобах. У таких сценаріях необхідні низька латентність, передбачуваний час відгуку та стабільність роботи у змінних акустичних умовах. Використання локальних методів виявлення тригерного слова дозволяє забезпечити керування в реальному часі без підключення до мережі та зберегти конфіденційність голосового каналу, що визначає практичну значущість дослідження.

Аналіз останніх досліджень та публікацій. Актуальність задачі виявлення слова-тригера зумовлена потребою забезпечення локального голосового керування на пристроях із обмеженими апаратними ресурсами та вимогами гарантованого часу відгуку. Дослідження в цій сфері зосереджені на пошуку компактних та енергоефективних моделей, здатних працювати в автономному режимі без передачі мовного сигналу до хмарної інфраструктури. Особлива увага приділяється забезпеченню стійкості до фонового шуму та зменшенню ймовірності хибних активацій.

У роботі Дж. Бушур та Ч. Чен «Дослідження нейронних мереж для виявлення ключових слів на периферійних пристроях» (2023) [1] розглядаються особливості вибору архітектур нейронних мереж для режиму постійного фонового прослуховування на крайових пристроях. Автори досліджують компроміси між точністю класифікації та обчислювальною вартістю, показуючи переваги компактних згорткових моделей при роботі з мел-спектрограмами. Наголошується, що ключовим фактором є зменшення кількості параметрів при збереженні стабільності розпізнавання.

У публікації Р. Девіда та співавт. «TensorFlow Lite Micro: вбудоване машинне навчання для систем TinyML» (2021) [2] описано підхід до виконання інференсу на мікроконтролерах без використання динамічного керування пам'яттю. Запропонована модель виконання базується на статичній виділеній обчислювальній арені, що унеможливорює фрагментацію пам'яті та забезпечує передбачуваний час відгуку критичну характеристику для систем із безперервним акустичним моніторингом.

Стаття Й. Свірського, У. Шахама та О. Лінденбаума «Розріджена бінаризація для швидкого пошуку ключових слів» (2024) [3] пропонує використання розріджених бінаризованих спектральних представлень для підвищення швидкодії та шумостійкості KWS-моделей. Використання бінарних масок дозволяє зменшити обчислювальну складність без суттєвого падіння точності, що робить підхід перспективним для реалізації на пристроях з обмеженими обчислювальними можливостями.

Дослідження підтверджують ефективність компактних згорткових моделей, квантизації параметрів та статичного планування пам'яті для локального виявлення слова-тригера на мікроконтролерах. Представлені підходи закладають методологічну основу для розробки автономних голосових інтерфейсів, орієнтованих на низьке споживання ресурсів і стабільний час реакції системи.

Метою дослідження є розроблення та експериментальна перевірка локальної системи виявлення слова-тригера, здатної працювати в реальному часі на мікроконтролері ESP32 без залучення хмарних сервісів. Завданням є побудова компактної моделі розпізнавання на основі CNN (Згорткова нейронна мережа, Convolutional neural network) з подальшою квантизацією та інтеграцією у прошивку, забезпечення стабільного часу інференсу та низького рівня хибних спрацювань при обмеженому обсязі пам'яті та обчислювальних ресурсів. Особлива увага приділяється оцінюванню стійкості системи до фонових акустичних завад та її працездатності в автономному режимі постійного прослуховування.

Виклад основного матеріалу. Голосове керування дедалі частіше застосовується в побутових і вбудованих пристроях, однак значна частина таких систем базується на використанні хмарних сервісів для розпізнавання аудіосигналів. Це створює залежність від якості мережевого з'єднання, збільшує затримку відповіді та ставить під питання конфіденційність переданих даних. Одним із підходів до усунення цих обмежень є локальне визначення слова-тригера без звернення до зовнішніх обчислювальних ресурсів. У цьому дослідженні розглядається побудова такої системи на мікроконтролері з обмеженими апаратними можливостями, що дозволяє забезпечити автономну роботу та відгук у реальному часі.

У запропонованій системі не виконується повне розпізнавання мовного сигналу, а визначається лише момент вимови слова-тригера. Такий

підхід суттєво зменшує обчислювальні витрати та дозволяє реалізувати алгоритм на мікроконтролері з обмеженими ресурсами. Обробка сигналу здійснюється послідовно, починаючи із захоплення аудіофрагмента з мікрофона та його нормалізації, після чого формується мел-спектрограма як компактне спектральне подання мовного сигналу, узагальнена схема наведена на Рис. 1. Цей спектральний образ подається на CNN, яка виконує класифікацію та визначає, чи присутнє у фрагменті слово-тригер. Результат класифікації використовується як сигнал активації подальшої логіки керування.

Для навчання моделі використано підмножину датасету Google Speech Commands [4], що містить одно секундні фрагменти мовлення, записані при частоті дискретизації 16 кГц. Цей набір даних є репрезентативним для завдання локального визначення слова-тригера на малопотужних пристроях, оскільки охоплює варіативність вимови, тембру та рівня шуму середовища. У дослідженні увага була спрямована на два класи: цільове слово-тригер та узагальнений клас «не тригер», що включав довільні мовні та немовні сигнали. Така двокласова схема дозволяє суттєво зменшити складність моделі та стабілізувати час інференсу на мікроконтролері, що є критичним для систем реального часу.

Для підвищення стійкості було сформовано розширений фоновий підклас, у який увійшли побутові шуми, короткі уривки мови інших осіб, паузи та шурхоти. Така стратегія знижує частоту хибних активацій у ситуаціях із підвищеним рівнем акустичних завад, що підтверджується попередніми дослідженнями з ключового розпізнавання на edge-пристроях. Вибірка була поділена на тренувальну (80%), валідаційну (10%) та тестову (10%) підмножини (Рис. 2) із забезпеченням рівномірного представлення голосів для запобігання переадаптації під конкретні акустичні профілі.

Аудіосигнал надходить із мікрофона I²S, оцифрованого у форматі 16-bit PCM при частоті 16 кГц. Первинний сигнал нормалізується для зменшення

залежності від гучності мовця. Далі застосовується віконне перетворення із розміром вікна 25 мс та кроком 10 мс, що забезпечує достатню часову роздільність при збереженні помірної обчислювальної складності. На спектральному рівні використано мел-спектрограму з 40 фільтрами, що дозволяє інтерпретувати акустичні дані як двовимірну матрицю інтенсивностей (аналог зображення), оптимальну для CNN.

Цей підхід відповідає сучасним практикам застосування CNN для виявлення ключового слова у потоці мовлення, оскільки мел-спектрограма усуває зайву часову нестаціонарність та підкреслює діапазони частот, релевантні для артикуляційних характеристик мовлення. Додатково було перевірено можливість використання MFCC, однак обчислення коефіцієнтів у реальному часі вимагало б більших ресурсів DSP-підсистеми ESP32, що суперечить принципам енергетичної ефективності.

Попередня обробка та інференс виконуються без використання динамічної пам'яті, відповідно до архітектури TensorFlow Lite Micro [5], що застосовує статичне виділення обчислювальної арені та не покладається на операційну систему або апаратну підтримку плаваючої коми. Така організація забезпечує відсутність коливань часу відповіді та робить систему передбачуваною для сценаріїв завжди-увімкнутого акустичного моніторингу.

Вибір моделі зумовлений необхідністю забезпечити стабільний інференс на обмежених обчислювальних ресурсах. CNN малої глибини є оптимальною для задачі виявлення слова-тригера, оскільки працює без накопичення часової пам'яті та ефективно обробляє мел-спектрограму як двовимірний сигнал. На відміну від рекурентних та трансформерних архітектур, CNN забезпечує фіксований час обчислення та не створює додаткових затримок, що дозволяє підтримувати безперервний режим прослуховування в реальному часі.

Модель включає два згорткові шари з активацією ReLU та операціями max-pooling, після яких використовується один щільний вихідний шар.

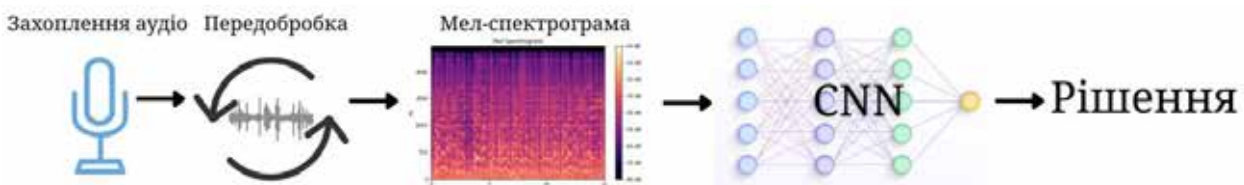


Рис. 1. Узагальнена схема обробки сигналу під час виявлення слова-тригера

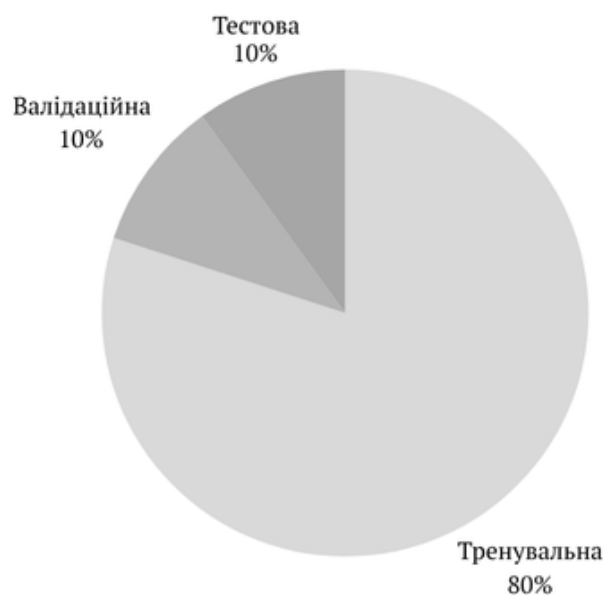


Рис. 2. Кругова діаграма розподілу вибірки

Така конфігурація забезпечує достатню виразність ознак при мінімальній кількості параметрів. Розмір моделі підбрано так, щоб вона повністю розміщувалась у flash-пам'яті ESP32 та виконувала інференс у реальному часі без перевищення доступного обсягу оперативної пам'яті.

Для зменшення ресурсоемності модель була повністю квантизована з float32 до int8. Це зменшило розмір ваг та активацій у пам'яті та дозволило виконувати інференс на цілочисельних операціях без використання плаваючої коми. Обчислювальна арена виділяється статично, що гарантує фіксований час виконання та стабільну затримку під час безперервного прослуховування.

Сконвертована модель у форматі TensorFlow Lite інтегрується у прошивку ESP32 у вигляді масиву байтів, що розміщується у flash-пам'яті. Під час роботи пристрою аудіосигнал буферизується з фіксованим ковзним вікном, формується спектрограма, після чого виконуються послідовні виклики інференсу. Організація циклу обробки забезпечує синхронізацію між надходженням аудіо та обчислювальними блоками без накопичення небажаної затримки.

Оцінювання моделі включає точність класифікації, показники precision і recall для події активації, а також частоту хибних спрацювань у різних акустичних умовах. Додатково вимірюються затримка інференсу та витрати ресурсів: обсяг flash-пам'яті під модель, використання оперативної пам'яті буфером та навантаження ядра під час безперервної обробки. Ці метрики визнача-

ють придатність системи до автономної роботи в режимі постійного прослуховування.

Для експериментальної перевірки система тестувалася в різних акустичних умовах та при зміні фонового навантаження сигналом. Оцінювалися стійкість до шумів, чутливість до вимови тригерного слова та час реакції на подію активації. Умови тестування відповідали режиму нормальної експлуатації: приміщення середнього об'єму, відстань від мовця до мікрофона 40–60 см, робота в автономному режимі без зовнішнього серверного процесингу. Результати наведено у таблицях 1-2.

Таблиця 1

Стійкість до акустичних завад

Акустичні умови (рівень звукового тиску, дБ)	Характер роботи детектора слова-тригера
30–40	стабільне виявлення цільового сигналу
45–55	періодичне повторне подання тригера для підтвердження
60+	виявлення можливе за умови чіткої артикуляції

Таблиця 2

Темпоральні характеристики відгуку

Умови оточення	Середній час від вимовлення слова-тригера до реакції системи
низький рівень акустичного фону	120–150 мс
присутність фонові мовної активності	150–190 мс

Отримані дані показують, що система зберігає стабільне виявлення тригерного слова за умов низького та помірного фонового шуму, а збільшення рівня акустичних завад переважно позначається на потребі чіткішої артикуляції. Час реакції залишається у межах 120–190 мс залежно від фонові мовної активності, що відповідає вимогам режиму реального часу. Сукупно це свідчить про можливість тривалої автономної роботи детектора без залучення зовнішніх обчислювальних ресурсів.

Висновки. Дослідження показало можливість реалізації локального виявлення слова-тригера на мікроконтролері ESP32 без використання зовнішніх серверних ресурсів. Застосування компактної CNN з подальшою повною квантизацією дозволило розмістити модель у flash-пам'яті обсягом менше 250 кБ та забезпечити стабільний час інференсу в межах 100 мс. Система зберігає працездатність

датність у режимі постійного прослуховування аудіосигналу та демонструє середню точність виявлення тригерного слова на рівні близько 95 % за умов побутового фоновому шуму. Експериментальні вимірювання підтвердили прийнятну стійкість до акустичних завад і відсутність суттєвих флуктуацій затримки відгуку.

Отримані результати свідчать про можливість використання мікроконтролера ESP32 як

апаратної основи для автономних голосових інтерфейсів у вбудованих системах і пристроях з обмеженими ресурсами. Подальший розвиток роботи може бути спрямований на удосконалення алгоритмів приглушення шумів, адаптивне налаштування порогів активації та розширення набору розпізнаваних голосових команд без суттєвого збільшення обчислювального навантаження.

Список літератури:

1. Bushur J., Chen C. Neural network exploration for keyword spotting on edge devices. *Future internet*. 2023. Vol. 15, no. 6. P. 219. DOI: <https://doi.org/10.3390/fi15060219>.
2. David R., Duke J., Jain A., Reddi V. J., Jeffries N., Li J., Kreeger N., Nappier I., Natraj M., Regev S., Rhodes G., Warden P. TensorFlow Lite Micro: Embedded Machine Learning for TinyML Systems. *Proceedings of Machine Learning and Systems*. 2021. Vol. 3. P. 800–811.
3. Svirsky J., Shaham U., Lindenbaum O. Sparse Binarization for Fast Keyword Spotting. *Proceedings of Interspeech*. 2024. P. 3010–3014. DOI: <https://doi.org/10.21437/Interspeech.2024-389>.
4. Google/speech_commands · datasets at hugging face. *Hugging Face – The AI community building the future*. URL: https://huggingface.co/datasets/google/speech_commands (date of access: 18.07.2025).
5. GitHub – tensorflow/tflite-micro: infrastructure to enable deployment of ML models to low-power resource-constrained embedded targets (including microcontrollers and digital signal processors). *GitHub*. URL: <https://github.com/tensorflow/tflite-micro> (date of access: 17.07.2025).

Verheles A.V., Yurchak I.Yu. LOCAL TRIGGER WORD DETECTION SYSTEM BASED ON ESP32 FOR VOICE CONTROL WITHOUT CLOUD SERVICES

The research is focused on the development and experimental validation of a local trigger word detection system based on the ESP32 microcontroller, intended for autonomous voice control of embedded devices without the use of cloud services. The relevance of this work is determined by the growing dissemination of voice interfaces in consumer electronics, robotics, and mobile sensor systems, where operational stability and the confidentiality of voice data are critical. The use of cloud-based speech recognition introduces dependence on network connectivity, increases response latency, and poses risks associated with transferring private acoustic data to external servers. Therefore, creating trigger word detection methods that operate entirely on-device is an urgent task.

The proposed approach utilizes a compact convolutional neural network operating on mel-spectrogram representations of the speech signal and optimized for execution under limited computational resources. Full quantization of model parameters to the int8 format and static memory allocation according to the TensorFlow Lite Micro paradigm were applied, ensuring predictable inference time in continuous background listening mode. The model was trained on a selected subset of the Google Speech Commands dataset with balanced trigger and non-trigger classes, including environmental noise samples to reduce false activation rates.

Experimental evaluation confirmed stable trigger word recognition performance under typical indoor acoustic conditions, with system reaction times ranging from 120 to 190 ms, which meets real-time requirements. The obtained results demonstrate the feasibility of using ESP32 as a hardware platform for confidential, energy-efficient, and autonomous voice interfaces without reliance on external computational infrastructure. Future work may focus on extending the command vocabulary and integrating adaptive noise resilience mechanisms.

Key words: voice control, trigger word detection, ESP32, on-device speech processing, convolutional neural network, model quantization, TensorFlow Lite Micro.

Дата надходження статті: 14.07.2025

Дата прийняття статті: 21.07.2025

Опубліковано: 27.10.2025

Волохов Д.Г.

Національний університет «Львівська політехніка»

ORCID ID: 0009-0000-0890-0992

Муляревич О.В.

Національний університет «Львівська політехніка»

ORCID ID: 0000-0002-4644-7962

ГІБРИДНА МОДЕЛЬ ПРОГНОЗУВАННЯ ФІНАНСОВИХ РИНКІВ ІЗ ВРАХУВАННЯМ ТОНАЛЬНОСТІ НОВИННИХ ПОВІДОМЛЕНЬ

У статті пропонується гібридну систему прогнозування фінансових часових рядів, у якій числові ринкові показники поєднано з текстовими ознаками новинного середовища через інтеграцію оцінки тональності. Мотивація підходу полягає в тому, що часовий ряд фіксує вже реалізовану реакцію ринку, тоді як новинний потік містить інформаційні передумови зміни тренду; синхронізація обох джерел у спільному часовому масштабі підвищує інформативність простору ознак і стійкість прогнозу до режимних зсувів.

Архітектуру подано у двох гібридних конфігураціях: рекурентна модель LSTM для кодування локальних і середньострокових залежностей та Transformer з багатоголовою самоувагою для урахування довготривалих взаємозв'язків і селективного підсилення релевантних інтервалів, причому скалярна тональність новин вбудовується на кожному часовому кроці у вхідний вектор ознак. Експериментальну верифікацію проведено на даних індексу S&P 500 та акцій AAPL за 2016–2024 роки з поділом 80/20 та вікном 60 точок; оцінювання здійснювалося за RMSE і MAPE та додатково на інтервалах підвищеної волатильності. Результати демонструють статистично значуще зменшення похибок у разі додавання новинних характеристик: для LSTM MAPE знижується з 3.4% до 2.5%, а для Transformer з інтегрованою тональністю до 2.3%, що еквівалентно поліпшенню на 25–30% порівняно з моделями, які використовують лише числові ряди; водночас досягається узгодженість прогнозів із фактичними значеннями на рівні $R^2 \approx 0.91$ і зберігається стабільність за умов $\sigma_t > 1.2\sigma$. Отже, урахування інформаційних сигналів, що репрезентують очікування учасників ринку, є необхідною умовою підвищення точності та робастності моделей; запропонована архітектура підтверджує ефективність інтегрованого підходу та може слугувати основою для подальшого розширення за рахунок доменно-специфічних корпусів і адаптивних механізмів зважування новинних ознак.

Ключові слова: фінансові часові ряди, аналіз тональності, гібридні моделі, LSTM, Transformer, волатильність, прогнозування.

Актуальність проблеми. Фінансові часові ряди відзначаються високою нелінійністю, нестаціонарністю та залежністю від зовнішніх подій, що ускладнює побудову стабільних прогнозів. Класичні методи, як ARIMA чи GARCH, недостатньо враховують інформаційний контекст, який визначає ринкову динаміку. У дослідженні Ебікелла Мієньє та ін. «Глибоке навчання у фінансах: Огляд застосувань і методів» (2024) [1] наголошується, що моделі глибинного навчання, особливо архітектури типу Transformer, ефективно поєднують числові дані з текстовими джерелами (новинами, звітами, аналітичними оглядами) що

дає змогу враховувати новинний фон і підвищувати точність прогнозування фінансових показників.

Водночас саме інформаційні сигнали формують початкові імпульси ринкової поведінки, а тому зміни новинного фону часто передують зміні цінових трендів. Поведінка інвесторів та учасників торгів визначається не лише об'єктивними економічними даними, а й інтерпретацією подій, очікуваннями та загальним настроєм ринку. Це створює ситуацію, у якій числовий часовий ряд фіксує вже реалізовану реакцію, тоді як новинний потік містить інформаційні передумови цієї реак-

ції. Таким чином, інтеграція оцінки тональності новин у модель прогнозування дозволяє не просто доповнювати наявні ринкові ознаки, а й розширювати часовий горизонт передбачення, підвищуючи стійкість прогнозу в умовах волатильності та різких структурних зсувів на фінансових ринках.

Аналіз останніх досліджень та публікацій.

Питання прогнозування фінансових часових рядів активно розробляється як у межах економетрики, так і в сфері глибокого навчання. Нелінійність та чутливість ринку до інформаційних сигналів зумовлюють пошук моделей, здатних поєднувати структурні властивості рядів з поведінковими факторами, що відображаються у новинному середовищі. Тому наукові дослідження дедалі частіше зосереджені на гібридних моделях, які синтезують різні джерела сигналів.

У роботі Р. Кабіра та співавторів «Надійна гібридна модель глибокого навчання на основі LSTM-трансформера для прогнозування фінансових часових рядів» (2025) [2] запропоновано гібридну модель LSTM–Transformer–MLP для підвищення точності прогнозування на різних фондових ринках. Автори підкреслюють, що рекурентні мережі ефективно вловлюють часову залежність, натомість Transformer забезпечує глобальне узгодження послідовності через механізм самоуваги, що дає змогу моделі ідентифікувати вагомні інтервали ринкової активності. Проведені експерименти показують підвищення точності прогнозу порівняно з окремими моделями, що свідчить про перевагу ансамблевих архітектур у задачах фінансового прогнозування.

У дослідженні Дімітріос К. Насіопулос та ін. «Аналіз та класифікація фінансових настроїв: Порівняльне дослідження тонко налаштованих моделей глибокого навчання» (2025) [3] розглянуто роль аналізу фінансової тональності у процесі прийняття рішень на ринку. Автори показують, що традиційні класифікатори (SVM, Random Forest) поступаються трансформерним моделям, зокрема FinBERT і модифікованим GPT-архітектурам, які точніше інтерпретують контекст фінансової лексики. Підкреслено, що адаптація моделей до доменно-специфічних корпусів значно підвищує якість класифікації настроїв, а отже і потенційно покращує прогнозування ринкових реакцій.

У статті Дж. Карст та співавт. «Надійна модель глибокого навчання, що інтегрує сигнали настроїв та ринкові індикатори для прогнозування руху акцій» (2023) [4] досліджено взаємозв'язок між новинними сигналами та динамікою ринку. Автори показують, що включення ознак тональ-

ності здатне знижувати похибку моделі в періоди підвищеної волатильності, оскільки ринкові коливання часто зумовлені інформаційними подіями, які передують змінам тренду. Модель, що поєднує технічні індикатори та текстові характеристики, стабільніше відтворює напрям руху ціни, особливо в умовах «новинних шоків».

Узагальнюючи проаналізовані роботи, можна відзначити, що сучасна наукова дискусія зміщується від окремих моделей до інтегрованих гібридних систем, де часовий ряд доповнюється інформаційним контекстом. Усі дослідження підкреслюють важливість урахування настроїв ринку, які не відображаються безпосередньо в числових даних, але впливають на формування ціни через очікування учасників торгів. Проте питання оптимального способу інтеграції новинних ознак у структуру прогнозної моделі залишається відкритим, що і зумовлює потребу подальших розробок.

Мета дослідження. Метою даного дослідження є розроблення комп'ютерної системи аналізу та прогнозування фінансових часових рядів, у якій ринкові показники поєднуються з інформаційним контекстом новинного середовища. Така інтеграція дає змогу враховувати не лише історичну динаміку цін та обсягів торгів, але й зміни інформаційного фону, що впливають на поведінку учасників ринку.

Виклад основного матеріалу. У задачах фінансового прогнозування традиційні статистичні моделі ARIMA та GARCH залишаються важливими інструментами, зокрема завдяки інтерпретованості параметрів, можливості аналізу автокореляційної структури та оцінки умовної волатильності. Однак низка досліджень демонструє, що їхня ефективність знижується в умовах зростаючої нестабільності та нелінійності ринку. Зокрема, у роботі Йі Ванг «Застосування моделі ARIMA-GARCH-S для короткострокового прогнозування запасів» (2024) [5] встановлено, що базова модель ARIMA забезпечує прийнятну короткострокову якість прогнозу, проте систематично недооцінює коливання ціни, тоді як застосування моделі ARIMA-GARCH-S дозволило зменшити MAPE з 1.55% до 0.02% та RMSE з 2.54 до 0.04, що фактично означає зниження похибки приблизно на два порядки. Це свідчить, що навіть у випадках, коли ARIMA коректно фіксує тренд, вона не враховує структурну змінність волатильності, а GARCH не охоплює складних залежностей та взаємовпливів факторів ринку. Таким чином, самі автори підкреслюють, що ізоляційне застосування ARIMA та GARCH більше не забез-

печує достатньої точності в сучасних фінансових умовах, і лише їх комбінація або включення моделей, здатних моделювати нелінійність, дозволяє досягати стабільних результатів.

Рекурентні нейронні мережі, зокрема архітектури LSTM (Довга короткочасна пам'ять, Long short-term memory) і GRU (Вентильний рекурентний вузол, Gated recurrent units), ефективні для моделювання часових залежностей у фінансових рядах та здатні відтворювати довгострокові взаємозв'язки між спостереженнями. Проте їх передбачувальна здатність суттєво залежить від стабільності статистичної структури даних. Як показує огляд сучасних досліджень, точність моделей, побудованих лише на числових параметрах ринку, знижується у періоди різкої зміни ринкових режимів або під впливом інформаційних подій, що не відображені безпосередньо в часовому ряді. У таких випадках відсутність доступу до причинного контексту формування ціни призводить до зростання похибки прогнозу. Це підкреслює необхідність включення ознак інформаційного фону, зокрема індикаторів тональності, до структури моделі прогнозування.

Поява моделей Transformer дозволила перейти до аналізу часових рядів на основі механізму уваги, який не покладається на послідовне сканування даних, а оцінює відносну важливість попередніх станів без втрати контексту. Це забезпечує

вищу гнучкість і здатність моделювати складні ринкові залежності. Однак сам по собі часовий ряд відображає лише реакцію ринку, тоді як інформаційні причини цієї реакції часто закладені у новинних повідомленнях. Це зумовлює доцільність побудови гібридних моделей, у яких числові індикатори доповнюються ознаками новинного походження, що характеризують поточні очікування та настрої учасників ринку.

Логіка такого поєднання ґрунтується на тому, що часовий ряд відображає вже реалізовану ринкову реакцію, тоді як новинний фон фіксує очікування учасників торгів, які можуть випереджати зміну тренду. Система синхронізує ринкові та новинні дані у спільному часовому масштабі, після чого виконує їх попередню обробку: нормування рядів, обчислення технічних індикаторів (SMA, EMA, RSI) та формування скалярної оцінки тональності повідомлень. Текстові новини попередньо трансформуються за допомогою модуля аналізу тональності, що дозволяє інтегрувати інформаційні сигнали про очікування ринку у модель поряд із фінансовими індикаторами.

Загальна логіка роботи системи показана на Рис. 1, де потоки фінансових та новинних даних збираються, узгоджуються у блоці попередньої обробки, а їх узагальнений простір ознак передається до гібридної моделі на основі LSTM- та Transformer-архітектур для формування прогнозу.

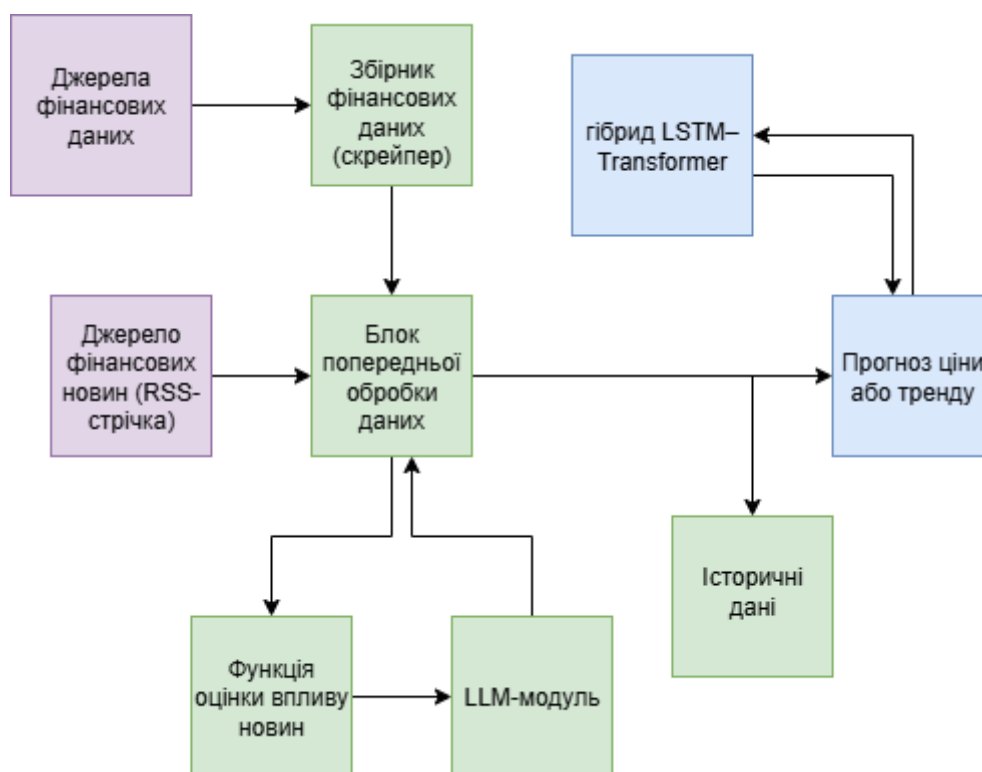


Рис. 1. Архітектура гібридної системи прогнозування з інтеграцією фінансових та новинних даних

Після обробки обидві групи ознак узгоджуються в єдиному часовому просторі, що уникає проблеми зміщення впливу новин щодо фактичних ринкових змін, і подаються у вигляді єдиного вектора стану. Узагальнений вектор стану формується як

$$x_t = [C_t, V_t, SMA_t, EMA_t, RSI_t, s_t], \quad (1)$$

де C_t – ціна закриття, V_t – обсяг торгів, SMA_t та EMA_t – ковзні середні, RSI_t – індекс відносної сили, s_t – скалярна оцінка тональності новин.

Така структура ознак дозволяє моделі одночасно враховувати поточний тренд, інтенсивність ринкової активності, імпульс зміни ціни та інформаційний настрій ринку, що підвищує стійкість прогнозу в умовах волатильності.

Для прогнозування фінансових часових рядів застосовано дві архітектури глибинного навчання, з яких перша базується на рекурентній моделі LSTM. Архітектура LSTM складається з трьох послідовних шарів із 128, 64 і 32 нейронами, що реалізують послідовне кодування довготривалих і короткотривалих залежностей між станами фінансового ряду. Для уникнення перенавчання між шарами використано коефіцієнт $\text{dropout} = 0.2$, а як функцію активації обрано $\tanh$, що забезпечує стабільність градієнтів під час тривалих послідовностей. Оптимізація параметрів виконується за допомогою алгоритму Adam з початковою швидкістю навчання $\alpha = 0.001$, що забезпечує швидку адаптацію ваг при зміні ринкових умов.

Кожен нейрон LSTM містить вхідний, забувальний і вихідний клапани, що дозволяють відокремлювати релевантні патерни ринку від шуму короткочасних коливань. Завдяки цьому модель формує стійку динамічну репрезентацію сигналу навіть у випадках волатильності чи різких змін тенденцій. Вихідна частина мережі складається з одного лінійного нейрона, який генерує прогнозоване значення ціни закриття $\hat{C}_{t+1}$. Навчання здійснюється на основі функції втрат у вигляді середньоквадратичної похибки

$$L = (1/n) \sum (\hat{C}_{t+1} - C_{t+1})^2, \quad (2)$$

що мінімізується для підвищення точності коротко- та середньострокових прогнозів.

Архітектура Transformer у запропонованій системі використовується для моделювання довгострокових залежностей у часовому ряді та інтеграції оцінки тональності новин у процес прогнозування. Вхідний вектор ознак, сформований на основі фінансових показників і оцінки тональ-

ності, лінійно проектується у простір сталої розмірності $d_{\text{model}} = 64$, що забезпечує узгоджене представлення даних для подальшого обчислення контекстних зв'язків. Оскільки механізм самоуваги не містить внутрішнього урахування порядку елементів послідовності, застосовується позиційне кодування, яке фіксує часову структуру ряду.

Центральним компонентом моделі є блок багатоголової самоуваги, у якому використано чотири незалежні канали уваги для оцінювання відносної важливості ознак у часовій послідовності. У межах кожного каналу окремо обчислюються матриці запитів (Q), ключів (K) та значень (V) розмірності 64×64 , на основі яких формується ваговий коефіцієнт впливу попередніх станів на поточний прогноз. Такий підхід дозволяє моделі вибірково підсилювати інтервали, у яких зміни ринкових показників корелюють із трансформаціями інформаційного фону. У результаті система здатна одночасно враховувати як короткочасні імпульсні коливання, так і довготривалі трендові залежності без деградації градієнтів при збільшенні довжини послідовності, що є критично важливим для аналізу фінансових часових рядів.

Після блоку самоуваги використовується двохшарова feed-forward мережа з розширенням розмірності внутрішнього простору ознак до $d_{\text{ff}} = 256$, що дає змогу моделі формувати нелінійні перетворення та уточнювати локальні патерни залежностей між ознаками. Для забезпечення стабільності навчання перед кожним підблоком застосовуються нормалізація шару (LayerNorm) та залишкові з'єднання, які запобігають накопиченню похибок і сприяють збереженню градієнтів при збільшенні глибини архітектури. Навчання параметрів здійснюється за алгоритмом Adam зі швидкістю навчання $\alpha = 0.001$ та розміром батчу 32 протягом 50 епох.

Оцінка тональності новин s_t інтегрується у вхідний простір ознак на кожному часовому кроці, що забезпечує прямий зв'язок між змінами інформаційного контексту та динамікою фінансового ряду. У такій постановці Transformer аналізує не лише історичні закономірності часової послідовності, а й адаптивно реагує на корекцію очікувань учасників ринку, що є критично важливим у періоди підвищеної волатильності та структурних зрушень у поведінці ринку.

Комп'ютерна система реалізована як локальний програмно-апаратний комплекс, що працює автономно без звернення до зовнішніх обчислювальних сервісів. Навчання та тестування вико-

Результати точності прогнозування для різних моделей

Модель	Дані	RMSE	MAPE, %	Поліпшення
LSTM	часові ряди	0.036	3.4	–
LSTM + NLP	ряди + новини	0.027	2.5	24 %
Transformer + NLP	ряди + новини	0.025	2.3	29 %

нувались на робочій станції з процесором Intel Core i5-11400, 8 ГБ ОЗП та графічним адаптером NVIDIA GTX 1650 (4 GB, CUDA). Використання GPU забезпечило прискорення обчислень під час тренування Transformer-моделі, зокрема при формуванні матриць уваги та обчисленні градієнтів. Для конфігурацій без GPU передбачено варіант виконання на пристроях із підтримкою OpenCL із пропорційним зменшенням розміру батчу.

Навчання моделей виконувалося протягом 50 епох із розміром батчу 32. Часовий ряд було поділено на навчальну й тестову вибірки у пропорції 80/20 із формуванням вхідних послідовностей довжиною 60 точок шляхом використання ковзного вікна. Для оцінки здатності моделі працювати за різних ринкових режимів додатково проведено тестування на інтервалах із підвищеною волатильністю, коли поточне відхилення σ_t перевищувало середнє значення σ більш ніж на 20 %. Це дозволило оцінити не лише базову точність прогнозу, а й стійкість моделі до аномальних коливань ринку.

Дані включали історичні ряди індексу S&P 500 та акцій Apple Inc. (AAPL) за період 2016–2024 рр. У конфігураціях із використанням новинних характеристик ознака тональності агрегувалася для кожного дня за середнім значенням оцінок окремих новинних повідомлень. Точність прогнозування оцінювалася за метриками RMSE (середньоквадратична похибка) та MAPE (середня абсолютна відносна похибка), що забезпечує можливість коректного порівняння результатів у як абсолютному, так і відносному масштабах.

Результати оцінки наведено в таблиці 1.

Отримані результати свідчать, що включення інформаційних сигналів у вигляді оцінки тональності новин забезпечує зниження похибки прогнозування на 25–30 % порівняно з моделями, які використовують лише числові ринкові показники. При цьому гібридна модель зберігає стабільність прогнозних оцінок у періоди підвищеної волатильності ($\sigma_t > 1.2\sigma$) та характеризується середнім значенням коефіцієнта детермінації $R^2 = 0.91$, що свідчить про високу відповідність між прогнозованими та фактичними значеннями. Таким чином, поєднання часових рядів із текстовими новинними ознаками підвищує інформативність вхідного простору та покращує стійкість моделі до структурних зрушень у ринковій динаміці.

Висновки. Отримані результати засвідчують, що включення оцінки тональності новин у модель прогнозування підвищує точність і стійкість передбачення в умовах ринкової волатильності. Додавання новинних характеристик до LSTM зменшило середню відносну похибку з 3.4% до 2.5%, а використання Transformer у поєднанні з текстовими ознаками дало подальше зниження до 2.3%. Це свідчить про ефективність гібридної архітектури, у якій числові показники ринку поєднуються з інформаційним контекстом, що відображає очікування учасників торгів. Урахування інформаційних сигналів є ключовою умовою формування надійних прогнозних моделей, особливо у періоди структурних зсувів та інформаційних шоків, що підтверджує доцільність подальшого розвитку таких систем і дослідження їх узагальнювальної здатності на ширших класах активів.

Список літератури:

1. Deep learning in finance: a survey of applications and techniques / E. Mienye et al. *Ai*. 2024. Vol. 5, no. 4. P. 2066–2091. DOI: <https://doi.org/10.3390/ai5040101>.
2. LSTM–Transformer–Based robust hybrid deep learning model for financial time series forecasting / M. R. Kabir et al. *Sci*. 2025. Vol. 7, no. 1. P. 7. DOI: <https://doi.org/10.3390/sci7010007>.
3. Financial sentiment analysis and classification: a comparative study of fine-tuned deep learning models / D. K. Nasiopoulos et al. *International journal of financial studies*. 2025. Vol. 13, no. 2. P. 75. DOI: <https://doi.org/10.3390/ijfs13020075>.
4. Evaluation of transformer models for financial targeted sentiment analysis in Spanish / R. Pan et al. *PeerJ computer science*. 2023. Vol. 9. P. e1377. DOI: <https://doi.org/10.7717/peerj-cs.1377>.
5. Wang Y. Application of ARIMA-GARCH-S model to short-term stock forecasting. *Finance & economics*. 2024. Vol. 1, no. 4. DOI: <https://doi.org/10.61173/y2k63978>.

Volokhov D.H., Muliarevych O.V. HYBRID MODEL FOR FINANCIAL MARKET FORECASTING WITH SENTIMENT-AWARE NEWS INTEGRATION

The article proposes a hybrid forecasting system for financial time series that combines numerical market indicators with textual characteristics of news environments through sentiment integration. The motivation of the approach lies in the fact that time series reflect already realized market reactions, while the news stream contains informational precursors of trend shifts. Synchronizing both sources in a unified temporal framework increases the informativeness of the feature space and improves forecast stability under regime changes.

The architecture is presented in two hybrid configurations: an LSTM recurrent model for encoding local and medium-term dependencies, and a Transformer model with multi-head self-attention for capturing long-range relationships and selectively amplifying relevant temporal intervals. In both cases, scalar news sentiment is embedded into the input feature vector at each time step. Experimental validation was conducted using S&P 500 and AAPL data for 2016–2024 with an 80/20 split and a sliding window of 60 points. Evaluation was performed using RMSE and MAPE, with additional testing under increased volatility conditions. The results demonstrate statistically significant error reduction when incorporating news sentiment: for LSTM, MAPE decreases from 3.4% to 2.5%, and for the sentiment-enhanced Transformer to 2.3%, corresponding to a 25–30% improvement compared to models using only numerical indicators. Forecast consistency with actual price movements reaches $R^2 \approx 0.91$, while stability is maintained when $\sigma_t > 1.2\bar{\sigma}$.

These findings indicate that incorporating informational signals reflecting market expectations is essential for increasing the accuracy and robustness of forecasting models. The proposed architecture confirms the effectiveness of the integrated approach and may serve as a basis for further development through domain-specific news corpora and adaptive sentiment weighting mechanisms.

Key words: financial time series, sentiment analysis, hybrid models, LSTM, Transformer, volatility, forecasting.

Дата надходження статті: 27.07.2025

Дата прийняття статті: 06.08.2025

Опубліковано: 27.10.2025

Відомості про авторів

Абрамова А.О. – кандидат технічних наук, доцент, доцент кафедри технічних та програмних засобів автоматизації Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»

Алтухов В.О. – магістр з комп'ютерних наук Донбаської державної машинобудівної академії

Алтухов О.В. – кандидат технічних наук, старший викладач кафедри комп'ютерних інформаційних технологій Донбаської державної машинобудівної академії

Амонс О.А. – кандидат технічних наук, доцент кафедри інформаційних систем та технологій Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»

Арамян А.М. – кандидат фізико-математичних наук, доцент кафедри інженерних технологій ПЗВО «Міжнародний класичний університет імені Пилипа Орлика»

Артеменко О.І. – кандидат технічних наук, доцент, завідувач кафедри комп'ютерних систем і технологій Приватного вищого навчального закладу «Буковинський університет»

Бабенко В.Г. – доктор технічних наук, професор, професор кафедри інформаційної безпеки та комп'ютерної інженерії Черкаського державного технологічного університету

Баглай О.П. – аспірант за спеціальністю 175 Інформаційно-вимірювальні технології Українського державного університету залізничного транспорту

Бандура В.М. – доктор технічних наук, доцент, професор кафедри інженерних технологій ПЗВО «Міжнародний класичний університет імені Пилипа Орлика»

Бачкала Б.О. – аспірант кафедри інформаційних систем та технологій Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»

Билбас Д.О. – студент кафедри технічних та програмних засобів автоматизації Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»

Білан М.В. – здобувач вищої освіти Університету митної справи та фінансів

Білий Д.О. – магістр кафедри інженерії програмного забезпечення Національного університету «Одеська політехніка»

Бісікало О.В. – доктор технічних наук, професор, завідувач кафедри автоматизації та інтелектуальних інформаційних технологій Вінницького національного технічного університету

Біщак Д.С. – аспірант, асистент кафедри програмного забезпечення Тернопільського національного технічного університету імені Івана Пулюя

Бойко О.В. – кандидат технічних наук, доцент, доцент кафедри інформаційних систем та технологій Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»

Большакова О.О. – кандидат технічних наук, доцент кафедри інженерних технологій ПЗВО «Міжнародний класичний університет імені Пилипа Орлика»

Брестовицький Р.М. – аспірант кафедри електронних обчислювальних машин Харківського національного університету радіоелектроніки

Бугаєва Л.М. – кандидат технічних наук, доцент, доцент кафедри технічних та програмних засобів автоматизації Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»

Бугрій А.М. – кандидат технічних наук, старший викладач кафедри електронних обчислювальних машин Харківського національного університету радіоелектроніки

Бурчак П.В. – аспірант кафедри програмного забезпечення комп'ютерних систем Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»

Васильєва Л.В. – кандидат технічних наук, доцент, доцент кафедри інформаційних систем в економіці Київського національного економічного університету імені Вадима Гетьмана

Вергелес А.В. – бакалавр кафедри електронних обчислювальних машин Національного університету «Львівська політехніка»

Веселовська Г.В. – кандидат технічних наук, доцент, доцент кафедри комп'ютерних систем та мереж Херсонського національного технічного університету

Вишневий С.В. – кандидат технічних наук, старший викладач кафедри радіотехнічних систем радіотехнічного факультету Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»

Волк М.О. – доктор технічних наук, професор, професор кафедри електронних обчислювальних машин Харківського національного університету радіоелектроніки

Волохов Д.Г. – бакалавр кафедри електронних обчислювальних машин Національного університету «Львівська політехніка»

Гончаров Д.С. – аспірант кафедри комп'ютерної інженерії Чорноморського національного університету імені Петра Могили

Горбаченко С.А. – доктор економічних наук, професор, завідувач кафедри кібербезпеки Національного університету «Одеська юридична академія»

Грезіна О.М. – доктор філософії, асистент кафедри інформаційних технологій Національного університету «Одеська юридична академія»

Гурін Є.П. – здобувач Національного університету «Одеська юридична академія»

Дадиверін В.В. – аспірант факультету автоматизації та інтелектуальних інформаційних технологій Вінницького національного технічного університету

Деркач Я.О. – здобувач вищої освіти Харківського національного університету радіоелектроніки

Дика А.І. – аспірант, асистент кафедри інформаційних технологій Національного університету «Одеська юридична академія»

Дідок В.Ю. – аспірант кафедри будівельної механіки будівельного факультету Київського національного університету будівництва і архітектури

Добровольський Ю.Г. – доктор технічних наук, професор кафедри програмного забезпечення комп'ютерних систем Чернівецького національного університету імені Юрія Федьковича

Дубко В.О. – доктор фізико-математичних наук, професор Таврійського національного університету імені В.І. Вернадського

Дячук Р.Л. – аспірант кафедри програмного забезпечення комп'ютерних систем Чернівецького національного університету імені Юрія Федьковича

Єганов О.Ю. – кандидат технічних наук, професор, професор кафедри інженерних технологій ПЗВО «Міжнародний класичний університет імені Пилипа Орлика»

Єфименко М.В. – доктор технічних наук, професор, професор кафедри інформаційних технологій електронних засобів Національного університету «Запорізька політехніка»

Журба А.В. – бакалавр кафедри радіотехнічних систем радіотехнічного факультету Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»

Здольник М.О. – аспірант кафедри автоматизації енергетичних процесів Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»

Зіноватна С.Л. – кандидат технічних наук, доцент кафедри інженерії програмного забезпечення Національного університету «Одеська політехніка»

Зінченко Д.Д. – аспірант кафедри автоматизації енергетичних процесів Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»

Іванченко Д.В. – здобувач другого освітнього ступеня спеціальності «Комп'ютерна інженерія» кафедри електронних обчислювальних машин Дніпровського інституту інфраструктури і транспорту Українського державного університету науки і технологій

Ільїн М.О. – аспірант кафедри програмного забезпечення комп'ютерних систем Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»

Калашник В.Ю. – кандидат технічних наук, старший викладач кафедри комп'ютерної інженерії та електромеханіки Київського національного університету технологій та дизайну

Ковтун Є.І. – аспірант кафедри електронних обчислювальних машин Харківського національного університету радіоелектроніки

Ковтюх В.А. – здобувач ступеня доктора філософії кафедри інформаційної безпеки та комп'ютерної інженерії Черкаського державного технологічного університету

Козуб Г.О. – кандидат технічних наук, доцент, доцент кафедри інформаційних технологій та систем Навчально-наукового інституту математики та інформаційних технологій ДЗ «Луганський національний університет імені Тараса Шевченка»

Комлева Н.О. – кандидат технічних наук, доцент, завідувач кафедри інженерії програмного забезпечення Національного університету «Одеська політехніка»

Короткий Т.К. – здобувач ступеня доктора філософії кафедри інформаційних технологій проектування Черкаського державного технологічного університету

Костенко В.П. – старший викладач кафедри комп'ютерних інформаційних технологій факультету інформаційних технологій Військового інституту телекомунікацій та інформатизації імені Героїв Крут

Кошкін В.К. – кандидат технічних наук, доцент кафедри інженерних технологій ПЗВО «Міжнародний класичний університет імені Пилипа Орлика»

Кравчук Я.Я. – веб-розробник Velar Voyage LLC USA

Крамар А.В. – магістр, викладач кафедри комп'ютерних систем і технологій Приватного вищого навчального закладу «Буковинський університет»

Кримська А.О. – кандидат технічних наук, доцент кафедри менеджменту, маркетингу і міжнародної логістики Чернівецького торговельно-економічного інституту Державного торговельно-економічного університету

Куляба-Харитонova Т.І. – старший викладач кафедри інформаційних технологій електронних засобів Національного університету «Запорізька політехніка»

Кухаренко С.В. – кандидат технічних наук, доцент, доцент кафедри кібербезпеки Національного університету «Одеська юридична академія»

Латипов О.В. – здобувач групи БК-613м Національного університету «Запорізька політехніка»

Лобода Ю.Г. – кандидат педагогічних наук, доцент, доцент кафедри інформаційних технологій Національного університету «Одеська юридична академія»

Логінова Н.І. – кандидат педагогічних наук, доцент, завідувачка кафедри інформаційних технологій Національного університету «Одеська юридична академія»

Манаків С.Ю. – кандидат технічних наук, доцент, доцент кафедри інформаційних технологій Національного університету «Одеська юридична академія»

Маріаш Ю.І. – доктор філософії, доцент кафедри автоматизації енергетичних процесів Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»

Мельник Г.В. – кандидат технічних наук, доцент, доцент кафедри комп'ютерної інженерії та електромеханіки Київського національного університету технологій та дизайну

Миронова Н.О. – кандидат технічних наук, доцент, доцент кафедри інформаційних технологій електронних засобів Національного університету «Запорізька політехніка»

Мовчан К.О. – старший науковий співробітник Українського науково-дослідного інституту спеціальної техніки та судових експертиз Служби безпеки України

Молочков Д.Є. – доктор філософії, старший викладач кафедри інформаційних технологій електронних засобів Національного університету «Запорізька політехніка»

Муляревич О.В. – кандидат технічних наук, доцент, доцент кафедри електронних обчислювальних машин Національного університету «Львівська політехніка»

Налісний Г.І. – викладач кафедри комп'ютерних інформаційних технологій факультету інформаційних технологій Військового інституту телекомунікацій та інформатизації імені Героїв Крут

Нікітченко М.І. – аспірант кафедри інженерії програмного забезпечення Національного університету «Одеська політехніка»

Овчинников М.Ю. – асистент кафедри кібербезпеки Національного університету «Одеська юридична академія»

Олещенко Л.М. – кандидат технічних наук, доцент кафедри програмного забезпечення комп'ютерних систем Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»

Онщенко І.О. – викладач інформатики Одеського ліцею «Надія» Одеської міської ради

Онщенко Т.В. – старший викладач кафедри інженерії програмного забезпечення Національного університету «Одеська політехніка»

Осадчук С.І. – викладач кафедри комп'ютерних систем і технологій Приватного вищого навчального закладу «Буковинський університет»

Пахомова В.М. – кандидат технічних наук, доцент, доцент кафедри електронних обчислювальних машин Дніпровського інституту інфраструктури і транспорту Українського державного університету науки і технологій

Петрик М.Р. – доктор фізико-математичних наук, професор, завідувач кафедри програмної інженерії Тернопільського національного технічного університету імені Івана Пулюя

Петросян А.Р. – аспірант, асистент кафедри комп'ютерних наук Державного університету «Житомирська політехніка»

Плахотнюк М.Ф. – аспірант кафедри комп'ютерних систем та мереж Херсонського національного технічного університету

Правосєров А.О. – аспірант Інституту атомної та теплової енергетики Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»

Прокопович-Ткаченко Д.І. – кандидат технічних наук, доцент, завідувач кафедри кібербезпеки та інформаційних технологій Університету митної справи та фінансів, старший науковий співробітник Державної наукової установи «Інститут інформації, безпеки і прав Національної академії правових наук України»

Ролік О.І. – доктор технічних наук, професор, завідувач кафедри інформаційних систем та технологій Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»

Рудницький В.М. – доктор технічних наук, професор, головний науковий співробітник Державного науково-дослідного інституту випробувань і сертифікації озброєння та військової техніки, професор кафедри інформаційної безпеки та комп'ютерної інженерії Черкаського державного технологічного університету

Рудницький С.В. – кандидат технічних наук, доцент, доцент кафедри інформаційних технологій проектування Черкаського державного технологічного університету, провідний науковий співробітник Державного науково-дослідного інституту випробувань і сертифікації озброєння та військової техніки

Савенко О.С. – доктор технічних наук, професор, професор кафедри комп'ютерної інженерії та інформаційних систем Хмельницького національного університету

Самойлов І.А. – аспірант кафедри електронних обчислювальних машин Харківського національного університету радіоелектроніки

Сичков В.С. – аспірант, асистент кафедри інженерії програмного забезпечення Національного університету «Одеська політехніка»

Соробей Б.В. – аспірант кафедри електронних обчислювальних машин Харківського національного університету радіоелектроніки

Стаценко В.В. – доктор технічних наук, доцент, професор кафедри комп'ютерної інженерії та електромеханіки Київського національного університету технологій та дизайну

Стаценко Д.В. – кандидат технічних наук, доцент, завідувач кафедри комп'ютерної інженерії та електромеханіки Київського національного університету технологій та дизайну

Степанець О.В. – кандидат технічних наук, доцент, доцент кафедри автоматизації енергетичних процесів Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»

Стецюк Ю.В. – аспірант, асистент кафедри комп'ютерної інженерії та інформаційних систем Хмельницького національного університету

Столярчук В.В. – доктор філософії, спеціаліст, інженер-випробувач Державного підприємства «Конструкторське бюро «Південне» імені М. К. Янгеля», аспірант кафедри двигунобудування фізико-технічного факультету Дніпровського національного університету імені Олеся Гончара

Стоцький І.В. – старший викладач кафедри комп'ютерних інформаційних технологій факультету Інформаційних технологій Військового інституту телекомунікацій та інформатизації імені Героїв Крут

Тимошин Ю.А. – кандидат технічних наук, доцент кафедри інформаційних систем та технологій Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»

Ульяницька К.О. – кандидат технічних наук, доцент кафедри інформаційних систем та технологій Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»

Ульяновська Ю.В. – кандидат технічних наук, доцент, завідувач кафедри комп'ютерних наук та інженерії програмного забезпечення Університету митної справи та фінансів

Устинов Д.А. – викладач кафедри комп'ютерних інформаційних технологій факультету інформаційних технологій Військового інституту телекомунікацій та інформатизації імені Героїв Крут

Фомін О.С. – кандидат технічних наук, доцент кафедри автоматики, електроніки та телекомунікацій Навчально-наукового інституту інформаційних технологій та робототехніки Національного університету «Полтавська політехніка імені Юрія Кондратюка»

Хрушков Б.С. – аспірант кафедри кібербезпеки та інформаційних технологій Університету митної справи та фінансів

Цимбал С.І. – асистент кафедри інформаційних систем та технологій Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»

Чепурна О.Є. – кандидат фізико-математичних наук, доцент, доцент кафедри кібербезпеки Національного університету «Одеська юридична академія»

Черевко Є.В. – кандидат фізико-математичних наук, доцент, доцент кафедри кібербезпеки Національного університету «Одеська юридична академія»

Черкаський О.В. – здобувач вищої освіти Університету митної справи та фінансів

Щербина Ю.В. – кандидат технічних наук, доцент, доцент кафедри інформаційних технологій Національного університету «Одеська юридична академія»

Юрчак І.Ю. – кандидат технічних наук, доцент, доцент кафедри систем автоматизованого проектування Національного університету «Львівська політехніка»

Яковенко В.О. – доктор технічних наук, доцент, професор кафедри комп'ютерних наук та інженерії програмного забезпечення Університету митної справи та фінансів

НОТАТКИ

Науковий журнал

**ВЧЕНІ ЗАПИСКИ
ТАВРІЙСЬКОГО НАЦІОНАЛЬНОГО УНІВЕРСИТЕТУ
ІМЕНІ В.І. ВЕРНАДСЬКОГО**

Серія: Технічні науки

Том 36 (75) № 4 2025

Частина 2

Коректура • *Н. Славогородська*

Комп'ютерна верстка • *Н. Кузнєцова*

Адреса редакції:

Таврійський національний університет імені В.І. Вернадського

м. Київ, вул. Джона Маккейна, 33

Електронна пошта: editor@tech.vernadskyjournals.in.ua

Сторінка журналу: www.tech.vernadskyjournals.in.ua

Формат 60×84/8. Гарнітура Times New Roman.

Папір офсетний. Цифровий друк. Обл.-вид. арк. 35,21. Ум. друк. арк. 42,32. Зам. № 0925/730

Підписано до друку 12.09.2025. Наклад 150 прим.

Видавництво і друкарня – Видавничий дім «Гельветика»

65101, м. Одеса, вул. Інглезі, 6/1

Телефони: +38 (095) 934 48 28, +38 (097) 723 06 08

E-mail: mailbox@helvetica.ua

Свідоцтво суб'єкта видавничої справи

ДК № 7623 від 22.06.2022 р.