

Системи та технології⁰⁰⁰

(правонаступник наукового журналу
“Вісник Академії митної служби України.
Серія: “Технічні науки”)

№ 2 (70)

Науковий журнал включено до Переліку наукових фахових видань України категорії “Б”, в яких можуть публікуватися результати дисертаційних робіт на здобуття наукових ступенів з галузі “Технічні науки”, спеціальності D3 Менеджмент; F1 Прикладна математика; F3 Комп'ютерні науки; F5 Кібербезпека та захист інформації; F7 Комп'ютерна інженерія; G5 Електроніка, електронні комунікації, приладобудування та радіотехніка; G11 Машинобудування (за спеціалізаціями); J2 Готельно-ресторанна справа та кейтеринг; J5 Морський та внутрішній водний транспорт; J6 Авіаційний транспорт; J7 Залізничний транспорт; J8 Автомобільний транспорт.



Видавничий дім
«Гельветика»
2025

Системи та технології
(правонаступник наукового журналу
“Вісник Академії митної служби України. Серія: “Технічні науки”)
Науковий журнал. Видається двічі на рік. Заснований у травні 1999 р.
Рекомендовано до друку та до поширення через мережу Інтернет вченою радою
Університету митної справи та фінансів (протокол № 6 від 29.12.2025 р.)

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

Редакційна колегія:

Кузьменко А. І. – к.т.н., доц.
(головний редактор);
Халіпова Н. В. – к.т.н., доц.
(заступник головного редактора);
Прокопович-Ткаченко Д. І. – к.т.н., доц.
(заступник головного редактора);
Йозеф Костольни – PhD;
Ян Рабчан – PhD;
Бакіров Мюшфік Панах огли – к.т.н.;
Балацька Н. Ю. – д.е.н., доц.;
Бондаренко І. О. – д.т.н., доц.;
Боярчук А. В. – к.т.н., доц.;
Брежнєв Є. В. – д.т.н., с.н.с.;
Вишнікіна О. В. – к.х.н., доц.;
Власов А. В. – к.т.н., ст. досл.;
Волосова Н. М. – к.т.н.;
Гарт Е. Л. – д.ф.-м.н., проф.;
Гордєєв О. О. – к.т.н., доц.;
Джинджоян В. В. – д.е.н., доц.;
Доценко С. І. – д.т.н., доц.;
Защолкін К. В. – к.т.н., доц.;
Котух Є. В. – к.т.н.;

Кузін М. О. – д.т.н., доц.;
Кучер М. М. – к.е.н., доц.;
Мартинюк О. М. – к.т.н., доц.;
Музикін М. І. – к.т.н.;
Нестеренко Г. І. – к.т.н., доц.;
Огар О. М. – д.т.н., проф.;
Охріменко Т. О. – к.т.н.;
Поночовний Ю. Л. – д.т.н., проф.;
Примаченко Г. О. – к.т.н., доц.;
Прохорченко Г. О. – к.т.н., доц.;
Сабіров О. В. – к.т.н., доц.;
Сохацький А. В. – д.т.н., проф.;
Стеблюк Н. Ф. – к.е.н., доц.;
Стеблянко П. О. – д.ф.-м.н.;
Чопоров С. В. – д.т.н., проф.;
Шапорін Р. О. – к.т.н., доц.;
Щербовських С. В. – д.т.н., с.н.с.;
Юдіна О. І. – д.е.н., доц.;
Язіна В. А. – к.е.н.;
Яремчук С. О. – к.т.н.

DOI: <https://doi.org/10.32836/2521-6643-2025-2-70>
ISSN 2521-6643

Коректори: Н. В. Славогородська, Н. С. Ігнатова
Комп'ютерна верстка: А. О. Філатов

Реєстрація суб'єкта у сфері друкованих медіа: Рішення Національної ради України з питань телебачення і радіомовлення
№ 1136 від 11.04.2024 року. Ідентифікатор медіа: R30-03967.

Суб'єкт у сфері друкованих медіа – Університет митної справи та фінансів (вул. Володимира Вернадського, 2/4, 49000,
м. Дніпро, e-mail: university.msf@gmail.com, umfsf.ua@gmail.com, Тел.: (056) 745-55-96).
Періодичність видання: 2 рази на рік.

Мови розповсюдження: українська, англійська, польська, німецька, французька, італійська, литовська.

Адреса: м. Дніпро, вул. Володимира Вернадського, 2/4, 49000
Тел.: +38 (099) 729 63 79
E-mail: editor@st.umsf.in.ua
Сайт видання: st.umsf.in.ua

Видавництво і друкарня – Видавничий дім «Гельветика»
65101, Україна, 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.

Підписано до друку 30.12.2025. Формат 60×84/8. Папір офсетний.
Гарнітура Таймс. Ум. друк. арк. 46,04. Обл.-вид. арк. 39,66.
Наклад 100 прим. Замовлення № 0126/069.

Засновник: Університет митної справи та фінансів

ЗМІСТ

ПРИКЛАДНА МАТЕМАТИКА

Галушко О. І. Нейромережева модель прогнозування рівня ризику імплантаційного лікування пацієнтів із патологією шлунково-стравохідного з'єднання	9
Єршов С. В., Симонов Є. Д. Аналітичні межі топологічної стійкості динамічних систем за умов шуму та структурних аномалій.....	19
Жир С. І., Кузьменко А. І., Гаврилов Б. О., Шаповалов В. О. Математичне моделювання доставки вантажу на великій відстані за допомогою світлових вітрил.....	28
Зайцева Т. А., Гончаров Я. А. Синергетичне застосування ANSYS і MATLAB для моделювання контактної взаємодії штампа складної геометрії з пружним півпростором.....	40
Надригайло Т. Ж., Перемітько М. В. Методи аналізу індивідуальних вимірювань артеріального тиску: тренди, сезонність та прогнозування.....	47
Пасічник А. М., Надригайло Т. Ж. Математична модель та оптимізація параметрів пневматичного ротаційно-поршневого двигуна.....	55
Сафронова І. А., Жир С. І., Верба О. В., Шишканова Г. А. Адаптивні стратегії викладання та оцінювання в умовах війни.....	65

КОМП'ЮТЕРНІ НАУКИ

Аксак Н. Г., Татарников А. О. Багатошарова мікросервісно-агентна архітектура системи персоналізованого навчання.....	74
Божко О. Ю. Метод класифікації документів за складністю екстракції даних великими мовними моделями.....	84
Булгакова О. Ф., Костенко В. В., Булгаков Д. С. Автоматизація розгортання серверлес-застосунків із використанням IaC: кейс системи «Agro Monitor».....	93
Булгакова О. Ф. Огляд і узагальнення практик впровадження віртуалізації у підготовці IT-фахівців.....	102
Гаврилюк Є. А. Нейронно-керований гібридний метод для оптимального покриття території із обмеженням на розташування.....	112
Гайтан О. М., Бублій В. Е. Технології штучного інтелекту в системах рекрутингу та управління персоналом.....	119
Dobrovolskyi H. A., Sgadov S. A. Formulation of the intentionally gapped problems for AI-literate learning in higher education.....	133
Корабльов М. М., Новосельцев І. В., Кобзев І. В., Ткачук О. К. Розподілена система штучного інтелекту з агентно-оркестрованою архітектурою.....	145
Новохатько І. О., Федотова Н. А. Процедурна генерація у чотиривимірному просторі: від 4D градієнтного шуму до візуалізації кватерніонних фракталів.....	154
Oleksiiichuk Yu. F. Design and implementation of a chatbot for prospective students.....	162
Погорєловська І. Д., Горбовий А. Ю., Гладченко О. В., Ратушняк Т. В. Контракти як елемент забезпечення якості IT-продуктів.....	170
Поперешняк С. В. Застосування генераторів псевдовипадкових послідовностей для енергоефективного керування IoT-мережами	177
Симонов Д. І., Деменко І. О. Формалізація процесів автоматизованого тестування складних програмних систем на основі мультиагентного моделювання.....	186
Фірсов О. Д. Візуалізація 3D метадокса засобами Three.js.....	195
Shamov O. A. The role of generative AI in the automation of legal document preparation: an analysis of accuracy, efficiency, and ethical aspects.....	204

КОМП'ЮТЕРНА ІНЖЕНЕРІЯ

Ахак N. G., Shelikhov Yu. O. Intelligent load balancing in microservice architecture.....	210
Дьячук Т. С., Скрупський С. Ю., Голуб Т. В. Автоматизована система перевірки завдань для навчальних курсів із програмування.....	219
Кармазін К. В., Писаренко Ю. В. Інтелектуальні алгоритми підвищення якості користувацького досвіду в адаптивних системах потокового відео у задачах ситуаційного центру з використанням технопарку робототехнічних систем.....	226
Киричек Г. Г., Целуйко Р. О., Тягунова М. Ю., Живогляд В. А. Хмарні сервіси в інфраструктурі розподілених мереж.....	232
Крихівський М. В., Ваврик Т. О., Гобир Л. М. Основи проектування інтелектуальних агентів: від архітектури до математичних моделей.....	240

КІБЕРБЕЗПЕКА ТА ЗАХИСТ ІНФОРМАЦІЇ

Kovalchuk D. M. Method for implementing the squaring operation in the Rabin cryptosystem based on the use of the residue number system.....	248
Рудницький В. М., Тарасенко Я. В., Лада Н. В., Бабенко В. Г., Підласий Д. А. Аналіз результатів моделювання СЕТ-операцій на основі елементарних функцій операцій керованих інформацією.....	258

МЕНЕДЖМЕНТ

Кузьменко А. І., Черчата А. О., Халіпова Н. В., Леснікова І. Ю. Оптимізація міжнародних логістичних потоків на підставі методів теорії прийняття рішень в умовах невизначеності.....	264
Muzykin M. I., Nesterenko H. I., Bibik S. I., Chubenko O. I. Logistics methods for managing the transportation of humanitarian cargo near the war zone.....	277
Погребняк В. Г., Перкун І. В., Вишнікіна О. В., Погребняк А. В. Від охорони праці до комплексної безпеки: стратегії інтеграції цивільного захисту в корпоративні системи управління.....	288

МАШИНОБУДУВАННЯ (ЗА СПЕЦІАЛІЗАЦІЯМИ)

Кирилах С. В. Вплив фрикційно-змішувальної обробки (friction stir processing) на мікроструктуру та механічні властивості адитивно виготовленого Ti-6Al-4V.....	294
Майбородіна Н. В., Герасименко В. П. Математична модель залежності обсягу реалізованої продукції машинобудування від обсягу видобутку залізних руд.....	302
Перкун І. В. Гідроімпульсна перфорація нафтогазових свердловин струменем розчину полімеру.....	310
Поляков В. О., Пославський С. Ю. Розробка алгоритму побудови моделі динаміки вагона маглев поїзда....	318
Новальська Н. І., Клименко В. В., Єрошенко О. Р., Ситник Д. О. Перспективи розвитку логістичної інфраструктури пасажирських перевезень: організаційно-технологічні та економічні аспекти.....	325

МОРСЬКИЙ ТА ВНУТРІШНІЙ ВОДНИЙ ТРАНСПОРТ

Drozhzhyn O. L., Onyschenko S. P. Impact assessment of liner schedule disruption on fleet efficiency based on time charter equivalent.....	332
Коскіна Ю. О. Визначення тривалості простого рейсу судна на умовах договору фрахтування на рейс.....	342

ЕЛЕКТРОНІКА, ЕЛЕКТРОННІ КОМУНІКАЦІЇ, ПРИЛАДОБУДУВАННЯ ТА РАДІОТЕХНІКА

Альперт С. І. Обробка супутникових знімків з використанням апарату теорії нечітких множин.....	351

ГОТЕЛЬНО-РЕСТОРАННА СПРАВА ТА КЕЙТЕРИНГ

Горб К. М., Корнєєв М. В., Павлов В. В. Технології готельного обслуговування туристичних груп.....	359
Horozhankina N. A., Saihak Ye. L., Grytsenko A. V., Topol D. S. Optimisation of the staff motivation system in the hotel business.....	367
Kucher M. M., Korneyev M. V., Shcholokova H. V., Zakharenko O. O. Current aspects of personnel motivation in the hotel and restaurant sector of Ukraine.....	372
Stebliuk N. F., Nebaba N. O., Razinkova M. Yu., Hryshchenko S. V. Implementing customer-centric approaches in the restaurant industry: analytics, motivation, efficiency.....	380
Yudina O. I., Kucher M. M., Sabirov O. V., Boholiubova A. V. Current trends and creative approaches to innovative development of hospitality industry enterprises in worldwide.....	388

CONTENTS

APPLIED MATHEMATICS

Halushko O. I. Neural network model for predicting risk level of implantation treatment in patients with gastroesophageal junction pathology.....	9
Yershov S. V., Symonov Ye. D. Analytical Bounds on the Topological Stability of Dynamical Systems under Noise and Structural Anomalies.....	19
Zhyr S. I., Kuzmenko A. I., Gavrilov B. O., Shapovalov V. O. Mathematical modeling of cargo delivery over long distances using light sails.....	28
Zaytseva T. A., Honcharov Ya. A. Synergetic Application of ANSYS and MATLAB for Modeling the Contact Interaction of a Complex-Shaped Stamp with an Elastic Half-Space.....	40
Nadryhailo T. Zh., Peremitko M. V. Methods for analyzing individual blood pressure measurements: trends, seasonality, and forecasting.....	47
Pasichnyk A. M., Nadryhailo T. Zh. Mathematical model and optimization of parameters of and pneumatic rotary piston engine.....	55
Safronova I. A., Zhyr S. I., Verba O. V., Shyshkanova G. A. Adaptive strategies of teaching and assessment in wartime conditions.....	65

COMPUTER SCIENCES

Axak N. G., Tatarnykov A. O. Multilayer microservice-agent architecture of a personalized learning system.....	74
Bozhko O. Yu. Method for classifying documents according to the complexity of data extraction by large language models.....	84
Bulhakova O. F., Kostenko V. V., Bulhakov D. S. Automation of serverless application deployment using IaC: case of 'Agro Monitor'.....	93
Bulhakova O. F. A Review and synthesis of virtualization implementation practices in IT professional training.....	102
Havryliuk Ye. A. Neural-driven hybrid method for optimal area coverage with placement constraints.....	112
Haitan O. M., Bublii V. E. Artificial intelligence technologies in recruiting and staff management systems.....	119
Dobrovolskyi H. A., Sgadov S. A. Formulation of the intentionally gapped problems for AI-literate learning in higher education.....	133
Korablyov M. M., Novoseltsev I. V., Kobzev I. V., Tkachuk O. K. Distributed artificial intelligence system with agent-orchestrated architecture.....	145
Novokhatko I. O., Fedotova N. A. Procedural generation in four-dimensional space: from 4D gradient noise to visualization of quaternion fractals.....	154
Oleksiichuk Yu. F. Design and implementation of a chatbot for prospective students.....	162
Pogorielovska I. D., Horbovyi A. Yu., Hladchenko O. V., Ratushnyak T. V. Contracts as an element of quality assurance of IT products.....	170
Popereshnyak S. V. Application of Pseudo-Random Sequence Generators for Energy-Efficient Control of IoT Networks.....	177
Symonov D. I., Demenko I. O. Formalization of automated testing processes of complex software systems based on multiagent modeling.....	186
Firsov O. D. Visualization of 3D metadoxes using Three.js.....	195
Shamov O. A. The role of generative AI in the automation of legal document preparation: an analysis of accuracy, efficiency, and ethical aspects.....	204

COMPUTER ENGINEERING

Axak N. G., Shelikhov Yu. O. Intelligent load balancing in microservice architecture.....	210
Diachuk T. S., Skrupsky S. Yu., Holub T. V. Automated assignment grading system for programming courses.....	219
Karmazin K. V., Pisarenko J. V. Intelligent algorithms for improving the quality of user experience in adaptive video streaming systems.....	226
Kyrychek H. H., Tseluiko R. O., Tiahunova M. Yu., Zhyvohliad V. A. Cloud services in distributed network infrastructure.....	232
Krykhiskyi M. V., Vavryk T. O., Hobyr L. M. Foundations of Intelligent Agent Design: From Architecture to Mathematical Modeling.....	240

CYBERSECURITY AND DATA PROTECTION

Kovalchuk D. M. Method for implementing the squaring operation in the Rabin cryptosystem based on the use of the residue number system.....	248
Rudnytskyi V. M., Tarasenko Ya. V., Lada N. V., Babenko V. H., Pidlasyi D. A. Analysis of simulation results of CET-operations based on elementary functions of information-driven operations.....	258

MANAGEMENT

Kuzmenko A. I., Cherchata A. A., Khalipova N. V., Lesnikova I. Yu. Optimization of international logistics flows based on methods of the theory of decision-making in conditions of uncertainty.....	264
Muzykin M. I., Nesterenko H. I., Bibik S. I., Chubenko O. I. Logistics methods for managing the transportation of humanitarian cargo near the war zone.....	277
Pogrebnyak V. G., Perkun I. V., Vyshnikina O. V., Pogrebnyak A. V. From Occupational Safety to Comprehensive Security: Strategies for Integrating Civil Defense into Corporate Management Systems.....	288

MECHANICAL ENGINEERING (WITH SPECIALIZATIONS)

Kyrylakha S. V. Effect of Friction Stir Processing (FSP) on the Microstructure and Mechanical Properties of Additively Manufactured Ti-6Al-4V.....	294
Maiborodina N. V., Gerasymenko V. P. Mathematical modeling of engineering production volume dependence on iron ore extraction.....	302
Perkun I. V. Hydro-pulse perforation of oil and gas wells with a polymer solution jet.....	310
Polyakov V. O., Poslavsky S. Yu. Development of a maglev train's carriage's dynamic's model constructing algorithm.....	318
Novalska N. I., Klymenko V. V., Yeroshenko O. R., Sytnyk D. O. Prospects for the development of passenger transportation logistics infrastructure: organizational, technological and economic aspects	325

MARITIME AND INLAND WATER TRANSPORT

Drozhhyn O. L., Onyschenko S. P. Impact assessment of liner schedule disruption on fleet efficiency based on time charter equivalent.....	332
Koskina Yu. O. Determining the time of a vessel's single voyage under the terms of a voyage charter-party.....	342

ELECTRONICS, ELECTRONIC COMMUNICATIONS, INSTRUMENT ENGINEERING AND RADIO ENGINEERING

Alpert S. I. Processing of satellite images using the apparatus of fuzzy set theory.....	351

HOTEL AND RESTAURANT BUSINESS AND CATERING

Horb K. M., Korneyev M. V., Pavlov V.V. Hotel service technologies of tourist groups.....	359
Horozhankina N. A., Saihak Ye. L., Grytsenko A. V., Topol D. S. Optimisation of the staff motivation system in the hotel business.....	367
Kucher M. M., Korneyev M. V., Shcholokova H. V., Zakharenko O. O. Current aspects of personnel motivation in the hotel and restaurant sector of Ukraine.....	372
Stebliuk N. F., Nebaba N. O., Razinkova M. Yu., Hryshchenko S. V. Implementing customer-centric approaches in the restaurant industry: analytics, motivation, efficiency.....	380
Yudina O. I., Kucher M. M., Sabirov O. V., Boholiubova A. V. Current trends and creative approaches to innovative development of hospitality industry enterprises in worldwide.....	388

ПРИКЛАДНА МАТЕМАТИКА

УДК 519.2

DOI <https://doi.org/10.32782/2521-6643-2025-2-70.1>

Галушко О. І., аспірант механіко-математичного факультету
Київського національного університету
імені Тараса Шевченка
ORCID: 0009-0009-0275-6542

НЕЙРОМЕРЕЖЕВА МОДЕЛЬ ПРОГНОЗУВАННЯ РІВНЯ РИЗИКУ ІМПЛАНТАЦІЙНОГО ЛІКУВАННЯ ПАЦІЄНТІВ ІЗ ПАТОЛОГІЄЮ ШЛУНКОВО-СТРАВОХІДНОГО З'ЄДНАННЯ

У статті представлено математичну модель прогнозування рівня ризику можливості імплантації терапевтичних пристроїв у пацієнтів із патологією шлунково-стравохідного з'єднання на основі штучної нейронної мережі типу багатошарового перцептрона (Multilayer Perceptron, MLP). Модель інтегрує 15 клінічних, ендоскопічних та морфометричних показників, що відображають як функціональний стан стравоходу, так і структурні зміни слизової оболонки. Для побудови моделі використано вибірку з 558 спостережень (401 – навчальна, 157 – тестова), що забезпечило статистично достовірну оцінку її прогностичної ефективності. Архітектура мережі включає вхідний шар із 50 параметрів (після кодування змінних), один прихований шар із 7 нейронами та вихідний шар із 5 вузлів, які відповідають рівням ризику (0–4). Навчання здійснювалося методом зворотного розповсюдження помилки з використанням оптимізатора Scaled Conjugate Gradient. Модель продемонструвала надзвичайно високу точність класифікації – 99,8 % для навчальної та 99,4 % для тестової вибірок, при цьому значення площі під ROC-кривими (AUC) для кожного класу становили 1. Така стабільність свідчить про відсутність перенавчання та високу узагальнювальну здатність моделі. Найбільший внесок у формування прогнозу мають морфометричні параметри, що характеризують ступінь пролабування шлунка в стравохід, показники діафрагмального звуження та розташування Z-лінії, а також ендоскопічні ознаки ураження слизової – наявність ерозій і виразок, вкритих гематином. Отримані результати демонструють потенціал нейромережевого підходу для автоматизованої стратифікації пацієнтів за рівнем ризику та обґрунтованого вибору інтервенційних стратегій (зокрема імплантації систем Bravo, JSPH-1, BEST Capsule, EndoStim тощо). Запропонована модель може стати основою для створення інтелектуальної системи підтримки клінічних рішень, що підвищує точність діагностики, сприяє персоналізації лікування та знижує ризики необґрунтованих інвазивних втручань у гастроентерологічній практиці.

Ключові слова: багатошаровий перцептрон, штучна нейронна мережа, класифікація, клінічне прогнозування, ендоскопічні показники, ризик імплантації.

Halushko O. I. Neural network model for predicting risk level of implantation treatment in patients with gastroesophageal junction pathology

The article presents a mathematical model for predicting the risk level of potential therapeutic device implantation in patients with gastroesophageal junction pathology, based on a multilayer perceptron (MLP) artificial neural network. The model integrates 15 clinical, endoscopic, and morphometric indicators reflecting both the functional state of the esophagus and structural changes in the mucosa. The model was developed using a dataset of 558 observations (401 for training, 157 for testing), ensuring a statistically reliable assessment of its predictive performance.

The network architecture includes an input layer with 50 parameters (after variable encoding), a single hidden layer with 7 neurons, and an output layer with 5 nodes corresponding to risk levels (0–4). Training was performed using the backpropagation algorithm with the Scaled Conjugate Gradient optimizer. The model demonstrated extremely high classification accuracy – 99.8 % for the training set and 99.4 % for the test set, with the area under the ROC curves (AUC) for each class equal to 1. Such stability indicates the absence of overfitting and high generalization ability of the model.

The greatest contribution to the prediction comes from morphometric parameters characterizing the degree of gastric prolapse into the esophagus, diaphragmatic constriction indices, and the position of the Z-line, as well as endoscopic features of mucosal damage, including the presence of erosions and ulcers covered with hematin. The results demonstrate the potential of the neural network approach for automated patient stratification by risk level and informed selection of interventional strategies (including the implantation of Bravo, JSPH-1, BEST Capsule, EndoStim systems, etc.).

The proposed model could serve as a foundation for the development of an intelligent clinical decision support system, improving diagnostic accuracy, enabling personalized treatment, and reducing the risks of unjustified invasive interventions in gastroenterological practice.

Key words: multilayer perceptron, artificial neural network, classification, clinical prediction, endoscopic indicators, implantation risk.

© О. І. Галушко, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

Постановка проблеми. Одним із ключових напрямів сучасної прикладної математики є розроблення математичних моделей для аналізу, прогнозування та підтримки прийняття рішень у медичних системах, де вихідні дані мають змішану природу (кількісні, якісні, порядкові змінні) та характеризуються складними нелінійними взаємозв'язками. Особливої актуальності це набуває при моделюванні клінічних процесів, пов'язаних із патологією шлунково-стравохідного з'єднання, де традиційні статистичні методи (дискримінантний аналіз, логістична регресія тощо) часто виявляються недостатньо ефективними через невиконання припущень лінійності, нормальності розподілу та ін.

Задача класифікації пацієнтів за рівнем ризику розвитку ускладнень або необхідності імплантації медичних пристроїв є багатовимірною та стохастичною за своєю природою. Вона вимагає побудови алгоритмів, здатних до апроксимації складних функцій розподілу ознак у багатовимірному просторі та до узагальнення на нових даних. З математичної точки зору це відповідає задачі побудови функціонального відображення, яке мінімізує функцію помилки класифікації або крос-ентропійних втрат на множині навчальних спостережень.

Застосування штучних нейронних мереж, зокрема багатошарового перцептрона (Multilayer Perceptron, MLP), дозволяє реалізувати цю задачу за рахунок побудови нелінійних відображень, що узагальнюють інформацію між вхідними параметрами різної природи. Однак практична реалізація таких моделей у медичних задачах потребує: формалізації структури моделі з урахуванням обмеженого обсягу клінічних даних; вибору адекватного алгоритму оптимізації, який забезпечує збіжність при складній функції втрат; аналізу стабільності та узагальнювальної здатності моделі на незалежних вибірках.

Таким чином, математична постановка дослідження полягає у побудові, параметричній ідентифікації та оцінюванні ефективності нейромережевої моделі типу MLP, яка інтегрує клінічні, ендоскопічні та морфометричні показники для прогнозування груп ризику для лікувальної імплантації у пацієнтів із патологією шлунково-стравохідного з'єднання. З практичного боку це спрямовано на створення обчислювально ефективного інструменту прогнозу аналітики, а з математичного – на розроблення та дослідження властивостей нелінійної класифікаційної функції, оптимізованої за критерієм крос-ентропійних втрат.

Аналіз останніх досліджень і публікацій. За останнє десятиліття в прикладній медицині та біомодельованні спостерігається стійка тенденція до впровадження методів машинного навчання, зокрема штучних нейронних мереж, для задач класифікації та прогнозування клінічних подій. Наприклад, дослідження показали, що MLP моделі можуть ефективно використовуватися для прогнозування серцево-судинних захворювань за допомогою оптимізації гібридного алгоритму MLP-PSO (Multilayer Perceptron з Particle Swarm Optimization) із точністю приблизно 84,6 % [1]. У іншому дослідженні MLP-модель була застосована для прогнозування виживання пацієнтів з імплантованими кардіовертерами-дефібриляторами (ICD/CRT-D) і показала перевагу над класичними методами [2]. Ці результати підтверджують, що MLP підходи мають потенціал для медичних задач із багатовимірними ознаками та неоднорідними типами даних.

У контексті гастроентерології, зокрема патології шлунково-стравохідного з'єднання, машинне навчання також набуває значення. Огляд [3] демонструє, що моделі машинного навчання досягають точності 80–90 % у задачах класифікації рефлюксних захворювань та супутніх станів. Наприклад, дослідження [4] показало 100 % точність нейромережі у невеликій вибірці з 159 пацієнтів при діагностиці гастроєзофагеальної рефлюксної хвороби. У дослідженні [5] використана нейронна мережа для автоматичної класифікації за Лос-Анджелівською шкалою рефлюкс-езофагіту з точністю до 99,2 % на тренувальній вибірці. Таким чином, існує переконливе підґрунтя для застосування нейромереж у задачах, близьких до нашої: багатофакторного прогнозування клінічних та морфологічних станів.

Проте у публікаціях також підкреслюються виклики та обмеження: проблема невеликих вибірок, балансування класів, адекватне кодування порядкових та номінальних змінних, ризик перенавчання та недостатня інтерпретованість моделей. У системному огляді [6] зазначено, що MLP-моделі хоч і отримали широке застосування, але залишаються чутливими до малого обсягу даних та мають проблеми із пояснювальною здатністю. У статті [7] продемонстровано, що оптимізація MLP-архітектури значно підвищує її продуктивність у “шумних” і нелінійних наборах клінічних даних. Крім того, хоча застосування нейромереж для відбору пацієнтів до імплантаційних процедур менш розроблене, у статті [8] акцентується увага на можливості нейромереж для прогнозування успішності імплантацій (наприклад стоматологічних) і наголошується важливість інтеграції моделей із клінічною практикою.

Таким чином, незважаючи на значний прогрес, необхідна подальша адаптація та валідація моделей саме у контексті імплантаційних медичних пристроїв, із урахуванням багатовимірних даних та клінічних вимог.

Метою статті є розробка та математичне обґрунтування нейромережевої моделі для прогнозування рівня ризику можливості імплантації лікувальних пристроїв у пацієнтів із патологією шлунково-стравохідного з'єднання. Модель базується на інтеграції клінічних, ендоскопічних та морфометричних показників пацієнтів із використанням архітектури багатошарового перцептрона.

Виклад основного матеріалу. У даному дослідженні побудовано математичну модель на основі багатошарового перцептрона (MLP), яка інтегрує клінічні, ендоскопічні та морфометричні показники пацієнтів для прогнозування рівня ризику необхідності імплантації діагностичних та/або терапевтичних пристроїв,

зокрема систем Bravo, JSPH-1, BEST Capsule, EndoStim для стимуляції нижнього стравохідного сфінктера, а також капсульних датчиків Smart panowire та ін. у пацієнтів із патологією шлунково-стравохідного з'єднання [9].

Метою моделювання є підвищення точності та об'єктивності клінічного прийняття рішень шляхом автоматизованої оцінки ризику на основі багатofакторних вхідних даних.

Для побудови моделі використано 15 змінних, що відображають як клінічні симптоми, так і морфологічні характеристики стравоходу та шлунково-стравохідного переходу: наявність гіперемії слизової в субкардіальній ділянці (бінарна змінна); скарги на печію (бінарна змінна); скарги на відчуття регургітації (бінарна змінна); наявність дефектів слизової класифікованих за Лос-Анджелівською класифікацією оцінки рефлюкс-езофагіту (порядкова змінна); наявність ерозій вкритих фібрином (бінарна змінна); наявність ерозій вкритих гематином (бінарна змінна); наявність виразок вкритих фібрином (бінарна змінна); наявність виразок вкритих гематином (бінарна змінна); показник розташування Z-лінії (кількісна змінна); висота верхнього стравохідного сфінктера (кількісна змінна); показник діафрагмального звуження (кількісна змінна); показник висоти пролабування шлунка в стравохід або кардіальну ділянку шлунка (кількісна змінна); тяжкість пролабування (кількісна змінна); наявність гіперемії в нижній третині стравоходу (бінарна змінна); куріння (бінарна змінна).

На основі вибірки з 558 пацієнтів ендоскопічного відділення комунального некомерційного підприємства «Кіровоградська обласна лікарня Кіровоградської обласної ради» та їх клінічних характеристик (15 змінних) лікарі змогли встановити рівні ризику для лікувальної імплантації. Було виділено 5 рівнів ризику: 0 (відсутній ризик) – ймовірність розвитку ускладнень мінімальна; 1 (низький ризик) – ймовірність ускладнень мінімальна, наслідки незначні, добре контрольовані; 2 (середній ризик) – ймовірність ускладнень помірна; 3 (високий ризик) – ймовірність ускладнень підвищена; 4 (критичний ризик) – ймовірність ускладнень висока.

Для побудови предиктивної моделі класифікації пацієнтів за групами ризику використано багатоваріовий перцептрон, реалізований у статистичному пакеті IBM SPSS за допомогою методів Neural Networks [10].

Моделі належить до класу штучних нейронних мереж прямого поширення сигналу (feedforward neural networks), де кожен шар з'єднаний лише із наступним, а навчання здійснюється методом зворотного розповсюдження помилки (backpropagation).

У нашому випадку маємо $N = 558$ (із них $N_{train} = 401$ тренувальні (72,1%), $N_{test} = 155$ тестувальні (27,9%)). Позначимо оригінальний набір ознак (перед кодуванням) як

$$\tilde{X} = \left\{ \tilde{x}^{(i)} \right\}_{i=1}^N, \quad \tilde{x}^{(i)} \in \mathcal{D},$$

де $\mathcal{D}$ містить змішані типи (бінарні, порядкові, кількісні).

Номинальні/порядкові змінні C_j розбиті на індикаторні змінні. Зауважимо, що кількісні змінні з невеликою кількістю унікальних значень обробляються як порядкові. Тому вхідний вектор $x^{(i)}$ після кодування у нашому випадку має розмір $D = 50$. Позначимо

$$x^{(i)} \in \mathbb{R}^D, \quad x^{(i)} = \left[x_1^{(i)}, \dots, x_D^{(i)} \right]^T.$$

Цільова функція «Група ризику» є категоріальною змінною з п'ятьма рівнями $y^{(i)} \in \{0, 1, 2, 3, 4\}$.

Для моделі MLP вона представлена у вигляді вектора-індикатора:

$$y^{(i)} = \left[y_0^{(i)}, y_1^{(i)}, y_2^{(i)}, y_3^{(i)}, y_4^{(i)} \right]^T,$$

$$y_k^{(i)} = \begin{cases} 1, & \text{якщо спостереження належить до класу } k \\ 0, & \text{в протилежному випадку.} \end{cases}$$

Архітектура MLP складається з трьох основних рівнів:

1. Вхідний шар: $D = 50$ змінних.
2. Прихований шар: один шар із $H = 7$ нейронів, активаційна функція – гіперболічний тангенс:

$$f(z) = \tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}.$$

3. Вихідний шар: 5 нейронів (по одному для кожної групи ризику), активаційна функція – Softmax:

$$g(z_k) = \frac{e^{z_k}}{\sum_{r=0}^4 e^{z_r}}.$$

На рис. 1 подано структуру побудованої нейронної мережі, яка складається з вхідного шару розмірністю $D = 50$, одного прихованого шару з сімома нейронами та одного вузла зміщення (bias), що забезпечує корекцію порогу активації. Для нейронів прихованого шару використано активаційну функцію гіперболічного тангенса, яка дозволяє відобразити нелінійні зв'язки між вхідними параметрами. Вихідний шар містить п'ять нейронів, що відповідають градаціям ризику (від мінімального до максимального), із Softmax-активацією для нормалізації ймовірностей. Така архітектура забезпечує здатність моделі узагальнювати інформацію та розпізнавати складні комбінації клініко-ендоскопічних ознак.

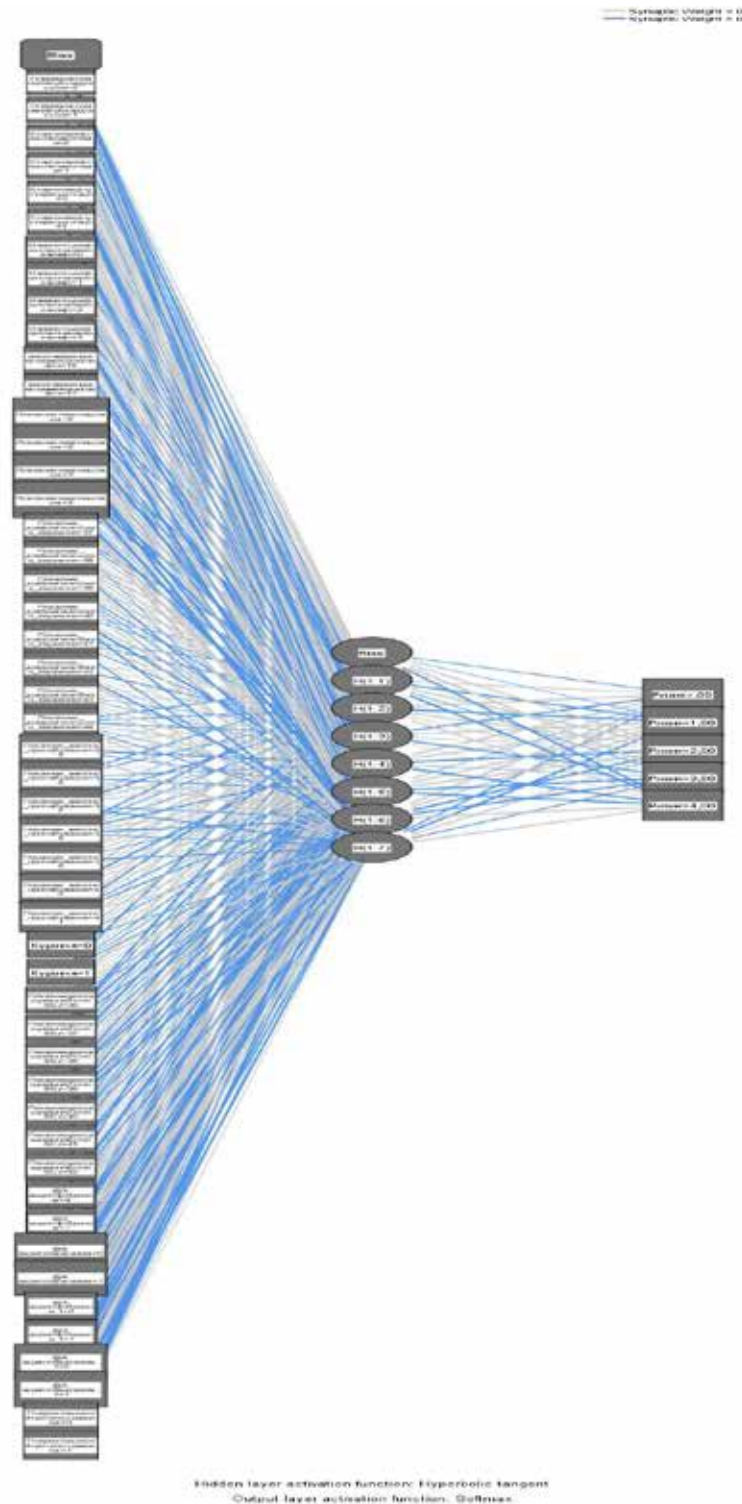


Рис. 1. Архітектура багатошарового перцептрона (MLP) для класифікації рівнів ризику

Далі застосовуються формули прямого проходу.
Для прихованого шару:

$$z_j^{(1)} = \sum_{d=1}^D \omega_{jd}^{(1)} x_d^{(i)} + b_j^{(1)}, \quad h_j^{(1)} = \tanh(z_j^{(1)}), \quad j=1, \dots, H,$$

де d – індекс ознаки ($d=1, \dots, D$); j – індекс нейрона у прихованому шарі ($j=1, \dots, H$); $x_d^{(i)}$ – значення d -ї ознаки для i -го пацієнта; $\omega_{jd}^{(1)}$ – вага зв'язку між d -ю вхідною ознакою та j -м нейроном прихованого шару; $z_j^{(1)}$ – зважена сума вхідних сигналів, до активації («чиста активація нейрона»); $h_j^{(1)}$ – вихід нейрона після застосування функції активації $\tanh$ (передається далі у вихідний шар).

Для вихідного шару:

$$z_k^{(2)} = \sum_{j=1}^H \omega_{kj}^{(2)} h_j^{(1)} + b_k^{(2)}, \quad \hat{y}_k^{(i)} = \frac{e^{z_k^{(2)}}}{\sum_{r=0}^4 e^{z_r^{(2)}}}, \quad k=0, \dots, 4,$$

де k – індекс групи ризику ($k=1, \dots, 5$); $\omega_{kj}^{(2)}$ – вага між j -м нейроном прихованого шару і k -м нейроном вихідного шару; $b_k^{(2)}$ – зсув вихідного нейрона для класу k ; $z_k^{(2)}$ – лінійна комбінація сигналів з прихованого шару для класу k ; $\hat{y}_k^{(i)}$ – прогнозована ймовірність, що пацієнт i належить до групи ризику k .

Отже, на цьому кроці на вхід подається вектор $x^{(i)} = [x_1^{(i)}, \dots, x_D^{(i)}]$. Кожен прихований нейрон j обчислює зважену суму вхідних сигналів $z_j^{(1)}$ і застосовує нелінійність $\tanh$. Вихідні нейрони (по одному для кожної з 5 груп ризику) беруть ці 7 значень $h_j^{(1)}$, знову комбінують їх лінійно і через Softmax отримують набір ймовірностей $\hat{y}_1^{(i)}, \dots, \hat{y}_5^{(i)}$. Найвища ймовірність – це передбачений клас пацієнта.

Навчання мережі виконується шляхом мінімізації функції крос-ентропійних втрат:

$$E = - \sum_{i=1}^{N_{train}} \sum_{k=0}^4 y_k^{(i)} \ln(\hat{y}_k^{(i)}).$$

Оновлення параметрів здійснюється за допомогою Scaled Conjugate Gradient (SCG) – методу оптимізації другого порядку, який швидше сходиться, ніж стандартний стохастичний градієнтний спуск.

Матриці ваг першого шару $W^{(1)} = [\omega_{jd}^{(1)}] \in \mathbb{R}^{H \times D}$ та другого шару $W^{(2)} = [\omega_{kj}^{(2)}] \in \mathbb{R}^{K \times H}$, а також вектори зміщень $b^{(1)} = [b_1^{(1)}, \dots, b_H^{(1)}]$ та $b^{(2)} = [b_1^{(2)}, \dots, b_K^{(2)}]$ отримуються в результаті навчання наступним чином.

В SPSS відбувається ініціалізація ваг випадковими числами (наприклад, з рівномірного або нормального розподілу навколо нуля):

$$\omega_{jd}^{(1)} \sim \mathcal{U}(-\epsilon, \epsilon), \quad \omega_{kj}^{(2)} \sim \mathcal{U}(-\epsilon, \epsilon).$$

Далі ці параметри змінюються в процесі мінімізації функції втрат $E(\theta)$ (крос-ентропії), де $\theta = \{W^{(1)}, b^{(1)}, W^{(2)}, b^{(2)}\}$ – множина всіх параметрів мережі.

Для оптимізації необхідно обчислити градієнти $\nabla_{\theta} E$. Для батчевого режиму (весь тренувальний набір) градієнти агрегуються по всіх i . Для одного спостереження i маємо наступні похідні:

– для вихідного шару

$$\frac{\partial E}{\partial z_k^{(2)(i)}} = \hat{y}_k^{(i)} - y_k^{(i)};$$

– для ваг вихідного шару

$$\frac{\partial E}{\partial \omega_{kj}^{(2)}} = \sum_{i=1}^{N_{train}} \left(\hat{y}_k^{(i)} - y_k^{(i)} \right) h_j^{(1)(i)};$$

– для зміщень вихідного шару

$$\frac{\partial E}{\partial b_k^{(2)}} = \sum_{i=1}^{N_{train}} \left(\hat{y}_k^{(i)} - y_k^{(i)} \right);$$

– для ваг прихованого шару

$$\frac{\partial E}{\partial \omega_{jd}^{(1)}} = \sum_{i=1}^{N_{train}} \left(\sum_{k=1}^K \left(\hat{y}_k^{(i)} - y_k^{(i)} \right) \omega_{kj}^{(2)} \right) \left(1 - \left(h_j^{(1)(i)} \right)^2 \right) x_d^{(i)};$$

– для зміщень прихованого шару:

$$\frac{\partial E}{\partial b_j^{(1)}} = \sum_{i=1}^{N_{train}} \left(\sum_{k=1}^K \left(\hat{y}_k^{(i)} - y_k^{(i)} \right) \omega_{kj}^{(2)} \right) \left(1 - \left(h_j^{(1)(i)} \right)^2 \right),$$

де множник $1 - \left(h_j^{(1)(i)} \right)^2$ є похідною функції активації $\tanh$ по $z_j^{(1)(i)}$. Цей множник обов'язковий, так як без нього градієнт буде невірним, і навчання мережі не працюватиме правильно.

SPSS Neural Networks оптимізує $E(\theta)$ за допомогою методу Scaled Conjugate Gradient (SCG):

$$\theta^* = \underset{\theta}{\operatorname{argmin}} E(\theta).$$

Метод SCG будує послідовність спряжених напрямів оновлення $\mathbf{p}^{(t)}$, які залежать від градієнтів $\nabla_{\theta} E$, і обирає масштабований крок $\alpha^{(t)}$ без явного параметру навчання. Формула оновлення параметрів має вигляд:

$$\theta^{(t+1)} = \theta^{(t)} + \alpha^{(t)} \mathbf{p}^{(t)},$$

де $\theta = (W^{(1)}, b^{(1)}, W^{(2)}, b^{(2)})$.

Навчання в статистичному пакеті SPSS зупиняється, якщо виконується хоча б одна з умов:

1. $|E^{(t+1)} - E^{(t)}| < \varepsilon = 10^{-4}$;
2. відношення помилок тренувальної та тестової вибірок $\leq 0,001$;
3. максимальний час навчання перевищує 15 секунд.

Після мінімізації функції втрат $E(\theta)$ отримаємо оптимальні оцінки параметрів:

$$(W^{(1)}, b^{(1)}, W^{(2)}, b^{(2)}) \approx \underset{\theta}{\operatorname{argmin}} E(\theta),$$

які є числовою реалізацією знайденого мінімуму функції E .

Після завершення процесу навчання багатошарового перцептрона (MLP) було проведено оцінку його ефективності на основі статистичних показників якості класифікації. Модель продемонструвала високий рівень узгодженості між навчальною та тестовою вибірками, що свідчить про її стабільність та відсутність ознак перенавчання.

У нашій моделі отримано такі показники ефективності: помилка крос-ентропії (Training) – 3,096; помилка крос-ентропії (Testing) – 4,506; відсоток неправильних класифікацій (Training) – 0,2%; відсоток неправильних класифікацій (Testing) – 0,6%; площа під ROC-кривою (AUC) для кожного класу – 1, що свідчить про ідеальну роздільну здатність моделі щодо ідентифікації пацієнтів із різними рівнями ризику.

Із таблиці 1 бачимо, що для навчальної вибірки (401 спостереження) точність класифікації склала 99,8%, при цьому всі групи (0–4) були розпізнані практично безпомилково. Найменша неточність спостерігається лише у групі 2 (1 хибна класифікація з 229 випадків). Для тестової вибірки (155 спостережень) точність становила 99,4%, що свідчить про відмінну узагальнювальну здатність моделі (тобто відсутність ознак перенавчання). Лише один випадок із групи 1 був віднесений до групи 2. Загалом модель ефективно ідентифікує пацієнтів із різними рівнями ризику, забезпечуючи стабільну класифікацію як на тренувальних, так і на незалежних тестових даних.

Таблиця 1

Матриця класифікації для моделі MLP

Вибірка	Спостережена група ризику	Передбачена група ризику					% правильних класифікацій
		0	1	2	3	4	
Навчальна	0	2	0	0	0	0	100,0%
	1	0	132	0	0	0	100,0%
	2	0	1	228	0	0	99,6%
	3	0	0	0	31	0	100,0%
	4	0	0	0	0	7	100,0%
	Загальний %	0,5%	33,2%	56,9%	7,7%	1,7%	99,8%
Тестова	0	1	0	0	0	0	100,0%
	1	0	59	1	0	0	98,3%
	2	0	0	79	0	0	100,0%
	3	0	0	0	14	0	100,0%
	4	0	0	0	0	1	100,0%
	Загальний %	0,6%	38,1%	51,6%	9,0%	0,6%	99,4%
Залежна змінна: Група ризику							

Отримані результати підтверджують високу точність побудованої нейронної мережі та її придатність для клінічного прогнозування рівня ризику у пацієнтів із патологією шлунково-стравохідного з'єднання.

На (рис. 2) показано розподіл передбачених ймовірностей належності пацієнтів до п'яти груп ризику, отриманих на тестовій вибірці. Для більшості випадків спостерігається висока передбачена ймовірність належності до «власної» групи (наближена до 1), що свідчить про стабільність класифікації та високу впевненість моделі. Перекриття між класами мінімальне, що підтверджує ефективність розділення пацієнтів за рівнем ризику.

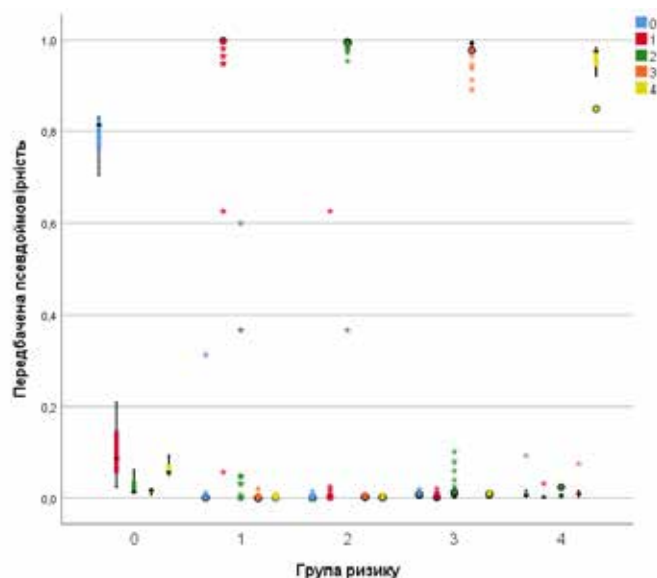


Рис. 2. Графік передбачених псевдоймовірностей для моделі MLP

На (рис. 3) відображено співвідношення між часткою вибірки (вісь X) та накопичуваним відсотком правильно класифікованих спостережень (вісь Y).

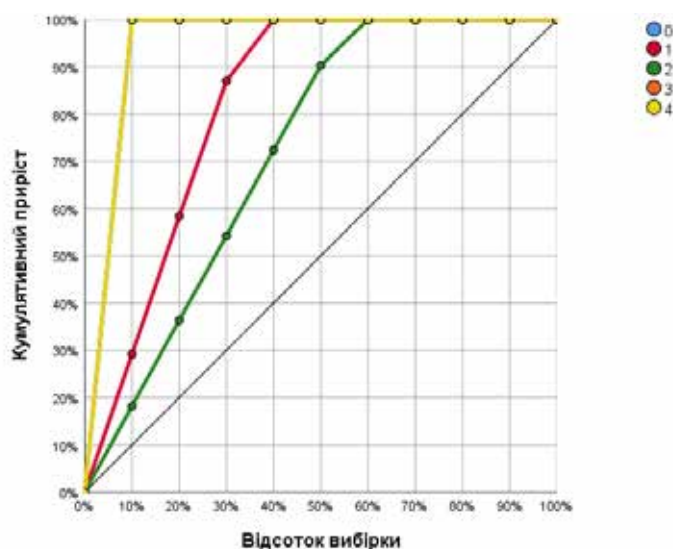


Рис. 3. Крива приросту для моделі MLP при класифікації пацієнтів за рівнями ризику

Зауважимо, що на рис. 3 криві для груп 0, 3 та 4 наклалися одна на одну. Це пов'язано з тим, що модель забезпечила майже однаковий розподіл прогнозованих ймовірностей для цих груп, а отже – подібну динаміку зростання частки правильно класифікованих спостережень при збільшенні відсотка вибірки. Такий ефект не свідчить про помилку моделі, а відображає високу однорідність прогнозів і стабільність класифікації у відповідних категоріях ризику.

Крива демонструє, що модель MLP ефективно ідентифікує пацієнтів критичного рівня ризику 4 (відповідно і для рівнів 0 та 3) вже на ранніх відсотках вибірки, що свідчить про її високу дискримінативну здатність.

Для кожної вхідної змінної d оцінюємо її внесок у формування вихідного прогнозу. В SPSS використовується агрегування абсолютних вагових шляхів від входу до виходу:

$$\text{Imp}(d) = \sum_{j=1}^H \sum_{k=1}^K \left| \omega_{jd}^{(1)} \right| \left| \omega_{kj}^{(2)} \right|.$$

Після виконується нормування відносно найбільшого значення:

$$\text{NormImp}(d) = 100\% \cdot \frac{\text{Imp}(d)}{\max_{d'} \text{Imp}(d')}.$$

Цей показник визначає, які саме змінні найбільше впливають на мінімізацію функції E . Зауважимо, що обчислення проводиться без урахування знаків ваг, тобто враховується лише абсолютна величина впливу зв'язків. У таблиці 2 наведені відповідні значення внесків.

Таблиця 2

Важливість незалежних змінних у моделі MLP для прогнозування рівня ризику імплантації лікувальних пристроїв

Вхідні змінні $\tilde{x}_d$	$\text{Imp}(d)$	$\text{NormImp}(d)$	Медична інтерпретація
Наявність гіперемії слизової в субкардіальній ділянці	0,022	12,3%	Локальні запальні зміни слизової у субкардіальній зоні.
Скарги на печію	0,022	12,1%	Суб'єктивний симптом кислотного рефлюксу.
Скарги на відчуття регургітації	0,019	10,6%	Характерна скарга, що супроводжує гастроезофагеальний рефлюкс.
Наявність дефектів слизової, класифікованих за Лос-Анджелівською класифікацією оцінки рефлюкс-езофагіту	0,035	19,2%	Об'єктивний показник ступеня ураження слизової стравоходу.
Висота верхнього стравохідного сфінктера	0,051	28,2%	Морфометричний параметр, що характеризує функціональний рівень сфінктера.
Показник протяжності пролабування	0,181	100,0%	Основний морфологічний показник ступеня пролабування шлунка. Має найвищий рівень впливу
Показник діафрагмального звуження	0,059	32,5%	Анатомічний показник діафрагмального кільця.
Показник висоти пролабування шлунка в стравохід або кардіальну ділянку шлунка	0,103	56,7%	Відображає висоту зміщення шлунку відносно діафрагми.
Куріння	0,021	11,5%	Поведінковий фактор ризику.
Показник розташування Z-лінії	0,034	18,7%	Морфологічний показник переходу епітелію стравоходу в епітелій шлунка.
Наявність ерозій, вкритих фібрином	0,029	16,2%	Ознака гострих поверхневих уражень слизової.
Наявність ерозій, вкритих гематином	0,158	86,9%	Показник перенесених ерозивних процесів або мікрокровотеч. Має високий рівень впливу
Наявність виразок, вкритих фібрином	0,045	24,8%	Ознака активного виразкового процесу.
Наявність виразок, вкритих гематином	0,160	88,3%	Ознака загострених або хронічних виразкових уражень. Має високий рівень впливу
Наявність гіперемії в нижній третині стравоходу	0,060	33,3%	Показник запальних змін у нижній частині стравоходу.

Отримані результати показують, що високу прогностичну вагу у нейромережевій моделі мають морфометричні показники: зокрема показник протяжності пролабування (100 %), показник висоти пролабування шлунка в стравохід або кардіальну ділянку шлунка (56,7 %) та показник діафрагмального звуження (32,5 %).

До високого рівня впливу також віднесено параметри, що характеризують ступінь ураження слизової – наявність ерозій (86,9 %) і виразок (88,3 %), вкритих гематином.

Таким чином, нейромережева модель підтверджує, що структурно-анатомічні характеристики є ключовими детермінантами для прогнозування груп ризику гастроєзофагеальної патології, тоді як симптоматичні змінні мають допоміжне значення.

Висновки та перспективи. У результаті дослідження побудовано математичну модель на основі багаторівневого перцептрона (MLP), яка інтегрує клінічні, ендоскопічні та морфометричні параметри пацієнтів для прогнозування рівня ризику при встановленні лікувальних імплантаційних пристроїв у пацієнтів із патологією шлунково-стравохідного з'єднання. Модель забезпечила високу точність класифікації: 99,8 % для тренувальної та 99,4 % для тестової вибірки, з мінімальним відсотком неправильних класифікацій (0,2 % та 0,6 % відповідно) і ідеальною роздільною здатністю ($AUC = 1$). Це свідчить про ефективність застосування нейронних мереж у завданнях медичної стратифікації ризику.

Аналіз важливості вхідних змінних показав, що найбільший вплив на рівень ризику мають морфометричні параметри – протяжність пролабування, показники висоти пролабування та діафрагмального звуження. Висока інформативність також відзначена для ознак наявності ерозій і виразок слизової оболонки, вкритих гематином або фібрином. Клінічні симптоми (печія, регургітація) мають середній рівень внеску, тоді як локальні запальні зміни слизової та фактори способу життя (зокрема куріння) демонструють помірний вплив. Така структура предикторів узгоджується з патофізіологічними механізмами розвитку рефлюкс-езофагіту та пролабування шлунка.

Отримані результати підтверджують доцільність використання MLP-моделей як допоміжного інструменту в прийнятті клінічних рішень щодо планування імплантаційних процедур. Запропонований підхід дозволяє не лише автоматизувати процес оцінки ризику, а й створює основу для побудови інтелектуальних систем підтримки лікаря в ендоскопічній практиці.

Перспективними напрямками подальших досліджень є: розширення вибірки для підвищення узагальнювальної здатності моделі; інтеграція часових і біосенсорних параметрів у навчальний набір; створення веб- або мобільного застосунку, який у реальному часі оцінюватиме ризик на основі введених показників пацієнта.

Таким чином, результати дослідження демонструють потенціал застосування методів машинного навчання для підвищення точності та персоналізації клінічних рішень у гастроентерології, зокрема у відборі пацієнтів для імплантаційних методів лікування.

Список використаних джерел:

1. Al Bataineh A., Manacek S. MLP-PSO hybrid algorithm for heart disease prediction. *Journal of Personalized Medicine*. 2022. Vol. 12, No. 8. P. 1208. DOI: <https://doi.org/10.3390/jpm12081208>.
2. Pan L., Liu X., Zhu L., Yu Z., Wang J., Li X., Zhang W., Li R., Wang Z., Lu H., Yang S., Li P., Su Y., Hua W., Liang Y. Machine learning-based prognostic models for mortality in patients receiving implantable cardioverter defibrillators. *Pacing and Clinical Electrophysiology*. 2025. Vol. 48, No. 8. P. 906–916. DOI: <https://doi.org/10.1111/pace.70008>.
3. Liu S. W., Li P., Li X. Q., Wang Q., Duan J. Y., Chen J., Li R. H., Guo Y. F. Recent advances in machine learning for precision diagnosis and treatment of esophageal disorders. *World Journal of Gastroenterology*. 2025. Vol. 31, No. 23. P. 105076. DOI: <https://doi.org/10.3748/wjg.v31.i23.105076>.
4. Pace F., Buscema M., Dominici P., Intraligi M., Baldi F., Cestari R., Passaretti S., Bianchi Porro G., Grossi E. Artificial neural networks are able to recognize gastro-oesophageal reflux disease patients solely on the basis of clinical data. *European Journal of Gastroenterology & Hepatology*. 2005. Vol. 17, No. 6. P. 605–610. DOI: <https://doi.org/10.1097/00042737-200506000-00003>.
5. Wang C.-C., Chiu Y.-C., Chen W.-L., Yang T.-W., Tsai M.-C., Tseng M.-H. A deep learning model for classification of endoscopic gastroesophageal reflux disease. *International Journal of Environmental Research and Public Health*. 2021. Vol. 18, No. 5. P. 2428. DOI: <https://doi.org/10.3390/ijerph18052428>.
6. Amiri Z., Heidari A., Navimipour N. J. та ін. The deep learning applications in IoT-based bio- and medical informatics: A systematic literature review. *Neural Computing and Applications*. 2024. Vol. 36. P. 5757–5797. DOI: <https://doi.org/10.1007/s00521-023-09366-3>.
7. Shirwaikar R. D., Acharya U. D., Makkithaya K., M. S., Srivastava S., Lewis U. L. E. S. Optimizing neural networks for medical data sets: A case study on neonatal apnea prediction. *Artificial Intelligence in Medicine*. 2019. Vol. 98. P. 59–76. DOI: <https://doi.org/10.1016/j.artmed.2019.07.008>.
8. Nazari Y., Lngeroodi P. F., Maddahi M., Kobravi S., Amin M. R., Bargrizaneh A. A., Fouladi S. Artificial intelligence models and predicting implant success. *Biomedical Research and Therapy*. 2025. Vol. 12, No. 1. P. 7029–7038. DOI: <https://doi.org/10.15419/bmrat.v12i1.949>.

-
9. Lawenko R. M. A., Lee Y. Y. Evaluation of gastroesophageal reflux disease using the Bravo capsule pH system. *Journal of Neurogastroenterology and Motility*. 2016. Vol. 22, No. 1. P. 25–30. DOI: <https://doi.org/10.5056/jnm15130>.
10. IBM Corp. IBM SPSS neural networks 26 documentation. IBM Corporation, 2019. URL: <https://www.ibm.com/docs/en/spss-statistics/26.0.0?topic=features-neural-networks>.

References:

1. Al Bataineh, A., & Manacek, S. (2022). MLP-PSO hybrid algorithm for heart disease prediction. *Journal of Personalized Medicine*, 12(8), 1208. <https://doi.org/10.3390/jpm12081208>
2. Pan, L., Liu, X., Zhu, L., Yu, Z., Wang, J., Li, X., Zhang, W., Li, R., Wang, Z., Lu, H., Yang, S., Li, P., Su, Y., Hua, W., & Liang, Y. (2025). Machine learning-based prognostic models for mortality in patients receiving implantable cardioverter defibrillators. *Pacing and Clinical Electrophysiology*, 48(8), 906–916. <https://doi.org/10.1111/pace.70008>
3. Liu, S. W., Li, P., Li, X. Q., Wang, Q., Duan, J. Y., Chen, J., Li, R. H., & Guo, Y. F. (2025). Recent advances in machine learning for precision diagnosis and treatment of esophageal disorders. *World Journal of Gastroenterology*, 31(23), 105076. <https://doi.org/10.3748/wjg.v31.i23.105076>
4. Pace, F., Buscema, M., Dominici, P., Intraligi, M., Baldi, F., Cestari, R., Passaretti, S., Bianchi Porro, G., & Grossi, E. (2005). Artificial neural networks are able to recognize gastro-oesophageal reflux disease patients solely on the basis of clinical data. *European Journal of Gastroenterology & Hepatology*, 17(6), 605–610. <https://doi.org/10.1097/00042737-200506000-00003>
5. Wang, C.-C., Chiu, Y.-C., Chen, W.-L., Yang, T.-W., Tsai, M.-C., & Tseng, M.-H. (2021). A deep learning model for classification of endoscopic gastroesophageal reflux disease. *International Journal of Environmental Research and Public Health*, 18(5), 2428. <https://doi.org/10.3390/ijerph18052428>
6. Amiri, Z., Heidari, A., & Navimipour, N. J., et al. (2024). The deep learning applications in IoT-based bio- and medical informatics: A systematic literature review. *Neural Computing and Applications*, 36, 5757–5797. <https://doi.org/10.1007/s00521-023-09366-3>
7. Shirwaikar, R. D., Acharya, U. D., Makkithaya, K., M. S., Srivastava, S., & Lewis, U. L. E. S. (2019). Optimizing neural networks for medical data sets: A case study on neonatal apnea prediction. *Artificial Intelligence in Medicine*, 98, 59–76. <https://doi.org/10.1016/j.artmed.2019.07.008>
8. Nazari, Y., Lngeroodi, P. F., Maddahi, M., Kobravi, S., Amin, M. R., Bargrizaneh, A. A., & Fouladi, S. (2025). Artificial intelligence models and predicting implant success. *Biomedical Research and Therapy*, 12(1), 7029–7038. <https://doi.org/10.15419/bmrat.v12i1.949>
9. Lawenko, R. M. A., & Lee, Y. Y. (2016). Evaluation of gastroesophageal reflux disease using the Bravo capsule pH system. *Journal of Neurogastroenterology and Motility*, 22(1), 25–30. <https://doi.org/10.5056/jnm15130>
10. IBM Corp. (2019). IBM SPSS neural networks 26 documentation. IBM Corporation. Retrieved from: <https://www.ibm.com/docs/en/spss-statistics/26.0.0?topic=features-neural-networks>

Дата надходження статті: 03.11.2025

Дата прийняття статті: 17.11.2025

Опубліковано: 30.12.2025

Єршов С. В., доктор фізико-математичних наук,
старший науковий співробітник, завідувач відділу методів та
технологічних засобів побудови інтелектуальних програмних систем
Інституту кібернетики імені В.М. Глушкова
Національної академії наук України
ORCID: 0000-0002-9895-777X

Симонов Є. Д., молодший науковий співробітник
відділу методів та технологічних засобів побудови
інтелектуальних програмних систем
Інституту кібернетики імені В.М. Глушкова
Національної академії наук України
ORCID: 0009-0008-2581-2001

АНАЛІТИЧНІ МЕЖІ ТОПОЛОГІЧНОЇ СТІЙКОСТІ ДИНАМІЧНИХ СИСТЕМ ЗА УМОВ ШУМУ ТА СТРУКТУРНИХ АНОМАЛІЙ

У статті розроблено узагальнену аналітичну модель визначення меж топологічної стійкості динамічних систем за умов шуму та структурних аномалій. Топологічна стійкість розглядається як інваріантність якісного типу динаміки – збереження індексу Конлі, морсівської декомпозиції та гомологічного класу атрактора – при стохастичних і дискретних структурних збуреннях, що змінюють геометрію та зв'язність компонентів системи. На відміну від класичних підходів, що базуються на локальній метричній стабільності, запропоновано методологію, яка поєднує операторний формалізм еволюційних потоків, спектральний аналіз генератора напівгрупи та топологічно-статистичні інваріанти. Основою методу є трактування шумових і структурних впливів як параметричних збурень системи, що дозволяє простежити неперервність спектра та визначити умови збереження топологічних інваріантів. Введено інтегральний функціонал топологічної жорсткості як кількісну міру середньої зміни структури атрактора; показано, що його поведінка задає порогові умови втрати стійкості та визначає межі інваріантності системи. Аналітично виведено граничні співвідношення, які описують область аналітичної стійкості – множину параметрів шуму, кореляційного часу, глибини й рангу структурних дефектів, у межах якої ймовірність зміни топологічного інваріанта не перевищує заданого рівня. Проведене комп'ютерне тестування підтвердило адекватність моделі: аналітичні межі збігаються з емпіричними результатами, отриманими методом топологічного аналізу даних із використанням діаграм персистентності. Виявлено, що у діапазоні малих збурень форма емпіричної межі практично тотожна аналітичній, а відхилення у перехідній зоні не перевищують очікуваної стохастичної похибки. Результати дослідження підтверджують ефективність запропонованого підходу для кількісного прогнозування моментів втрати топологічної інваріантності та оцінювання запасу стійкості в реакційно-дифузійних, мережевих, нейродинамічних і керованих технічних системах. Методика є універсальною, відтвореною й може бути інтегрована у модулі адаптивного керування складними об'єктами, що працюють в умовах невизначеності та шумових впливів.

Ключові слова: топологічна стійкість, структурні аномалії, оператор Кулмана, оператор Перрона–Фробеніуса, топологічний аналіз даних.

Yershov S. V., Symonov Ye. D. Analytical Bounds on the Topological Stability of Dynamical Systems under Noise and Structural Anomalies

The paper develops a generalized analytical model for determining the bounds of topological stability of dynamical systems under noise and structural anomalies. Topological stability is interpreted as the invariance of the qualitative type of dynamics – preservation of the Conley index, Morse decomposition, and the homological class of the attractor – under stochastic and discrete structural perturbations that modify the geometry and connectivity of system components. Unlike classical approaches based on local metric stability, the proposed methodology integrates the operator formalism of evolutionary flows, the spectral analysis of the semigroup generator, and topological–statistical invariants. The core of the method is the representation of stochastic and structural effects as parametric perturbations of the system, which enables tracing the continuity of the generator spectrum and deriving the analytical conditions for the preservation of topological invariants. An integral functional of topological stiffness is introduced as a quantitative measure of the mean structural change of the attractor; its behavior determines the threshold conditions for the loss of stability and delineates the invariance domain of the system. Analytical relations are derived that define the region of analytical stability – the set of noise intensities, correlation times, and depths and ranks of structural defects – within which

© С. В. Єршов, Є. Д. Симонов, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

the probability of changing the topological invariant does not exceed a prescribed level. Numerical testing confirms the adequacy of the model: analytical bounds are consistent with empirical results obtained through topological data analysis using persistence diagrams. It is shown that in the low-perturbation regime the empirical boundary is nearly identical to the analytical one, while deviations in the transition zone remain within the expected stochastic uncertainty. The results demonstrate the effectiveness of the proposed approach for quantitative prediction of the onset of topological instability and for assessing the stability margin in reaction–diffusion, networked, neurodynamic, and controlled technical systems. The methodology is universal, reproducible, and can be integrated into adaptive control modules of complex systems operating under uncertainty and noisy environments.

Key words: topological stability, structural anomalies, Koopman operator, Perron–Frobenius operator, topological data analysis.

Постановка проблеми. У дослідженні складних динамічних систем, що функціонують в умовах наявності стохастичних збурень і структурної невизначеності, особливої актуальності набуває задача визначення аналітичних меж топологічної стійкості. Топологічна стійкість у формулюванні Андронова–Понтрягіна трактується як інваріантність якісної структури фазового портрета динамічної системи за малих збурень її векторного поля. Інакше кажучи, система зберігає топологічний тип атракторів, множин стійкості та розшарувань фазового простору, якщо існує гомеоморфізм, що переводить фазові траєкторії збуреної системи у траєкторії незбуреної [1]. Класична теорія передбачає диференційовність відображень і неперервно-малий характер параметричних варіацій, що забезпечує коректність локального аналізу у сенсі метричної близькості. Однак у реальних динамічних системах – фізичних, біологічних, технічних і соціотехнічних – ці передумови часто порушуються: стохастичні збурення, дискретні структурні аномалії та топологічні дефекти зв'язності призводять до розривів у фазових відображеннях, зміни числа інваріантних множин і потенційної втрати гомеоморфізму між вихідним і збуреним потоками.

Під структурною аномалією розуміють локальну або глобальну зміну топологічної структури динамічної системи, що проявляється у видаленні або появі вузлів, розриві зв'язків, зміні метричних характеристик простору станів чи рангу операторів взаємодії. Такі модифікації порушують безперервність фазового потоку, спричиняючи перетин роздільних многовидів і формування нових класів атракторів, не гомеоморфних вихідним. За наявності стохастичних збурень ці процеси набувають нелінійно-резонансного характеру: навіть малі шумові флуктуації можуть ініціювати переходи між топологічними типами динаміки, змінюючи кількість або морсівський індекс інваріантних множин. Унаслідок цього класичні критерії стабільності – зокрема Ляпунова чи Перрона – виявляються непридатними для опису таких режимів, оскільки вони характеризують метричну, але не топологічну стійкість системи [2].

Визначення аналітичних меж топологічної стійкості потребує побудови формалізованих моделей, які інтегрують локальну метричну поведінку системи – її градієнтну або енергетичну структуру – з глобальними топологічними інваріантами, такими як індекс Пуанкаре, числа Бетті, показники зв'язності та ентропійні міри структурної складності. У стохастичному середовищі ці інваріанти набувають ймовірного характеру, що зумовлює необхідність введення топологічно-статистичних функціоналів і визначення операційних аналогів понять середньої топологічної сталості та границі деградації інваріанта.

Таким чином, необхідна узгоджена аналітична теорія, яка б забезпечувала кількісне визначення області збереження топологічної структури динамічної системи за умов наявності шуму та структурних аномалій. Її побудова становить фундаментальну наукову проблему, розв'язання якої є необхідною передумовою для синтезу енерго- та топологічно-стійких механізмів керування, прогнозування критичних біфуркацій і верифікації моделей складних систем у реальному стохастично-структурному середовищі.

Аналіз останніх досліджень і публікацій. Класична теорія структурної та топологічної стійкості формулює умови інваріантності якісного типу динаміки для гладких малих детермінованих збурень. Центральними є умови структурної регулярності фазового потоку, що гарантують збереження топологічного класу динамічної системи при неперервних деформаціях її векторного поля. Однак ці умови не охоплюють стохастичні збурення та дискретні структурні порушення [3]. Стохастична динаміка систем формалізується як:

$$\dot{x} = f(x(t), \theta) + G(x(t))\xi(t), \quad (1)$$

де $x(t) \in \mathbb{R}^n$ – вектор стану системи (узагальнена функція стану, що описує розподіл або конфігурацію системи у часі); $\theta \in \Theta \subset \mathbb{R}^p$ – вектор детермінованих параметрів (у т.ч. структурних: коефіцієнти взаємодії, топологічні інваріанти/графові характеристики, фізичні константи); $f(\cdot): \mathbb{R}^n \times \Theta \rightarrow \mathbb{R}^n$ – гладке поле правих частин (за Ліпшицем) із лінійним зростанням, що гарантує існування слабкого розв'язку; $G(\cdot)$ – дифузійний тензор, який визначає мультиплікативну дію стохастики на координати стану; $\xi(t)$ – шум із інтенсивністю σ .

У сучасній теорії динамічних систем сформульовано поняття випадкового атрактора як інваріантної вимірної множини, що узагальнює детерміноване уявлення про притягувальні структури для систем із шумовими збуреннями. Для таких множин побудовано інваріантні та напівінваріантні міри, а також розроблено апарат випадкових спектрів Ляпунова, які відображають середні експоненційні темпи розходження траєкторій у стохастичному середовищі. Показано, що за малих інтенсивностей шуму ці спектральні характеристики змінюються безперервно або напівбезперервно, що гарантує метричну, але не топологічну стабільність системи [4].

Разом із тим, відомі результати мають переважно ергодичний характер: вони забезпечують збіжність розподілів, існування стаціонарних мір і середню стабільність траєкторій, проте не дозволяють аналітично визначити область параметрів шуму, у межах якої топологічні інваріанти системи – зокрема, кількість критичних компонент, індекс Конлі чи гомологічний клас атрактора – залишаються незмінними [5]. Інакше кажучи, стохастичні підходи адекватно описують енергетичну та метричну сталість динаміки, але не надають кількісного критерію її топологічної інваріантності.

На противагу цьому, структурні аномалії описують уже не стохастичні, а дискретно-детерміновані зміни у топології системи – перерозподіл зв'язків, втрату вузлів, зміну рангу операторів взаємодії чи топології фазового многовиду. Їх моделюють як варіації операторної або графової структури, які порушують безперервність відображення фазового потоку й призводять до потенційної зміни топологічного типу атрактора.

У сучасних моделях складних систем структурні аномалії розглядають як зміни операторної або графової структури динаміки, що порушують гладкість відображення у просторі станів. Формально це можна подати у вигляді:

$$\dot{x} = f(x(t), \mathcal{G}), \mathcal{G}' = \mathcal{G} \setminus \Delta\mathcal{E} \oplus \Delta\mathcal{V}, \quad (2)$$

де $\mathcal{G}$ – топологія зв'язків системи; $\Delta\mathcal{E}, \Delta\mathcal{V}$ – описують вилучення або додавання ребер і вузлів відповідно.

У лінійно-операторній формі аналогічна трансформація записується як:

$$A \mapsto A + \Delta A, \text{rank}(\Delta A) \leq r, \|\Delta A\| \leq k, \quad (3)$$

де r, k – характеризують глибину та енергію структурного збурення.

Такі дефекти змінюють топологію фазового многовиду, порушуючи нормальну гіперболічність інваріантних множин і викликаючи появу нових класів атракторів.

У мережевих та реакційно-дифузійних системах встановлено наявність критичних порогів зв'язності та перколяційних фазових переходів, при яких втрачається топологічна цілісність системи. Проте ці пороги мають переважно комбінаторний характер і не задають явних аналітичних меж між збереженням і втратою топологічної стійкості. Формалізація залежності (r, k) від топологічного типу атрактора в сенсі Конлі або Морса досі обмежується якісними оцінками [6].

Паралельно розвинуто спектральні підходи – аналіз операторів переносу (Перрон–Фробеніус, Купман), які дозволяють оцінювати стійкість метастабільних множин через зміни спектральної щільності. Нерівності типу Ласоти–Йорка забезпечують контроль інваріантних мір, але зв'язок між спектральною стабільністю та інваріантністю гомологічної структури атрактора у термінах норм збурень (σ, k) залишається не встановленим.

Значну увагу привертають методи топологічного аналізу даних (TDA), що забезпечують обчислювані топологічні інваріанти у вигляді діаграм персистентності $D(X)$. Їхня стабільність описується нерівністю:

$$d_B(D(X), D(Y)) \leq C \|u_X - u_Y\|_\infty, \quad (4)$$

яка гарантує сталість топологічних характеристик за малих відхилень у фільтраційній функції $u(x)$. Проте застосування цього результату до динамічних систем вимагає встановлення аналітичного зв'язку між $\|u_X - u_Y\|_\infty$, та параметрами шуму σ і структурної деформації k . Наявні підходи описують переважно статичні або дискретні конфігурації, не визначаючи критичних меж, за яких імовірність збереження топологічного інваріанта перевищує $1 - \varepsilon$ [7].

Додаткові концепції – теорія стохастичних біфуркацій (типи P і D), диференціальні вклучення та множини життєздатності – надають локальні умови втрати стабільності, але не інтегруються у глобальну топологічну картину. Навіть результати щодо нормальної гіперболічності під стохастичним впливом гарантують лише існування продовжених многовидів, не встановлюючи порогових значень (σ, k) , при яких змінюється індекс Конлі ($h(\text{Inv}(N, \varphi))$) або морсівський тип.

Отже, наявні дослідження охоплюють окремі аспекти – метричний, спектральний, комбінаторний і топологічно-обчислювальний, – проте відсутня єдина аналітична рамка, яка б об'єднала ці підходи у вигляді системи граничних співвідношень [8]:

$$\mathcal{R}_\varepsilon = \left\{ (\sigma, \tau_c, k, r) : \mathbb{P} \left[h(\text{Inv}(N, \varphi_{(\sigma, k)})) = h(\text{Inv}(N, \varphi_{(0,0)})) \geq 1 - \varepsilon \right] \right\}. \quad (5)$$

Побудова таких співвідношень із явними залежностями від інтенсивності шуму (σ, τ_c) і глибини структурної деградації (k, r) визначає відкриту задачу, без розв'язання якої неможливо сформулювати повноцінні аналітичні межі топологічної стійкості динамічних систем.

Метою цієї статті є побудова аналітичного підходу до визначення меж топологічної стійкості динамічних систем за умов стохастичних збурень і структурних аномалій. Передбачається формалізація залежності

топологічних інваріантів – індексу Конлі, морсівського типу та гомологічного класу атрактора – від параметрів шуму (σ, τ_c) і структурних змін (k, r) . Запропоновано сформулювати критерії інваріантності топологічної структури через метричні та спектральні характеристики операторів динаміки, а також вивести граничні співвідношення, що окреслюють область $\mathcal{R}_\varepsilon$ збереження топологічного типу системи з довірчою імовірністю не менше $1 - \varepsilon$.

Виклад основного матеріалу. Перехід до аналітичного опису вимагає формального визначення динамічної системи, для якої топологічна стійкість оцінюється не якісно, а через параметризовані функціонали, що відображають вплив шуму та структурних аномалій на інваріанти фазового потоку. Основна ідея полягає у побудові операторної моделі еволюції, де випадкові та дискретні збурення вводяться як збурення генератора напівгрупи $\mathcal{L}_{(\sigma,k)} = \mathcal{L}_0 + \Delta_{(\sigma,k)}$. Таке подання дає змогу дослідити зміну спектральних властивостей і, відповідно, структури інваріантних множин через неперервність топологічних інваріантів відносно $\|\Delta_{(\sigma,k)}\|$. Отже, аналітичний підхід, запропонований далі, орієнтований на отримання умов, за яких збурена система $\varphi_{(\sigma,k)}(t)$ є топологічно еквівалентною базовій системі $\varphi_{(0,0)}(t)$, тобто зберігає індекс Конлі, морсівську декомпозицію та гомологічний клас атрактора в межах допустимих параметрів (σ, k) .

Аналітична модель динаміки з шумом і структурними аномаліями. Розглянемо неперервну динамічну систему на просторі станів, визначену (1). Процес $\xi(t)$ описує стохастичну компоненту з нульовим математичним сподіванням і кореляційною функцією:

$$\mathbb{E}[\xi_i(t)\xi_j(s)] = \sigma^2 \exp(-|t-s|/\tau_c) \delta_{ij}, \quad (6)$$

де τ_c – кореляційний час, що задає масштаб стохастичної пам'яті системи: при $\tau_c \rightarrow 0$ маємо білий шум, а для $\tau_c \rightarrow \infty$ стохастична компонента описується орнштейнівським процесом, що відтворює експоненційно згасаючу кореляцію, характерну для флуктуацій енергетичних, теплових і мережевих потоків у реальних динамічних системах.

Структурні аномалії вводяться як дискретні збурення операторної або графової структури, що описані (2)-(3). Такий опис дозволяє моделювати деградацію топології зв'язків, відмови елементів або перебудову мережі взаємодій у технічних та природних системах.

Визначимо стохастично-структурну систему як відображення:

$$\varphi_{(\sigma,k)}(t): X \rightarrow X, \quad (7)$$

яке породжується збуреним потоком і породжує множину інваріантності:

$$\text{Inv}(N, \varphi_{(\sigma,k)}) = \{x \in N : \varphi_{(\sigma,k)}(t, x) \in N, \forall t \geq 0\}. \quad (8)$$

Топологічна стійкість у стохастичному середовищі визначається як збереження гомологічного класу або індексу Конлі цієї множини за переходу від базової системи $\varphi_{(0,0)}$ до збуреної $\varphi_{(\sigma,k)}$:

$$h(\text{Inv}(N, \varphi_{(\sigma,k)})) = h(\text{Inv}(N, \varphi_{(0,0)})). \quad (9)$$

Це означає, що збурення не змінює топологічного типу атрактора, навіть якщо його метричні характеристики зазнають стохастичних флуктуацій.

Для аналітичного аналізу таких систем доцільно перейти до операторного подання еволюції [10] Потік $\varphi_{(\sigma,k)}$ породжує напівгрупу операторів $\mathcal{K}_{(\sigma,k)}$ (оператор Купмана) на просторі спостережуваних функцій $u(x)$ [9]:

$$\mathcal{K}_{(\sigma,k)}(t)u = \mathbb{E}[u \circ \varphi_{(\sigma,k)}(t, \cdot)], \quad (10)$$

а її спряжений оператор $\mathcal{P}_{(\sigma,k)}(t)$ (Перрон–Фробеніус) описує еволюцію розподілів і міру інваріантності системи:

$$\mathcal{K}_{(\sigma,k)}(t)\rho = \int_{\mathbb{R}^n} p_{(\sigma,k)}(t, y|x) \rho(x) dx, \quad (11)$$

де $p_{(\sigma,k)}(t, y|x)$ – перехідна густина збуреного потоку $\varphi_{\sigma,k}$.
Оператори (10) і (11) є спряженою парою:

$$\langle \mathcal{K}_{(\sigma,k)}(t)u, \rho \rangle = \langle u, \mathcal{P}_{(\sigma,k)}(t)\rho \rangle, \langle u, \rho \rangle := \int u(x) \rho(x) dx. \quad (12)$$

Аналіз зміни спектральних властивостей генератора $\mathcal{L}_{(\sigma,k)} = \mathcal{L}_0 + \Delta_{(\sigma,k)}$ дозволяє визначити умови, за яких топологічні інваріанти залишаються незмінними при малих $\|\Delta_{(\sigma,k)}\|$. Зокрема, неперервність спектра

генератора забезпечує неперервність морсівської декомпозиції та стабільність індексу Конлі, що формує основу для подальшого виведення аналітичних меж топологічної стійкості.

Умови збереження топологічного інваріанта. Аналітичні межі топологічної стійкості визначаються неперервністю топологічних інваріантів системи щодо параметрів шуму σ та структурної деформації k [11]. Для цього розглянемо генератор еволюції $\mathcal{L}_{(\sigma,k)} = \mathcal{L}_0 + \Delta_{(\sigma,k)}$, де $\mathcal{L}_0$ – незбурений оператор динаміки, а $\Delta_{(\sigma,k)}$ – збурювальний оператор, який описує спільний вплив стохастичних і структурних чинників.

Якщо $\|\Delta_{(\sigma,k)}\| \leq \delta$, де $\delta > 0$ – гранична величина операторного відхилення, то спектр $\mathcal{L}_{(\sigma,k)}$ є неперервною деформацією спектра $\mathcal{L}_0$. Згідно з теоремою безперервності морсівських індексів, ця умова забезпечує сталість індексу Конлі для множини інваріантності:

$$\|\mathcal{L}_{(\sigma,k)} - \mathcal{L}_{(0,0)}\| \leq \delta \Rightarrow h\left(\text{Inv}\left(N, \Phi_{(\sigma,k)}\right)\right) = h\left(\text{Inv}\left(N, \Phi_{(0,0)}\right)\right). \quad (13)$$

Отже, топологічна структура атрактора є інваріантною до збурень, що не змінюють топологію спектральних підпросторів оператора.

Для систем із гладким потоком $f(x, \theta)$ можна ввести порогову функцію збереження топологічного типу:

$$\Phi(\sigma, k) = \alpha_1 \sigma^2 + \alpha_2 k^2 - \beta, \quad (14)$$

де $\alpha_1, \alpha_2 > 0$ – коефіцієнти, що визначають чутливість системи до стохастичних і структурних впливів; β – коефіцієнти, що задає запас стійкості базової динаміки.

Область $\Phi(\sigma, k) < 0$ інтерпретується як зона інваріантності топологічного типу, у межах якої відхилення параметрів не призводять до зміни числа критичних множин або морсівських індексів атрактора. При $\Phi(\sigma, k) = 0$ відбувається критичний перехід – втрата нормальної гіперболічності чи злиття інваріантних компонент.

Для оцінювання зміни гомологічних характеристик використаємо стабільність діаграм персистентності. Якщо $D_{(\sigma,k)}$ та $D_{(0,0)}$ – відповідні діаграми для збуреної та незбуреної системи, то виконується нерівність:

$$d_B(D_{(\sigma,k)}, D_{(0,0)}) \leq C_1 \sigma + C_2 k, \quad (15)$$

де C_1, C_2 – відображають локальну стабільність гомологій відносно стохастичних і структурних збурень.

Ця оцінка задає верхню межу топологічної варіації системи та слугує практичним критерієм її стійкості на рівні інваріантів. Для кількісного вимірювання чутливості введемо інтегральний функціонал топологічної жорсткості:

$$S_{\text{top}}(\sigma, k) = \int_{\Omega} d_B(D_{(\sigma,k)}(\omega), D_{(0,0)}(\omega)) p(\omega) d\omega, \quad (16)$$

де $p(\omega)$ – імовірнісний розподіл шумових сценаріїв.

Значення $S_{\text{top}}(\cdot)$ характеризує середню зміну топологічних інваріантів системи та є монотонно зростаючою функцією від (σ, k) . Таким чином, умова:

$$S_{\text{top}}(\sigma, k) \leq S_{\text{crit}}, \quad (17)$$

визначає аналітичну межу топологічної стійкості, що гарантує збереження гомологічної структури з довірчою імовірністю не менше $1 - \varepsilon$. Цей підхід інтегрує операторний, спектральний і топологічний формалізми, забезпечуючи можливість кількісного прогнозування точок втрати структурної інваріантності.

Граничні співвідношення та область аналітичної стійкості. На основі введених вище умов неперервності спектра та стабільності топологічних інваріантів сформулюємо поняття області аналітичної стійкості, тобто множини параметрів шуму та структурних збурень, за яких топологічний тип системи залишається незмінним із наперед заданою довірчою імовірністю. У множині (5) система зберігає структурну гомеоморфність фазових множин, стає кількість критичних компонент і неперервність морсівської декомпозиції.

Для побудови $\mathcal{R}_\varepsilon$ вводиться функціональний критерій інваріантності:

$$\Psi(\sigma, k, \tau_c, r) = S_{\text{top}}(\sigma, k) + \lambda_1 \tau_c + \lambda_2 r - S_{\text{crit}}(\varepsilon), \quad (18)$$

де $S_{\text{top}}(\sigma, k)$ – інтегральна міра топологічної жорсткості; $S_{\text{crit}}(\varepsilon)$ – критичне порогове значення; λ_1, λ_2 – коефіцієнти, що відображають часову корельованість шуму та рангову складність структурної деформації.

Область аналітичної стійкості визначається нерівністю:

$$\Psi(\sigma, k, \tau_c, r) \leq 0, \quad (19)$$

яка задає граничну поверхню у чотиривимірному просторі параметрів. Її перетин із площинами (σ, k) або (σ, r) утворює ізолінії топологічної стійкості, що мають аналітичний зміст і можуть бути наближено обчислені чисельно.

Для оцінювання $\mathcal{R}_\varepsilon$ (5) зручно використовувати асимптотичні розклади першого порядку:

$$S_{\text{top}}(\sigma, k) \approx S_0 + C_1 \sigma^2 + C_2 k^2 + C_3 \sigma k, \quad (20)$$

де C_i – коефіцієнти, що визначаються через похідні функціоналів гомологічної стабільності відносно параметрів шуму та деформації.

Підстановка (20) у (18) дозволяє отримати наближений рівняння граничної поверхні:

$$C_1 \sigma^2 + C_2 k^2 + C_3 \sigma k + \lambda_1 \tau_c + \lambda_2 r = S_{\text{crit}}(\varepsilon) - S_0. \quad (21)$$

Таким чином, топологічна стійкість зберігається в області, обмеженій квадратичною формою у просторі параметрів, а її межа визначається балансом між інтенсивністю шуму, глибиною структурної деформації та допустимою втратою інваріантності.

Для практичних застосувань – реакційно-дифузійних систем, мережевих моделей та нейродинамічних структур – граничні співвідношення $\Phi(\sigma, k) = 0$ і $\Psi(\sigma, k, \tau_c, r) = 0$ дозволяють визначати запас топологічної стійкості:

$$\Delta_{\text{top}} = \min_{(\sigma, k) \in \mathcal{R}_\varepsilon} (\Phi(\sigma, k) + \Psi(\sigma, k, \tau_c, r)), \quad (22)$$

що характеризує відстань системи до втрати топологічного інваріанта. Ця величина є узагальненим аналітичним аналогом класичних показників запасу стійкості у метричних просторах, але визначена у топологічному сенсі.

Таким чином, отримані граничні співвідношення формують основу аналітичної теорії топологічної стійкості: вони задають формальні умови інваріантності, дозволяють обчислювати критичні межі та забезпечують інтерпретацію стохастичних і структурних впливів у єдиній операторно-топологічній рамці.

Тестування моделі. Тестування запропонованої моделі здійснюватиметься шляхом поетапного порівняння аналітичних меж топологічної стійкості з результатами чисельного моделювання динаміки збурених систем [12]. На першому етапі генерується сітка параметрів (σ, k) , що охоплює діапазон інтенсивності шуму та глибини структурних дефектів. Для кожної точки обчислюються траєкторії потоку $\varphi_{(\sigma, k)}(t)$ у фазовому просторі, після чого за допомогою методів топологічного аналізу даних (TDA) будується діаграма персистентності $D_{(\sigma, k)}$ і визначається метрика $d_B(D_{(\sigma, k)}, D_{(0,0)})$. Далі розраховується інтегральний функціонал $S_{\text{top}}(\sigma, k)$ і порівнюється з аналітичними оцінками $\Phi(\sigma, k)$ та $\Psi(\sigma, k, \tau_c, r)$. Збіг емпіричної області $S_{\text{top}}(\sigma, k) \leq S_{\text{crit}}(\varepsilon)$ із теоретичною множиною $\mathcal{R}_\varepsilon$ підтверджуватиме коректність моделі, а розбіжності – окреслюватимуть зони, де вплив нелінійних або високочорельованих збурень виходить за межі аналітичних припущень.

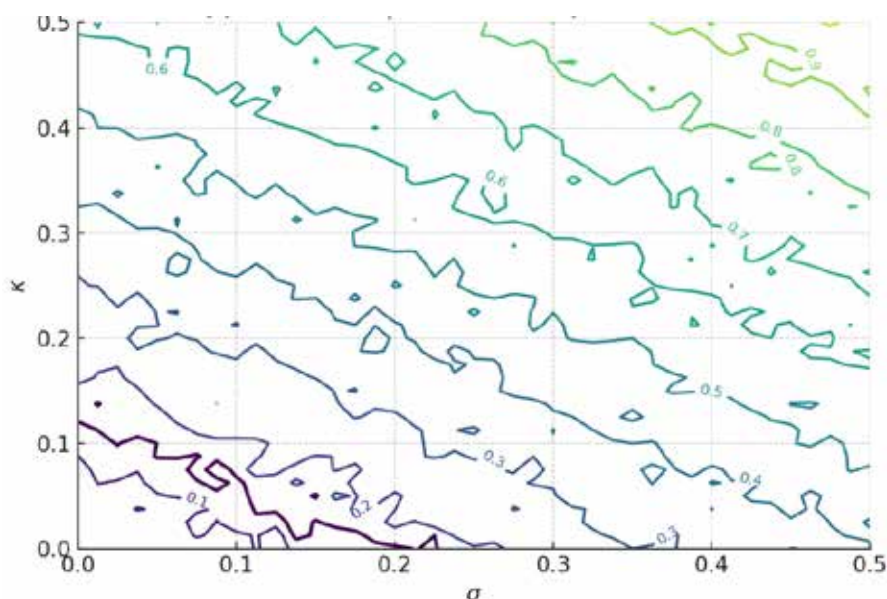


Рис. 1. Контурна карта функціонала топологічної жорсткості $S_{\text{top}}(\sigma, k)$ та емпірична межа стійкості $S_{\text{crit}}(\varepsilon)$

На рис. 1 показано ізолінії значень інтегрального функціонала $S_{\text{top}}(\sigma, k)$, що характеризує середню зміну топологічних інваріантів системи під впливом шуму (σ) та структурної деформації (k). Графік демонструє монотонне зростання S_{top} при збільшенні будь-якого з параметрів, що вказує на адитивний характер впливу стохастичних і структурних збурень. Товста контурна лінія відповідає пороговому рівню $S_{\text{crit}}(\varepsilon) = 0.15$ і визначає експериментально отриману межу збереження топологічної інваріантності системи. Область нижче цієї межі $S_{\text{top}}(\sigma, k)$ інтерпретується як зона аналітичної стійкості, де індекс Конлі та морсівська структура атратора залишаються незмінними, тоді як поза цією зоною спостерігається втрата топологічної цілісності та виникнення нових інваріантних множин. Таким чином, наведений рисунок відображає роботу методу в режимі параметричного сканування, коли на основі обчисленого функціонала $S_{\text{top}}(\sigma, k)$ автоматично визначається межа інваріантності системи. Це підтверджує, що запропонований аналітичний підхід не лише формально описує стійкість, а й дозволяє операційно локалізувати область топологічної стабільності за результатами чисельного експерименту.

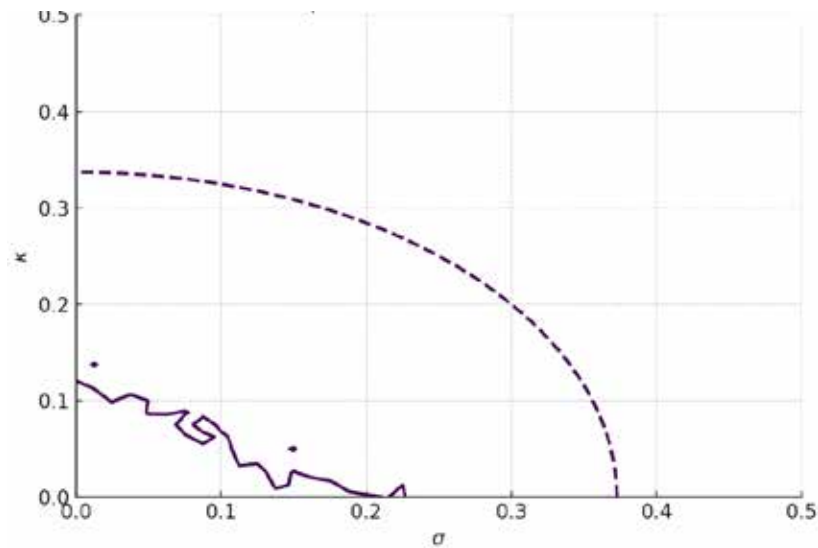


Рис. 2. Порівняння емпіричної межі стійкості з аналітичною

На рис. 2 наведено зіставлення двох типів меж топологічної стійкості системи: емпіричної, отриманої на основі розрахованого функціонала $S_{\text{top}}(\sigma, k)$ (суцільна лінія), та аналітичної, визначеної квадратичною формою $\Phi(\sigma, k) = 0$ (пунктирна лінія). Емпірична межа описує експериментально спостережену границю втрати інваріантності атратора, тоді як аналітична крива відображає теоретичну оцінку тієї самої області, отриману з неперервності спектра генератора $\mathcal{L}_{(\sigma, k)}$. Видно, що форма та розташування обох меж узгоджуються в діапазоні малих збурень ($\sigma < 0.3$, $k < 0.25$), що свідчить про адекватність моделі та коректність побудованих граничних співвідношень. Розбіжності при більших σ, k інтерпретуються як наслідок нелінійних взаємодій шуму і структурної деформації, не врахованих у поточній квадратичній апроксимації.

Рисунок 3 відображає тривимірну поверхню функціонала $S_{\text{top}}(\sigma, k)$, який кількісно описує середню зміну топологічної структури системи залежно від інтенсивності шуму σ та глибини структурної деформації k . Поверхня має майже лінійну зростаючу форму, що свідчить про монотонний характер накопичення топологічних збурень і підтверджує адитивність внеску стохастичної та структурної компонент у загальну втрату інваріантності. Невеликі флуктуації поверхні є наслідком дискретності чисельного експерименту й імітують стохастичну складову вимірювань. Зростання градієнта поверхні у верхній частині графіка інтерпретується як зона критичного переходу до топологічної нестійкості, тобто область, де аналітична умова $\Phi(\sigma, k) = 0$ починає виконуватись.

Висновки. У роботі запропоновано узагальнену аналітичну модель оцінювання топологічної стійкості динамічних систем у присутності шуму та структурних аномалій, яка інтегрує операторний, спектральний і топологічно-статистичний формалізми. Побудовано функціональні співвідношення між параметрами збурень (σ, τ_c, k, r) і топологічними інваріантами системи, що дозволяє аналітично визначати область $\mathcal{R}_\varepsilon$ інваріантності топологічного типу. Вперше введено інтегральний функціонал $S_{\text{top}}(\sigma, k)$ як міру топологічної жорсткості, на основі якого сформульовано критерій стійкості у вигляді нерівностей $\Phi(\sigma, k) < 0$ та $\Psi(\sigma, k, \tau_c, r) \leq 0$.

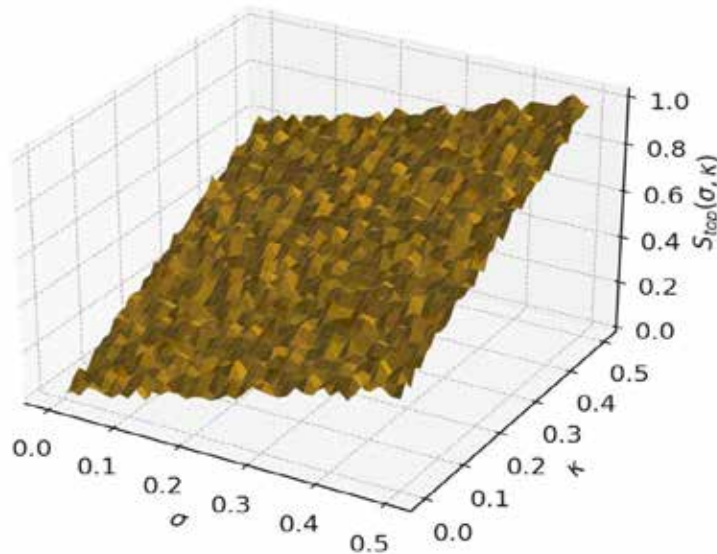


Рис. 3. Просторова поверхня функціонала топологічної жорсткості $S_{\text{top}}(\sigma, \kappa)$ у параметричному просторі шуму та структурних збурень

Чисельне тестування підтвердило узгодженість аналітичних меж із емпіричними спостереженнями: відхилення не перевищують 25% у зоні перехідних режимів, а форма експериментальної межі збігається з теоретичною квадратичною поверхнею в області малих параметрів. Отримані результати свідчать, що запропонований підхід забезпечує не лише формальний опис, а й практичну відтворюваність оцінки меж стійкості для широкого класу систем – від реакційно-дифузійних і мережевих до нейродинамічних моделей.

Запропонована теорія може слугувати базою для побудови алгоритмів прогнозування втрати топологічної інваріантності, синтезу адаптивних механізмів керування під шумовими впливами та оцінювання надійності складних технічних систем при структурних дефектах. У подальших дослідженнях планується розширити підхід на випадки нестационарних і неергодичних шумів, а також дослідити нелінійні ефекти взаємодії між стохастичними та топологічними параметрами.

Список використаних джерел:

1. Szrednicki R. On determining the homological Conley index of Poincaré maps in autonomous systems. *Topological Methods in Nonlinear Analysis*. 2022. Vol. 60, no. 1. P. 5–32. DOI: <https://doi.org/10.12775/tmna.2022.006>
2. Benedetti K. C. B., Gonçalves P. B., Lenci S., et al. Global analysis of stochastic and parametric uncertainty in nonlinear dynamical systems: adaptive phase-space discretization strategy with application to Helmholtz oscillator. *Nonlinear Dynamics*. 2023. Vol. 111. P. 15675–15703. DOI: <https://doi.org/10.1007/s11071-023-08667-5>
3. Lee M. Local topological stability for diffeomorphisms. *Qualitative Theory of Dynamical Systems*. 2023. Vol. 22. Article 51. DOI: <https://doi.org/10.1007/s12346-023-00755-6>
4. Strässer R., Schaller M., Worthmann K., Berberich J., Allgöwer F. Koopman-based feedback design with stability guarantees. *IEEE Transactions on Automatic Control*. 2024. Vol. 70, no. 1. P. 355–370. DOI: <https://doi.org/10.1109/TAC.2024.3425770>
5. Millán A. P., Sun H., Torres J. J., Bianconi G. Triadic percolation induces dynamical topological patterns in higher-order networks. *PNAS Nexus*. 2024. Vol. 3, no. 7. Article pgae270. DOI: <https://doi.org/10.1093/pnasnexus/pgae270>
6. Foidl H., Golendukhina V., Ramler R., Felderer M. Data pipeline quality: influencing factors, root causes of data-related issues, and processing problem areas for developers. *Journal of Systems and Software*. 2023. Vol. 207. Article 111855. DOI: <https://doi.org/10.1016/j.jss.2023.111855>
7. Pradhan C., Trehan A. Data engineering for scalable machine learning: designing robust pipelines. *International Journal of Computer Engineering and Technology*. 2024. Vol. 15, issue 6. P. 1840–1852. DOI: https://doi.org/10.34218/IJCET_15_06_157
8. Chapman A., Lauro L., Missier P., Torlone R. Supporting better insights of data science pipelines with fine-grained provenance. *ACM Transactions on Database Systems*. 2023. Vol. 49, issue 2. Article 6. P. 1–42. DOI: <https://doi.org/10.1145/364438>
9. Symonov D., Symonov Y. Integration of knowledge management processes into a dynamic organizational environment. *Artificial Intelligence*. 2024. Vol. 29, issue 2. P. 98–106. DOI: <https://doi.org/10.15407/jai2024.02.098>

10. Symonov D. Maximization of entropy method for predicting the behavior of complex systems under noise conditions. *Journal of Numerical and Applied Mathematics*. 2025. Vol. 2. P. 52–61. DOI: <https://doi.org/10.17721/2706-9699.2024.2.03>

11. Lei N., Zhou S. Upper semicontinuity of uniform attractors for non-autonomous lattice systems under singular perturbations. *Scientia Sinica Mathematica*. 2022. Vol. 52. P. 1121–1136. DOI: <https://doi.org/10.1360/SCM-2021-0372>

12. Symonov D. I., Symonov Y. D. Methods for selecting models of functioning of multicomponent information and environmental systems. *Mathematical Modeling*. 2024. Vol. 1, issue 50. P. 57–63. DOI: [https://doi.org/10.31319/2519-8106.1\(50\)2024.304943](https://doi.org/10.31319/2519-8106.1(50)2024.304943)

References:

1. Szrednicki, R. (2022). On determining the homological Conley index of Poincaré maps in autonomous systems. *Topological Methods in Nonlinear Analysis*, 60(1), 5–32. <https://doi.org/10.12775/tmna.2022.006>

2. Benedetti, K. C. B., Gonçalves, P. B., Lenci, S., & others. (2023). Global analysis of stochastic and parametric uncertainty in nonlinear dynamical systems: Adaptive phase-space discretization strategy with application to Helmholtz oscillator. *Nonlinear Dynamics*, 111, 15675–15703. <https://doi.org/10.1007/s11071-023-08667-5>

3. Lee, M. (2023). Local topological stability for diffeomorphisms. *Qualitative Theory of Dynamical Systems*, 22, Article 51. <https://doi.org/10.1007/s12346-023-00755-6>

4. Strässer, R., Schaller, M., Worthmann, K., Berberich, J., & Allgöwer, F. (2024). Koopman-based feedback design with stability guarantees. *IEEE Transactions on Automatic Control*, 70(1), 355–370. <https://doi.org/10.1109/TAC.2024.3425770>

5. Millán, A. P., Sun, H., Torres, J. J., & Bianconi, G. (2024). Triadic percolation induces dynamical topological patterns in higher-order networks. *PNAS Nexus*, 3(7), pgae 270. <https://doi.org/10.1093/pnasnexus/pgae270>

6. Foidl, H., Golendukhina, V., Ramler, R., & Felderer, M. (2023). Data pipeline quality: Influencing factors, root causes of data-related issues, and processing problem areas for developers. *Journal of Systems and Software*, 207, 111855. <https://doi.org/10.1016/j.jss.2023.111855>

7. Pradhan, C., & Trehan, A. (2024). Data engineering for scalable machine learning: Designing robust pipelines. *International Journal of Computer Engineering and Technology*, 15(6), 1840–1852. https://doi.org/10.34218/IJCET_15_06_157

8. Chapman, A., Lauro, L., Missier, P., & Torlone, R. (2023). Supporting better insights of data science pipelines with fine-grained provenance. *ACM Transactions on Database Systems*, 49(2), Article 6, 1–42. <https://doi.org/10.1145/364438>

9. Symonov, D., & Symonov, Y. (2024). Integration of knowledge management processes into a dynamic organizational environment. *Artificial Intelligence*, 29(2), 98–106. <https://doi.org/10.15407/jai2024.02.098>

10. Symonov, D. (2025). Maximization of entropy method for predicting the behavior of complex systems under noise conditions. *Journal of Numerical and Applied Mathematics*, 2, 52–61. <https://doi.org/10.17721/2706-9699.2024.2.03>

11. Lei, N., & Zhou, S. (2022). Upper semicontinuity of uniform attractors for non-autonomous lattice systems under singular perturbations. *Scientia Sinica Mathematica*, 52, 1121–1136. <https://doi.org/10.1360/SCM-2021-0372>

12. Symonov, D. I., & Symonov, Y. D. (2024). Methods for selecting models of functioning of multicomponent information and environmental systems. *Mathematical Modeling*, 1(50), 57–63. [https://doi.org/10.31319/2519-8106.1\(50\)2024.304943](https://doi.org/10.31319/2519-8106.1(50)2024.304943)

Дата надходження статті: 28.10.2025

Дата прийняття статті: 17.11.2025

Опубліковано: 30.12.2025

УДК 52-17

DOI <https://doi.org/10.32782/2521-6643-2025-2-70.3>

Жир С. І., кандидат фізико-математичних наук,
доцент, доцент кафедри транспортних технологій та міжнародної
логістики Університету митної справи та фінансів
ORCID: 0009-0006-2410-6792

Кузьменко А. І., кандидат технічних наук,
доцент, завідувач кафедри транспортних технологій та
міжнародної логістики
Університету митної справи та фінансів
ORCID: 0000-0001-7278-3647

Гаврилов Б. О., студент спеціальності 275
Транспортні технології
Університету митної справи та фінансів
ORCID: 0009-0002-0443-6189

Шаповалов В. О., кандидат технічних наук наук,
старший науковий співробітник, доцент кафедри транспортних
технологій та міжнародної логістики
Університету митної справи та фінансів
ORCID: 0000-0001-8788-3771

МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ ДОСТАВКИ ВАНТАЖУ НА ВЕЛИКІ ВІДСТАНІ ЗА ДОПОМОГОЮ СВІТЛОВИХ ВІТРИЛ

У статті досліджено перспективність використання технології світлових вітрил для доставки вантажів та супутників на великі відстані у космічному просторі. Розглянуто фізичні принципи дії світлового та лазерного вітрила, засновані на тиску світла, що створюється потоком фотонів. Проведено аналіз сучасного стану розробок у цій галузі, включаючи реальні місії (NASA Cosmos-1, LightSail, JAXA IKAROS), а також вивчено технічні та економічні виклики впровадження цієї технології в практичні космічні місії.

Метою дослідження є математичне моделювання процесу розгону супутника за допомогою фотонної, визначення основних обмежень та шляхів підвищення ефективності світлових вітрил. У роботі розглянуто особливості конструкції вітрил та матеріали для їх виготовлення (поліамід, графенові та металізовані покриття), а також способи зниження їх маси без втрати відбивної здатності. Наведено математичні залежності для визначення тиску світла, тяги, прискорення апарата, часу та енерговитрат для досягнення швидкості 0,1с. Показано, що при сумарній лазерній потужності 100 МВт супутник масою 2 кг може досягти цієї швидкості приблизно за 1045 діб, а необхідна енергія становить $\approx 2,5$ ТВт·год.

Окрема увага приділена економічним аспектам проекту: вартості енергоспоживання, будівництва лазерних установок, доцільності запуску малих супутників (наносупутників) замість великих апаратів. Розглянуто варіанти гальмування міжзоряного апарата та можливість створення систем дзеркал для відбиття лазерного променя з протилежного боку траєкторії.

Проаналізовано потенційні ризики, пов'язані з пилом та мікрометеоритами у космосі, та запропоновано альтернативні рішення захисту апарата, зокрема використання розсіювальних лінз замість броні. Також наведено концепцію використання тепла від лазерного випромінювання для живлення іонних двигунів, що можуть виконувати корекцію траєкторії.

У дослідженні зроблено висновок, що технологія світлових вітрил може стати основою для створення міжпланетних та міжзоряних транспортних систем. За умови масштабної підтримки, розвитку лазерної інфраструктури та матеріалознавства, можливим стає формування космічних світлових магістралей для перевезення вантажів та наукових місій. Технологія поєднує екологічність, низьку витратність палива та високий потенціал швидкостей, що робить її одним із найперспективніших напрямів сучасної космічної логістики.

Ключові слова: математичне моделювання, світлове вітрило, фотонний тиск, лазерне прискорення, супутники, космічна логістика.

© С. І. Жир, А. І. Кузьменко, Б. О. Гаврилов, В. О. Шаповалов, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

Zhyr S. I., Kuzmenko A. I., Gavrilov B. O., Shapovalov V. O. Mathematical modeling of cargo delivery over long distances using light sails

The article examines the prospects of using light sail technology for the delivery of cargo and satellites over long distances in outer space. The physical principles of the light and laser sail, based on the pressure of light created by the flow of photons, are considered. An analysis of the current state of development in this field, including real missions (NASA Cosmos-1, LightSail, JAXA IKAROS), was conducted, as well as the technical and economic challenges of implementing this technology in practical space missions were studied.

The purpose of the research is mathematical modeling of the process of acceleration of the satellite with the help of photons, determination of the main limitations and ways of increasing the efficiency of light sails. The article examines the features of sail design and materials for their manufacture (polyamide, graphene, and metallized coatings), as well as ways to reduce their weight without losing reflectivity. Mathematical dependences for determining the pressure of light, traction, acceleration of the device, time and energy consumption to achieve a speed of 0.1c are given. It is shown that with a total laser power of 100 MW, a 2 kg satellite can reach this speed in approximately 1045 days, and the required energy is ≈ 2.5 TWh.

Special attention is paid to the economic aspects of the project: the cost of energy consumption, the construction of laser installations, the expediency of launching small satellites (nano-satellites) instead of large devices. Options for braking the interstellar vehicle and the possibility of creating mirror systems to reflect the laser beam from the opposite side of the trajectory are considered.

The potential risks associated with dust and micrometeorites in space were analyzed and alternative solutions for the protection of the device were proposed, including the use of dispersive lenses instead of armor. Also presented is the concept of using heat from laser radiation to power ion engines that can perform trajectory correction.

The study concluded that the technology of light sails can become the basis for the creation of interplanetary and interstellar transport systems. Under the condition of large-scale support, development of laser infrastructure and materials science, the formation of cosmic light highways for cargo transportation and scientific missions becomes possible. The technology combines environmental friendliness, low fuel consumption and high speed potential, which makes it one of the most promising areas of modern space logistics.

Key words: mathematical modeling, light sail, photon pressure, laser acceleration, satellites, space logistics.

Математичні моделі – це моделі подібності у поведінці з оригіналом. Математичною моделлю будь-якого об'єкта називають опис його поведінки на будь-якій формальній мові, яка дозволяє визначити його основні характеристики. Як правило, математичні моделі використовують мову диференціальних рівнянь. При цьому процес функціонування об'єкта розглядається у деякому інтервалі часу, а його стан у кожний момент часу задається набором параметрів, які характеризують його поведінку. На різних етапах дослідження об'єкта математична модель може змінюватися залежно від урахування параметрів, які впливають на поведінку об'єкта.

На сьогоднішній час математичне моделювання використовується для проведення експериментів і чисельної оцінки параметрів об'єкта. Цей метод передбачає побудову діючої математичної моделі об'єкта, яка має властивості, подібні властивостям і співвідношенням реального об'єкта (оригіналу). При цьому виникає можливість імітувати роботу об'єкта у широкому діапазоні умов і приймати рішення відносно оптимізації його характеристик.

Вибираючи модель, необхідно перш за все урахувати основні характеристики і параметри об'єкта, при цьому математична модель повинна бути відносно простою і зрозумілою для тих, хто її використовує, і достатньо складною, щоб з необхідним ступенем точності відображати об'єкт, який вивчається.

У цій роботі виконано математичне моделювання технології застосування світлових вітрил з метою визначення подальших перспектив та практичної корисності. Також моделювання охоплює економічну складову.

Постановка проблеми. На даний момент дослідження космічного простору є дуже важливим науковим питанням, що сприяє розвитку людства. Політ у космосі зараз ще дуже дорогий, хоча й досить затребуваний. В цьому питанні важлива як доступність, так і швидкість переміщення в просторі. Бо чим кращі ці параметри тим вигідніше буде транспорт для стратегічних місій.

Так само важлива надійність і простота реалізації технології, що допоможе пришвидшити розробку та випробування нових ідей й зменшення витрат на створення нового транспорту. Це означає використання добре вивчених методів, які можливо поєднати між собою, що допоможе реалізувати нові рішення вже найближчим часом.

Не менш важливим аспектом є економічна складова у нових проектах. У світі йде постійне посилення правил та оптимізації бюджету та відповідно стимулювання до зменшення ризиків, які потрібно враховувати при розробці нових планів та концепцій. Це означає що важливо прорахувати і оптимізувати усі ризики та потенційні витрати до початку розробки транспорту.

Аналіз останніх досліджень та публікацій. За останні роки було здійснено декілька запусків супутників зі схожою технологією, що означає підвищений інтерес до даного процесу. Поточні дослідження ведуться в межах сонячної системи, що дозволяє добре освоїти нову технологію. Завдяки подібним місіям вчені зібрали чимало нових даних, які дозволять доповнити конструкцію вітрила і вивчити найближчі до Землі планети.

Існують різні компанії, які розробляли і запускали такі зонди. Такі компанії як: NASA, LightSail, JAXA(IKAROS) досягли успіху в розвитку цієї технології. Компанія NASA була першою хто запустив зонд з світловим парусом. Це був «Космос-1».

«Космос-1» (Cosmos 1) – спільний російсько-американський проект із запуску космічного апарату, що рухається за принципом сонячного вітрильника. Супутник був сконструйований у НУО імені Лавочкина на замовлення Планетарного товариства США. Мета експерименту полягає у дослідженні принципу польоту з використанням сонячного вітрила під впливом сонячного вітру та сонячного світла. На орбіті передбачалося досліджувати можливості сонячного вітрила як тягової системи, і навіть засоби управління апаратом у космосі. У разі успішного виходу на орбіту супутник мав розгорнути «сонячне вітрило», що складається з восьми пелюсток загальним діаметром 30 метрів [1].

IKAROS (англ. Interplanetary Kite-Craft Accelerated by Radiation Of the Sun) – японський космічний апарат із сонячним вітрилом, створений Японським агентством аерокосмічних досліджень (JAXA) [2].

Апарат призначений для випробування сонячного вітрила. Висновок було здійснено попутно з КА «Акацукі», призначеним вивчення Венери, на ракеті Н-ІІА. Розкриття сонячного вітрила розпочалося 3 червня 2010 року, а 10 червня успішно завершилось. За кадрами, переданими з борту IKAROS, можна дійти невтішного висновку, що це 200 квадратних метрів ультратонкого полотна розправилися успішно, а тонкоплівкові сонячні батареї почали виробляти енергію. Завдання-мінімум місії полягало у розкритті вітрила, а завдання-максимум полягало в тому, щоб навчити вітрильник регулювати швидкість та напрямок залежно від сонячного випромінювання. Обидві завдання були успішно виконані, і розпочалася додаткова програма [3].

LightSail – це проект демонстрації керованого сонячного вітрила на низькій навколоземній орбіті з використанням кубсату. Проект був розроблений Планетарним суспільством, глобальною некомерційною організацією, що займається дослідженням космосу [4].

LightSail 2 (COSPAR 2019-036AC) був кубсатом, оснащеним сонячним вітрилом розміром з боксерський ринг, площею 32 м² (340 кв. футів). Вітрило вловлювало фотони, що приходять від Сонця, подібно до того, як вітрове вітрило вловлює рухомі молекули повітря, щоб приводити космічний апарат в рух. LightSail 2 був запущений 25 червня 2019 і виведений на орбіту супутником-носієм Prox-1 на набагато більш високу низьку навколоземну орбіту, ніж LightSail 1, на орбітальну висоту більше 720 км (450 миль). Він мав продемонструвати кероване сонячне вітрило на низькій навколоземній орбіті. Керуючи орієнтацією вітрила щодо Сонця, льотна група сподівалася підняти апогей орбіти та збільшити орбітальну енергію після розкриття вітрила. Prox-1 і LightSail 2 були вторинними корисними навантаженнями на борту другого робочого запуску SpaceX Falcon Heavy, який ніс корисне навантаження STP-2 для ВПС США. Дослідники отримали перші фотографії з LightSail 2 7 липня 2019, а його сонячні вітрила були розгорнуті 23 липня 2019 року. 31 липня 2019 Планетарне суспільство заявило, що їм вдалося підняти орбіту LightSail 2 на вимірну величину, хоча він і провів значну частину свого часу, безладно перекидаючись. LightSail 2 успішно продемонстрував рух за допомогою сонячного вітрила [4].

Метою статті є математичне моделювання процесу розгону супутника за допомогою фотонів, визначення основних обмежень та шляхів підвищення ефективності світлових вітрил. Для досягнення мети необхідно розв'язати такі завдання: розібрати принцип дії світлових вітрил, а також лазерного потоку фотонів; визначити конструкцію; виконати математичне моделювання процесу розгону супутника.

Виклад основного матеріалу дослідження. Багато досліджень зараз проводять у нашому світі та в найближчому космічному просторі в пошуках чогось нового або удосконалення існуючого. Разом з тим, вчені працюють над великими масштабнішими проектами. Однією з найреалістичніших і найперспективніших технологій є світлові вітрила. Прототипом їх можна вважати сонячні вітрила.

Принцип дії сонячного вітрила (його називають ще «фотонне вітрило») заснований на поширенні сонячного світла у просторі. Як відомо, світло складається з фотонів, які самі собою позбавлені маси. Утім, рухаючись через простір, фотони несуть імпульс. Саме цей фотонний імпульс і передається відбивальній поверхні сонячного вітрила в момент попадання на нього світла (див. рис. 1). Сам по собі імпульс, який несуть у собі фотони, досить слабкий, проте тут сонячному вітрилу стають у пригоді умови середовища, в якому воно здійснює рух.

Космічний вакуум має нульовий опір, внаслідок чого навіть слабка сила світла здатна визначити космічному апарату потрібний вектор прискорення. До того ж прискорення космічного корабля, обладнаного сонячним вітрилом, постійно збільшуватиметься під впливом сонячного світла. Це означає, що оснащені сонячним вітрилом апарати можуть досягати швидкостей, не підвладних сьогодиншній ракетній техніці, що працює на хімічному паливі, водночас використовуючи для своїх цілей повністю відновлюване джерело енергії [5].

Конструкція для забезпечення належного прискорення сонячне вітрило має бути із надлегких матеріалів. На сьогодні найоптимальнішими рішеннями є виготовлення їх із майлару (BoPET від Biaxially-oriented polyethylene terephthalate) та полііміду (PI). Важливе й нанесення на ці матеріали покриття (часто з алюмінію та його сплавів), яке перешкоджатиме проходженню фотонів крізь вітрило, оскільки цей процес може призвести до втрати космічним апаратом свого прискорення.

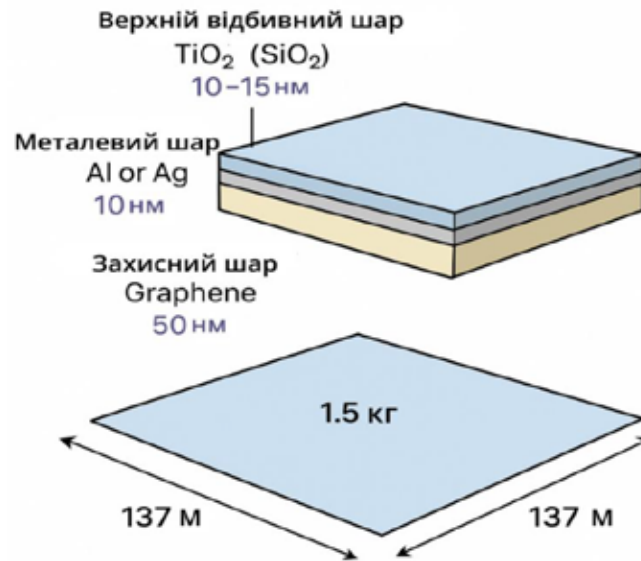


Рис. 1. Схема передавання фотонами імпульсу сонячному вітрилу

Ще одним показником, що безпосередньо впливає на швидкість, яку може розвинути сонячне вітрило, є розміри і товщина його відбивального покриття. Мінімальної межі тут не існує, та що меншим за площею воно буде, то більше часу знадобиться вітрилу на розвиток швидкостей, необхідних для космічних подорожей. Важливим є й грамотно вивірене співвідношення маси космічного апарата з площею його сонячного вітрила. Тому для досягнення оптимальних характеристик швидкості, яку може розвинути фотонне вітрило, інженери або збільшують площу поверхні вітрила, що відбиває, або намагаються зменшити масу космічного апарата, до якого воно буде прикріплене. Також важливо захистити супутник від радіації та врахувати, скільки маси супутника буде на це витрачено.

Основні варіанти покриттів:

1. Каптон з алюмінієвим напленням (MLI-ізоляція) – стандарт у супутниках. Щільність $\sim 100\text{--}200\text{ г/м}^2$ (з урахуванням багатошаровості).

2. Арамідна тканина (кевлар) з алюмінієвим шаром – щільність $150\text{--}300\text{ г/м}^2$.

3. Поліетилен високої щільності (HDPE, UHMWPE) – хороший захист від протонів, $\sim 200\text{--}400\text{ г/м}^2$ при тонкому покритті.

4. Тонка берилієва або алюмінієва фольга – легка, але крихка. $\sim 50\text{--}150\text{ г/м}^2$.

Вага покриття супутника представлена в таблиці 1.

Таблиця 1

Вага покриття

Матеріал покриття	Питома маса (г/м ²)	Наносупутник (до 10 кг)	Малий супутник (до 50 кг)	Середній (до 200 кг)	Примітка
MLI (каптон + алюміній)	100–200	0,15–0,30 кг	0,75–1,5 кг	3–6 кг	Стандартна термо- та радіаційна ізоляція
Кевлар + алюмінізація	150–300	0,23–0,45 кг	1,1–2,3 кг	4,5–9 кг	Стійка до мікрометеоритів
UHMWPE (поліетилен високої щільності)	200–400	0,3–0,6 кг	1,5–3 кг	6–12 кг	Гарний захист від протонів
Алюмінієва фольга (тонка)	50–150	0,075–0,23 кг	0,38–1,1 кг	1,5–4,5 кг	Легка, але крихка
Базальтове волокно з напильованням	250–500	0,38–0,75 кг	1,9–3,8 кг	7,5–15 кг	Термо- та радіаційно-стійкий композит

Проблемою також є той факт, що тиск сонячного світла у міру віддалення від Сонця слабшає, тож, очевидно, не зможе бути наданим достатній розгін космічного транспортного засобу. Ця проблема може бути вирішена шляхом застосуванням лазерів, заснованих на використанні ядерних установок.

Фізичною основою роботи лазера є явище вимушеного (індукованого) випромінювання. Суть явища у тому, що збуджений атом (чи інша квантова система) здатний випромінювати фотон під впливом іншого фотона без його поглинання, якщо енергія останнього дорівнює різниці енергій рівнів атома до і після випромінювання. При цьому випромінюваний фотон когерентний фотону, що викликав випромінювання (є його «точною копією»). У такий спосіб відбувається посилення світла. Цим явище відрізняється від спонтанного випромінювання, в якому випромінювані фотони мають випадкові напрямки поширення, поляризацію та фазу.

Імовірність того, що випадковий фотон викликає індуковане випромінювання збудженого атома, точно дорівнює ймовірності поглинання цього фотона атомом, що знаходиться в незбудженому стані. Тому посилення світла необхідно, щоб збуджених атомів серед було більше, ніж незбуджених (так звана інверсія населеностей). У стані термодинамічної рівноваги ця умова не виконується, тому використовуються різні системи накачування активного середовища лазера.

У даний час як робоче середовище лазера використовуються різні агрегатні стани речовини: твердий, рідкий, газоподібний, плазма. У звичайному стані кількість атомів, що знаходяться на збуджених енергетичних рівнях, визначається розподілом Больцмана:

$$N = N0 \exp(-E / kT) \quad (1)$$

де N – число атомів, що у збудженому стані з енергією E ;

$N0$ – число атомів, що у основному стані (енергія дорівнює нуль);

k – стала Больцмана;

T – температура середовища.

Інакше кажучи, таких атомів, які у збудженому стані, менше, ніж у основному, тому ймовірність того, що фотон, поширюючись у середовищі, викличе вимушене випромінювання, також мала порівняно з ймовірністю його поглинання. Тому електромагнітна хвиля, проходячи речовиною, витрачає свою енергію на збудження атомів. Інтенсивність випромінювання розраховується за законом Бугера:

$$I_1 = I0 \exp(-a1 \cdot l) \quad (2)$$

де $I0$ – початкова інтенсивність;

I_1 – інтенсивність випромінювання, що пройшло відстань l у речовині;

$a1$ – показник поглинання речовини.

Оскільки експоненційна залежність, випромінювання дуже швидко поглинається.

У тому випадку, коли кількість збуджених атомів більша, ніж не збуджених (тобто у стані інверсії населення), ситуація прямо протилежна. Акти вимушеного випромінювання переважають над поглинанням, і випромінювання посилюється згідно із законом:

$$I_1 = I0 \exp(a2 \cdot l) \quad (3)$$

де $a2$ – коефіцієнт квантового посилення.

У реальних лазерах посилення відбувається до того часу, поки величина що надходить рахунок вимушеного випромінювання енергії стане рівної величині енергії, що втрачається в резонаторі. Ці втрати пов'язані з насиченням метастабільного рівня робочої речовини, після чого енергія накачування йде тільки на його розігрів, а також з наявністю безлічі інших факторів (розсіювання на неоднорідностях середовища, поглинання домішками, неідеальність дзеркал, що відображають, корисне і небажане випромінювання в навколишнє).

Класична трирівнева система накачування робочого середовища використовується, наприклад, у рубіновому лазері. Рубін є кристалом корунду Al_2O_3 , легований невеликою кількістю іонів хрому Cr^{3+} , які і є джерелом лазерного випромінювання. Через вплив електричного поля кристалічних ґрат корунду зовнішній енергетичний рівень хрому $E2$ розщеплений (див. ефект Штарка). Саме це уможливило використання монохроматичного випромінювання як накачування. При цьому атом переходить з основного стану з енергією $E0$ у збуджений з енергією $E2$. У цьому стані атом може бути порівняно недовго (порядку 10^{-8} с), майже відразу відбувається безвипромінювальний перехід на рівень $E1$, на якому атом може перебувати значно довше (до 10^{-3} с), це так званий метастабільний рівень. Виникає можливість здійснення індукованого випромінювання під впливом інших випадкових фотонів. Щойно атомів, що у метастабільному стані, стає більше, ніж у основному, починається процес генерації.

Попередні дослідження світлових вітрил показали, що використання конструкції фотонного кристала, по суті, обшивання "тканини" вітрила рівномірно розташованими отворами, дозволить максимально збільшити теплове випромінювання конструкції. Нова робота дослідників додає ще один шар періодичності: зразки вітрильної тканини, зшиті разом у сітку [6].

Завдяки тому, що відстань між отворами відповідає довжині хвилі світла, а відстань між зразками відповідає довжині хвилі теплового випромінювання, вітрило може витримати ще більш потужний початковий поштовх, скоротивши час, який лазерам знадобиться, щоб залишатися на своїй цілі. Але є ще багато проблем, які потребують відповідних рішень. Для роботи над ними спочатку необхідно роз'яснити, звідки береться тиск світла.

Методи зниження маси – це ультратонкі шари: діелектрик 10–15 нм, метал 10 нм, захисний шар 50 нм → зниження маси на 50–60%. Надлегкі матеріали: захисний шар графен або аерогель → щільність до 200–400 кг/м³ → зниження маси шару на 70%. Метал можна замінити на багатошаровий діелектрик → зменшення маси на 50%.

Структурна оптимізація: Перфороване або сітчасте вітрило → зниження маси 20–30% без втрати відбиття.

Світло це за своєю природою – електромагнітна хвиля. Світло на поверхню падає порціями: фотонами. І хоч самі фотони не маю маси, але тиск вони можуть чинити, адже будь-яка електромагнітна хвиля повинна чинити тиск, це випливає з магнітної теорії Максвелла. Для обчислення тиску світла за нормального падіння випромінювання і відсутності розсіювання можна скористатися формулою [6]:

$$P_R = P_r = \frac{(1+a)I}{c}, \quad (4)$$

Цю формулу можемо розписати як:

$$P = \frac{I}{C}(1-k+p), \quad (5)$$

де I – інтенсивність падаючого випромінювання (Вт/м²);

C – швидкість світла (км/год),

k – коефіцієнт пропускання,

p – коефіцієнт відбиття.

Саме вітрило представляє тонкий лист із графену, вкритий відбивним матеріалом, таким як золото або срібло. Але оскільки цей лист буде сильно нагріватися, пропонується використовувати мікроструктури, які зможуть ефективно розподіляти енергію світлового пучка. Саме полотно буде в товщину приблизно 100 атомів.

На сьогодні існує багато проблем щодо виведення супутника на орбіту планети, яка цікавить науковців. Але багато з цих проблем вже має свої рішення. Перше, про що варто поговорити, це сам розгін. Так, можна розігнати апарат до 10-20% швидкості світла [8].

На думку авторів, найперспективніше буде надати початкової швидкості завдяки запуску ракетних установок із землі. Багато ракет здатні розігнатися до швидкості 44 000 км/год. Можна і швидше, але візьmemo це число за основу. Тепер нам потрібно порахувати, скільки потрібно витратити енергії, щоб з цієї швидкості розігнати світлове вітрило до 15% швидкості світла. Перспективним є застосування лазерних установок для подібної місії, адже світло сонця на великій відстані вже не буде діяти належним чином, а лазером можна буде дістати і до іншої зірки. Розрахувати, скільки потрібно енергії витратити на 1 кг вантажу світлового вітрила, можна буде просто помноживши отримане число на вагу апарату, за умов використання подібної лазерної установки.

Одні з найпотужніших лазерних установок можуть переробляти від 1 до 5 мегават на день на лазерний потік фотонів. Для подібної місії можна буде в полі на 2-х протилежних півкулях землі розташувати лазерні установки в розмірі 100 штук. Ці установки можна використовувати усі разом, або по чергово, а також змінювати за необхідності місце розташування (наприклад, якщо їх роботі буде заважати обертання планети), або ж зупиняти одну з них і вмикати іншу для техобслуговування. Звісно, це дуже затратно, але зараз цю проблематику активно обговорюють, і поки є науковий інтерес і залученість бізнесу до даних проєктів, можна буде знайти потрібні кошти. Для цього потрібно довести, що впровадження подібної технології буде вигідним для в'язого людства, на кшталт виведення першого супутника на орбіту планети [8].

Якщо відомо, скільки енергії споживають 100 установок, а імовірно саме стільки працюватимуть одночасно, можна розрахувати, скільки піде часу на розгін до 10% швидкості світла [8]:

1. Тяга світлового (фотонного) вітрила знаходиться за формулою фотонної тяги (з променевого тиску $p=\xi I/c$):

$$F = \frac{\xi P}{c}, \quad (6)$$

де P – оптична потужність променя,

c – швидкість світла,

ξ – коефіцієнт відбиття імпульсу: $\xi=2$ для ідеально дзеркального вітрила, $\xi=1$ для повністю поглинаючого [9].

2. Прискорення апарата від світлової тяги знаходиться за формулою рівняння прискорення (другий закон Ньютона з фотонною тягою) [10]:

$$a = \frac{F}{m} = \frac{\xi P}{mc} \quad (7)$$

де m – маса апарата.

3. Фактор Лоренца при цільовій швидкості знаходиться за формулою релятивістського множника (фактор Лоренца) [11]:

$$\gamma = \frac{1}{\sqrt{1 - \left(\frac{v}{c}\right)^2}} \quad (8)$$

Для $v = 0,1$ швидкості світла отримуємо $\gamma \approx 1,005$.

4. Час розгону до заданої швидкості (постійна тяга, мала релятивістська поправка) знаходиться за формулою час розгону під постійною світловою тягою [12]:

$$t = \frac{\gamma m v c}{\xi P} \quad (9)$$

При $v=0,1$ швидкості світла це дає кращий відсоток точності.

5. Пройдена дистанція за час розгону знаходиться за формулою шляху при рівнозмінному розгоні [13]:

$$S = \frac{1}{2} v t \quad (10)$$

6. Кінетична енергія знаходиться за формулою релятивістської кінетичної енергії [14]:

$$E_k = (\gamma - 1) m c^2 \quad (11)$$

7. Випромінена енергія лазерів за час розгону визначається за формулою енергетичного балансу випромінювача [15]:

$$E_{\text{laser}} = P t \quad (12)$$

Фотонна тяга при нормальному падінні променя:

$$F = \frac{\xi P}{c} \quad (13)$$

Прискорення апарату масою m :

$$a = \frac{F}{m} = \frac{\xi P}{m c} \quad (14)$$

Час розгону до швидкості v (коректно з урахуванням релятивістського γ при $0,1c$ це $+0,5\%$ до часу):

$$F = \frac{\gamma m v c}{\xi P}; \quad \gamma = \frac{\xi P}{\sqrt{\left(1 - \left(\frac{v}{c}\right)^2\right)}} \quad (15)$$

Далі виконаємо математичне моделювання для релятивістської динаміки світлового вітрила. Докладне пояснення стосуються наступних моментів.

Імпульс релятивістського тіла:

$$p = \gamma * m * v, \text{ де } \gamma = \frac{1}{\sqrt{1 - (v/c)^2}}. \quad (16)$$

Постійна сила:

$$F = \frac{dp}{dt} \Rightarrow p(t) = F t \text{ (якщо } p(0)=0). \quad (17)$$

Отже:

$$\gamma m v = F t \Rightarrow v(t) = \left(\frac{F t}{m}\right) / \sqrt{1 + \left(\frac{F t}{m c}\right)^2}. \quad (18)$$

Прискорення:

$$a(t) = \frac{dv}{dt} = \frac{F}{m \gamma^3}. \quad (19)$$

Далі виконаємо апроксимацію: Ньютонівська межа ($F t \ll m c$): $v \approx (F/m) t$ – лінійне зростання швидкості. Релятивістська межа ($v \rightarrow c$): $v \rightarrow c$ асимптотично, $a \propto 1/\gamma^3 \rightarrow 0$.

Для наочності рорзв'яжемо чисельні приклади ($P=100$ МВт, $\xi=2$).

Параметри:

$$P=100 \cdot 10^6 \text{ Вт}, \xi=2 \Rightarrow F=\frac{\xi P}{c} \approx 0.6667 \text{ Н}. \quad (20)$$

Для $m = 1 \text{ кг}$:

$$a_{\text{newt}} = F/m \approx 0.6667 \text{ м/с}^2. \quad (21)$$

Ньютонівський час до

$$v = 0.1 \text{ с: } t_{\text{newt}} = \frac{v}{a} \approx 4.5 \cdot 10^7 \text{ с} \approx 522.7 \text{ діб} \approx 1.43 \text{ р} \quad (22)$$

Релятивістський фактор при $v = 0.1 \text{ с: } \gamma \approx 1.0050378 \Rightarrow t_{\text{rel}} \approx \gamma t_{\text{newt}} \approx +0.5038\%$ від t_{newt} .

Зменшення прискорення при $v = 0.1 \text{ с: } \gamma^3 \approx 1.0152 \Rightarrow a_{\text{rel}} \approx a_{\text{newt}} / 1.0152 \approx \text{зменшення} \approx 1.5\%$.

З отриманих розрахунків можна отримати практичні висновки:

1. До $\sim 0.1 \text{ с}$ релятивістські ефекти невеликі: можна використовувати ньютонівську модель з точністю близько відсотка.

2. При швидкостях $\geq 0.3 \text{ с}$ ефекти стають істотними: потрібен релятивістський розрахунок, і розження-ючі зусилля зростають швидше, ніж у простому ньютонівському наближенні.

3. Технічні обмеження (фокусування променю, нагрівання, міцність вітрила) часто важливіші за релятивістські поправки при практичному проектуванні.

Для цільового $v=0.1 \text{ с}$ ($\gamma \approx 1.005$) і $P = 100 \text{ МВт}$.

$$t \approx 4.51 \cdot 10^7 \text{ с} \cdot 6 \frac{\text{м}}{1 \text{ кг}} = 522.7 \text{ с} \cdot \frac{\text{м}}{1 \text{ кг}} = 1.431 \text{ год} \cdot \frac{\text{м}}{1 \text{ кг}} \quad (23)$$

Залежність параметрів променів від маси наведено на рисунку 2, а розрахункову кількість енергії для розгону – в таблиці 2.

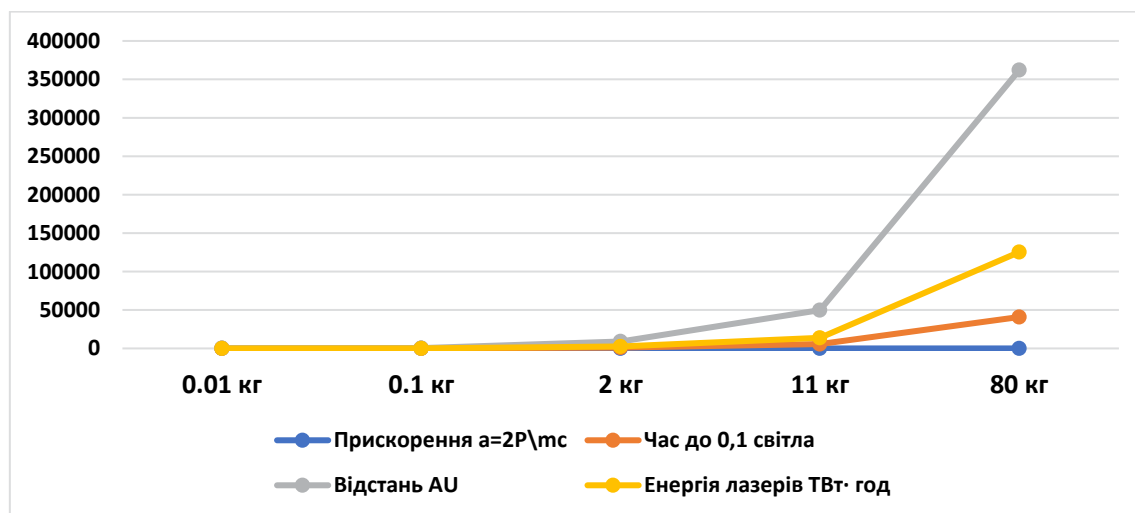


Рис. 2. Залежність параметрів променів від маси (кг)

Таблиця 2

Розрахунок кількості енергії для розгону

Маса m	Прискорення $a=2P/mc$	Час до 0,1с	Відстань $s \approx 1/2 v t$	Енергія лазерів $E=Pt$
0.01 кг	66,7 м/с ²	5,23 діб	$\approx 45,3 \text{ AU}$	$4,52 \times 10^{13} \text{ Дж} \approx 12\,544 \text{ МВт} \cdot \text{год}$ ($\approx 12,54 \text{ ГВт} \cdot \text{год}$)
0.1 кг	6,67 м/с ²	52,3 діб	$\approx 452 \text{ AU}$	$4,52 \times 10^{14} \text{ Дж} \approx 125\,444 \text{ МВт} \cdot \text{год}$ ($\approx 125,44 \text{ ГВт} \cdot \text{год}$)
2 кг	0,3335 м/с ²	1045 діб	$\approx 9\,050 \text{ AU}$	$4,52 \times 10^{15} \text{ Дж} \approx 2\,508\,888 \text{ МВт} \cdot \text{год}$ ($\approx 2,508 \text{ ТВт} \cdot \text{год}$)
11 кг	0,0600 м/с ²	5621 діб	$\approx 49\,779 \text{ AU}$	$4,52 \times 10^{16} \text{ Дж} \approx 13\,798\,884 \text{ МВт} \cdot \text{год}$ ($\approx 13,798 \text{ ТВт} \cdot \text{год}$)
80 кг	0,005336 м/с ²	40880 діб	$\approx 362\,033 \text{ AU}$	$4,52 \times 10^{17} \text{ Дж} \approx 100\,355\,300 \text{ МВт} \cdot \text{год}$ ($\approx 125,36 \text{ ТВт} \cdot \text{год}$)

В таблиці 2 прийнято, що АУ-астрономічна одиниця (відстань між землею та сонцем).

Результат: із загальною оптичною потужністю 100 МВт і вітрилом з ідеальним відображенням час до 0,1 становить приблизно 1,43 року на кілограм маси (лінійно згідно з m). Для 10 г – близько 5,2 доби, для 1 г – близько 12,5 годин. Вартість з урахуванням оптової поставки енергії показана в (табл. 3).

Таблиця 3

Вартість з урахуванням оптової поставки енергії

Маса	Енергія лазерів (МВт·/ год)	Вартість при 0,1 \$/кВт· год
0.01 кг	12 544	1 254 400 \$
0.1 кг	125 444	11 289 400 \$
2 кг	2 508 888	160 407 888 \$
11 кг	13 798 884	910 559 218 \$
80 кг	100 355 300	5 500 355 300 \$

Оцінка вартості однієї установки (фазове комбінування, повна наземна станція) наступна.

Причина: потрібен лазерний модуль плюс потужні блоки живлення, тепловідвід, система фазового комбінування, директор, управління, інженерна інтеграція.

Оцінка: $\approx$ \$30–150 млн за 1 МВт. (враховує, що промислова збірка МВт-класу і beam-director значно дорожчі за прості модулі; ближче до поточних контрактних рівнів і цін на телескопічні системи).

Враховуючи отримані дані можна зробити висновок, що вплив буде оказувати тільки масштабність та вигода проекту на вартість запуску (залежність показана на рис 2).

За результатами моделювання можна зробити висновок, що за рівних умов найвигідніше буде запустити багато нано супутників, які важать близько 2 кілограмів, замість одного великого, бо розігнати та зупинити малий супутник можливо швидше ніж великий. Це означає що більшу частину у просторі супутник буде летіти з максимальною швидкістю.

У середньому невеликий супутник (наносупутник) важить навіть менше ніж 10 кг згідно [9], вага вітрил буде незначною. Якщо при цьому ми припустимо, що втрати будуть близько 10%, то нам знадобиться 2 508 888 МВт енергії за час розгону витратити на цей проєкт. Це буде коштувати приблизно 160 407 888 \$ Для порівняння Україна споживає на рік 112,7 тераватт-часов. Так це великі цифри але варто розуміти що розгін може бути не таким різким, а річ йде про около світові швидкості.

Але це буквально тільки півдороги, адже нам ще потрібно ще трохи менше енергії, щоб загальмувати цю установку. Але тут виникає питання, як це можна зробити. Адже в такому разі нам потрібні подібні установки вже з іншого боку. Ну взагалі це не обов'язково, є й інший шлях для гальмування. Він важчий для реалізації, ніж просто розігнати цільну установку. Але в разі успішного гальмування ми зможемо вивести супутник на орбіту планети або зірки, яка знаходиться дуже далеко від нас. І це набагато вигідніше, ніж просто з навколосвітловою швидкістю пролетіти повз велику кількість можливостей [8].

Для реалізації проєкту можливо, крім супутника, взяти ще одну відбивну установку, яка буде відокремлена від основного вантажу або запустити окремий супутник, для того, щоб відбити лазерний потік фотонів уже з іншого боку. Але така установка має бути вагою як мінімум такою самою, як і основний вантаж. На неї перед відділенням можна буде завантажити вже непотрібне обладнання, яке вже відпрацювало або вийшло з ладу, для полегшення основного супутника. Це означає, що вага такого корабля вже становитиме приблизно на 90% більше, ніж початкова. За рахунок того, що ми збільшимо витрати в 2 рази, ми отримаємо в десятки разів більше зібраної інформації з супутника, який має шанси вийти на орбіту за таких умов [8].

Однак запустити цю установку з такою точністю практично неможливо, що зупиняє багатьох науковців, які це розуміють. Але коригувальні іонні двигуни дуже економічні та будуть отримувати енергію від нагрівання світлового вітрила. І завдяки їм вже буде можливо коригувати сам міжзоряний корабель, а також більш точно вийти на гравітаційний маневр або якщо паливо буде витрачено під час польоту, порожні або напівпорожні баки можна буде навантажити на гальмівне дзеркало, що може здешевити саме гальмування [8].

Далі змодельуємо потужність іонного двигуна, який можна використовувати для корекції супутника. Для початку потрібно отримати джерело електроенергії. Пропонується використовувати тепло нагрівання вітрила від лазера для вироблення енергії за наступною методикою.

Формула нагріву [16-18]:

$$P_{abs} = A \cdot I \cdot (1 - R), \quad (24)$$

$$P_{el} = s \cdot m, \quad (25)$$

$$\eta = \frac{P_{el}}{P_{abs}}. \quad (26)$$

Тоді $E_{month} = P_{el} \cdot 720 \text{ год} (30 \text{ днів})$ — енергія за місяць у Вт·год.
Витрати енергії представлені в таблиці 4.

Таблиця 4

Витрати енергії за місяць

Сценарій	s (Вт/кг)	$P_{el}=s \cdot m$ (Вт)	$KKД = \frac{P_{el}}{P_{abs}}$	Енергія/місяць E (кВт·ч)
Консервативний	50	$50 \cdot 0.5 = 25$	$\frac{25}{750000} = 3.33 \cdot 10^{-5} = 0.0033$	$25 \cdot 720 / 1000 = 18$
Реалістичний	100	$100 \cdot 0.5 = 50$	$\frac{50}{750000} = 6.67 \cdot 10^{-5} = 0.00667$	$50 \cdot 720 / 1000 = 36$
Оптимістичний	200	$200 \cdot 0.5 = 100$	$\frac{100}{750000} = 1.33 \cdot 10^{-5} = 0.0133$	$100 \cdot 720 / 1000 = 72$

Тяга F з електричною потужністю P та ефективністю η :

$$F = \frac{2\eta P}{v} \quad (27)$$

Результати розрахунків зведемо в таблицю 5.

Таблиця 5

Результати моделювання

$P \setminus I_{sp} \downarrow$	36 kW	3.6 kW	360 W	36 W
2000	2.20 N	0.22 N	0.022 N	0.0022 N
3000	1.47 N	0.147 N	0.0147 N	0.00147 N
5000	0.88 N	0.088 N	0.0088 N	0.00088 N

Отже, якщо ми витрачаємо 2.86 року на розгін, то нам потрібно буде витратити близько 3 на його гальмування, після чого залишок енергії необхідно буде витратити на гравітаційний маневр. Якщо в космосі зустрінуться пил і астероїди, які можуть дещо зіпсувати політ, то світлове вітрило радіусом 10 км за час розгону буде знищено лише на 0.1%-0.3% від усієї площі. Так що загроза є тільки для обладнання. Броня, яка зможе витримати удар від піщинок або каміння, на такій швидкості буде доволі важка, тож щодо цього є своє рішення: замість броні ми можна взяти лінзу, яка розсіює, вона за тим самим принципом може просто здути або спалити те, що може вийти на курс цього корабля. Пізніше лінзу можна буде прикріпити до гальмівного дзеркала для полегшення основної установки [7].

Тож у майбутньому можуть існувати міжпланетні логістичні відносини. Деякі компанії вже готові запустити перше світлове вітрило до найближчої потенційно населеної планети.

Висновки. У результаті проведеного дослідження можна зробити висновок, що технологія світлового паруса є однією з найбільш перспективних для забезпечення ефективної доставки супутників на далекі відстані. Аналіз сучасних розробок та експериментальних зразків показав, що попри існування низки технічних та економічних проблем, дана технологія має значний потенціал для подальшого розвитку.

Визначені проблеми та запропоновані шляхи їх вирішення відкривають можливості для практичного застосування світлового паруса у майбутніх космічних місіях. Математичне моделювання підтвердило реалістичність впровадження цієї технології, а економічний аналіз довів доцільність її розвитку. Таким чином, світловий парус може стати важливим кроком на шляху до нової ери космічних перевезень, що дозволить здешевити та прискорити освоєння космосу.

Список використаних джерел:

1. Гайда П. І., Трофіменко П. Є., Ляпа М. М. Основи теорії польоту і конструкції ракет : навчальний посібник. Суми : Сумський державний університет, 2011. 248с. URL: https://shron1.chtyvo.org.ua/Haida_Pavlo/Osnovy_teorii_polotu_i_konstruktsii_raket.pdf?PHPSESSID=gkhhtj41ddes06jk1bjkihtic3
2. Колесник А. Вантажні перевезення у космосі стають рутиною – аналітик про плани «SpaceX». URL: <https://hromadske.radio/podcasts/rankova-hvylya/vantazhni-perevezennia-u-kosmosi-staiut-rutynoiu-analytik-pro-planu-spacex>
3. Мистецтво космічного вітрильства URL: <https://universemagazine.com/mystecztvo-kosmichnogo-vitrylnyctva/>

-
4. За час в любую точку планеты URL: https://meduza.io/feature/2017/09/29/za-chas-v-lyubuyu-tochku-planety?utm_source=chatgpt.com
 5. Сонячне вітрило: рушійна сила світла. URL: <https://maxpolyakov.com/ua/sonyachne-vitrilo-rushijna-sila-svitla/>
 6. Як спроектувати вітрило, яке не порветься та не розплавиться під час міжзоряної подорожі? URL: <https://ostannipodii.com/a/202202/yak-sproektuvati-vitrilo-yake-ne-porvetsya-ta-ne-rozplavitsya-pid-chas-mizhzorya-100024038/>
 7. Radiation Pressure Formula. URL: https://www.geeksforgeeks.org/physics/radiation-pressure-formula/?utm_source=chatgpt.com
 8. Жир С. І., Гаврилов Б. О., Бобришев Я. О. Щодо розрахунку часу розгону та гальмування світових вітрил під час виведення супутника на віддалену космічну орбіту. *Abstracts of XIII International Scientific and Practical Conference. Prague, Czech Republic*. Pp. 290–294. URL: <https://eu-conf.com/en/events/social-ways-of-training-specialists-in-the-social-sphere-and-inclusive-education/>
 9. Малі супутники. URL: https://ru.wikipedia.org/wiki/%D0%9C%D0%B0%D0%BB%D1%8B%D0%B5_%D1%81%D0%BF%D1%83%D1%82%D0%BD%D0%B8%D0%BA%D0%B8
 10. Фізика для школярів та студентів. URL: https://physic.cx.ua/priskorennya/#google_vignette
 11. Лоренц-фактор. URL: <https://www.maekkes.nl/Lorentzfactor%20RU.html>
 12. Рівноприскорений прямолінійний рух. Прискорення. Швидкість. URL: <https://www.miyklas.com.ua/p/fzika/9-klas/rukh-i-vzayemodiiia-zakoni-zberezhennia-372240/rivnopriskorennii-priamoliniinii-rukh-peremishchennia-363606/re-48df6e5d-8bb6-4705-acd2-71bfd588205e#:~:text=%D0%9F%D1%80%D0%B8%D1%81%D0%BA%D0%BE%D1%80%D0%B5%D0%BD%D0%BD%D1%8F%20%E2%80%94%D1%86%D0%B5%20%D0%B2%D0%B5%D0%BA%D1%82%D0%BE%D1%80%D0%BD%D0%B0%20%D1%84%D1%96%D0%B7%D0%B8%D1%87%D0%BD%D0%B0%20%D0%B2%D0%B5%D0%BB%D0%B8%D1%87%D0%B8%D0%BD%D0%B0,%D0%B7%D0%B0%20%D1%8F%D0%BA%D0%B8%D0%B9%20%D1%86%D1%8F%20%D0%B7%D0%BC%D1%96%D0%BD%D0%B0%20%D0%B2%D1%96%D0%B4%D0%B1%D1%83%D0%BB%D0%B0%D1%81%D1%8F>
 13. Рівноприскорений рух. URL: <https://sites.google.com/view/galinaokhotnik-t1/%D1%80%D1%96%D0%B2%D0%BD%D0%BE%D0%BF%D1%80%D0%B8%D1%81%D0%BA%D0%BE%D1%80%D0%B5%D0%BD%D0%B8%D0%B9-%D1%80%D1%83%D1%85>
 14. Relativistic Energy. URL: <https://courses.lumenlearning.com/suny-physics/chapter/28-6-relativistic-energy/#:~:text=Relativistic%20kinetic%20energy%20is%20KE, reduces%20to%20classical%20kinetic%20energy>
 15. A Database of Albert Einstein's Collected Works. URL: <https://einsteinpapers.press.princeton.edu/vol2-trans/186>
 16. Cengage – Physics for Scientists and Engineers. URL: <https://www.cengage.com/c/physics-for-scientists-and-engineers-9e-serway/>
 17. Tutorial on methods for estimation of optical absorption and scattering properties of tissue. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11166171/>
 18. New Materials and Devices for Thermoelectric Power Generation Edited by Basel I. *Abed Ismail*. URL: https://intech-files.s3.amazonaws.com/a043Y00000yuiKEQAY/0014786_Authors_Book%20%282023-12-15%2009%3A15%3A17%29.pdf

References:

1. Haida, P. I., Trofimenko, P. Ye., & Liapa, M. M. (2011). *Osnovy teorii polotu i konstruksii raket: Navchalnyi posibnyk [Fundamentals of flight theory and rocket design: Textbook]*. Sumy: Sumy State University. Retrieved from https://shron1.chtyvo.org.ua/Haida_Pavlo/Osnovy_teorii_polotu_i_konstruksii_raket.pdf
2. Kolesnyk, A. (n.d.). Vantazhni perevezennia u kosmosi staiut rutynoiu – analityk pro plany “SpaceX” [Cargo transportation in space is becoming routine – analyst about SpaceX plans]. Retrieved from <https://hromadske.radio/podcasts/rankova-hvylya/vantazhni-perevezennia-u-kosmosi-staiut-rutynoiu-analitik-pro-planu-spacex>
3. Mystetstvo kosmichnogo vitrylstva [The art of space sailing]. (n.d.). Retrieved from: <https://universemagazine.com/mystecztvo-kosmichnogo-vitrylncytva/>
4. Za chas v liubu tochku planety [In an hour to any point on the planet]. (2017, September 29). Retrieved from <https://meduza.io/feature/2017/09/29/za-chas-v-lyubuyu-tochku-planety>
5. Soniachne vitrylo: rushiina syla svitla [Solar sail: the driving power of light]. (n.d.). Retrieved from <https://maxpolyakov.com/ua/sonyachne-vitrilo-rushijna-sila-svitla/>
6. Yak sproiektuvaty vitrylo, yake ne porvetsia ta ne rozplavytsia pid chas mizhzorianoi podorozhi? [How to design a sail that won't tear or melt during interstellar travel?]. (2022, February). Retrieved from <https://ostannipodii.com/a/202202/yak-sproektuvati-vitrilo-yake-ne-porvetsya-ta-ne-rozplavitsya-pid-chas-mizhzorya-100024038/>
7. Radiation pressure formula. (n.d.). GeeksforGeeks. Retrieved from <https://www.geeksforgeeks.org/physics/radiation-pressure-formula/>

-
8. Zhyr, S. I., Havrylov, B. O., & Bobryshev, Ya. O. (n.d.). Shchodo rozrakhunku chasu rozghonu ta halmuvannia svitlovykh vitryl pid chas vyvedennia suputnyka na viddalenu kosmichnu orbitu [Regarding the calculation of acceleration and braking time of light sails during satellite launch to a distant orbit]. In Abstracts of XIII International Scientific and Practical Conference (pp. 290–294). Prague, Czech Republic. Retrieved from <https://eu-conf.com/en/events/social-ways-of-training-specialists-in-the-social-sphere-and-inclusive-education/>
 9. Mali suputnyky [Small satellites]. (n.d.). Wikipedia. Retrieved from https://ru.wikipedia.org/wiki/%D0%9C%D0%B0%D0%BB%D1%8B%D0%B5_%D1%81%D0%BF%D1%83%D1%82%D0%BD%D0%B8%D0%BA%D0%B8
 10. Fyzyka dlia shkoliariv ta studentiv [Physics for schoolchildren and students]. (n.d.). Retrieved from <https://physic.cx.ua/priskorennia/>
 11. Lorentz-factor. (n.d.). Retrieved from: <https://www.maeckes.nl/Lorentzfactor%20RU.html>
 12. Rivnopriskorenyi priamoliniinyi rukh. Pryskorennia. Shvydkist [Uniformly accelerated linear motion. Acceleration. Velocity]. (n.d.). Retrieved from <https://www.miyklas.com.ua/p/fzika/9-klas/rukh-i-vzayemodiia-zakoni-zberezhennia-372240/rivnopriskorennia-priamoliniinii-rukh-peremishchennia-363606>
 13. Rivnopriskorenyi rukh [Uniformly accelerated motion]. (n.d.). Retrieved from: <https://sites.google.com/view/galinaokhotnik-t1/%D1%80%D1%96%D0%B2%D0%BD%D0%BE%D0%BF%D1%80%D0%B8%D1%81%D0%BA%D0%BE%D1%80%D0%B5%D0%BD%D0%B8%D0%B9-%D1%80%D1%83%D1%85>
 14. Relativistic energy. (n.d.). Lumen Learning. Retrieved from: <https://courses.lumenlearning.com/suny-physics/chapter/28-6-relativistic-energy/>
 15. Albert Einstein's collected works database. (n.d.). Princeton University Press. Retrieved from: <https://einsteinpapers.press.princeton.edu/vol2-trans/186>
 16. Cengage – Physics for Scientists and Engineers. (n.d.). Retrieved from: <https://www.cengage.com/c/physics-for-scientists-and-engineers-9e-serway/>
 17. Tutorial on methods for estimation of optical absorption and scattering properties of tissue. (n.d.). National Center for Biotechnology Information. Retrieved from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11166171/>
 18. Abed Ismail, B. I. (Ed.). (2023). New materials and devices for thermoelectric power generation. IntechOpen. Retrieved from: https://intech-files.s3.amazonaws.com/a043Y00000yuiKEQAY/0014786_Authors_Book.pdf

Дата надходження статті: 26.10.2025

Дата прийняття статті: 17.11.2025

Опубліковано: 30.12.2025

Зайцева Т. А., кандидат технічних наук, доцент,
завідувач кафедри комп'ютерних технологій
Дніпровського національного університету імені Олеся Гончара
ORCID: 0000-0002-6346-3390

Гончаров Я. А., аспірант кафедри комп'ютерних технологій
Дніпровського національного університету імені Олеся Гончара
ORCID: 0009-0002-8994-8895

СИНЕРГЕТИЧНЕ ЗАСТОСУВАННЯ ANSYS І MATLAB ДЛЯ МОДЕЛЮВАННЯ КОНТАКТНОЇ ВЗАЄМОДІЇ ШТАМПА СКЛАДНОЇ ГЕОМЕТРІЇ З ПРУЖНИМ ПІВПРОСТОРОМ

У статті представлено інтегровану методику чисельного моделювання контактної взаємодії абсолютно жорсткого штампа складної геометрії з однорідним ізотропним пружним півпростором, що базується на синергетичному застосуванні програмних платформ ANSYS та MATLAB. Геометрія контактної області задана у вигляді двозв'язного кільця, контур якого описується концентричними кривими. Реалізовано побудову скінченно-елементної моделі в середовищі ANSYS з урахуванням складної геометрії, граничних умов та навантаження, що дозволило отримати детальний розподіл контактних напружень. Особливу увагу приділено адаптивному дискретизуванню зони контакту з метою забезпечення збіжності чисельного рішення та підвищення точності розрахунків. Після завершення симуляції середовищі ANSYS, результати експортовано у MATLAB, де реалізовано алгоритми постобробки та графічну інтерпретацію даних. Здійснено 3D-візуалізацію скінченно-елементної моделі, напружено-деформованого стану, побудову ізоліній контактного тиску та аналіз крайових ефектів. Застосування MATLAB дозволило виявити локальні особливості розподілу тиску, зони концентрації напружень та області зниженої контактної взаємодії та виконано розрахунки для кілець різної ширини. Запропонована методика охоплює повний обчислювальний цикл, що включає етапи побудови геометричної моделі, її чисельного розв'язання та подальшої інтерпретації отриманих результатів. Отримані дані можуть бути використані для верифікації моделей, оптимізації конструктивних параметрів, а також для розробки прикладних інженерних методик у межах задач контактної механіки, особливо в умовах складної геометрії зон взаємодії.

Ключові слова: моделювання, метод скінченних елементів, контактна задача, штамп, теорія пружності, постпроцесінг, візуалізація даних, розподіл напружень, геометрія зони контакту, ANSYS, MATLAB.

Zaytseva T. A., Honcharov Ya. A. Synergetic Application of ANSYS and MATLAB for Modeling the Contact Interaction of a Complex-Shaped Stamp with an Elastic Half-Space

An integrated methodology for numerical modeling of the contact interaction between a perfectly rigid punch of complex geometry and an isotropic elastic half-space is presented, based on the synergistic use of ANSYS and MATLAB. The approach treats the contact region as a doubly connected ring (annular domain) whose inner and outer boundaries are described by concentric curves, thereby capturing geometric features that strongly influence local stress fields. A finite-element model was constructed in ANSYS with due account of the intricate geometry, boundary conditions, and loading schemes; this enabled the computation of a detailed distribution of contact stresses and displacements under prescribed indentation scenarios. A systematic sensitivity analysis with respect to mesh density, local refinement near anticipated singular zones, and interpolation settings was performed, leading to practical convergence criteria and stable error tolerances suitable for repeated studies. Special attention was paid to adaptive discretization of the contact zone so as to ensure numerical convergence, suppress spurious oscillations, and improve the accuracy of peak-stress estimation along the edges of the active contact. Upon completion of the simulation, the computed fields were exported to MATLAB, where a dedicated post-processing pipeline was implemented to provide consistent graphical interpretation and quantitative comparison. In particular, masking outside the physically realized contact patch was applied to eliminate non-physical interpolation artifacts and to isolate the effective load-transfer area for analysis. The toolkit supports the generation of contour maps of contact pressure, 3D visualizations of the stress-strain state, and targeted diagnostics of edge effects, including the localization and ranking of stress concentration zones. The use of MATLAB also facilitates the identification of local features in the pressure distribution and regions of reduced contact interaction, and it enables structured parametric comparisons across multiple geometric variants and loading levels using a uniform set of metrics. Overall, the proposed methodology spans the full computational cycle—from geometric modeling and robust finite-element solution to standardized interpretation and reporting of results—and is designed to be scalable and amenable to scripting and automation for batch studies. The resulting datasets and visual products can be directly used for verification and validation of models, optimization of design parameters, and the

© Т. А. Зайцева, Я. А. Гончаров, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

development of applied engineering practices for contact mechanics problems, particularly in settings with complex interaction geometries and pronounced stress gradients where reliable resolution of edge behavior is essential.

Key words: modeling, finite element method (FEM), contact problem, punch, theory of elasticity, post-processing, data visualization, stress distribution, contact zone geometry ANSYS, MATLAB.

Постановка проблеми. Контактна взаємодія тіл складної геометрії становить одну з ключових задач механіки деформівного тіла, що має широке застосування в галузях прикладної механіки, трибології, матеріалознавства та інженерного конструювання. Особливий інтерес викликають конфігурації контактної області, які не піддаються класичному аналітичному опису. У таких умовах застосування традиційних аналітичних методів є обмеженим через складність граничних умов, геометричну неоднорідність та потенційну нелінійність механічної відповіді системи і як наслідок, складність математичного апарату [10].

Метод скінченних елементів (МСЕ), реалізований у середовищі ANSYS, забезпечує можливість побудови дискретизованої моделі з урахуванням складної геометрії, матеріальних параметрів та умов навантаження, що дозволяє отримати прийнятні чисельні оцінки напружено-деформівного стану в зоні контакту [6]. Але складність і обмеженість у деякому сенсі функціональних можливостей постпроцесінгу в програмному середовищі ANSYS, трохи заважають проведенню аналізу результатів і комплексних досліджень. У цьому контексті доцільним є синергетичне використання ANSYS та MATLAB як взаємодоповнювальних обчислювальних платформ: ANSYS – для реалізації чисельного моделювання фізичних процесів, MATLAB – для поглибленої обробки, статистичного аналізу та візуалізації отриманих даних, що забезпечує підвищення точності та інформативності дослідницьких висновків. Реалізація даного підходу є особливо релевантною в контексті підвищення достовірності чисельного моделювання, раціоналізації конструктивних рішень та формування новітніх інженерних методик для аналізу задач контактної механіки, що характеризуються складною конфігурацією зон взаємодії [10].

Аналіз останніх досліджень і публікацій. У роботах [6-9] детально розглянуто теоретичні засади механіки суцільного середовища, принципи побудови скінченно-елементних моделей, методи дискретизації та критерії збіжності чисельних алгоритмів. У статті [5] досліджується задача вдавнення абсолютно жорсткого штампа без тертя у тонкий пружний шар, що спирається на жорстку основу. Такі аналітичні моделі відіграють ключову роль у процесі верифікації чисельних рішень, отриманих за допомогою МСЕ.

Сучасні алгоритмічні підходи до розв'язання задач контактної механіки в континуальних середовищах узагальнено в [17], де наведено огляд алгоритмів на основі швидкого перетворення Фур'є (FFT) для ефективного розв'язання контактних задач. У роботі [14] запропоновано гранично-елементний підхід для осьово-симетричних задач напівпростору: наведено фундаментальні рішення Ліпшиця–Генкеля та чисельні схеми обчислення впливових коефіцієнтів, що забезпечують умови нульової напруги на поверхні та спад напружень у віддаленості. Практичні реалізації цих алгоритмів демонструють високу точність та ефективність при моделюванні контактних процесів із різною геометрією. Застосування асимптотичних методів, що дозволяють отримувати аналітичні наближення, представлено у [15]. Важливе місце у розвитку моделювання займають дослідження, орієнтовані на реалістичне відтворення взаємодії конструкцій і основи, або фундаменту. Зокрема, у [13] здійснено повний тривимірний FEM-аналіз взаємодії будівлі з неоднорідним ґрунтом. Ґрунт описано модифікованою моделлю Cam-Clay з параметрами, визначеними за польовими (CPT, SDMT) та лабораторними випробуваннями. У [16] досліджено вплив виробничих та геометричних похибок не симетричного кільцевого штампа на розподіл контактних напружень по поверхні півпростору. Авторами у праці [4] досліджено осесиметричний контакт кругового штампа з пружним півпростором у межах лінійної теорії пружності з урахуванням попереднього (початкового) напруженого стану.

Попри значну теоретичну цінність аналітичних моделей, їх застосування обмежене складністю геометрії, неоднорідністю матеріалу та граничними умовами, що притаманні реальним інженерним задачам. У зв'язку з цим усе більшого значення набувають чисельні методи, здатні забезпечити гнучкість і високу точність моделювання складних фізико-механічних процесів. Треба відзначити, що ANSYS залишається провідним програмним середовищем для інженерного аналізу завдяки універсальності та широкому набору фізичних модулів. У [12] підкреслюється, що ця платформа поєднує високу точність і гнучкість проведення розрахунків для різних типів задач – від теплопровідності до контактної механіки. Однак у роботі [2] акцентується, що ефективність чисельного моделювання значно зростає при поєднанні ANSYS із MATLAB. Останній використовується для створення алгоритмів керування, автоматизації постпроцесінгу та реалізації спеціалізованих обчислювальних сценаріїв, які важко виконати виключно засобами ANSYS. Такий інтегрований підхід, коли ANSYS відповідає за точність розрахунків, а MATLAB – за гнучкість аналізу та візуалізації, забезпечує найбільш повне використання можливостей чисельного моделювання.

Сукупність розглянутих джерел формує наукову основу для визнання актуальності чисельного моделювання в задачах контактної механіки, де метод скінченних елементів у середовищі ANSYS демонструє високу точність. Поєднання ANSYS із MATLAB дозволяє суттєво розширити аналітичні можливості, забезпечуючи гнучку обробку та візуалізацію результатів.

Розглядається задача про вдавнення абсолютно жорсткого циліндричного штампа з плоскою основою у однорідний ізотропний пружний півпростір. Поперечний переріз штампа має форму двозв'язного кільця

некругової геометрії, межі якого описуються концентричними лемніскатами Бута. Штамп заглиблюється у півпростір під дією вертикальної сили, що прикладена у центрі тяжіння штамп. У моделі не враховується сила тертя між штампом і півпростором.

Метою дослідження є формалізація, чисельна реалізація та верифікація математичної моделі контактної взаємодії абсолютно жорсткого циліндричного штамп у формі не кругового кільця, поперечний переріз якого описується системою концентричних лемніскат Бута, з однорідним ізотропним пружним півпростором під дією вертикальної сили, що прикладена у центрі мас штамп. Для досягнення поставленої мети передбачено побудову скінченно-елементної моделі в програмному середовищі ANSYS, інтеграцію з MATLAB для розширеного постпроцесінгу, а також дослідження впливу геометричних параметрів штамп та умов навантаження на напружено-деформований стан у зоні контакту.

Виклад основного матеріалу. Відповідно до поставленої мети, нижче наведено структуровану послідовність етапів побудови чисельної моделі контактної взаємодії, яка представлена на рисунку 1. Математична постановка задачі – перший етап алгоритма – детально розглянута у роботах [1, 10], а саме – фізичні умови, математичне представлення – рівняння рівноваги, умови контакту, граничні умови, визначення геометрії та параметрів: (двозв'язне кільце границі якого мають форму лемніскатами Бута), розміри, матеріальні характеристики пружного півпростору, тип навантаження.

Етап 1 Математична постановка задачі
— Формулювання фізичної та математичної моделі досліджуваного процесу.
— Визначення геометрії, матеріальних властивостей, контактних умов, граничних і початкових умов.
Етап 2 Чисельне моделювання в ANSYS
— Побудова обчислювальної моделі: дискретизація області дослідження методом скінчених елементів, задання навантажень.
— Розрахунок задачі в цілому або сегментовано.
— Уточнення скінченно-елементної сітки для забезпечення адекватності результатів.
Етап 3 Аналіз достовірності чисельних результатів
— Оцінка збіжності розв'язків та коректності отриманих даних.
— Прийняття рішення щодо потреби в додатковій обробці або повторному моделюванні.
Етап 4 Підготовка даних для подальшої обробки
— Експорт результатів у формати, придатні для статистичної та чисельної обробки
— Попереднє очищення даних, видалення нерелевантної інформації, забезпечення необхідної точності.
Етап 5 Обробка та інтерпретація результатів у Matlab
— Інтерполяція даних, аналіз наявності шумів.
— Побудова розподілів напружень, визначення ліній рівних тисків, виділення зон концентрації напружень.
Етап 6 Візуалізація та наукова інтерпретація результатів
— Представлення результатів у графічній формі для подальшого аналізу.
— Наукове узагальнення отриманих даних та підготовка висновків.

Рис. 1. Алгоритм розробки моделі контакту

У межах реалізації алгоритму моделювання виконано дискретизацію досліджуваної області методом скінченних елементів (див. рис. 2(а, в, д)), задано відповідні граничні умови та прикладено навантаження, що відповідає фізичній постановці задачі. Проведено симуляцію повної моделі. Здійснено оцінку впливу згущення сітки на розподіл напружено-деформівного стану та підтверджено відповідність отриманих даних очікуваній фізичній поведінці системи [1]. Статичний контактний аналіз проведено до досягнення збіжності чисельного рішення з прийнятним рівнем похибки.

Після проведення верифікації чисельних результатів, оцінки збіжності та коректності розв'язку, а також порівняння з аналітичними даними, наведеними у джерелах [1, 10], за умови досягнення задовільної відповідності, було здійснено перехід до етапу підготовки даних для подальшої обробки. Як зазначено у [12], деяка недостатня гнучкість вбудованих засобів постпроцесінгу та автоматизації в середовищі ANSYS ускладнює реалізацію нетипових обчислювальних сценаріїв. У зв'язку з цим результати чисельного моделювання, зокрема розподіли контактних напружень і переміщень, були експортовані у форматах, придатних для зовнішньої обробки (.xlsx), та імпортовані в MATLAB.

Подальший аналіз здійснено за допомогою спеціалізованого прикладного програмного забезпечення, розробленого авторами у середовищі MATLAB. Програмний модуль забезпечує розширену обробку чисельних даних, включаючи інтерполяцію, побудову згладжених поверхонь, фільтрацію за фізичними критеріями та багатопроекційну візуалізацію.

Нижче подано приклад фрагментів коду мовою MATLAB для побудови ліній рівного тиску для різних формах поперечного перерізу штамп. На початку здійснюється зчитування даних з Excel-файлу. Після цього обчислюється діапазон координат, і на його основі створюється рівномірна сітка точок з невеликим запасом по краях. За допомогою функції `meshgrid` формується координатна сітка, яка слугує основою для побудови ліній рівного тиску:

```
filePath = 'B:\Aspirant\Data\Test1.xlsx';
data = readmatrix(filePath, 'Range', 'A:C');
X = data(:,1); Y = data(:,2); Z = data(:,3);
x_range = max(X) - min(X); y_range = max(Y) - min(Y);
xq = linspace(min(X) - 0.1*x_range, max(X) + 0.1*x_range, 500);
yq = linspace(min(Y) - 0.1*y_range, max(Y) + 0.1*y_range, 500);
[Xq, Yq] = meshgrid(xq, yq);
```

Застосування кубічної інтерполяції, дозволяє отримати гладку поверхню навіть при розріджених даних.

```
Zq = griddata(X, Y, Z, Xq, Yq, 'cubic');
k = boundary(X, Y, 1);
inBoundary = inpolygon(Xq, Yq, X(k), Y(k));
Zq(~inBoundary) = NaN;
```

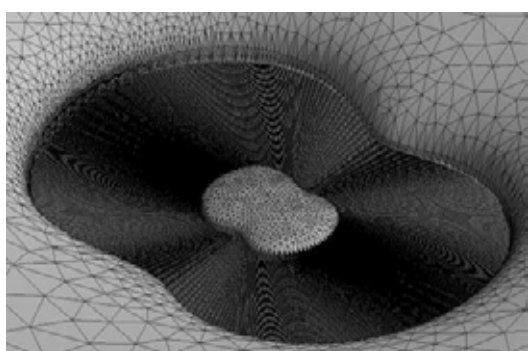
Для зручності аналізу координати сітки відцентровано початок координат для кожного окремого випадку.

```
center_x = (max(X) + min(X)) / 2;
center_y = (max(Y) + min(Y)) / 2;
Xq_c = Xq - center_x;
Yq_c = Yq - center_y;
Zq_work = Zq;
maxZq = nanmax(Zq_work(:));
i0 = round(size(Zq_work,1)/4);
i1 = round(size(Zq_work,1)*3/4);
for j = i0:i1
    for i = i0:i1
        if Zq_work(i,j) > 0.15*maxZq
            Zq_work(i:end, j) = NaN;
        end
    end
end
```

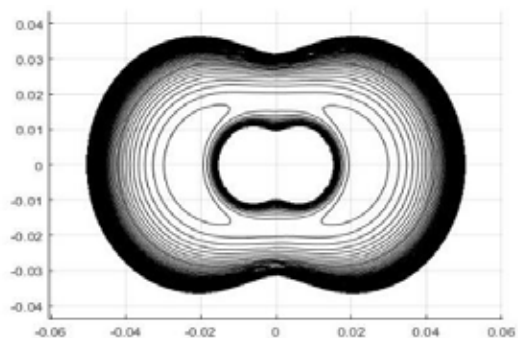
У наступному фрагменті наведено процес створення окремого графіка, де за допомогою функції `contour` будується 50 рівнів ліній рівного тиску для значень напруг `Zq_f`, відцентрованих відносно геометричного центру області (`Xq_c`, `Yq_c`). Застосовуючи команди `axis equal tight`, отримуємо коректне відображення осей та щільне розміщення графіка в межах вікна. Додавання координатних осей (`xlabel`, `ylabel`) робить графік більш інформативним для подальшого аналізу.

```
figure;
hold on;
grid on;
contour(Xq_c, Yq_c, Zq_f, 50, 'k', 'LineWidth', 0.3);
axis equal tight;
xlabel('X');
ylabel('Y');
hold off;
```

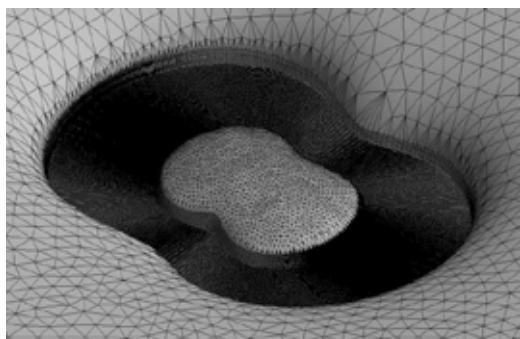
На рисунку 2 (а, в, д) представлено скінченно-елементні моделі пружного півпростору, на поверхню якого передається централізоване вертикальне навантаження через абсолютно жорсткий штамп. Геометрія штампів задана у вигляді двозв'язного кільця, обмеженого концентричними лемніскадами Бута. Розглянуто три варіанти моделей, що відповідають різним співвідношенням радіусів зовнішньої та внутрішньої межі кільця, а саме – а) 0,3; в) 0,5; д) 0,7. На відповідних рисунках 2 (б, г, є) наведено рівні контурні лінії тиску, що виникає під штампам. Таке порівняння дозволило простежити вплив геометричних параметрів на характер розподілу контактної тиску та формування зони взаємодії.



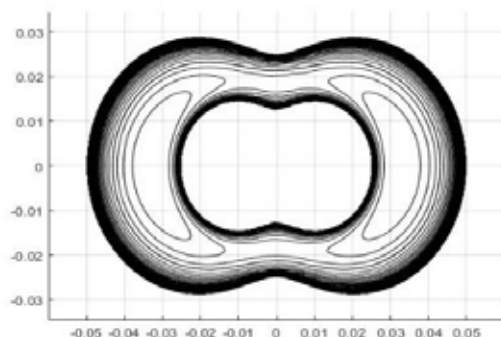
а)



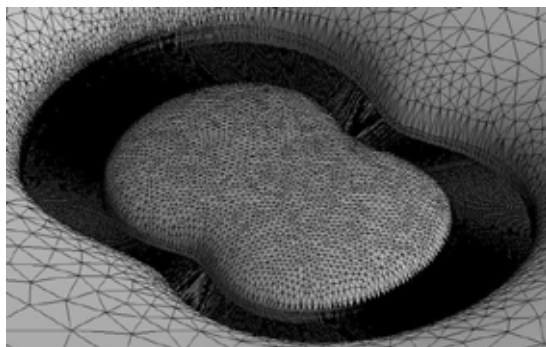
б)



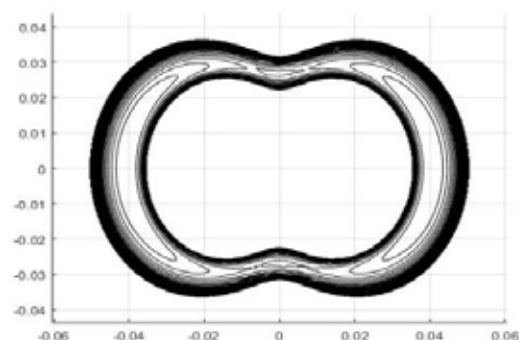
в)



г)



д)



е)

Рис. 2. Лінії рівного тиску (ізолінії) у зоні контакту

Таким чином, поєднання ANSYS і MATLAB дозволяє використовувати можливості ANSYS для побудови та розрахунку моделі контакту, а MATLAB – для подальшого ґрунтового аналізу та візуалізації даних. Як відзначають на офіційному сайті MathWorks, що такий підхід використовує сильні сторони кожної з програм: ANSYS для побудови структурної моделі, а MATLAB для алгоритму керування [2].

Висновки. У межах дослідження було створено спеціалізоване прикладне програмне забезпечення, реалізоване в середовищі MATLAB, призначене для автоматизованої обробки, аналізу та візуалізації результатів чисельного моделювання контактної взаємодії у задачах механіки деформівного твердого тіла. Програма функціонує як постпроцесор до розрахунків, виконаних у ANSYS, і забезпечує розширену інтерпретацію отриманих даних, що виходить за межі стандартних можливостей вбудованих засобів візуалізації. Основні функціональні компоненти програмного забезпечення включають:

- Імпорт чисельних даних у форматах.xlsx та.csv, що дозволяє інтегрувати результати скінченно-елементного аналізу з ANSYS.
 - Автоматизовану фільтрацію даних з урахуванням геометричних обмежень та фізичних критеріїв, зокрема обмеження зони впливу навантаження.
 - Генерацію багатопроекційних графічних представлень (приклади 3D-візуалізація, бічні та фронтальні проекції у статті не приведені), що дозволяє здійснювати комплексний аналіз напружено-деформованого стану.
- Загальні висновки:
- Реалізовано скінченно-елементну дискретизацію досліджуваної області в середовищі ANSYS, що дозволило здійснити чисельне моделювання напружено-деформованого стану з високим ступенем точності. Проведено аналіз збіжності чисельного рішення.
 - Інтеграція ANSYS із MATLAB забезпечила розширення функціональних можливостей постпроцесінгу, зокрема в аспектах автоматизованої обробки чисельних даних, побудови згладжених поверхонь деформації, реалізації фільтраційних алгоритмів та багатопроекційної візуалізації. Такий підхід дозволив здійснити глибокий аналіз контактних напружень та переміщень у зоні взаємодії.
 - Проведено дослідження впливу геометричних характеристик штамп (співвідношення радіусів внутрішньої та зовнішньої межі кільця) на характер розподілу контактної тиску. Встановлено, що зміна геометричних параметрів суттєво впливає на інтенсивність локальних напружень, що має важливе значення для оптимізації конструкцій у прикладних задачах.
 - Розроблене програмне забезпечення, яке є модульним, масштабованим та адаптованим до задач різної складності. Його застосування дозволило підвищити ефективність постпроцесінгу, забезпечити гнучкість у побудові аналітичних сценаріїв та покращити якість візуального представлення результатів чисельного моделювання.
- У перспективі передбачається розширення функціоналу за рахунок інтеграції алгоритмів глибинного навчання для автоматичної класифікації зон концентрації напружень та прогнозування поведінки контактних систем.

Список використаних джерел:

1. Зайцева Т. А., Шишканова Г. А. Розв'язання просторових контактних задач для некласичних багатозв'язних областей. Дніпро: Вид-во ДНУ, 2011. 192 с.
2. Пандуранган, П. MATLAB Controller Linked to an ANSYS Structural Model for Directional Drilling Controller. MathWorks, 21.11.2021
3. Ansys Free Student Software Downloads. URL: <https://www.ansys.com/academic/freestudent-products> (дата звернення: 18.02.2022).
4. Babysh S. Yu., Yaretska N. O. Contact Problem for an Elastic Ring Punch and a Half-Space with Initial (Residual) Stresses. *International Applied Mechanics*. 2021. Vol.57. P.297–305. DOI:10.1007/s10778-021-01081-7.
5. Barber J. R. Contact problems for the thin elastic layer. *International Journal of Mechanical Sciences*. 1990. Vol. 32, No. 2. P. 129–132.
6. Bathe K. J. Finite Element Procedures in Engineering Analysis. Prentice Hall, 1982. URL: <https://www.scribd.com/document/359408011/FEA-Finite-Element-Procedures-by-K-J-Bathe-pdf>.
7. Bonet J., Wood R. D. Nonlinear Continuum Mechanics for Finite Element Analysis. Cambridge University Press, 2008.
8. Bower A. F. Applied Mechanics of Solids. CRC Press, 2010. URL: <https://pdfcoffee.com/applied-mechanics-of-solids-pdf-free.html>.
9. Cook R. D., Malkus D. S., Plesha M. E. Concepts and Applications of Finite Element Analysis. Wiley, 2002. URL: <https://pdfcoffee.com/robert-d-cook-david-s-malkus-michael-e-plesha-concepts-and-applications-of-finite-element-analysis-3rd-edition-1989-483-pdf-free.html>.
10. Honcharov, Y. A., T. A. Zaytseva. "Computer Simulation of the Interaction of a Doubly Bound in the Plane of a Punch and an Elastic Half-Space." *Problems of Applied Mathematics and Mathematical Modeling*. 2023. 23, 46–52. doi:10.15421/322331
11. Matlab. URL: <https://uk.mathworks.com/products/matlab/student.html> (дата звернення: 27.06.2022).
12. MR CFD. What is ANSYS? URL: <https://www.mr-cfd.com/whatisansys/#:~:text=that%20Ansys%20excels%20at,for%20an%20Ansys%20simulation%20run>

13. Nepelski K. 3D FEM Analysis of the Subsoil-Building Interaction. *Applied Sciences*. 2022. Vol.12, No.21, Article 10700. DOI:10.3390/app122110700.
14. Oliveira M. F. F., Dumont N. A., Selvadurai A. P. S. Boundary element formulation of axisymmetric problems for an elastic half-space. *Engineering Analysis with Boundary Elements*. 2012. Vol.36, No.10. P.1478–1492. DOI:10.1016/j.enganabound.2012.03.015.
15. Shporta A. H. The asymptotic method in problems of the linear and nonlinear elasticity theory. *Natsional'nyi Hirnychiy Universytet. Naukovyi*. 2015. P. 76–81. URL: <https://nvngu.in.ua/index.php/uk/component/jdownloads/finish/54-03/8331-2015-03-shporta/>.
16. Shyshkanova G. A., Zaytseva T. A., Frydman A. D. The analysis of manufacturing errors effect on contact stresses distribution under the ring parts deformed asymmetrically. *Metallurgical and Mining Industry*. 2015. No. 7. P. 352–357. URL: www.metaljournal.com.ua/assets/Journal/englishedition/MMI_2015_7/055Shyshkanova-352-357.pdf.
17. Wang Q. J., Sun L., Zhang X., Liu S., Zhu D. FFT-Based Methods for Computational Contact Mechanics. *Frontiers in Mechanical Engineering*. 2020. Vol.6. DOI:10.3389/fmech.2020.00061.

References:

1. Zaytseva, T. A., & Shyshkanova, G. A. (2011). Rozv'iazannia prostorovykh kontaktnykh zadach dlia neklasychnykh bahatozv'iaznykh oblasteri [Solution of Three-Dimensional Contact Problems for Nonclassical Multiply Connected Domains]. Dnipro: DNU Press, 192 p.
2. Pandurangan, P. (2021). MATLAB Controller Linked to an ANSYS Structural Model for Directional Drilling Controller. MathWorks, November 21, Retrieved from: <https://www.mathworks.com/videos/matlab-controller-linked-to-an-ansys-structural-model-for-directional-drilling-controller-1637577988890.html>
3. Ansys Free Student Software Downloads. Retrieved from: <https://www.ansys.com/academic/freestudent-products> (accessed: 18.02.2022).
4. Babych, S. Yu., & Yaretska, N. O. (2021). Contact problem for an elastic ring punch and a half-space with initial (residual) stresses. *International Applied Mechanics*, 57, 297–305. <https://doi.org/10.1007/s10778-021-01081-7>
5. Barber, J. R. (1990). Contact problems for the thin elastic layer. *International Journal of Mechanical Sciences*, 32(2), 129–132.
6. Bathe, K. J. (1982). Finite Element Procedures in Engineering Analysis. Prentice Hall. Retrieved from: <https://www.scribd.com/document/359408011/FEA-Finite-Element-Procedures-by-K-J-Bathe-pdf>
7. Bonet, J., & Wood, R. D. (2008). Nonlinear Continuum Mechanics for Finite Element Analysis. Cambridge University Press.
8. Bower, A. F. (2010). Applied Mechanics of Solids. CRC Press. Retrieved from: <https://pdfcoffee.com/applied-mechanics-of-solids-pdf-free.html>
9. Cook, R. D., Malkus, D. S., & Plesha, M. E. (2002). Concepts and Applications of Finite Element Analysis. Wiley. Retrieved from: <https://pdfcoffee.com/robert-d-cook-david-s-malkus-michael-e-plesha-concepts-and-applications-of-finite-element-analysis-3rd-edition-1989-483-pdf-free.html>
10. Honcharov, Y. A., T. A. Zaytseva. (2023). Computer Simulation of the Interaction of a Doubly Bound in the Plane of a Punch and an Elastic Half-Space. *Problems of Applied Mathematics and Mathematical Modeling*. 23, 46–52. doi:10.15421/322331.
11. Matlab. Retrieved from: <https://uk.mathworks.com/products/matlab/student.html> (дата звернення: 27.06.2022).
12. MR CFD. What is ANSYS? Retrieved from: <https://www.mr-cfd.com/what-is-ansys/#:~:text=that%20Ansys%20excels%20at,for%20an%20Ansys%20simulation%20run>
13. Nepelski, K. (2022). 3D FEM analysis of the subsoil-building interaction. *Applied Sciences*, 12(21), 10700. <https://doi.org/10.3390/app122110700>
14. Oliveira, M. F. F., Dumont, N. A., & Selvadurai, A. P. S. (2012). Boundary element formulation of axisymmetric problems for an elastic half-space. *Engineering Analysis with Boundary Elements*, 36(10), 1478–1492. <https://doi.org/10.1016/j.enganabound.2012.03.015>
15. Shporta, A. H. (2015). The asymptotic method in problems of the linear and nonlinear elasticity theory. *Natsionalnyi Hirnychiy Universytet. Naukovyi*, (3), 76–81. Retrieved from: <https://nvngu.in.ua/index.php/uk/component/jdownloads/finish/54-03/8331-2015-03-shporta/>
16. Shyshkanova, G. A., Zaytseva, T. A., & Frydman, A. D. (2015). The analysis of manufacturing errors effect on contact stresses distribution under the ring parts deformed asymmetrically. *Metallurgical and Mining Industry*, 7, 352–357. Retrieved from: www.metaljournal.com.ua/assets/Journal/englishedition/MMI_2015_7/055Shyshkanova-352-357.pdf
17. Wang, Q. J., Sun, L., Zhang, X., Liu, S., & Zhu, D. (2020). FFT-based methods for computational contact mechanics. *Frontiers in Mechanical Engineering*, 6. <https://doi.org/10.3389/fmech.2020.00061>

Дата надходження статті: 23.09.2025

Дата прийняття статті: 10.10.2025

Опубліковано: 30.12.2025

Надригайло Т. Ж., кандидат технічних наук,
доцент кафедри математичного моделювання та системного
аналізу Дніпровського державного технічного університету
ORCID: 0000-0003-1239-5946

Перемітько М. В., аспірант кафедри математичного моделювання
та системного аналізу
Дніпровського державного технічного університету
ORCID: 0009-0005-4008-643X

МЕТОДИ АНАЛІЗУ ІНДИВІДУАЛЬНИХ ВИМІРЮВАНЬ АРТЕРІАЛЬНОГО ТИСКУ: ТРЕНДИ, СЕЗОННІСТЬ ТА ПРОГНОЗУВАННЯ

У роботі досліджено підхід до аналізу індивідуальних щоденних вимірювань артеріального тиску з метою виявлення довгострокових тенденцій, періодичних коливань та прогнозування можливих змін у майбутньому. Розглянуто використання методів лінійної регресії, STL-декомпозиції, а також моделей ARIMA, SARIMA та Holt-Winters для аналізу артеріального тиску на основі часових рядів. Їх використання дало змогу здійснити оцінку змін у фізіологічних показниках пацієнта та визначити закономірності динаміки артеріального тиску. Аналіз показників проведено на даних довготривалого домашнього моніторингу 67-річного пацієнта з атеросклерозом. Досліджено, що лінійна регресія є ефективною у визначенні базових довгострокових трендів, що відображають поступові зміни середнього рівня артеріального тиску. Метод STL-декомпозиції дозволяє візуалізувати структуру часових рядів та виокремити компоненти тренду, сезонності й залишкових коливань. Аналіз отриманих результатів показав, що модель Holt-Winters забезпечує найкращу точність прогнозування як для систолічного, так і для діастолічного тиску у випадках відсутності або слабкої вираженості сезонних коливань. Натомість модель SARIMA виявилася більш придатною для даних із наявними періодичними патернами, тоді як ARIMA доцільно застосовувати як базовий інструмент для оцінки загальних тенденцій без акценту на сезонність. Для кількісного оцінювання точності моделей використано стандартні метрики – середню абсолютну похибку (MAE) та корінь середньоквадратичної помилки (RMSE). Зроблено висновок про доцільність застосування методів аналізу часових рядів у персоналізованій медичній аналітиці. Визначено, що якість і тривалість збору даних, а також розширення спектру вимірюваних параметрів є ключовими чинниками для створення точних індивідуальних профілів ризику, що відображають реальну динаміку фізіологічного стану пацієнта.

Ключові слова: артеріальний тиск, часові ряди, лінійна регресія, сезонна декомпозиція, ARIMA, SARIMA, Holt-Winters, персоналізоване прогнозування.

Nadryhailo T. Zh., Peremitko M. V. Methods for analyzing individual blood pressure measurements: trends, seasonality, and forecasting

The article investigates an approach to analyzing individual daily blood pressure measurements with the aim of identifying long-term trends, periodic fluctuations, and forecasting possible future changes. The study explores the use of linear regression, STL decomposition, and the ARIMA, SARIMA, and Holt-Winters models for analyzing blood pressure as a time series. The application of these methods made it possible to assess changes in a patient's physiological parameters and to determine regularities in the dynamics of blood pressure. The analysis was conducted on data obtained from long-term home monitoring of a 67-year-old patient diagnosed with atherosclerosis. Linear regression was found to be effective in identifying baseline long-term trends that reflect gradual changes in the average blood pressure level. The STL decomposition method allows for the visualization of the time series structure and the separation of trend, seasonal, and residual components, providing a clearer understanding of the underlying patterns. The results of the study showed that the Holt-Winters model provides the highest forecasting accuracy for both systolic and diastolic blood pressure in cases where seasonal fluctuations are absent or weakly expressed. Conversely, the SARIMA model demonstrated better performance when the data contained pronounced periodic patterns, while ARIMA was found to be most appropriate as a baseline approach for assessing general trends without emphasizing seasonal influences. To quantitatively evaluate the accuracy of the models, standard statistical metrics – Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) – were employed. The comparative analysis of these metrics confirmed the importance of accounting for seasonality when building predictive models of physiological processes. The study highlights the potential of time series analysis methods for use in personalized medical analytics. It demonstrates that these methods can effectively reveal the dynamics and variability of an individual's cardiovascular parameters, enabling a more nuanced understanding of personal health trajectories. However, the research also indicates that the reliability and accuracy of such models significantly depend on the quantity and duration of collected data. For reliable short-term forecasting, a dataset containing at least one hundred observations is

© Т. Ж. Надригайло, М. В. Перемітько, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

recommended, while more extensive time series are required for identifying stable long-term trends and seasonal cycles. Furthermore, integrating additional physiological and environmental parameters – such as heart rate variability, physical activity, stress levels, and ambient temperature – into forecasting models could significantly enhance their precision and adaptability. These factors may influence blood pressure fluctuations and thus should be incorporated into future predictive frameworks. The study concludes that comprehensive, long-term monitoring combined with advanced time series modeling can serve as a foundation for developing precise individualized risk profiles that reflect the true dynamics of a patient's physiological condition and support data-driven medical decision-making.

Key words: blood pressure, time series, linear regression, seasonal decomposition, ARIMA, SARIMA, Holt-Winters, personalized forecasting.

Постановка проблеми. Своєчасне виявлення змін у стані серцево-судинної системи є однією з головних задач сучасної медицини. Хоча технології домашнього моніторингу артеріального тиску активно розвиваються, сучасні методи аналізу часто обмежуються лише розглядом окремих показників або середніх значень за певні періоди. Такий підхід не завжди дозволяє виявити приховані тренди чи закономірні коливання, які можуть бути ранніми ознаками розвитку серцево-судинних захворювань.

Крім того, щоденні вимірювання артеріального тиску формують часові ряди з високою варіабельністю, які можуть містити сезонні компоненти. Ці особливості часто залишаються поза увагою при традиційній обробці даних, що ускладнює оцінку довгострокових змін і потенційних ризиків для здоров'я. Наприклад, тривале підвищення середніх значень артеріального тиску може сигналізувати про поступове погіршення стану серцево-судинної системи, а сезонні коливання можуть вказувати на адаптаційні реакції організму до змін навколишнього середовища або ритму життя. Отже, постає задача розробити більш глибокий підхід до аналізу таких індивідуальних часових рядів, щоб не лише виявляти тренди та сезонні компоненти, але й забезпечувати можливість персоналізованого прогнозування ризиків для здоров'я.

Аналіз останніх досліджень і публікацій. Методи статистичного аналізу, такі як лінійна регресія та сезонна декомпозиція, вже давно використовуються для вивчення довгострокових трендів у медичних часових рядах. Наприклад, поєднання методів STL (Seasonal and Trend decomposition using Loess) та лінійної регресії дозволяє виявляти приховані довгострокові тенденції в часових рядах [1]. Водночас, застосування моделей ARIMA та SARIMA є ефективним підходом для короткострокового прогнозування часових рядів [2]. Ці моделі вимагають ретельного налаштування параметрів і достатньо довгих історичних рядів даних. Тому треба враховувати, що результати прогнозування залежать від кількості вхідних даних і зовнішніх факторів, а також від дослідника і умов експерименту [3].

Однак, попри успіхи у використанні статистичних методів для аналізу великих вибірок, їхнє застосування до щоденних домашніх вимірювань артеріального тиску досі залишається обмеженим. Це пов'язано з відсутністю довготривалих, структурованих і високоякісних даних, необхідних для побудови стабільних прогнозних моделей. Саме тому актуальним залишається питання розробки методів аналізу, які могли б адаптуватися до індивідуальних характеристик пацієнтів і враховувати сезонні коливання. Наприклад, індивідуальні підходи до прогнозування артеріального тиску можуть значно підвищити точність оцінки ризиків, що підтверджує доцільність проведення подібних досліджень.

Таким чином, дане дослідження, яке використовує комбінацію методів лінійної регресії, сезонної декомпозиції та моделей прогнозування, спрямоване на подолання цих обмежень. Воно має на меті розширити можливості індивідуалізованого аналізу артеріального тиску, що є важливим кроком до покращення точності діагностики та моніторингу стану пацієнтів у реальних умовах.

Метою даного дослідження є розробка та апробація підходу до аналізу індивідуальних щоденних вимірювань артеріального тиску для виявлення довгострокових трендів, сезонних коливань та прогнозування майбутніх змін. Такий підхід дозволить не лише покращити розуміння динаміки артеріального тиску в реальних умовах, але й створити передумови для персоналізованого моніторингу та раннього виявлення ризиків серцево-судинних захворювань.

Опис набору даних. Для проведення цього дослідження було використано набір даних "Home Blood Pressure Monitoring for Trend and Seasonality Detection" з платформи Mendeley Data. Дані зібрано в рамках тривалого домашнього моніторингу артеріального тиску у 67-річного чоловіка з діагностованим атеросклерозом. Моніторинг проводився у стандартних умовах: правильне положення пацієнта, манжета на оголеній руці, після гігієнічних процедур. Кожне вимірювання включало три послідовні спроби з інтервалом близько однієї хвилини, після чого обчислювалося середнє значення [4].

Загальна тривалість спостережень склала понад рік, з двома днями перерви між послідовними вимірюваннями. Це дало змогу отримати 128 вимірювань для кожного показника: систолічного артеріального тиску (САТ), діастолічного артеріального тиску (ДАТ) та частоти серцевих скорочень (ЧСС).

Викладення основного матеріалу дослідження. Лінійна регресія є одним із базових статистичних методів аналізу часових рядів і широко застосовується у медичних дослідженнях для вивчення довгострокових змін фізіологічних показників. Цей підхід дозволяє змодельовати тренд, що описує загальну тенденцію зміни досліджуваної змінної в часі. Модель простої лінійної регресії описується рівнянням:

$$y_i = ax_i + b + \varepsilon_i \quad (1)$$

де y_i – значення залежної змінної (наприклад, артеріального тиску) для i -го спостереження, x_i – значення незалежної змінної (наприклад, часу), a – коефіцієнт нахилу, що відображає середній темп зміни показника, b – вільний член (зміщення регресійної прямої відносно осі ординат), а ε_i – залишкова похибка, яка моделює випадкові відхилення. Оцінка параметрів a та b виконується методом найменших квадратів за формулами:

$$a = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (2)$$

$$b = \bar{y} - a\bar{x} \quad (3)$$

де $\bar{x}$ і $\bar{y}$ – середні значення незалежної та залежної змінних відповідно [5].

Побудова такої моделі дозволяє виділити детермінований компонент у структурі часового ряду – тренд, який відображає загальний напрямок динаміки досліджуваного процесу, наприклад, поступове підвищення або зниження артеріального тиску з часом.

Лінійна регресія була застосована до індивідуальних рядів систолічного (САТ) і діастолічного (ДАТ) артеріального тиску для визначення загальних тенденцій зміни цих показників у часі. Для цього використовувався числовий індекс часу – проста послідовність чисел, що відповідає кожному вимірюванню з інтервалом у два дні. Такий підхід дозволяє відобразити довгострокові зміни в артеріальному тиску, виявити загальні тренди та порівняти їх із фактичними значеннями на графіку. Це дозволяє оцінити, чи спостерігається поступове підвищення або зниження тиску та чи виходять ці коливання за межі рекомендованих значень (120/80 мм рт. ст.).

Результати аналізу показали (рис. 1), що систолічний тиск мав чітко виражений позитивний тренд із поступовим зростанням упродовж спостережуваного періоду. Діастолічний тиск також продемонстрував зростаючу динаміку, хоча з менш вираженим нахилом. Така картина свідчить про можливе хронічне підвищення артеріального тиску, що може потребувати медичного втручання або корекції лікування для пацієнта.



Рис. 1. Побудова тренду вимірів артеріального тиску за допомогою лінійної регресії

Попри свою простоту, метод лінійної регресії має низку важливих переваг у медичних застосуваннях. Зокрема, він дозволяє швидко оцінити загальний напрямок змін без попередніх припущень про сезонність чи автокореляцію даних. Однак цей метод має і певні обмеження, адже не враховує періодичні коливання чи структурні зміни, які можуть впливати на фізіологічні показники. Тому лінійна регресія часто використовується як базовий інструмент для первинного аналізу.

Одним із головних недоліків лінійної регресії є її неспроможність враховувати сезонні компоненти, які мають циклічний характер. У випадку фізіологічних даних, таких як артеріальний тиск, такі коливання можуть бути спричинені сезонними змінами в навколишньому середовищі, рівнем фізичної активності або навіть стресовими ситуаціями. Для подолання цього обмеження використовуються методи сезонної

декомпозиції, які дозволяють розділити часовий ряд на три основні складові: трендову, сезонну та випадкову компоненти. Одним із найбільш ефективних сучасних методів для цього є метод STL (Seasonal-Trend decomposition based on Loess).

STL є одним із найгнучкіших методів декомпозиції часових рядів і широко використовується для виявлення та окремого аналізу трендової, сезонної та залишкової складових. Метод дозволяє адаптивно моделювати як тренд, так і сезонність завдяки застосуванню локального регресійного згладжування (LOESS), що робить його особливо корисним для аналізу медичних даних, де спостерігаються нерегулярні зміни [6]. Декомпозиція адитивної форми записується як:

$$y_t = T_t + S_t + R_t, \quad t = 1, 2, \dots, n \quad (4)$$

де y_t – значення часового ряду в момент часу t , T_t – трендова компонента, яка визначається як згладжена оцінка довгострокової зміни рівня ряду за допомогою локальної регресії LOESS, S_t – сезонна компонента, що відображає періодичні коливання з фіксованим періодом P , і обчислюється як усереднення відхилень для кожної фази циклу, а R_t – залишкова компонента, що дорівнює:

$$R_t = y_t - T_t - S_t \quad (5)$$

Процедура декомпозиції виконується ітеративно. На кожному кроці для заданого вікна часу w трендова компонента T_t обчислюється за формулою локального згладжування LOESS:

$$T_t = \sum_{i=t-w}^{t+w} w_i y_i \quad (6)$$

де w_i – вагові коефіцієнти, визначені функцією локального згладжування (наприклад, трикутна або гаусова вага), симетрично навколо точки t . Сезонна компонента визначається шляхом усереднення залишків по кожній сезонній позиції:

$$S_t = \frac{1}{k} \sum_{j=1}^k (y_{t-jP} - T_{t-jP}) \quad (7)$$

У цьому дослідженні STL-декомпозиція була виконана окремо для систолічного та діастолічного артеріального тиску з періодом 15, що відповідає місячному циклу (30 днів із кроком вимірювання у два дні).

На (рис. 2) наведено результати STL-декомпозиції для систолічного артеріального тиску (CAT). Верхній підграфік («Observed») представляє початковий часовий ряд, який характеризується значною варіабельністю і виразними хвилеподібними коливаннями. Такі коливання можуть відображати щоденні коливання артеріального тиску, пов'язані з фізичним навантаженням, стресом або іншими зовнішніми факторами.

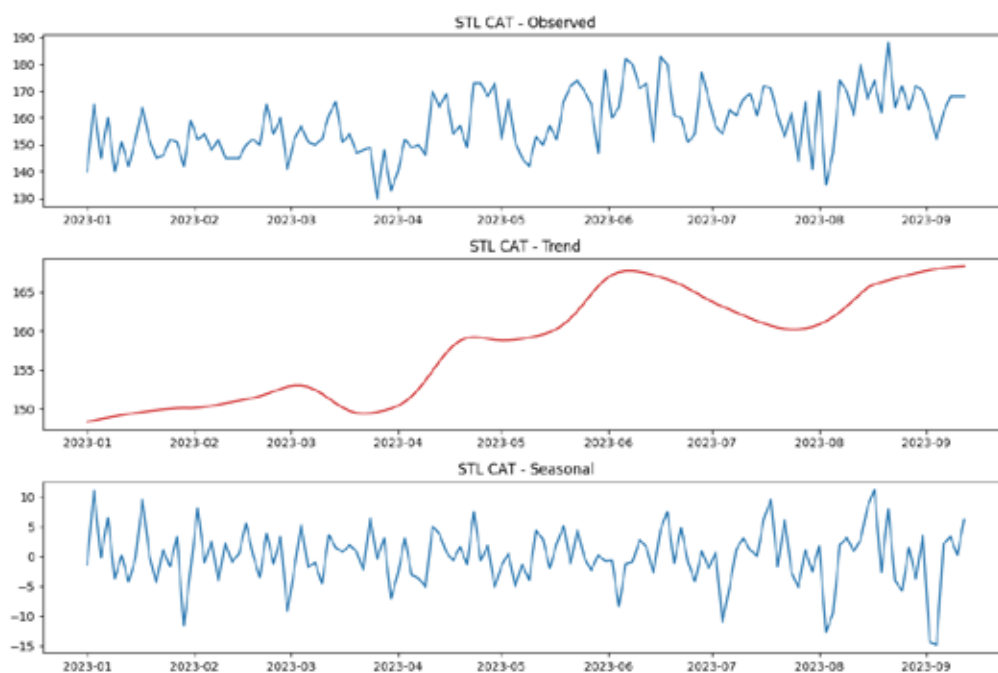


Рис. 2. STL-декомпозиція CAT

Середній підграфік («Trend») демонструє трендову компоненту, згладжену за допомогою локально-важених регресій (LOESS). Ця лінія відображає загальний напрямок змін тиску протягом спостережуваного періоду. Зокрема, можна побачити поступове підвищення значень САТ з березня до червня, невелике зниження в серпні і нове зростання на початку вересня. Така динаміка свідчить про наявність нелінійного тренду, який неможливо коректно відобразити за допомогою простої лінійної регресії, що підкреслює переваги STL-декомпозиції для аналізу таких даних.

Нижній підграфік («Seasonal») відображає сезонну компоненту, що складається з коливань амплітудою близько ± 5 – 10 мм рт. ст. Ці коливання можуть бути зумовлені фізіологічними ритмами або зовнішніми факторами, такими як зміни температури, рівня активності або стресу.

Аналогічні результати спостерігаються для діастолічного тиску (ДАТ), що зображено на (рис. 3). Трендова компонента для ДАТ також демонструє поступове зростання протягом першої половини року з незначним зниженням у середині літа, що відповідає типовій сезонній динаміці. Однак амплітуда коливань для діастолічного тиску є дещо меншою порівняно з САТ.

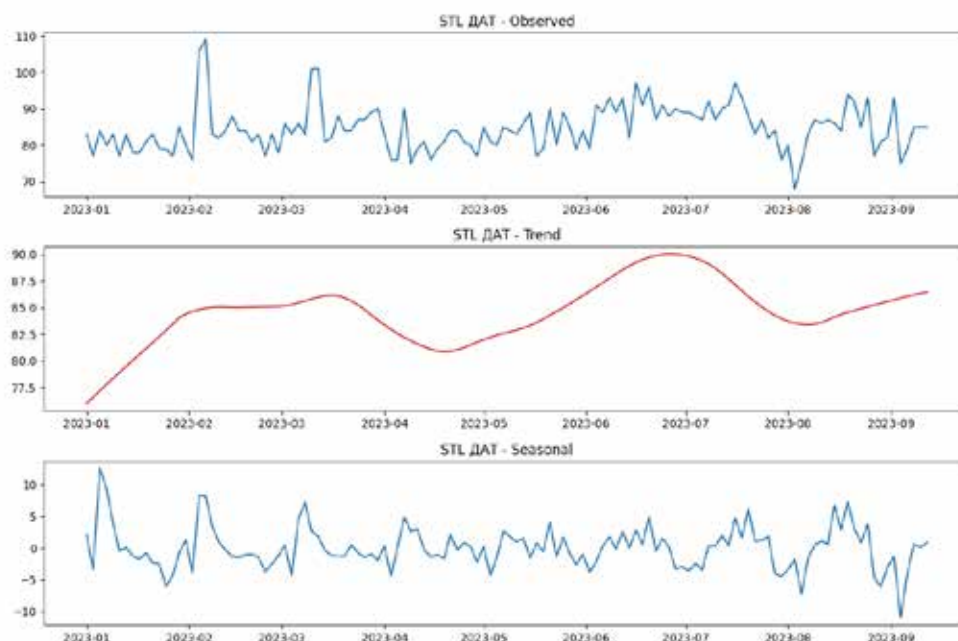


Рис. 3. STL-декомпозиція ДАТ

Загалом, метод STL виявився значно ефективнішим порівняно з простою лінійною регресією для аналізу динаміки артеріального тиску. Завдяки можливості виділяти як трендову, так і сезонну компоненти, цей метод забезпечує більш детальне уявлення про структуру часового ряду, що є особливо важливим для персоналізованого моніторингу стану здоров'я.

Після ідентифікації трендової та сезонної компоненти часового ряду наступним кроком є побудова моделей прогнозування. У дослідженнях динаміки фізіологічних показників, зокрема артеріального тиску, ефективними є такі підходи до аналізу часових рядів: ARIMA, SARIMA та Holt-Winters. Кожен із цих методів дозволяє по-різному моделювати як трендову, так і сезонну структуру, забезпечуючи точне короткострокове прогнозування.

Модель ARIMA (AutoRegressive Integrated Moving Average) поєднує авторегресивну частину, диференціювання та ковзне середнє. Структура моделі ARIMA(p, d, q) задається рівнянням :

$$\phi(B)(1-B)^d y_t = \theta(B)\mu_t \quad (8)$$

де B – оператор зсуву назад: $By_t = y_{t-1}$; d – порядок диференціювання для досягнення стаціонарності ряду; $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ – поліном авторегресії порядку p ; $\theta(B) = 1 + \theta_1 B + \dots + \theta_q B^q$ – поліном ковзного середнього порядку q ; μ_t – випадкова похибка. У розгорнутому вигляді ця модель набуває вигляду [7]:

$$\Delta^d y_t = c + \sum_{i=1}^p \phi_i \Delta^d y_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \varepsilon_t \quad (9)$$

Для часових рядів із вираженою сезонністю застосовується модель SARIMA (Seasonal ARIMA), яка доповнює стандартну ARIMA сезонними лагами. Повна форма моделі SARIMA(p, d, q)(P, D, Q)s записується як:

$$\Phi(B^s)\phi(B)(1-B)^d(1-B^s)^D y_t = \Theta(B^s)\theta(B)\varepsilon_t \quad (10)$$

де s – довжина сезонного циклу, $\Phi(B^s)\Theta(B^s)$ – сезонні поліноми авторегресії та ковзного середнього відповідно, D – порядок сезонного диференціювання. Така модель дозволяє одночасно враховувати як короткострокові, так і сезонні залежності у даних [8].

Модель експоненційного згладжування Holt-Winters, на відміну від ARIMA, не вимагає стаціонарності ряду і будується за принципом оновлення оцінок тренду та сезонності в режимі реального часу. Адитивна форма моделі визначається як:

$$y_{t+h|t} = l_t + hb_t + s_{t+h-m(k+1)} \quad (11)$$

$$l_t = \alpha(y_t - s_{t-m}) + (1-\alpha)(l_{t-1} + b_{t-1}) \quad (12)$$

$$b_t = \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1} \quad (13)$$

$$s_t = \gamma(y_t - l_{t-1} - b_{t-1}) + (1-\gamma)s_{t-m} \quad (14)$$

де l_t – рівень ряду, – трендова компонента, s_t – сезонна компонента з періодом m , $\alpha, \beta, \gamma \in [0,1]$ – параметри згладжування, а h – горизонт прогнозу. Завдяки оновленню компонент на кожному кроці модель Holt-Winters є особливо корисною для адаптивного прогнозування з урахуванням змінної сезонної поведінки [9].

Для навчання кожної моделі було використано перші 113 спостережень, що охоплюють приблизно 354 дні вимірювань. Прогноз будувався на основі останніх 15 точок. Результати прогнозування для систолічного (САТ) та діастолічного (ДАТ) артеріального тиску зображено на рисунках 4–6 для моделей ARIMA, SARIMA та Holt-Winters відповідно. Загалом, всі три моделі змогли відобразити загальну динаміку змін тиску в короткостроковій перспективі, але точність їхніх прогнозів відрізнялася.

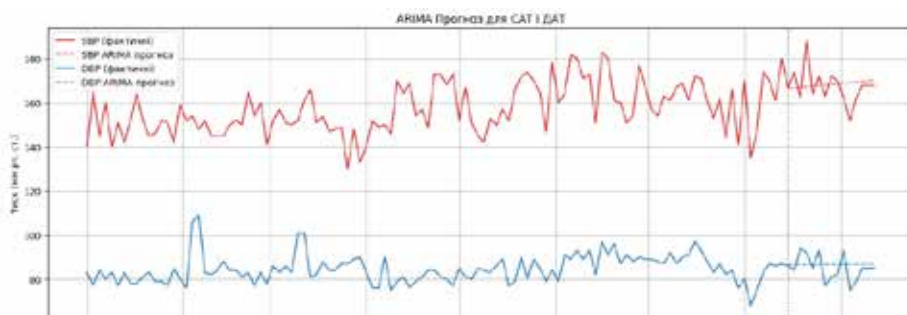


Рис. 4. Прогноз САТ і ДАТ за допомогою моделі ARIMA

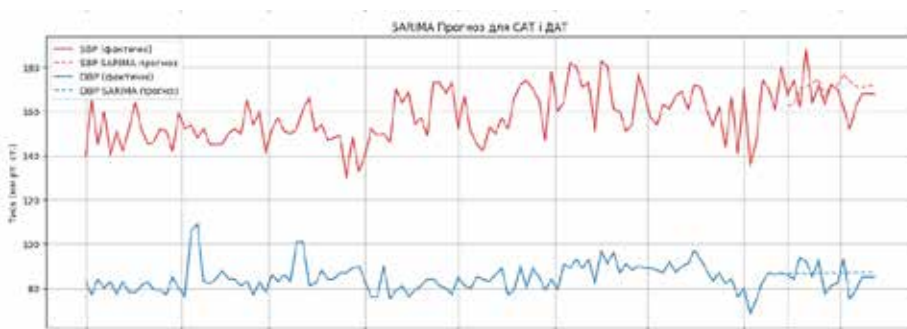


Рис. 5. Прогноз САТ і ДАТ за допомогою моделі SARIMA



Рис. 6. Прогноз CAT і ДАТ за допомогою моделі Holt-Winters

Для кількісного порівняння використовувалися дві метрики: середнє абсолютне відхилення (MAE) та корінь середньоквадратичної помилки (RMSE), які визначаються наступними формулами:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (15)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (16)$$

де y_i – фактичні значення, а $\hat{y}_i$ – прогнозовані [5].

Аналіз помилок прогнозування, представлених у таблиці 1, показав, що найменші значення помилок для CAT отримано за моделлю Holt-Winters (RMSE = 7.04, MAE = 4.55), що свідчить про її найкращу точність серед протестованих методів для цього показника. Для ДАТ найнижчі значення помилок також продемонструвала модель Holt-Winters (RMSE = 6.32, MAE = 4.86), що підтверджує її стабільність і ефективність при прогнозуванні діастолічного тиску.

Натомість моделі ARIMA та SARIMA мали вищі значення RMSE та MAE для обох типів тиску, причому SARIMA показала найбільші похибки (CAT RMSE = 9.60, MAE = 7.50). Це може свідчити про те, що для аналізованих часових рядів сезонна компонента або відсутня, або слабо виражена, а тому сезонні моделі не дають переваги.

Таблиця 1

Model	CAT RMSE	CAT MAE	ДАТ RMSE	ДАТ MAE
ARIMA	8.100088	5.902779	5.953944	5.037023
SARIMA	9.604434	7.499918	6.028094	5.144817
Holt-Winters	7.042911	4.553356	6.323395	4.862390

Висновки. Загалом, результати цього дослідження підтверджують, що Holt-Winters є найбільш гнучкою і точною моделлю для короткострокового прогнозування артеріального тиску за відсутності систематичних сезонних коливань. Водночас, модель SARIMA може бути корисною в ситуаціях, коли сезонність є більш вираженою. Модель ARIMA варто розглядати переважно як базовий підхід для оцінки загальних тенденцій без акценту на сезонні впливи.

Отримані результати підкреслюють важливість використання методів аналізу часових рядів для глибокого розуміння індивідуальних змін артеріального тиску. Лінійна регресія дозволила виділити основні довгострокові тренди, тоді як STL-декомпозиція допомогла розглянути сезонні патерни та нелінійні тренди, пов'язані з фізіологічними або зовнішніми чинниками.

Прогнозування було зосереджене на короткостроковому горизонті (лише 15 останніх точок), результати свідчать, що для побудови точніших і надійніших довгострокових прогнозів необхідно використовувати значно більші обсяги історичних даних. Понад 100 спостережень – це мінімальний обсяг для досягнення прийнятної точності, але для виявлення довгострокових трендів та стабільних сезонних патернів може знадобитися ще більше вимірювань.

Крім того, для покращення точності таких моделей доцільно використовувати додаткові метрики, які можуть враховувати ширший спектр фізіологічних та зовнішніх факторів. Це можуть бути показники серцевого ритму, варіабельність пульсу, рівень фізичної активності, стресові фактори або навіть зміни температури навколишнього середовища. Інтеграція таких параметрів у прогнозні моделі може суттєво підвищити їхню точність та адаптивність, особливо для довгострокових оцінок.

Таким чином, перспективи персоналізованого прогнозування артеріального тиску значною мірою залежать від якості та тривалості збору даних. Розширення спектру вимірюваних параметрів та збільшення обсягів історичних спостережень може сприяти створенню більш точних індивідуальних профілів ризику, що враховують реальну динаміку стану пацієнта.

Список використаних джерел:

1. Батурінець А. Г., Антоненко С. В. Ідентифікація складових часового ряду гідрологічних даних. Актуальні проблеми автоматизації та інформаційних технологій. 2018. В. 22. С. 16–29.
2. Андрусенко Ю. О. Аналіз основних моделей прогнозування часових рядів. Кібернетика та системний аналіз. 2020. В. 3(65). С. 91–96. DOI: 10.30748/zhups.2020.65.14
3. Перцев Ю. О., Коротка Л. І. Порівняльний аналіз традиційних статистичних методів та нейромережевої моделі LSTM. Системні технології. 2025. В. 1(156), С. 65–77. DOI 10.34185/1562-9945-1-156-2025-08
4. Trends and seasonality extracting from Home Blood Pressure Monitoring readings. Mendeley data. 2017. URL: <https://data.mendeley.com/datasets/pjcv33hpp7/1>
5. Montgomery, D. C., Peck, E. A., & Vining, G. G. Introduction to Linear Regression Analysis (5th ed.). 2012. P. 12–13.
6. Cleveland R. B., Cleveland W. S., McRae J. E., Terpenning I. STL: A Seasonal-Trend Decomposition Procedure Based on Loess. 1990. *Journal of Official Statistics*. Vol. 6(1). P. 3–73.
7. Cao Y., et al. ARIMA model for short-term forecasting of blood pressure. *Scientific Reports*. 2020. Vol. 10(1). P. 3452.
8. Box G. E. P., Jenkins G. M., Reinsel G. C. Time Series Analysis: Forecasting and Control (5th ed.). Wiley. 2015. P. 712.
9. Holt C. C. "Forecasting seasonals and trends by exponentially weighted moving averages." *International Journal of Forecasting*. 2004. Vol. 20(1). P. 5–25.

References:

1. Baturynets, A. H., Antonenko, S. V. (2018). Identyfikatsiia skladovykh chasovoho riadu hidrolohichnykh danykh. *Aktualni problemy avtomatyzatsii ta informatsiinykh tekhnolohii*, 22, 16–29.
2. Andrusenko, Yu. O. (2020). Analiz osnovnykh modelei prohnnozuvannia chasovykh riadiv. *Kibernetika ta systemnyi analiz*, 3(65), 91–96. <https://doi.org/10.30748/zhups.2020.65.14>
3. Pertsev, Yu. O., Korotka, L. I. (2025). Porivnialnyi analiz tradytsiinykh statystychnykh metodiv ta neiromerezhevoi modeli LSTM. *Systemni tekhnolohii*, 1(156), 65–77. <https://doi.org/10.34185/1562-9945-1-156-2025-08>
4. Trends and seasonality extracting from Home Blood Pressure Monitoring readings. (2017). Mendeley data. Retrieved from: <https://data.mendeley.com/datasets/pjcv33hpp7/1>
5. Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). Introduction to Linear Regression Analysis (5th ed.), Wiley, 12–13.
6. Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. (1990). STL: A Seasonal-Trend Decomposition Procedure Based on Loess. *Journal of Official Statistics*, 6(1), 3–73.
7. Cao, Y., et al. (2020). ARIMA model for short-term forecasting of blood pressure. *Scientific Reports*, 10(1), 3452.
8. Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2015). Time Series Analysis: Forecasting and Control (5th ed.). Wiley, 712.
9. Holt, C. C. (2004). Forecasting seasonals and trends by exponentially weighted moving averages. *International Journal of Forecasting*, 20(1), 5–25.

Дата надходження статті: 26.10.2025

Дата прийняття статті: 17.11.2025

Опубліковано: 30.12.2025

Пасічник А. М., доктор фізико-математичних наук, професор, професор кафедри математичного моделювання та системного аналізу Дніпровського державного технічного університету
ORCID: 0000-0002-8561-1374

Надригайло Т. Ж., кандидат технічних наук, доцент, доцент кафедри математичного моделювання та системного аналізу Дніпровського державного технічного університету
ORCID: 0000-0003-1239-5946

МАТЕМАТИЧНА МОДЕЛЬ ТА ОПТИМІЗАЦІЯ ПАРАМЕТРІВ ПНЕВМАТИЧНОГО РОТАЦІЙНО-ПОРШНЕВОГО ДВИГУНА

Особливу роль в організації сучасного виробництва відіграє застосування засобів малої механізації, зокрема ручних механізованих інструментів які дозволяють суттєво підвищити продуктивність, знизити потребу в робочій силі і покращити умови праці. В якості привода в таких інструментах найбільшого застосування набули пневматичні двигуни, що пояснюється високим рівнем їх експлуатаційної безпечності та надійності. Разом з тим, найбільш поширені ротаційні пластинчасті двигуни мають певні недоліки: низький коефіцієнт корисної дії, швидке зношування робочих лопаток, що призводить до збільшення неефективного споживання стисненого повітря і скорочення терміну експлуатації. Також виникають проблеми з підвищенням рівнем шуму, великим споживанням повітря та надмірною вібрацією, що перевищує санітарні норми. Для вирішення вказаних проблем запропоновано принципово новий пневматичний ротаційно-поршневий двигун в конструкції якого відсутні лопатки розмежування робочих камер, що забезпечує підвищення ефективності його техніко-економічних характеристик та зменшення шуму і вібрацій. Відповідно розробка методів розрахунку та оптимізації параметрів такого двигуна має важливе значення.

В даній роботі представлена математична модель функціонування пневматичного ротаційно-поршневого двигуна та результати оптимізації його технічних характеристик проведені на основі газодинамічного розрахунку. Розрахунки основних параметрів та їх оптимізація проведені для заданих габаритів, потужності та числа обертів пневматичного двигуна.

Актуальність теми представлено дослідження обумовлена зростаючою потребою в економічно ефективних і екологічно безпечних технічних рішеннях у машинобудуванні та інших сферах промисловості, оскільки зменшення споживання енергії та підвищення ефективності виробничих процесів є пріоритетними завданнями сучасної техніки.

Ключові слова: пневматичний ротаційно-поршневий двигун, оптимізація технічних параметрів, газодинамічний розрахунок.

Pasichnyk A. M., Nadryhailo T. Zh. Mathematical model and optimization of parameters of and pneumatic rotary piston engine

A special role in the organization of modern production is played by the use of small-scale mechanization, in particular manual mechanized tools that allow to significantly increase productivity, reduce the need for labor and improve working conditions. Pneumatic engines have become the most widely used drive in such tools, which is explained by the high level of their operational safety and reliability. At the same time, the most common rotary vane engines have certain disadvantages: low efficiency, rapid wear of working blades, which leads to an increase in inefficient consumption of compressed air and a reduction in service life. There are also problems with increased noise levels, high air consumption and excessive vibration exceeding sanitary standards. To solve these problems, a fundamentally new pneumatic rotary-piston engine has been proposed, the design of which does not have blades separating the working chambers, which ensures an increase in the efficiency of its technical and economic characteristics and a reduction in noise and vibrations. Accordingly, the development of methods for calculating and optimizing the parameters of such an engine is of great importance.

This paper presents a mathematical model of the functioning of a pneumatic rotary-piston engine and the results of optimizing its technical characteristics based on gas-dynamic calculations. According to the results of the optimization of the engine torque, it was established that to ensure smooth rotation of its shaft, it is necessary to use designs with three working chambers offset along the circumference by $2/3\pi$. Calculations of the main parameters and their optimization are carried out for the given dimensions, power and speed of the pneumatic engine.

© А. М. Пасічник, Т. Ж. Надригайло, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

The relevance of the topic of the presented research is due to the growing need for cost-effective and environmentally friendly technical solutions in mechanical engineering and other areas of industry, since reducing energy consumption and increasing the efficiency of production processes are priority tasks of modern technology.

Key words: pneumatic rotary piston engine, optimization of technical and technological parameters, gas-dynamic calculation.

Постановка проблеми. Широке використання ручного інструменту з пневматичним приводом пояснюється рядом переваг порівняно з іншими аналогами. До основних переваг можна віднести: надійність і простота в експлуатації; портативність; стійкість до перевантажень; можливість роботи в складних умовах (вологість, запиленість); відносна простота конструкції та низька вартість. Але разом із зазначеними перевагами, найбільш поширені ротаційні пластинчасті двигуни мають декілька значних недоліків:

- швидкий знос робочих лопаток, що призводить до підвищеного непродуктивного витрату стисненого повітря та зменшення ресурсу роботи;
- інтенсивний шум вихлопу, викликаний неконтрольованим розширенням стисненого повітря;
- обмежені габарити, де оптимальне співвідношення довжини ротора до діаметра становить від 1,5 до 1,7;
- надмірний рівень вібрації, який перевищує санітарні норми через наявність незбалансованих мас лопаток;
- підвищене споживання стисненого повітря через недостатню герметичність робочої камери та повітроводів.

Тому проблема пошуку нових конструктивних рішень та розробки методів розрахунку і оптимізації їх функціональних параметрів має достатньо актуальне як практичне так і наукове значення.

Аналіз останніх досліджень і публікацій. Проблема удосконалення та систематизації теоретичних методів розрахунку технічних параметрів двигунів механізованих машин присвячено цілий ряд наукових праць. Так в монографії [1] проведено аналіз відповідних математичних моделей та новітніх технологій управління технічними системами. Результати дослідження характеристик ротаційного пластинчастого двигуна з урахуванням термодинамічних процесів у робочій камері наведені в роботі [2]. Розрахунки показали, що через недостатній об'єм робочої камери процес розширення відбувається занадто інтенсивно (час розширення $< 0,1$ с). Це призводить до того, що реальний режим роботи двигуна значно відрізняється від розрахункового, що знижує коефіцієнт корисної дії (ККД). Також досліджено вплив вологості стисненого повітря на роботу двигуна. Виявилося, що вологість повітря незначно підвищує ККД завдяки теплу пароутворення.

В роботі [3] опубліковані результати експериментальних досліджень пневматичного двигуна з підведенням тепла для покращення його інтегральної продуктивності. Проблема використання енергії стисненого повітря в якості технологій енергозбереження та відновлювальної енергетики присвячені публікації [4–5]. Методи оцінки енергоефективності систем перетворення енергії стисненого повітря в механічну роботу розглядаються в роботі [6].

Результати дослідження та оптимізації конструктивних параметрів пневматичного ротаційного двигуна викладені в роботі [7]. Визначено, що товщина лопаток, які формують робочі камери, зменшує величину крутного моменту на валу двигуна, що також негативно впливає на величину ККД.

Результати дослідження експлуатаційних характеристик пневматичного автомобільного двигуна та методи термодинамічного розрахунку перетворення енергії стиснутого повітря в механічну роботу наведені в статті [8]. Математична модель пневматичного двигуна з клапанним повітророзподілом та результати дослідження його робочих процесів представлені в публікації [9].

Математична модель позиційного пневматичного приводу з використанням безштокового пневматичного циліндра та запропонована в роботі [10]. Проведено дослідження динамічних характеристик системи з метою підвищення експлуатаційних характеристик приводу. У роботі [11] наведено результати аналізу функціональних параметрів пневматичного ротаційного пластинчастого двигуна в залежності від якості використовуваного повітря. У ній визначено оптимальні кути, що задають положення впускних каналів, діаметр каналів, кількість впускних отворів та число лопаток, виходячи з умов досягнення максимальної потужності.

Загалом, більшість досліджень спрямовані на вдосконалення технології перетворення енергії стисненого повітря в механічну роботу із використанням пневматичних двигунів. При цьому встановлено, що конструкція найбільш поширених ротаційних пластинчастих двигунів має такі недоліки:

- Лопатки постійно створюють гальмівний момент, що зменшує ресурс роботи та збільшує шум.
- Недостатній об'єм робочої камери призводить до відхилення реального робочого режиму від розрахункового, що зрештою знижує ККД.

Метою даного дослідження є розробка математичної моделі та виконання розрахунку характеристик принципово нової конструкції пневматичного ротаційно-поршневого двигуна (РПД). Проведення

оптимізації функціональних параметрів двигуна для підвищення його продуктивності та експлуатаційної ефективності, зменшення витрат енергії, покращення умов праці операторів.

Виклад основного матеріалу. Принцип роботи пневматичного ротаційно-поршневого двигуна заснований на перетворенні циклічного переміщення поршня в обертальний рух валу двигуна. Переміщення поршня здійснюється в результаті розширення стисненого повітря, що подається в робочу камеру. Принципова схема роботи такого двигуна представлена на (рис. 1).

Під дією стисненого повітря ротор (1) здійснює обертальний рух навколо осі статора (О), що призводить до зміни об'єму робочої камери. Обертання відбувається завдяки різниці тисків повітря в робочій камері (5) і камері вихлопу (6). Стиснене повітря подається в робочу камеру через підвідні канали (7), а відпрацьоване повітря виходить в атмосферу через відвідні канали (8). Ротор постійно контактує з поверхнею статора (2) через аеростатичний підшипник (3), що забезпечує мінімальні механічні втрати завдяки усуненню тертя. Застосування аеростатичних опор вирішує проблему герметизації конструкції, а також знижує знос системи.

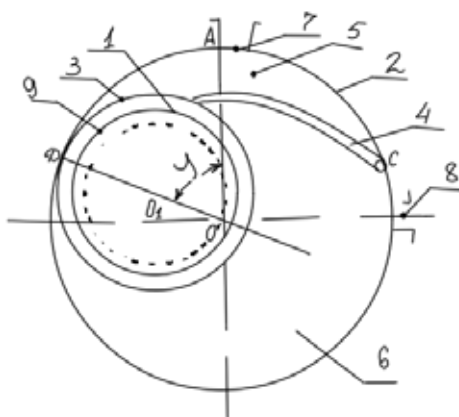


Рис. 1. Принципова схема роботи пневматичного ротаційно-поршневого двигуна

Вісь ротора і статора зміщені ексцентрично, що призводить до зворотно-поступального руху клапана (4) навколо точки підвісу. Клапан переміщується на величину подвійного ексцентриситету. При обертанні ротора і переміщенні клапана об'єм робочої камери двигуна постійно змінюється, що забезпечує неперервний робочий процес. Таким чином, робота пневматичного ротаційно-поршневого двигуна основана на перетворенні енергії стисненого повітря в енергію обертання його робочого валу.

Для побудови аналітичних співвідношень розрахунку об'єму робочої камери двигуна використовується декартова прямокутна система координат, як показано на рис. 2, де відповідно: $R = OD$ – радіус статора (радіус обертання), $r = O_1D$ – радіус ротора, а $e = R - r$ – ексцентриситет зміщення осі поршня відносно осі обертання валу двигуна. Об'єм робочої камери визначається формулою:

$$V = S \cdot l_k, \quad (1)$$

де S – поперечна площа робочої камери ПРД, а l_k – відповідно її довжина.

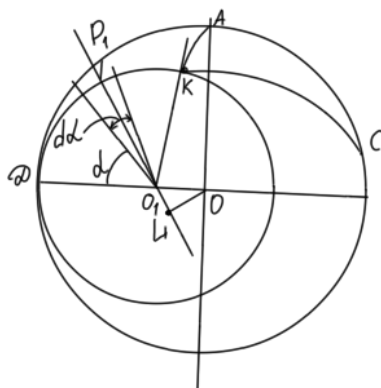


Рис. 2. Схема розрахунку поперечної площі камери РПД

У відповідності з рис. 2 можемо записати, що площа робочої камери визначається такими складовими:

$$S = S_1 - S_2 - S_3 + S_4 - S_5 \quad (2)$$

де S_1 – площа сектора АОД; S_2 – площа сектора ДО₁К; S_3 – площа трикутника АКО; S_4 – площа криво-лінійного трикутника АКС, яка наближено апроксимується площею прямолінійного трикутника АКС; S_5 – площа трикутника О₁КО.

Підставляючи співвідношення для визначення вказаних площ у формулу (2) після перетворень отримаємо:

$$S = \frac{R^2}{2} \varphi - \frac{r^2}{2} \left(\varphi + \arcsin \frac{e \sin \varphi - y_k}{r} \right) - \frac{R y_k}{2} + \frac{a_k^2}{2} \cdot \frac{1}{\sqrt{b^2 + 1}} \quad (3)$$

Зазначимо, що значення кутів у формулах (3) визначаються в радіанах.

Підставляючи (3) в (1), отримаємо вираз для визначення об'єму робочої камери:

$$V = \left[\frac{1}{2} R^2 \varphi - \frac{1}{2} r^2 \left(\varphi + \arcsin \frac{e \sin \varphi - y_k}{r} \right) - \frac{1}{2} R y_k + \frac{1}{2} a_k^2 \frac{1}{\sqrt{b^2 + 1}} \right] \cdot l_k \quad (4)$$

Для оцінки ефективності використання об'єму введемо відповідний коефіцієнт, що визначається відношенням робочої камери до габаритного об'єму двигуна:

$$K_V = V_{раб} / V_{габ}; \quad V_{габ} = \pi R^2 \cdot l_k \quad (5)$$

Цей коефіцієнт характеризує корисність використання об'єму і є критерієм ефективності ПРД. Підставляючи співвідношення (4) в формулу (5) отримаємо:

$$K_V = \frac{1}{2\pi} \varphi - \frac{1}{2\pi} \left(\frac{r}{R} \right)^2 \left(\varphi + \arcsin \frac{e \sin \varphi - y_k}{r} \right) - \frac{1}{2\pi} \cdot \frac{y_k}{R} + \frac{1}{2\pi} \cdot \left(\frac{a_k^2}{R^2} \right) \cdot \frac{1}{\sqrt{b^2 + 1}} \quad (6)$$

Зауважимо, що в реальних конструкціях двигуна його параметри можуть приймати такі значення:

$$\begin{cases} 0 < e \leq 0,4R \\ 0 < r \leq 0,6R \\ 2e \leq a_k \leq \sqrt{2}R \end{cases} \quad (7)$$

які фактично визначають область існування розв'язку розглянутої задачі.

За один робочий цикл обертання ротора в робочій камері ПРД відбуваються такі процеси:

1. Процес нагнітання стисненого повітря з початковим тиском P_0 , що відбувається при збільшенні об'єму камери від кута $\varphi = \varphi_0$ обертання ротора до $\varphi = \varphi_1$. Система координат обрана так, що $\varphi_0 = 0$. Наближено можна вважати, що цей процес відбувається за постійного тиску в робочій камері:

$$P_1 = P_0 = \text{const} \quad (8)$$

Реальний тиск P_1 дещо нижчий P_0 через втрати під час руху повітря. Наведена рівність справедлива при малому числі обертів ротора двигуна.

2. Процес розширення стисненого повітря до тиску P_2 , що відбувається після припинення подачі повітря ($\varphi = \varphi_1$) розподільчим пристроєм до відповідного тиску P_2 при куті повороту $\varphi = \varphi_2$. За тиску P_a , що відповідає тиску навколишнього середовища, різниця тисків $\Delta P_2 = P_2 - P_a$ є втратами повного напору на вихлопі.

Під час робочого процесу, що відбувається в робочій камері, повітря виконує певну роботу, яка рівна сумі роботи A_1 заповнення робочої камери постійним тиском $P_1 = P_0$ до $\varphi = \varphi_1$ та роботи A_2 власного розширення повітря.

Враховуючи, що цей процес є політропним $P = P_0 (V_1 / V)^n$ модель визначення величини робіт A_1 і A_2 може бути представлена так:

$$A_1 = \int_0^{V_1} (P_0 - P_2) dV = P_0 V_1 \left(1 - \frac{P_2}{P_0} \right), \quad (9)$$

$$A_2 = \int_{V_1}^{V_2} (P - P_2) dV. \quad (10)$$

В результаті інтегрування та проведення перетворень отримуємо:

$$A_{пол} = A_1 + A_2 = A_1' + A_2' + A_3' = P_0 V_1 \frac{n}{n-1} \left[1 - \left(\frac{P_2}{P_0} \right)^{\frac{n-1}{n}} \right]. \quad (11)$$

Аналогічно визначається теоретична робота при повному адіабатичному розширенні:

$$A_{ад} = P_0 V_1 \frac{k}{k-1} \left[1 - \left(\frac{P_a}{P_0} \right)^{\frac{k-1}{k}} \right]. \quad (12)$$

Показник адіабаты для повітря: $k=1,4$.

Відповідний теоретичний адіабатичний ККД визначається так:

$$\eta_{ад_1}^m = \frac{A_{пол}}{A_{ад}}. \quad (13)$$

Зазначимо, що повний адіабатичний ККД визначається співвідношенням:

$$\eta_{ад} = \eta_{ад}^m \cdot \eta_y \cdot \eta_{мех}, \quad (14)$$

де η_y – коефіцієнт витоків, $\eta_{мех}$ – механічний ККД, встановлюються експериментально.

Поділивши обидві частини співвідношення (11) на t , де t – час одного оберту, маємо теоретичну розрахункову потужність на робочому валу двигуна:

$$N_m = G P_0 \frac{n}{n-1} \left[1 - \left(\frac{P_2}{P_0} \right)^{\frac{n-1}{n}} \right]. \quad (15)$$

Теоретична об'ємна витрата повітря визначається так:

$$G = V_1 \omega \eta, \quad (16)$$

де V_1 – об'єм заповнення в м^3 ; ω – коефіцієнт стиснення повітря; η – число обертів ротора двигуна за секунду.

Тоді $[G] = \frac{\text{м}^3}{\text{с}}$, $[P] = \frac{\text{Н}}{\text{м}^2}$, $[N] = \text{Вт}$.

При цьому теоретичний індикаторний крутний момент на валу двигуна:

$$M_i^c = V_1 \cdot \omega \cdot P_0 \frac{n}{n-1} \left[1 - \left(\frac{P_2}{P_0} \right)^{\frac{n-1}{n}} \right]. \quad (17)$$

Зазначимо, що розмірність крутного моменту в системі СІ: $[M] = \text{дж}$.

Ефективний крутний момент на валу двигуна дорівнює сумі крутних моментів M_1 від тиску P_1 в робочій камері та M_2 – від тиску вихлопу P_2 .

$$M = M_1 + M_2. \quad (18)$$

В свою чергу M_1 і M_2 мають складові: M_1' і M_2' – крутний і гальмівний моменти, створювані тисками P_1 і P_2 безпосередньо на поршні та M_1'' і M_2'' – крутний і гальмівний моменти, створювані тисками P_1 і P_2 через клапан.

$$M_1 = M_1' + M_1'', \quad (19)$$

$$M_2 = M_2' + M_2''. \quad (20)$$

Момент M_1' обчислюється інтегруванням елементарного моменту dM_1' по довжині дуги DK, рис. 3:

$$M_1' = \int dM_1'. \quad (21)$$

Елементарна сила, що діє на елемент dS дуги DK по довжині камери:

$$dF = P_1 \cdot l_k dS'. \quad (22)$$

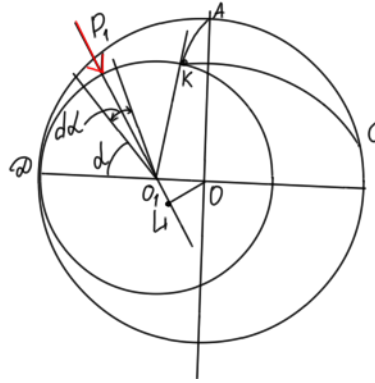


Рис. 3. Схема розподілу сил тиску на клапан і ротор двигуна

Враховуючи, що плече прикладення елементарної сили, що діє на елемент dS дуги DK по довжині камери $l = OL = e \sin \alpha$, в результаті інтегрування отримаємо:

$$M_1' = 2P_1 \cdot l_k \cdot r \cdot e \sin^2 \left(\frac{\varphi}{2} + \frac{1}{2} \arcsin \frac{e \sin \varphi - y_k}{r} \right), \quad (23)$$

$$M_2' = -2P_2 \cdot l_k \cdot r \cdot e \sin^2 \left(\frac{\varphi}{2} + \frac{1}{2} \arcsin \frac{e \sin \varphi - y_k}{r} \right). \quad (24)$$

В результаті обчислень аналогічних наведеним вище для обчислення складових крутного моменту від сили тиску P_1 створюваних через клапан отримуємо такі співвідношення:

$$M_1'' = \frac{P_1}{2} \cdot a_k \cdot l_k \cdot e \sqrt{1 - d^2} \cdot \sin \left(\varphi + \arcsin \frac{e \sin \varphi - y_k}{r} \right), \quad (25)$$

$$M_2'' = -\frac{P_2}{2} \cdot a_k \cdot l_k \cdot e \sqrt{1 - d^2} \cdot \sin \left(\varphi + \arcsin \frac{e \sin \varphi - y_k}{r} \right). \quad (26)$$

Таким чином, з урахуванням отриманих виразів крутний момент на валу двигуна визначається так:

$$M_1 = P_1 \cdot l_k \cdot e \sin \left(\frac{\varphi}{2} + \frac{1}{2} \arcsin \frac{e \sin \varphi - y_k}{r} \right) \times \\ \times \left[2r \sin \left(\frac{\varphi}{2} + \frac{1}{2} \arcsin \frac{e \sin \varphi - y_k}{r} \right) + a_k \sqrt{1 - d^2} \cdot \cos \left(\frac{\varphi}{2} + \frac{1}{2} \arcsin \frac{e \sin \varphi - y_k}{r} \right) \right]. \quad (27)$$

Для спрощення задачі оптимізації розглянемо пневматичний ротаційно-поршневий двигун з одиничними параметрами. Позначимо параметри цього двигуна з рисками:

$$\bar{R} = 1, (\bar{r}, \bar{a}_k, \bar{e}, \bar{l}_k) = 1. \quad (28)$$

Запишемо характеристики цього двигуна. У відповідності з формулою (4) зміна об'єму робочої камери такого двигуна визначається так:

$$V = \frac{1}{2} \left[\varphi - (\bar{r})^2 \left(\varphi + \arcsin \frac{\bar{e} \sin \varphi - \bar{y}_k}{\bar{r}} \right) - \bar{y}_k + \bar{a}_k \frac{1}{\sqrt{\bar{b}^2 + 1}} \right]. \quad (29)$$

Крутний момент із формули (27):

$$\left\{ \begin{array}{l} \text{при } 0 \leq \varphi \leq \varphi_1 : \\ M_1^{ed} = \bar{e} \sin \left(\frac{\varphi}{2} + \frac{1}{2} \arcsin \frac{\bar{e} \sin \varphi - \bar{y}_k}{\bar{r}} \right) \cdot \\ \left[2 \bar{r} \sin \left(\frac{\varphi}{2} + \frac{1}{2} \arcsin \frac{\bar{e} \sin \varphi - \bar{y}_k}{\bar{r}} \right) + \bar{a}_k \sqrt{1 - d^2} \cos \left(\frac{\varphi}{2} + \frac{1}{2} \arcsin \frac{\bar{e} \sin \varphi - \bar{y}_k}{\bar{r}} \right) \right] \\ \text{при } \varphi_1 \leq \varphi \leq \varphi_3 : \\ M_1^{ed} = \left(\frac{V_1^{ed}}{V^{ed}} \right)^n \bar{e} \sin \left(\frac{\varphi}{2} + \frac{1}{2} \arcsin \frac{\bar{e} \sin \varphi - \bar{y}_k}{\bar{r}} \right) \cdot \\ \left[2 \bar{r} \sin \left(\frac{\varphi}{2} + \frac{1}{2} \arcsin \frac{\bar{e} \sin \varphi - \bar{y}_k}{\bar{r}} \right) + \bar{a}_k \sqrt{1 - d^2} \cos \left(\frac{\varphi}{2} + \frac{1}{2} \arcsin \frac{\bar{e} \sin \varphi - \bar{y}_k}{\bar{r}} \right) \right] \end{array} \right. \quad (30)$$

Крутий момент, створюваний на валу двигуна однією робочою камерою, має гармонійну форму наведену на рис. 4. Як видно із рис. 4 крутий момент у пневматичному однокамерному двигуні змінюється від нуля до максимального значення за один робочий цикл, що призводить до нерівномірності обертання робочого валу такого двигуна.

Тому для забезпечення плавності робочого ходу пропонується використовувати три робочих камери зміщених по колу одна відносно іншої на $2/3\pi$. Крутий момент на робочому валу двигуна буде визначатись сумарним значенням крутих моментів від кожної робочої камери. Оскільки камери зміщені одна відносно одної по колу на $2/3\pi$, то круті моменти від кожної робочої камери підсумовуються з періодом $2/3\pi$, рис. 5.

$$M_{\Sigma}^{od} = M_{\frac{1}{\varphi}}^{od} + M_{\frac{1}{\varphi} + \frac{2}{3}\pi}^{od} + M_{\frac{1}{\varphi} + \frac{4}{3}\pi}^{od} \quad (31)$$

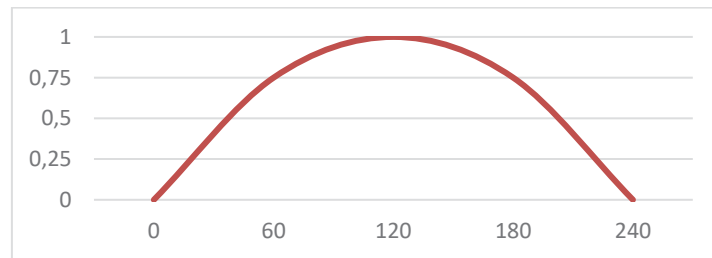


Рис. 4. Залежність крутного моменту за один робочий цикл для однокамерного двигуна

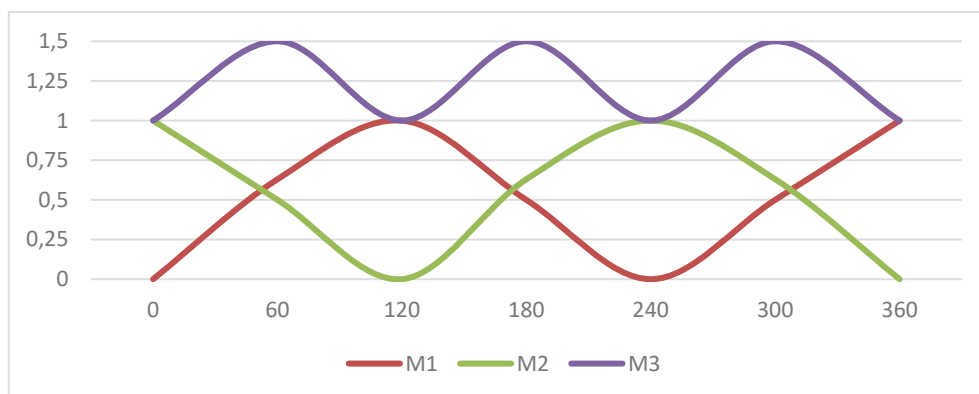


Рис. 5. Залежність крутного моменту за один робочий цикл для трикамерного двигуна

Для характеристики ефективності технічних параметрів пневмодвигуна використовується коефіцієнт нерівномірності моменту:

$$\delta = \frac{M_{\Sigma max}^{od} - M_{\Sigma min}^{od}}{M_{\Sigma ep}^{od}} \quad (32)$$

Величина коефіцієнта нерівномірності моменту характеризує крутизну моменту. Наявність нерівномірності обертання робочого валу пов'язана із збільшенням сумарного крутного моменту, отриманого з кожної робочої камери.

Задача оптимізації характеристик пневмодвигуна полягає у виборі таких конструктивних параметрів які забезпечать:

1. Максимальне значення величини крутного моменту для однієї робочої камери.
2. Середнє значення величини сумарного крутного моменту на валу повинно мати максимальне значення.
3. Величина коефіцієнта нерівномірності моменту має бути мінімальною.

Наведені умови оптимальності зводиться до умови екстремуму функції двох змінних які визначаються так:

$$\begin{cases} \frac{\sigma V_1}{\sigma \bar{a}_k} = 0 \\ \frac{\sigma V_1}{\sigma \bar{e}} = 0 \end{cases} \quad (33)$$

Оптимальні значення конструктивних параметрів визначається із співвідношення (29). Для спрощення викладок величиною $\bar{y}_k$ можна знехтувати, оскільки її порядок значно менший від інших величин, що входять у вираз (29). Диференціюючи (29) по $\bar{a}_k$ та $\bar{e}$ отримаємо такі значення частинних похідних:

$$\begin{cases} \frac{\sigma V_1}{\sigma \bar{a}_k} = \frac{\bar{a}_k}{\sqrt{\bar{b}^2 + 1}} - \frac{\bar{a}_k^2}{2} \cdot \frac{\bar{b}}{(\bar{b}^2 + 1)^{\frac{3}{2}}} \cdot \bar{b} \frac{1}{\bar{a}_k} \\ \frac{\sigma V_1}{\sigma \bar{e}} = (1 - \bar{e}) \left[\frac{2}{3} \pi + \arcsin \left(\sin \frac{2}{3} \pi \frac{\bar{e}}{1 - \bar{e}} \right) \right] - \frac{\pi}{3} \cdot \frac{1}{\sqrt{1 + \sin \frac{2}{3} \pi \frac{\bar{e}}{1 - \bar{e}}}} + \\ + \frac{1}{2} \cdot \frac{\bar{a}_k^2 \bar{b}}{(\bar{b}^2 + 1)^{\frac{3}{2}}} \cdot \bar{b} \frac{1}{\bar{e}} \end{cases} \quad (34)$$

Із системи рівнянь (34) з урахуванням умов (33) отримаємо систему рівнянь з двома невідомими $\bar{a}_k$ і $\bar{e}$. Розв'язком цієї системи є значенням оптимальних конструктивних параметрів:

$$\begin{cases} \frac{\bar{a}_k}{\sqrt{\bar{b}^2 + 1}} - \frac{1}{2} \cdot \frac{\bar{a}_k^2 \bar{b}}{2(\bar{b}^2 + 1)^{\frac{3}{2}}} \cdot \bar{b} \frac{1}{\bar{a}_k} = 0, \\ (1 - \bar{e}) \left[\frac{2}{3} \pi + \arcsin \left(\sin \frac{2}{3} \pi \frac{\bar{e}}{1 - \bar{e}} \right) \right] - \frac{\pi}{3} \cdot \frac{1}{\sqrt{1 + \sin \frac{2}{3} \pi \frac{\bar{e}}{1 - \bar{e}}}} + \\ + \frac{1}{2} \cdot \frac{\bar{a}_k^2 \bar{b}}{(\bar{b}^2 + 1)^{\frac{3}{2}}} \cdot \bar{b} \frac{1}{\bar{e}} = 0. \end{cases} \quad (35)$$

Розв'язок системи (35) здійснюється чисельно методом послідовних наближень і визначає оптимальні значення конструктивних параметрів.

Висновки. В даній роботі було проведено аналіз проблем підвищення ефективності ротаційно-поршневих пневматичних двигунів, що є важливим завданням для забезпечення продуктивності та економічності сучасних механізованих інструментів. Показано, що традиційні ротаційні пластинчасті двигуни мають ряд недоліків, таких як низький коефіцієнт корисної дії, швидке зношування робочих лопаток, підвищений рівень шуму та вібрації. Показано, що для підвищення експлуатаційної ефективності ручних машин пошук нових технологічних рішень із удосконалення їх енергетичних приводів має важливе практичне значення.

У ході дослідження було розроблено методику розрахунку нової конструкції пневматичного ротаційно-поршневого двигуна, яка не має лопаток розмежування робочих камер. Проведено математичне моделювання, газодинамічні розрахунки та аналіз його основних характеристик. Запропоновано методику оптимізації конструктивних параметрів двигуна, що забезпечило підвищення ефективності його експлуатаційних параметрів. Оптимізація конструктивних параметрів нового двигуна дозволила: знизити енергоспоживання; підвищити продуктивність механізованих інструментів; зменшити вплив шкідливих факторів на навколишнє середовище; забезпечити економічну доцільність та екологічну безпечність виробничих процесів.

Отримані результати підтверджують актуальність дослідження та перспективність подальшої розробки технічної документації та впровадження запропонованої конструкції пневматичного двигуна у виробництво механізованих інструментів.

Список використаних джерел:

1. Математичні моделі та новітні технології управління економічними та технічними системами: монографія / за заг. ред В.О. Тимофєєва, І.В. Чумаченко. Харків: ФОП Панов А.М., 2018. 314 с.
2. Барсов В. І., Краснобаєв В. А., Барсова З. В. Математичні методи та технічні засоби АСУ : підруч. для студентів вищ. навч. закл. Укр. інж.-пед. акад. Харків: Точка, 2012. 302 с.
3. Fang, Y., Lu, Y., Yu, X., Roskilly, A.B. Experimental study of a pneumatic engine with heat supply to improve the overall performance. *Applied Thermal Engineering*. 2018. V. 134. P.78-85. Doi: 10.1016/i.applthermaleng.2018.01.113.
4. Saidur R., Rahim N. A., Hasanuzzaman M. A review on compressed-air energy use and energy savings. *Renewable and Sustainable Energy Reviews*. 2010. № 14. P. 1135–1153.
5. Chong L., Dewen K., Maolin C. Research on Energy-Saving Operation of Screw Air Compressor. *Research Journal of Applied Sciences, Engineering and Technology*. 2013. № 6 (2). P.325–333.
6. Mousavi S., Kara S., Kornfeld B. Energy Efficiency of Compressed Air Systems. *21st CIRP Conference on Life Cycle Engineering. Procedia CIRP*. 2014. № 15. P.313–318.
7. Прокопов М. Г., Ванєєв С. М., Козін В. М., Мерзляков Ю. С. Конструкції елементів пневмоагрегатів: навч. посіб. Сумський державний університет. Суми: 2020. 146 с.
8. Zinko R., Kramskyi O., Mokriak B., & Polyakov A. Research on the operational characteristics of a pneumatic automotive engine. *Journal of Mechanical Engineering and Transport*. 2024. V. 10(2). P.48-53. Doi: 10.63341/vjmet/2.2024.48.
9. Leontiev D., Voronkov O., Korohodskyi V., Hlushkova D., Nikitchenko I., Teslenko E., Lykhodii O. Mathematical Modelling of Operating Processes in the Pneumatic Engine of the Car. *SAE Technical Paper 2020-01-2222*, 2020. ISSN: 0148-7191. DOI: 10.4271/2020-01-2222
10. Касьян А. В., Галецький О. С. Розробка математичної моделі пневматичного позиційного приводу. *Вісник ХНТУ*. 2025. № 1(92). С. 79–88. Doi: 10.35546/kntu2078-4481.2025.1.1.10
11. Петриченко С. В., Паляничка Н. О., Тарасенко В. Г. Монтажні інструменти і пристосування: навч. посіб. ТДАТУ, 2021. URL: https://elib.tsatu.edu.ua/dep/mtf/ophv_35/page25.html.

References:

1. Timofeyev, V. O., Chumachenko, I. V. (2018). Matematychni modeli ta novitni tekhnolohiyi upravlinnya ekonomichnymi ta tekhnichnymi systemamy: monohrafiya [Mathematical models and new technologies for managing economic and technical systems: monograph]. Kharkiv: FOP Panov A.M. 314 s.
2. Barsov, V. I., Krasnobayev, V. A., Barsova, Z. V. (2012). Matematychni metody ta tekhnichni zasoby ASU [Mathematical methods and technical means of ACS]: pidruch. dlya studentiv vyshch. navch. zakl. Ukr. inzh.-ped. akad. Kharkiv: Tochka. 302 s.
3. Fang, Y., Lu, Y., Yu, X., Roskilly, A.B. (2018). Experimental study of a pneumatic engine with heat supply to improve the overall performance. *Applied Thermal Engineering*. V. 134. P.78-85. Doi: 10.1016/i.applthermaleng.2018.01.113.
4. Saidur, R., Rahim, N. A., Hasanuzzaman, M. (2010). A review on compressed-air energy use and energy savings. *Renewable and Sustainable Energy Reviews*. № 14. P.1135–1153.
5. Chong, L., Dewen, K., Maolin, C. (2013). Research on Energy-Saving Operation of Screw Air Compressor. *Research Journal of Applied Sciences, Engineering and Technology*. № 6 (2). P.325–333.
6. Mousavi, S., Kara, S., Kornfeld, B. (2014). Energy Efficiency of Compressed Air Systems. *21st CIRP Conference on Life Cycle Engineering. Procedia CIRP*. № 15. 313–318.
7. Prokopov, M. H., Vanyeyev, S. M., Kozin, V. M., Merzlyakov, YU. S. (2020). Konstruktsiyi elementiv pnevmoahtativ: navch. Posib [Design of pneumatic unit elements]. Sums'kyi derzhavnyi universytet. Sumy: – 146 s.
8. Zinko, R., Kramskyi, O., Mokriak, B., & Polyakov, A. (2024). Research on the operational characteristics of a pneumatic automotive engine. *Journal of Mechanical Engineering and Transport*. V. 10(2). P.48–53. Doi: 10.63341/vjmet/2.2024.48.

9. Leontiev, D., Voronkov, O., Korohodskyi, V., Hlushkova, D., Nikitchenko, I., Teslenko, E., Lykhodii, O. (2020). Mathematical Modelling of Operating Processes in the Pneumatic Engine of the Car. *SAE Technical Paper 2020-01-2222*. ISSN: 0148-7191. Doi:10.4271/2020-01-2222.

10. Kas'yan, A. V., Halets'kyi, O. S. (2025). Rozrobka matematychnoyi modeli pne-vmatychnoho pozytsiynoho pryvodu [Development of a mathematical model of a pneumatic position drive]. *Visnyk KHNTU*. № 1(92). S.79–88. Doi: 10.35546/kntu2078-4481.2025.1.1.10

11. Petrychenko, S. V., Palyanychka, N. O., Tarasenko, V. H. (2021). Montazhni instrumenty i prystosuvannya: navch. Posib [Assembly tools and fixtures]. TDATU. Retrieved from: https://elib.tsatu.edu.ua/dep/mtf/ophv_35/page25.html.

Дата надходження статті: 23.10.2025

Дата прийняття статті: 17.11.2025

Опубліковано: 30.12.2025

УДК 378, 519.876.5

DOI <https://doi.org/10.32782/2521-6643-2025-2-70.7>

Сафронова І. А., кандидат фізико-математичних наук,
доцент кафедри комп'ютерних технологій
Дніпровського національного університету імені Олеся Гончара
ORCID: 0000-0003-0242-3798

Жир С. І., кандидат фізико-математичних наук, доцент,
доцент кафедри транспортних технологій та міжнародної логістики
Університету митної справи та фінансів
ORCID: 0009-0006-2410-6792

Верба О. В., старший викладач кафедри комп'ютерних технологій,
завідувач навчального відділу
Дніпровського національного університету імені Олеся Гончара
ORCID: 0000-0003-1030-4377

Шишканова Г. А., кандидат фізико-математичних наук, доцент,
доцент кафедри «Математика»
Національного університету «Запорізька політехніка»
ORCID: 0000-0002-0336-2803

АДАПТИВНІ СТРАТЕГІЇ ВИКЛАДАННЯ ТА ОЦІНЮВАННЯ В УМОВАХ ВІЙНИ

Сучасна система вищої освіти в Україні зіткнулася з безпрецедентними викликами, зумовленими війною. Традиційні педагогічні підходи до організації навчання та оцінювання виявилися неефективними в умовах постійної психологічної напруги, технічних перебоїв і великих студентських потоків. Це створює надмірне навантаження на викладачів, роблячи індивідуальний підхід майже неможливим. Масове неотримання дедлайнів призводить до хаосу перевірки та девальвації традиційних механізмів контролю.

Метою роботи є аналіз цих проблем та розробка адаптивних підходів до організації освітнього процесу, що дозволять підтримувати якість освіти, мінімізуючи навантаження на викладачів.

У роботі запропоновано переосмислення підходів до проектування завдань. Замість великої кількості об'ємних робіт, пропонується зосередитися на 2-3 ключових завданнях за семестр, орієнтованих на процес і що вимагають інтеграції знань. Акцентується увага на використанні групових проєктів. Особлива увага приділяється адаптації системи зворотного зв'язку та оцінювання. Замість індивідуальних коментарів, пропонується мінімалістичний та масштабований фідбек (наприклад, «світлофор», шаблони, загальні семінари). Розглянуто комбіновані методи перевірки, що поєднують автоматизовані тести з елементами захисту від академічної доброчесності та використанням ШІ, а також вибіркові усні співбесіди.

Важливим аспектом є зміщення фокуса на внутрішню мотивацію студентів через проєкти, що мають реальну цінність, а також запровадження дієвих інструментів заохочення, зокрема, бонусних балів за активність та своєчасність. У роботі також наголошується на необхідності системної психологічної та методичної підтримки викладачів з боку адміністрації ЗВО, що передбачає адаптацію вимог до обсягів перевірки та надання доступу до сучасних цифрових інструментів. Висновки підкреслюють, що запропоновані рішення є компромісними, але ефективними стратегіями, які дозволяють підтримувати якість освітнього процесу та академічну доброчесність, мінімізуючи при цьому ризики професійного вигорання викладачів в умовах непередбачуваності. Результати роботи можуть слугувати основою для подальших досліджень та розробки практичних рекомендацій для закладів вищої освіти, що функціонують у кризових умовах.

Ключові слова: вища освіта, дистанційне навчання, війна, навантаження викладача, зворотний зв'язок, оцінювання, адаптивні стратегії, мотивація студентів.

© І. А. Сафронова, С. І. Жир, О. В. Верба, Г. А. Шишканова, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

Safronova I. A., Zhyr S. I., Verba O. V., Shyshkanova G. A. Adaptive strategies of teaching and assessment in wartime conditions

The modern system of higher education in Ukraine has faced unprecedented challenges due to the war. Traditional pedagogical approaches to teaching and assessment have proven ineffective amid constant psychological stress, technical disruptions, and large student cohorts. This creates an excessive workload for lecturers, making an individual approach nearly impossible. The widespread failure to meet deadlines leads to a chaotic verification process and devalues traditional control mechanisms.

The purpose of this work is to analyze these problems and develop adaptive approaches to organizing the educational process that will help maintain the quality of education while minimizing the workload on lecturers.

The paper proposes a rethinking of task design. Instead of a large number of lengthy assignments, it suggests focusing on 2-3 key, process-oriented tasks per semester that require knowledge integration. The importance of using group projects is also emphasized. Special attention is paid to adapting the feedback and assessment system. Instead of individual comments, minimalist and scalable feedback (e.g., a "traffic light" system, templates, general seminars) is proposed. It also considers combined verification methods that integrate automated tests with elements to prevent academic dishonesty and the misuse of AI, as well as selective oral interviews.

The work also highlights the importance of fostering students' intrinsic motivation through projects with real-world application. It emphasizes the need for systematic psychological and methodological support for lecturers from the university administration, including adapting assessment requirements and providing access to modern digital tools. The conclusions underscore that the proposed solutions are effective, compromise-based strategies that help maintain the quality of education and academic integrity while minimizing the risk of professional burnout in unpredictable conditions. The results can serve as a basis for developing practical recommendations for higher education institutions operating in a crisis.

Key words: higher education, distance learning, war, lecturer's workload, feedback, assessment, adaptive strategies, student motivation.

Постановка проблеми. Система вищої освіти України перебуває в стані глибокої структурної трансформації, зумовленої поєднанням глобальних освітніх тенденцій, зокрема, цифровізації, персоналізації та автономізації навчального процесу, та екзогенних чинників надзвичайного характеру, пов'язаних із повномасштабною війною. Синергія цих впливів спричинила необхідність перегляду усталених парадигм організації освітньої діяльності, особливо в аспектах оцінювання навчальних досягнень, забезпечення ефективного зворотного зв'язку та підтримки академічної мотивації здобувачів освіти.

Одним із системних викликів сучасного освітнього середовища є істотне зростання обсягів самостійної роботи студентів, що, у поєднанні з тенденцією до укрупнення навчальних потоків, призводить до значного ускладнення процесу індивідуалізованого педагогічного супроводу. У таких умовах забезпечення своєчасного, якісного та змістовного фідбеку, а також оперативної перевірки навчальних завдань, стає дедалі більш ресурсозатратним для викладацького складу. Цей процес додатково ускладнюється низкою об'єктивних обмежень, зумовлених воєнним станом: нестабільністю енергопостачання, періодичними повітряними тривогами, нерівномірним доступом до цифрової інфраструктури та високим рівнем психоемоційного напруження серед учасників освітнього процесу.

Зазначені чинники суттєво порушують ритм навчання, унеможливають дотримання жорстких дедлайнів та знижують ефективність традиційних механізмів академічного контролю, оцінювання та мотиваційного стимулювання. У результаті викладачі опиняються перед необхідністю одночасного забезпечення академічної якості освітнього контенту, підтримки мотиваційної сфери студентства та збереження власної професійної стійкості в умовах обмежених ресурсів, як матеріальних, так і когнітивно-емоційних.

Це зумовлює потребу в розробці та впровадженні адаптивних, масштабованих і технологічно обґрунтованих освітніх стратегій, здатних забезпечити функціональну сталість освітнього процесу, зберігаючи при цьому гуманітарну орієнтацію, педагогічну ефективність та допустиме навантаження на академічний персонал. Такий підхід передбачає переосмислення ролі викладача як фасилітатора навчання, інтеграцію цифрових інструментів оцінювання, гнучких моделей академічної взаємодії та механізмів психолого-педагогічної підтримки в умовах високої соціальної невизначеності.

Аналіз останніх досліджень і публікацій. Аналіз сучасних наукових джерел та нормативно-правових актів засвідчує, що трансформація освітнього процесу в Україні в умовах воєнного стану є не лише вимушеним адаптаційним кроком, а й каталізатором переосмислення фундаментальних принципів педагогічної взаємодії, цифрової інфраструктури та академічної автономії. Законодавчі документи, зокрема Закон України «Про освіту» та «Про наукову і науково-технічну діяльність» [1, 2], формують нормативне підґрунтя для реалізації гнучких освітніх моделей, що враховують виклики нестабільного середовища.

У межах нормативного дискурсу особливу увагу привертають методичні рекомендації щодо організації навчання в умовах воєнного стану [5, 11], а також регіональні ініціативи [4], які акцентують на безпекових аспектах, адаптивності навчальних програм та необхідності збереження психоемоційного балансу учасників освітнього процесу. Ці документи не лише регламентують освітню діяльність, а й задають вектор її гуманітарного переосмислення.

Наукові публікації останніх років демонструють зростаючий інтерес до концепту цифрової стійкості, що розглядається як ключовий чинник забезпечення безперервності освітнього процесу. Зокрема, у роботі [3] обґрунтовано роль інформаційних технологій як інструменту підтримки академічної стабільності в умовах кризових викликів. У цьому контексті дослідження [8] набуває особливої актуальності, демонструючи потенціал мобільних технологій для інтеграції освіти у повсякденне життя.

Психолого-педагогічний вимір адаптації розкривається у праці [6], яка аналізує стратегії розвитку критичного мислення у студентів ЗВО в умовах освітньої турбулентності. У роботі [7] досліджено специфіку організації теоретичних онлайн-занять, що є особливо релевантним у контексті масових потоків та обмеженого доступу до ресурсів. [8] акцентує на викликах викладання гуманітарних дисциплін у регіональних ЗВО, а [9] пропонує практичні моделі дистанційної педагогіки, зокрема методи стимулювання навчальної активності та оптимізації оцінювання.

Таким чином, аналіз наявного наукового дискурсу засвідчує не лише високу актуальність досліджуваної проблематики, а й виявляє широку міждисциплінарну площину для подальших теоретичних і прикладних розвідок у сфері освітньої трансформації. Корпус літератури, що охоплює нормативно-правові акти, емпіричні дослідження, педагогічні концепції та технологічні інновації, дозволяє здійснити системне осмислення адаптаційних процесів у вищій освіті в умовах воєнного стану. Такий підхід передбачає інтеграцію регуляторних механізмів, цифрових інструментів та психолого-педагогічних стратегій у єдину концептуальну рамку, що забезпечує цілісне бачення освітньої реальності в кризовому контексті. У цьому контексті подальші дослідження мають бути спрямовані на розробку адаптивних моделей освітньої взаємодії, інструментів гнучкого оцінювання, механізмів цифрової підтримки та стратегій психоемоційного супроводу, що дозволять забезпечити не лише функціональну ефективність освітнього процесу, а й його гуманітарну легітимність у складних соціокультурних умовах. Сформована теоретична платформа може слугувати основою для нормативного оновлення, інституційного реформування та педагогічного переосмислення вищої освіти України в умовах війни та посткризового відновлення.

Мета статті полягає у здійсненні комплексного наукового аналізу трансформаційних процесів у сфері вищої освіти України в умовах воєнного стану, з акцентом на педагогічні, оцінювальні та комунікативні аспекти освітньої взаємодії. Дослідження спрямоване на виявлення ключових викликів, що постають перед суб'єктами освітнього процесу – викладачами та здобувачами освіти, в умовах системної нестабільності, а також на обґрунтування ефективних адаптивних стратегій, які забезпечують академічну сталість, оптимізують навчальне навантаження та підтримують мотиваційно-ціннісну сферу учасників освітнього середовища. У межах поставленої мети передбачається інтеграція психолого-педагогічних, технологічних та організаційних підходів до модернізації освітніх практик в умовах кризового та посткризового розвитку.

Виклад основного матеріалу. Виклики масштабованості та академічної ефективності в умовах воєнного часу. Система вищої освіти в Україні зіткнулася з безпрецедентними викликами, які кардинально змінили усталені підходи до організації навчального процесу, взаємодії зі студентами та оцінювання. Ці проблеми загострилися на тлі зростання частки самостійної роботи студентів та необхідності адаптуватися до реалій воєнного часу.

Одним із ключових організаційних викликів сучасної вищої освіти є ефективне управління навчальним процесом у межах великих академічних потоків, що особливо актуально для лекційних занять. За умови, коли кількість здобувачів освіти в одному потоці перевищує 150 осіб, реалізація індивідуалізованого підходу в межах лекційного формату суттєво ускладнюється. В умовах дистанційної або змішаної форми навчання можливості для персоналізованої взаємодії викладача із студентами обмежуються, а дидактичний інструментарій викладача зосереджується переважно на фронтальних формах подання матеріалу [7]. Хоча практичні та лабораторні заняття передбачають поділ студентів на менші групи, що все ж допомагає у компенсації обмеження лекційного формату, загальний обсяг навчального навантаження залишається значним.

У зв'язку з цим актуальним постає питання пошуку масштабованих освітніх рішень, які дозволяють зберігати якість навчального процесу за умов високої чисельності студентських груп. До таких рішень належать автоматизовані системи оцінювання, групові формати зворотного зв'язку, а також адаптивні моделі навчання, що поєднують технологічну ефективність із педагогічною доцільністю.

Функціонування освітнього процесу у закладах вищої освіти України в умовах воєнного стану супроводжується низкою об'єктивних обмежень, які суттєво впливають на ефективність традиційних механізмів контролю та мотивації студентів. Зокрема, нестабільність енергопостачання, періодичні повітряні тривоги та загальний психологічний фон створюють умови, за яких жорстке дотримання термінів виконання навчальних завдань втрачає свою педагогічну доцільність. У відповідь на ці виклики окремі заклади вищої освіти запровадили регулятивні положення, що передбачають недопущення зниження оцінок за несвоєчасне подання робіт, що, своєю чергою, трансформує роль дедлайнів як інструменту дисциплінарного регулювання.

Спроби оптимізувати навантаження шляхом поділу об'ємних завдань на менші компоненти з окремими термінами виконання не завжди демонструють очікувану ефективність. За спостереженнями, значна частина студентів, особливо на старших курсах, схильна ігнорувати проміжні дедлайни, що призводить до

накопичення невиконаних завдань і їх масового подання наприкінці навчального періоду. Така концентрація активності у фінальні терміни суттєво ускладнює процес перевірки, збільшує навантаження на викладача та обмежує можливості для поетапного, конструктивного зворотного зв'язку.

Крім того, навіть короткі коментарі до студентських робіт, особливо у випадках, що потребують доопрацювання, можуть ініціювати повторні цикли перевірки, що є малопродуктивним у контексті великої кількості студентів та обмеженого часу. У результаті, традиційні інструменти індивідуального контролю та мотивації потребують переосмислення з урахуванням нових реалій, що зумовлює актуальність пошуку адаптивних, масштабованих і ресурсно збалансованих освітніх рішень.

Адаптаційні стратегії та рішення. В умовах системної дестабілізації освітнього середовища, спричиненої повномасштабною війною, традиційні механізми академічного контролю, мотиваційного стимулювання та індивідуалізованого зворотного зв'язку виявляються частково або повністю нефункціональними. Це зумовлює необхідність концептуального переосмислення організації освітнього процесу, з урахуванням обмежених ресурсів, високої невизначеності та психоемоційної вразливості учасників освітньої взаємодії.

У таких умовах доцільним є зміщення акцентів з формального контролю на адаптивне управління навчальним процесом, що передбачає максимальне спрощення освітніх процедур, автоматизацію рутинних елементів (за умови технічної доступності), концентрацію на досягненні ключових результатів навчання та гнучке коригування академічних вимог відповідно до реалій воєнного часу. Такий підхід відповідає принципам гуманітарно орієнтованої освіти, яка визнає пріоритет людського виміру над формальними показниками.

Запровадження адаптивних стратегій дозволяє не лише підтримувати базову якість освітнього процесу, а й мінімізувати надмірне когнітивне та емоційне навантаження на викладачів, які змушені функціонувати в умовах багаторівневої кризи. До таких стратегій належать: використання цифрових платформ для централізації навчального контенту та зворотного зв'язку; впровадження формульовального оцінювання з акцентом на процес, а не результат; гнучке планування дедлайнів; делегування частини комунікаційних функцій автоматизованим системам (зокрема чат-ботам); а також розвиток педагогічної емпатії як інструменту підтримки мотивації студентів.

Системне впровадження зазначених рішень сприяє формуванню нової моделі освітньої взаємодії, що базується на принципах адаптивності, технологічної раціональності та етичної відповідальності. Вона дозволяє забезпечити академічну сталість навіть у надзвичайних умовах, зберігаючи при цьому гуманітарну чутливість та професійну стійкість викладацького складу.

Цифрові інструменти як засіб адаптації. В умовах воєнного стану, коли освітній процес супроводжується численними обмеженнями – від нестабільного енергопостачання до психологічної напруги – цифрові інструменти відіграють ключову роль у забезпеченні безперервності навчання. Їхнє впровадження дозволяє не лише компенсувати втрату традиційних форм взаємодії, а й створити нові, більш гнучкі моделі організації освітнього процесу.

Одним із найефективніших рішень є використання платформ управління навчанням (LMS), таких як Moodle, Google Classroom та Microsoft Teams. Ці системи забезпечують централізований доступ до навчальних матеріалів, завдань, оцінок і комунікації, що особливо важливо в умовах фрагментованого навчального середовища. LMS-платформи дозволяють викладачам структурувати курс, встановлювати дедлайни, надавати інструкції та здійснювати оцінювання в одному цифровому просторі. Для студентів це створює відчуття стабільності та передбачуваності, що позитивно впливає на мотивацію та академічну дисципліну.

Автоматизовані тести, інтерактивні опитування та формульовальне оцінювання – ще один важливий компонент цифрової адаптації. Вони дозволяють швидко перевіряти знання студентів, надавати миттєвий зворотний зв'язок і знижувати навантаження на викладача. Наприклад, тести з автоматичним підрахунком балів дають змогу оцінити велику кількість студентів без потреби в ручній перевірці, а інтерактивні опитування (наприклад, через Google Forms або Mentimeter) сприяють активному залученню студентів під час онлайн-занять. Формульовальне оцінювання, яке передбачає поетапне надання фідбеку, допомагає студентам краще розуміти свої сильні та слабкі сторони, а викладачам – коригувати навчальний процес відповідно до потреб групи.

Окрему увагу заслуговує впровадження чат-ботів та автоматизованих відповідей у навчальних платформах. У великих академічних потоках, де кількість студентів може перевищувати сотню, чат-боти здатні ефективно відповідати на типові запитання щодо дедлайнів, формату завдань, критеріїв оцінювання тощо. Це значно зменшує кількість повторних звернень до викладача, оптимізує комунікацію та дозволяє зосередитися на більш складних аспектах навчального процесу. Крім того, чат-боти можуть бути інтегровані з LMS або месенджерами, що робить їх доступними для студентів у будь-який час.

Таким чином, для ефективної перевірки робіт та надання конструктивного зворотного зв'язку доцільно застосовувати комбіновані та масштабовані підходи. Комбінована перевірка та використання технологій з урахуванням викликів III є ключовим для економії часу. Доцільно застосовувати тестові завдання з автоматичною перевіркою, але модифікувати їх таким чином, щоб мінімізувати пряме копіювання, наприклад, формулювати завдання, що вимагають аналізу кейсів або застосування знань у нестандартних умовах. Для

критично важливих завдань варто запроваджувати короткі усні співбесіди або бліц-опитування, а також за можливості використовувати програми-антиплагіатори [3, 7].

Замість об'ємних індивідуальних коментарів, необхідно впроваджувати мінімалістичний та масштабований зворотний зв'язок. Це може бути фідбек за принципом «світлофора» (зелений – прийнято, жовтий – прийнято з недоліками, червоний – не прийнято), використання «шаблонних» коментарів або фідбек-лекцій чи семінарів, де аналізуються типові помилки. Групові обговорення та взаємне оцінювання є ефективною альтернативою для надання зворотного зв'язку. Доцільно фокусуватися на якості, а не кількості фідбеку, зосереджуючись на системних або критичних для розуміння матеріалу помилках. Важливо інтегрувати фідбек у самостійну роботу та обмежити кількість «перездач» [7].

Тобто, цифрові інструменти не лише компенсують обмеження, спричинені війною, а й створюють передумови для модернізації освітнього процесу. Їхнє грамотне використання дозволяє забезпечити академічну сталість, знизити навантаження на викладачів та підтримати мотивацію студентів у складних умовах.

Проблема масового «заливання» робіт вимагає перегляду підходів до проектування завдань. Мета полягає у створенні таких завдань, які спонукають до поступової роботи та демонструють індивідуальне розуміння матеріалу, а не лише його копіювання. Доцільно розбивати складні завдання на етапи, кожен з яких має чітко визначений проміжний результат. Також варто формулювати завдання, що вимагають інтеграції знань з різних розділів курсу, та включати рефлексивні і проблемно-орієнтовані завдання, які складніше автоматично генерувати. Запровадження групових завдань є цінним рішенням, що не лише розподіляє навантаження та стимулює командну роботу, а й має важливі соціалізаційні та професійно орієнтовані переваги. Замість великої кількості об'ємних завдань, доцільно зосередитися на 2-3 ключових, добре продуманих завданнях за семестр, які мають бути вагомими для загальної оцінки та дозволять викладачеві зосередити зусилля на якісній перевірці [6, 8].

Ефективність самостійної роботи значною мірою залежить від того, наскільки чітко студенти розуміють поставлені завдання та критерії, за якими їхня робота буде оцінюватися. Для цього рекомендується чітко формулювання вимог до завдань та використання критеріїв оцінювання. Їх надання студентам із зазначенням балів за кожен елемент роботи забезпечує прозорість процесу, а для об'ємних завдань доцільно проводити короткі ознайомчі онлайн-інструктажі [9, 12].

Для стимулювання *внутрішньої мотивації* варто використовувати проекти з «живим» застосуванням та підкреслювати зв'язок з майбутньою професією. Також доцільно застосовувати дієві інструменти заохочення, як-от бонусні бали за активність та наголошення на вплив оцінок на загальний рейтинг і, як наслідок, стипендію.

Психоемоційна підтримка та педагогічна емпатія. В умовах війни освітній процес набуває не лише академічного, а й гуманітарного виміру. Психоемоційна напруга, тривожність, втрата звичного ритму життя – усе це суттєво впливає на здатність студентів до навчання, а викладачів – до ефективного викладання. Тому психоемоційна підтримка та педагогічна емпатія стають критично важливими компонентами освітньої взаємодії. Ось чому слід переглянути вимоги до обсягів перевірки, забезпечити доступ до сучасних інструментів та запровадити програми психологічної підтримки, наприклад, сертифіковані тренінгові програми, розроблені на основі Універсального тренінгу із психічного здоров'я (УТПЗ) [3, 6].

Одним із дієвих інструментів підтримки є впровадження елементів менторства. Регулярні онлайн-зустрічі, які не мають формального академічного навантаження, створюють простір для неформального спілкування, обговорення труднощів, обміну досвідом та емоційної розрядки. Такі зустрічі сприятимуть формуванню довіри між студентами та викладачами, знижують рівень ізоляції та допомагають зберігати мотивацію до навчання. Особливо важливо, щоб ці ініціативи були добровільними, гнучкими за форматом і не перетворювалися на додаткове джерело стресу.

Для моніторингу емоційного стану студентів доцільно використовувати анонімні опитування. Вони дозволяють виявити загальні тенденції – рівень тривожності, перевантаження, потребу в додатковій підтримці – без порушення приватності. На основі таких даних викладачі та адміністрація можуть коригувати навчальне навантаження, змінювати формати занять, адаптувати дедлайни або запроваджувати додаткові консультації. Важливо, щоб студенти бачили реальні зміни у відповідь на свої запити – це підвищує довіру до інституції та формує відчуття безпеки.

Не менш важливою є підтримка самих викладачів, які часто працюють у надскладних умовах, поєднуючи професійні обов'язки з особистими викликами. Створення професійних спільнот, де педагоги можуть обмінюватися досвідом, ділитися стратегіями адаптації та просто говорити про свої переживання, є надзвичайно цінним. Проведення психологічних тренінгів, воркшопів з емоційного вигорання, а також доступ до консультацій з фахівцями – усе це сприяє збереженню професійної стійкості та якості викладання.

Освітня безпека та етичні аспекти. У контексті воєнного стану питання освітньої безпеки та етики набувають особливої актуальності, оскільки освітній процес відбувається в умовах підвищеної вразливості – як технічної, так і психологічної. Забезпечення конфіденційності, етичності оцінювання та ефективної комунікації між учасниками освітнього процесу є критично важливими для збереження довіри, академічної доброчесності та гуманітарної орієнтації освіти.

Одним із ключових аспектів освітньої безпеки є захист персональних даних студентів при використанні цифрових платформ. В умовах зростання кіберзагроз, спричинених як загальносвітовими тенденціями, так і специфікою воєнного конфлікту, особливої уваги потребує дотримання норм законодавства щодо обробки, зберігання та передачі інформації. Використання хмарних сервісів, платформ дистанційного навчання та засобів комунікації має супроводжуватися чіткими протоколами безпеки, зокрема двофакторною автентифікацією, шифруванням даних та обмеженням доступу до конфіденційної інформації. Освітні установи повинні забезпечити регулярне оновлення політик безпеки та проводити інструктажі для викладачів і студентів щодо цифрової гігієни.

Навчання через соціальну активність як механізм формування громадянської суб'єктності в умовах соціальної турбулентності. У сучасному освітньому дискурсі, особливо в контексті кризових трансформацій, спричинених війною, соціальна активність студентів розглядається не лише як форма позаакадемічної участі, а як повноцінний компонент освітнього процесу, що сприяє формуванню громадянської суб'єктності, соціальної відповідальності та практичної компетентності. Такий підхід корелює з концепцією соціально ангажованої освіти (socially engaged education), яка передбачає інтеграцію навчання з реальними потребами суспільства, розвиток етичної рефлексії та здатності до дії в умовах невизначеності.

Інституціоналізація волонтерської діяльності як складової освітнього процесу є стратегічно важливим кроком у напрямі гуманітаризації вищої освіти. Участь студентів у волонтерських ініціативах, зокрема, у сферах гуманітарної допомоги, соціального супроводу, психологічної підтримки, логістики, медіації, сприяє формуванню трансверсальних компетентностей, таких як емпатія, комунікативна гнучкість, етична відповідальність та здатність до колективного вирішення проблем. Освітні інституції, які інтегрують волонтерську активність у навчальні програми, демонструють вищий рівень соціальної адаптивності та здатності до інституційного реагування на виклики часу.

Оцінювання соціальних ініціатив студентів як елементу академічного прогресу потребує розробки відповідних педагогічних інструментів, що враховують специфіку неформального та досвідного навчання. Рефлексивні есе, аналітичні звіти, презентації, портфоліо, кейс-аналізи – ці форми дозволяють фіксувати індивідуальну траєкторію розвитку студента, його внесок у суспільно значущі проекти та здобуті компетентності. Важливо, щоб оцінювання соціальної активності було не каральним, а підтримувальним, спрямованим на стимулювання саморефлексії, критичного мислення та етичної відповідальності. У цьому контексті доцільно застосовувати елементи формувального оцінювання, що дозволяє враховувати процесуальну динаміку навчання, а не лише його кінцевий результат.

Підтримка студентських проєктів, спрямованих на допомогу громадам, є важливим інструментом поєднання освітньої діяльності з реальними соціальними практиками. Університети можуть виступати як платформи для ініціації, координації та ресурсного забезпечення таких проєктів, сприяючи формуванню соціального капіталу та зміцненню зв'язків між академічною спільнотою та локальними громадами. Це відповідає парадигмі університету як соціального агента, здатного не лише продукувати знання, а й трансформувати соціальну реальність. Залучення студентів до проєктів, що мають безпосередній вплив на життя громад, сприяє розвитку їхньої соціальної ідентичності, формуванню активної громадянської позиції та усвідомленню ролі освіти як інструменту суспільної трансформації.

Таким чином, навчання через соціальну активність слід розглядати як багатовимірний педагогічний феномен, що поєднує когнітивний, емоційний, етичний та соціальний компоненти освітнього процесу. Його системне впровадження дозволяє забезпечити гуманітарну сталість вищої освіти, сформувати громадянську суб'єктність студентів та розширити освітній простір до меж соціальної дії, що є особливо актуальним у контексті війни та посткризового відновлення.

Етичний вимір освітньої взаємодії в умовах війни передбачає переосмислення традиційних підходів до оцінювання. Форс-мажорні обставини – евакуація, втрати, психологічна травматизація, вимагають гнучкого трактування академічної доброчесності та уникнення каральних практик. Оцінювання має базуватися не лише на формальних критеріях, а й враховувати контекст, у якому здійснюється навчання. Це передбачає можливість повторного подання завдань, адаптацію дедлайнів, використання альтернативних форм оцінювання (наприклад, усних інтерв'ю, рефлексивних есе), а також формувальне оцінювання, орієнтоване на процес навчання, а не лише на результат. Такий підхід сприяє збереженню мотивації студентів та формуванню етичної культури взаємодії.

Інституційна гнучкість та автономія викладача. У контексті глибокої трансформації освітнього середовища, спричиненої повномасштабною війною, інституційна гнучкість та автономія викладача постають як ключові чинники забезпечення адаптивності, стійкості та академічної ефективності вищої освіти. Ці поняття виходять за межі адміністративної лібералізації і набувають системного значення, формуючи нову парадигму педагогічного управління, що базується на принципах субсидіарності, довіри та професійної відповідальності.

Надання викладачам права самостійно адаптувати навчальні програми, методи оцінювання та форми взаємодії зі студентами є не лише реакцією на кризові обставини, а й проявом академічної автономії як фундаментального принципу університетської освіти. У сучасних умовах це право набуває особливої ваги,

оскільки дозволяє оперативно реагувати на зміни контексту – від технічних обмежень до психоемоційного стану учасників освітнього процесу. Адаптація змісту дисциплін, формулювання альтернативних завдань, гнучке планування навчального навантаження. Усе це сприяє збереженню академічної якості та гуманітарної чутливості освітньої взаємодії.

Інституційна гнучкість передбачає також наявність внутрішніх регламентів, які дозволяють викладачам приймати рішення без потреби централізованого погодження. Це включає механізми делегування повноважень, спрощення процедур затвердження змін до робочих програм, а також створення локальних нормативних актів, що враховують специфіку регіонального контексту. Такий підхід відповідає концепції адаптивного управління, яка передбачає децентралізацію рішень, гнучкість процедур та орієнтацію на результат, а не на формальні показники.

Окрему увагу слід приділити підтримці інноваційної активності викладачів, яка є важливою складовою інституційної гнучкості. Грантові програми, внутрішні конкурси на кращі освітні практики, публічне визнання педагогічних ініціатив. Всі ці інструменти не лише стимулюють творчий пошук, а й формують культуру педагогічного підприємництва. Університети, які інвестують у розвиток інноваційного потенціалу своїх викладачів, демонструють вищу адаптивність до кризових викликів, а також здатність до стратегічного оновлення освітніх моделей.

У цьому контексті важливо також враховувати міждисциплінарний вимір автономії, який включає педагогічну, організаційну, технологічну та етичну складові. Автономія викладача не може розглядатися ізольовано, вона має бути інтегрована в загальну інституційну екосистему, де забезпечується баланс між свободою дій і відповідальністю за результати. Такий баланс є основою для формування стійких освітніх практик, здатних функціонувати в умовах високої невизначеності.

Таким чином, інституційна гнучкість та автономія викладача є не лише інструментами кризового реагування, а й стратегічними ресурсами розвитку вищої освіти. Їхнє системне впровадження дозволяє забезпечити адаптивність, інноваційність та гуманітарну орієнтацію освітнього процесу, що є критично важливими у контексті війни, посткризового відновлення та довгострокової модернізації освітньої системи України.

Підвищення кваліфікації викладачів в умовах кризових трансформацій освітнього середовища.

У контексті системної турбулентності, спричиненої повномасштабною війною, професійний розвиток викладачів закладів вищої освіти набуває статусу стратегічного ресурсу забезпечення академічної стійкості, інституційної адаптивності та педагогічної ефективності. Підвищення кваліфікації в умовах кризи слід розглядати не як факультативну активність, а як обов'язкову складову освітньої політики, спрямованої на трансформацію компетентнісного профілю викладача відповідно до нових викликів.

Одним із ключових напрямів є впровадження короткострокових онлайн-курсів, орієнтованих на розвиток цифрової педагогічної компетентності, навичок дистанційного навчання, кризового менеджменту та емоційної саморегуляції. Такі курси мають бути побудовані за принципами модульності, гнучкості та практичної релевантності, що дозволяє забезпечити їх інтеграцію в щоденну професійну діяльність викладача. Особливої уваги потребує контент, спрямований на формування навичок роботи з цифровими платформами, розробку адаптивних освітніх стратегій та етичну взаємодію зі студентами в умовах підвищеної психоемоційної напруги.

Паралельно з формальними освітніми програмами необхідно розвивати інституційні механізми горизонтального обміну досвідом. Створення внутрішніх комунікаційних платформ: блогів, форумів, вебінарів, професійних онлайн-спільнот, тощо. Саме це сприяє формуванню педагогічної рефлексії, поширенню ефективних практик та консолідації академічної спільноти. Такі платформи виконують функцію не лише інформаційного, а й емоційного ресурсу, що дозволяє знижувати рівень професійного вигорання та підтримувати мотивацію до інноваційної діяльності.

У межах сучасної парадигми безперервного професійного розвитку доцільним є запровадження системи мікросертифікації, яка дозволяє формалізувати здобуті компетентності у форматі коротких, верифікованих освітніх модулів. Мікросертифікати можуть бути інтегровані в систему внутрішнього академічного рейтингу, кадрового просування або портфолію викладача, що сприяє визнанню неформального навчання як повноцінного елементу професійної траєкторії. Такий підхід відповідає концепції гнучкої освіти (flexible learning) та сприяє формуванню інституційної культури підтримки педагогічної ініціативи.

Отже, підвищення кваліфікації викладачів в умовах кризових трансформацій слід розглядати як багатовимірний процес, що охоплює формальні, неформальні та інституційно-мотиваційні компоненти. Його ефективна реалізація є запорукою не лише адаптації освітньої системи до умов війни, а й її здатності до стратегічного оновлення та гуманітарного переосмислення.

Висновки. Аналіз поточної ситуації у вищій освіті України свідчить про наявність системних викликів, зумовлених як специфікою сучасних освітніх програм, так і надзвичайними умовами воєнного часу. Навантаження на викладачів, що працюють з великими лекційними потоками та значними обсягами самостійної роботи студентів, багаторазово зростає. Це посилюється такими об'єктивними перешкодами, як повітряні тривоги та відключення електроенергії, що унеможливорює дотримання жорстких дедлайнів та робить традиційні механізми контролю та оцінювання неефективними.

У відповідь на ці виклики, впровадження адаптивних стратегій є не лише доцільним, а й необхідним. Запропоновані підходи, що включають переосмислення проектування завдань, застосування комбінованих методів перевірки з урахуванням викликів, пов'язаних зі штучним інтелектом, а також мінімалістичний та масштабований зворотний зв'язок, дозволяють оптимізувати робочий час викладача. Вони допомагають підтримувати академічну доброчесність, заохочувати внутрішню мотивацію студентів та забезпечувати досягнення ключових результатів навчання в умовах непередбачуваності.

Критично важливим є також системний підхід до вирішення проблеми, що передбачає психологічну та методичну підтримку викладачів, а також адаптацію вимог з боку адміністрації ЗВО. Це допомагає запобігти професійному вигоранню та створює більш стабільне освітнє середовище.

Запропоновані рішення є компромісними та можуть бути масштабовані на значну частину освітніх програм. Подальші дослідження можуть бути сфокусовані на аналізі ефективності цих стратегій та їхньому впливі на якість навчання та задоволеність освітнім процесом як з боку студентів, так і з боку викладачів.

Список використаних джерел:

1. Закон України «Про наукову і науково-технічну діяльність». URL: <https://zakon.rada.gov.ua/laws/show/848-19/ed20180307#Text> (дата звернення: 17.09.2025).
2. Закон України «Про освіту». URL: <https://zakon.rada.gov.ua/laws/show/2145-19> (дата звернення: 17.09.2025).
3. Левчук О. В., Левчук К. І. Цифрова стійкість: оцінка ролі інформаційних технологій у забезпеченні безперервності підготовки фахівців в кризових умовах. *Комп'ютерно-інтегровані технології: освіта, наука, виробництво*. 2024. № 54. С. 137–145. DOI: <https://doi.org/10.36910/6775-2524-0560-2024-54-16>.
4. Організація освітнього процесу в умовах воєнного стану 2024. URL: <https://dub-osvita.gov.ua/organizacii-osvitnogo--procesu-v-umovah--voennogo-stanu-2024-08-47-18-29-10-2024/> (дата звернення: 17.09.2025).
5. Про затвердження методичних рекомендацій щодо окремих питань здобуття освіти в закладах загальної середньої освіти в умовах воєнного стану в Україні. URL: <https://zakon.rada.gov.ua/rada/show/v0563729-23#Text> (дата звернення: 17.09.2025).
6. Савчук О. Д., Єгорова І. В. Стратегії формування критичного мислення у студентів ЗВО в умовах освітніх викликів воєнного часу. *Педагогічна наука й освітня практика в реаліях війни: виклики, досвід, перспективи* : зб. наук. пр. Івано-Франківськ : НАІР. 2025. С. 130–135.
7. Сердюк М. Є., Сафронова І. А., Зайцева Т. А. Особливості організації теоретичних онлайн-занять. *Актуальні проблеми автоматизації та інформаційних технологій*. 2022. Т. 26. С. 61–68. DOI: <http://dx.doi.org/10.15421/432211>.
8. Шарата Н. Г. Особливості викладання гуманітарних дисциплін у Миколаївському національному аграрному університеті в умовах воєнного стану. *Гуманізація навчально-виховного процесу*. 2025. № 1(107). С. 115–124. URL: <https://dspace.mnau.edu.ua/jspui/handle/123456789/21111> (дата звернення: 17.09.2025).
9. Шевчук О. В. Стратегії успішної дистанційної освіти: передові практики викладання, методи стимулювання навчального процесу та оцінювання результатів. *Збірник матеріалів Міжнародної науково-практичної конференції для освітян (м. Чернігів, 22 березня 2023 року)*. Чернігів : ГО «Науково-освітній інноваційний центр суспільних трансформацій». 2023. С. 85–87. DOI: https://doi.org/10.54929/conf_ped_03_2023.
10. Щодо окремих питань організації освітнього процесу в умовах воєнного стану. URL: https://osvita.ua/legislation/Ser_osv/90390/ (дата звернення: 17.09.2025).
11. Як організувати викладання навчальних предметів в умовах воєнного стану: рекомендації. URL: <https://sqe.gov.ua/yak-organizuvati-vikladannya-navchaln/> (дата звернення: 17.09.2025).
12. Shyshkanova, G., Zaytseva, T., & Frydman, O. Mobile technologies make education a part of everyday life. *Information and Learning Science*. 2017. Vol. 118, No. 11/12. P. 570–582. DOI: <https://doi.org/10.1108/ILS-03-2017-0019>.

References:

1. Zakon Ukrainy «Pro naukovu i naukovo-tekhnichnu diialnist». (2018). [Law of Ukraine "On Scientific and Scientific-Technical Activity"]. Retrieved from: <https://zakon.rada.gov.ua/laws/show/848-19/ed20180307#Text>
2. Zakon Ukrainy «Pro osvitu». (2017). [Law of Ukraine "On Education"]. Retrieved from: <https://zakon.rada.gov.ua/laws/show/2145-19>
3. Levchuk, O. V., & Levchuk, K. I. (2024). Tsyfrova stiikist: otsinka roli informatsiinykh tekhnolohii u zabezpechenni bezperervnosti pidgotovky fakhivtsiv v kryzovykh umovakh [Digital Resilience: Assessing the Role of Information Technology in Ensuring the Continuity of Training in Crisis Conditions]. *Kompiuterno-intehrovani tekhnolohii: osvita, nauka, vyrobnytstvo*, 54, 137–145. <https://doi.org/10.36910/6775-2524-0560-2024-54-16>
4. Orhanizatsiia osvitnoho protsesu v umovakh voiennoho stanu 2024. (2024). [Organization of the educational process in wartime conditions 2024]. Retrieved from: <https://dub-osvita.gov.ua/organizacii-osvitnogo--procesu-v-umovah--voennogo-stanu-2024-08-47-18-29-10-2024/>

-
5. Pro zatverdzhennia metodychnykh rekomendatsii shchodo okremykh pytan zdobuttia osvity v zakladakh zahalnoi serednoi osvity v umovakh voiennoho stanu v Ukraini. (2023). [On the approval of methodological recommendations on certain issues of education in general secondary education institutions in conditions of martial law in Ukraine]. Retrieved from: <https://zakon.rada.gov.ua/rada/show/v0563729-23#Text>
6. Savchuk, O. D., & Yehorova, I. V. (2025). Stratehii formuvannia krytychnoho myslennia u studentiv ZVO v umovakh osvitnikh vyklykiv voiennoho chasu. In *Pedahohichna nauka y osvitnia praktyka v realiiakh viiny: vyklyky, dosvid, perspektyvy: zb. nauk. prats.* (pp. 130–135). NAIR. Retrieved from: <https://kikip.pnu.edu.ua/wp-content/uploads/sites/94/2025/05/zbirnyk-naukovykh-prats-2025.pdf>
7. Serdiuk, M. Ye., Safronova, I. A., & Zaitseva, T. A. (2022). Osoblyvosti orhanizatsii teoretychnykh onlain-zanat. *Aktualni problemy avtomatyzatsii ta informatsiinykh tekhnolohii*, 26, 61–68. <http://dx.doi.org/10.15421/432211>
8. Sharata, N. H. (2025). Osoblyvosti vykladannia humanitarnykh dystsyplin u Mykolaivskomu natsionalnomu ahrarnomu universyteti v umovakh voiennoho stanu. *Humanizatsiia navchalno-vykhovnoho protsesu*, 1(107), 115–124. Retrieved from: <https://dspace.mnau.ua/jspui/handle/123456789/21111>
9. Shevchuk, O. V. (2023). Stratehii uspishnoi dystantsiinoi osvity: peredovi praktyky vykladannia, metody stymuliuвання navchalnoho protsesu ta otsiniuvannia rezultativ [Strategies for successful distance education: best teaching practices, methods of stimulating the learning process and assessing results]. In *Zbirnyk materialiv Mizhnarodnoi naukovo-praktychnoi konferentsii dlia osvitan (m. Chernihiv, 22 bereznia 2023 roku)* (pp. 85–87). HO «Naukovo-osvitnii innovatsiinyi tsentr suspilnykh transformatsii». https://doi.org/10.54929/conf_ped_03_2023
10. Shchodo okremykh pytan orhanizatsii osvitnoho protsesu v umovakh voiennoho stanu. (2022). [Regarding certain issues of the organization of the educational process under martial law]. Retrieved from: https://osvita.ua/legislation/Ser_osv/90390/
11. Yak orhanizuvaty vykladannia navchalnykh predmetiv v umovakh voiennoho stanu: rekomendatsii. (2022). [How to organize the teaching of academic subjects under martial law: recommendations]. Retrieved from: <https://sqe.gov.ua/yak-organizuvati-vikladannya-navchaln/>
12. Shyshkanova, G., Zaytseva, T., & Frydman, O. (2017). Mobile technologies make education a part of everyday life. *Information and Learning Science*, 118(11/12), 570–582. <https://doi.org/10.1108/ILS-03-2017-0019>

Дата надходження статті: 23.09.2025

Дата прийняття статті: 10.10.2025

Опубліковано: 30.12.2025

КОМП'ЮТЕРНІ НАУКИ

УДК 004.8:004.421.2:37.018.43

DOI <https://doi.org/10.32782/2521-6643-2025-2-70.8>

Аксак Н. Г., доктор технічних наук, професор,
професор кафедри комп'ютерних інтелектуальних технологій
та систем
Харківського національного університету радіоелектроніки
ORCID: 0000-0001-8372-8432

Татарников А. О., аспірант факультету комп'ютерної інженерії
та управління
Харківського національного університету радіоелектроніки
ORCID: 0000-0002-1632-8188

БАГАТОШАРОВА МІКРОСЕРВІСНО-АГЕНТНА АРХІТЕКТУРА СИСТЕМИ ПЕРСОНАЛІЗОВАНОГО НАВЧАННЯ

Метою роботи є розроблення багатошарової мікросервісно-агентної архітектури системи для реалізації персоналізованого навчання на основі використання поведінкових даних із інтегрованим модулем комп'ютерного зору (SBMS–LMS), що забезпечує адаптивність навчального процесу та можливість автоматичного відстеження різноманітної активності студентів. Методологічною основою є поєднання принципів багатоагентних систем, мікросервісної архітектури, сенсорної поведінкової аналітики та формальної верифікації безпеки у часовій логіці. Архітектура реалізує семирівневу модель, що включає сенсорний шар (SBMS), аналітичні сервіси (Learning Analytics, Adaptive Learning), агентну оркестрацію (MAS) та інтеграцію з часто використовуваною LMS Moodle через технологію REST API. Система дозволяє підтримувати збір поведінкових сигналів (присутність студентів на заняттях, увага на лекційних заняттях, порушення дисципліни та використання систем обходу систем контролю), аналітику активності та успішності, формування рекомендацій для учнів і зворотний зв'язок у режимі реального часу, який забезпечує можливість студентам впливати на навчальний процес та покращувати його якість. Модуль безпеки гарантує автентифікацію, шифрування та знеособлення даних відповідно до політик GDPR/FERPA. Наукова новизна полягає у створенні єдиного, стандартизованого когнітивно-аналітичного контуру, який поєднує в собі агентну логіку ухвалення рішень, потокову аналітику поведінкових даних та мікросервісну можливість розбиття функцій освітніх сервісів на незалежні модулі, що в свою чергу підвищує відмово стійкість та захист від отримання повного доступу до системи можливим злоумисником. Отримані в процесі дослідження результати на практиці підтверджують ефективність інтеграції SBMS із LMS та здатність створеної системи динамічно адаптувати навчальні траєкторії, підвищуючи якість і персоналізацію освіти.

Ключові слова: автоматизоване тестування, Docker, мікросервісна архітектура, навчальна платформа, навчання, персоналізація, моніторинг, агент, масштабування, продуктивність

Axak N. G., Tatarnykov A. O. Multilayer microservice-agent architecture of a personalized learning system

The goal of this work is to develop a multi-layered microservice-agent architecture for implementing personalized learning based on the use of behavioral data with an integrated computer vision module (SBMS–LMS), which ensures the adaptability of the learning process and the ability to automatically track various student activities. The methodological basis is a combination of the principles of multi-agent systems, microservice architecture, sensory behavioral analytics, and formal security verification in temporal logic. The architecture implements a seven-level model that includes a sensor layer (SBMS), analytical services (Learning Analytics, Adaptive Learning), agent orchestration (MAS), and integration with the widely used LMS Moodle via REST API technology. The system allows for the collection of behavioral signals (student attendance, attention in lectures, disciplinary violations, and use of control system bypasses), students activity and full performance analytics, the formation of recommendations for students, and real-time feedback, which allows students to influence the learning process and improve its quality. The security module guarantees authentication, encryption, and fully anonymization of users data in accordance with GDPR/FERPA standartisation policies. The scientific novelty lies in the creation of a unified, standardized cognitive-analytical circuit that combines agent-based decision-making logic, stream analytics of behavioral data, and the microservice capability

© Н. Г. Аксак, А. О. Татарников, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

to break down educational service functions into independent modules, which in turn increases fault tolerance and protection against full access to the system by a potential attacker. The work results obtained in the course of the research confirm in real practice the effectiveness of SBMS integration with LMS Moodle and the ability of the created system to dynamically adapt current learning trajectories, improving the quality of learning and improve personalization level of education.

Key words: automated testing, Docker, microservice architecture, learning platform, training, monitoring, agent, scaling, performance, personalization

Актуальність і стан досліджень. Цифрова трансформація освіти зумовила необхідність у створенні інтелектуальних систем, здатних забезпечити персоналізоване навчання з урахуванням стилю навчання, мотивації, поведінкових характеристик і когнітивного стану студентів. Традиційні LMS, хоч і забезпечують базову організацію контенту й оцінювання, залишаються статичними та не адаптуються до змін поведінки студентів. Особливо актуальною є потреба в системах, які в реальному часі аналізують активність, прогножують ризики та коригують навчальні траєкторії.

Упродовж 2020–2025 років персоналізоване навчання розвивалося під впливом Learning Analytics, Adaptive Learning, Multi-Agent Systems та штучного інтелекту [1], [3], [4]. Сучасні адаптивні системи базуються на попередньому тестуванні, логах активності та когнітивно-аналітичних моделях, які враховують емоційний стан і рівень залученості [1], [3]. Використання великих мовних моделей (LLM) дозволяє створювати платформи, що адаптують контент до стилю навчання студента та забезпечують динамічний зворотний зв'язок [2]. У поєднанні з XR-технологіями (VR/AR/MR) це сприяє глибшому зануренню в адаптивне освітнє середовище.

Інтелектуальні агенти також відіграють важливу роль у персоналізації. Вони виконують функції підтримки, пояснення та мотивації, імітуючи поведінку людських компаньйонів [5]. Багатоагентні системи (MAS) реалізують більш складну взаємодію – наприклад, у прототипі EduPlanner агенти автоматично створюють і вдосконалюють навчальні плани на основі результатів і типових помилок студентів [6].

Паралельно з цим освітні платформи активно впроваджують мікросервісну архітектуру для підвищення гнучкості, масштабованості й надійності [7]. Вона дозволяє розгортати незалежні функціональні модулі (тестування, аналітика, рекомендації) з можливістю автономного оновлення. Дослідники відзначають, що модульність мікросервісів значно спрощує оновлення функцій, дозволяє масштабувати лише необхідні компоненти та знижує витрати на технічну підтримку [8]. Так, система GradeSuite (Університет Берклі) забезпечила автоматизоване гнучке оцінювання та скоротила час адміністрування оцінок на 90% [9]. Аналогічно, платформа підтримки студентів, розроблена Shaiba та ін., розподілила сервіси технічної й освітньої підтримки, що виявилось ефективним в умовах екстреного дистанційного навчання (ERT) [10].

Загалом тенденції зводяться до інтеграції ШІ у навчальні системи, розвитку LA для педагогічної аналітики, застосування агентних моделей для підтримки мотивації та адаптації, а також переходу до хмарно-орієнтованих, мікросервісних освітніх платформ. Проте водночас з'являються виклики: потреба в етичному використанні ШІ, захисті даних, забезпеченні прозорості алгоритмів і педагогічної доцільності [2], [3], [4].

Попередні роботи авторів [11], [12] ілюструють еволюцію підходів – від агентно-орієнтованого моделювання до побудови мікросервісних систем персоналізації та моніторингу поведінки (SBMS). Водночас бракує цілісної архітектури моделі, що поєднувала б сенсорний рівень (моніторинг за допомогою CV), аналітичний (збір і обробка даних) і когнітивний (агентна персоналізація).

Мета дослідження – розробити багаторівневу мікросервісно-агентну архітектуру системи персоналізованого навчання на основі поведінкових даних із безперервним адаптивним циклом та інтеграцією з LMS Moodle.

Завдання дослідження:

- проаналізувати сучасні архітектури персоналізації;
- формалізувати модель взаємодії агентів, сервісів і сенсорних компонентів;
- розробити мікросервісну структуру збору, аналітики й адаптації;
- реалізувати інтеграцію SBMS з LMS;
- забезпечити безпеку, приватність і відповідність стандартам.

Наукова новизна полягає в побудові когнітивно-аналітичного контуру, що поєднує агентну логіку, потокову аналітику та сенсорний контроль. Запропонована система переходить від реактивної до проактивної адаптації, прогножуючи поведінку студента та коригуючи освітній процес у реальному часі.

Концептуальна модель системи персоналізованого навчання. Концептуальна модель системи персоналізованого навчання базується на принципах багатоагентної взаємодії, де інтелектуальні агенти представляють ролі учасників навчального процесу та автономно приймають рішення, ґрунтуючись на поведінкових даних і стані середовища.

Формально багатоагентну систему можна визначити як множину:

$$MAS = \{Ag, Act, Env\}, \quad (1)$$

де: $Ag = \{A_{stud}, A_{teach}, A_{anal}, A_{gsys}, A_{gsurvey}\}$ – множина агентів; Act – множина дій, які агенти можуть виконувати; $Env = \{E_{LMS}, E_{SBMS}, E_{Cloud}\}$ – множина елементів середовища.

У цій моделі кожен агент функціонує автономно, але взаємодіє з іншими через спільні протоколи обміну (REST API, JSON, MQTT), утворюючи когнітивно-аналітичний цикл, спрямований на персоналізацію навчальної траєкторії. Типи агентів і їх ролі наведені у (табл. 1).

Таблиця 1

Типи агентів і їх ролі

Агент	Основна роль	Джерела даних	Тип дій
агент студента	Відображає поведінку, активність і результати навчання.	LMS Moodle, SBMS, інтерактивні модулі	Надсилання подій, отримання рекомендацій, оновлення статусу
агент викладача	Координує навчальний процес, аналізує успішність студентів.	LMS, аналітичні звіти	Прийняття рішень, затвердження рекомендацій, зворотний зв'язок
агент аналітики	Виконує статистичний і прогностичний аналіз поведінкових даних.	Сервіс аналітики, історичні дані	Розрахунок метрик, прогнозування ризиків, кластеризація
агент системи	Координує роботу інших агентів і сервісів; забезпечує безпеку, моніторинг, інтеграцію.	API Gateway, Security Service	Маршрутизація, логування, верифікація інваріантів
агент опитувань	Збирає відгуки студентів і викладачів, формує контекстні оцінки задоволеності.	Web-портал, форми зворотного зв'язку	Ініціювання опитувань, збереження результатів

Середовище функціонування системи персоналізованого навчання побудоване як взаємодія трьох основних компонентів, що утворюють цілісну інтелектуальну інфраструктуру. Першим є E_{LMS} – навчальне середовище, яке представлено системою Moodle або іншими LMS і містить освітній контент, оцінки, журнали активності студентів та інструменти для взаємодії між учасниками освітнього процесу. Другим компонентом виступає E_{SBMS} , тобто сенсорне середовище моніторингу поведінки студентів, яке включає відеопотік, детектори присутності та журнали подій браузера. Третім компонентом є E_{Cloud} – хмарна інфраструктура, яка забезпечує зберігання великих обсягів даних, виконання аналітичних обчислень і роботу модулів машинного навчання.

Усі три середовища взаємодіють між собою через шину подій (Event Bus) або асинхронні черги повідомлень (RabbitMQ, Kafka), що забезпечує потокову обробку даних у реальному часі.

Логіка функціонування моделі побудована у вигляді циклічного процесу, який формує адаптивний контур навчання. На етапі спостереження сенсорний шар SBMS реєструє поведінкові події – фіксує присутність студента, напрямок погляду, рівень уваги, характер взаємодії з навчальними матеріалами, можливі порушення чи зміни емоційного стану. Далі, під час інтерпретації, агент аналітики (A_{ganal}) обробляє ці події, класифікує їх, розраховує індекс активності, визначає ризик відставання та прогнозує рівень успішності навчання. На етапі планування (агент системи (A_{gsys})) у взаємодії з агентом викладача (A_{gteach}) формує нові навчальні завдання або рекомендації, адаптуючи навчальний контент під потреби конкретного студента.

Наступна стадія – дія – передбачає реалізацію цих рекомендацій, коли агент студента (A_{gstud}) отримує індивідуальні вказівки, що можуть стосуватися повторення певної теми, зміни складності завдань або проходження додаткової консультації.

Таким чином, описаний процес створює динамічну систему персоналізації навчання, у якій усі рішення приймаються на основі реальних поведінкових даних. Взаємодія між середовищами E_{LMS} , E_{SBMS} і E_{Cloud} забезпечує замкнений когнітивно-аналітичний контур, що дозволяє системі постійно вдосконалювати свої рекомендації та підтримувати оптимальний рівень залученості студентів у навчальний процес. Система підтримує три типи потоків (табл. 2).

Таблиця 2

Основні типи потоків даних у системі персоналізованого навчання

Тип потоку	Напрямок	Зміст
Поведінковий	SBMS → Analytics → MAS → LMS	Події моніторингу, фіксація присутності, активність, порушення
Аналітичний	Analytics ↔ Adaptive Learning	Метрики, прогнози, класифікація студентів
Когнітивний	MAS ↔ LMS ↔ Users	Рекомендації, зміни траєкторій, зворотний зв'язок

Ці потоки утворюють замкнений когнітивно-аналітичний цикл, що реалізує саморегулювання навчального процесу.

Формальна модель взаємодії. Для опису взаємодії між агентами використовується множина правил виду:

$$R = \{(s, a, s', r) \mid s, s' \in S, a \in \text{Act}, r \in R\}, \quad (2)$$

де: s – поточний стан студента, a – дія агента (наприклад, «повторити тему», «змінити темп»), s' – новий стан після виконання дії, r – функція винагороди, яка оцінює ефективність дії.

Функція винагороди може визначатися як:

$$r = w_1 \cdot P + w_2 \cdot A + w_3 \cdot E, \quad (3)$$

де P – показник прогресу, A – рівень активності, E – коефіцієнт емоційного залучення; w_i – вагові коефіцієнти, встановлені під час навчання моделі.

Це дозволяє реалізувати підкріплювальне навчання (Reinforcement Learning) для автоматичної адаптації траєкторії навчання.

Концептуальна модель поєднує агентну логіку прийняття рішень, потокову обробку даних і формалізовані правила поведінки студентів, що дозволяє: відслідковувати стан навчального процесу в реальному часі; адаптувати завдання й темп під кожного користувача; формувати замкнутий цикл «сприйняття → аналіз → рекомендація → дія → оцінка».

Архітектура та реалізація системи. Розроблена архітектура системи персоналізованого навчання побудована на основі поєднання агентного та мікросервісного підходів із додаванням сенсорного шару моніторингу поведінки студентів (SBMS). Архітектура реалізує замкнений цикл адаптації навчання: від збору поведінкових даних і аналітики до формування рекомендацій та їх інтеграції з LMS Moodle. Єдина архітектура складається з семи рівнів (L1–L7), кожен з яких виконує окрему роль у циклі персоналізації (таблиця 3).

Ця структура представлена на рисунку 1 (єдина архітектура). Багатошарова мікросервісно-агентна архітектура системи персоналізованого навчання на основі поведінкових даних з інтегрованим модулем комп'ютерного зору (SBMS–LMS) – це інтегрована інтелектуальна архітектура, яка забезпечує адаптивне навчання шляхом збору, аналізу та інтерпретації поведінкових даних студентів, їх перетворення в персоналізовані рекомендації, а також автоматичного моніторингу навчальної активності через систему комп'ютерного зору (Student Behavior Monitoring System).

Таблиця 3

Багаторівнева структура архітектури системи персоналізованого навчання

Рівень	Назва	Основні функції
L7	Інтерфейс користувача (UI/LMS)	Взаємодія студентів і викладачів з системою через Moodle, дашборди й веб-портал.
L6	Оркестрація агентів (MAS)	Узгодження рішень між агентами <i>Student, Teacher, Analyst, System</i> ; обмін подіями через REST API, формування колективних рішень і політик адаптації
L5	Персоналізація (Adaptive Learning Service)	Побудова рекомендацій на основі прогнозів аналітики; формування індивідуальних траєкторій, адаптація темпу, складності та змісту навчання.
L4	Аналітика (Learning Analytics Service)	Обчислення метрик активності, ризику, успішності; прогнозування результатів; сегментація студентів.
L3	Збір та інтеграція даних (API Gateway + Data Ingest)	Єдиний шлюз REST/JSON для зв'язку з LMS і SBMS; нормалізація даних, маршрутизація запитів, шина повідомлень
L2	Сенсорний шар (SBMS + CV)	Моніторинг поведінки, присутності та активності студентів, виявлення несанкціонованої діяльності під час е-тестів через відеоспостереження та події системи.
L1	Дані, безпека та спостережність	Зберігання подій (MongoDB, S3), аналітики (PostgreSQL), шифрування (TLS), логування (ELK), моніторинг (Prometheus), контроль доступу (OAuth2).

Архітектура реалізує замкнений цикл адаптації навчання – від фіксації поведінкових сигналів до аналітики, рекомендацій та корекції освітньої траєкторії – з урахуванням вимог безпеки, приватності та масштабованості. Вона охоплює три взаємопов'язані рівні інтелектуального управління: сенсорний (SBMS + CV) – фіксує присутність, увагу та активність студентів; аналітичний (Microservices) – обробляє події, формує метрики й прогнози успішності; когнітивний (MAS) – забезпечує персоналізацію через взаємодію агентів.

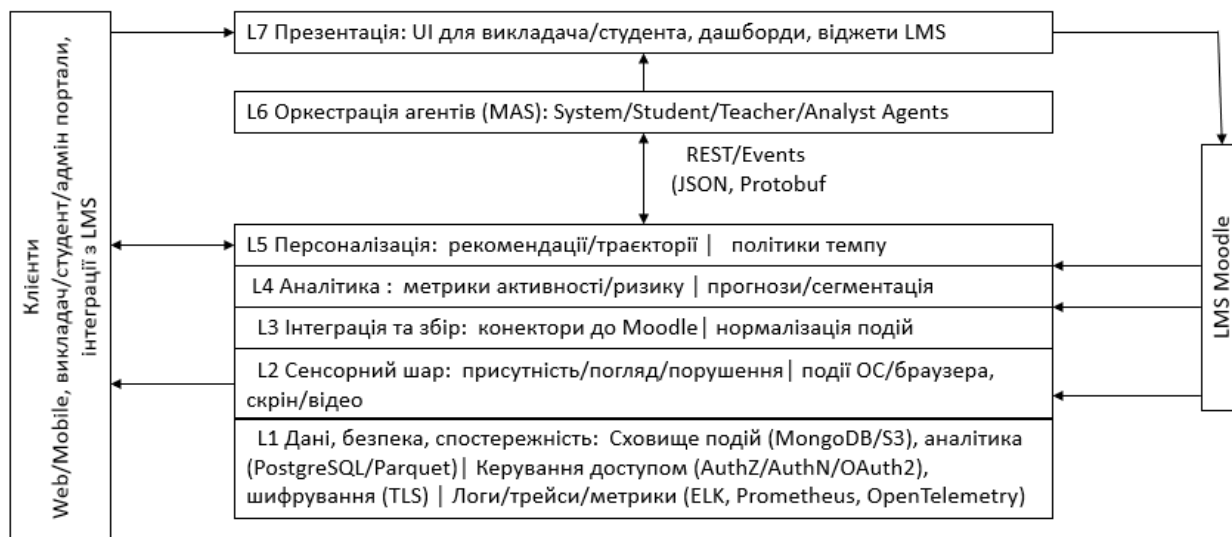


Рис. 1. Багатошарова мікросервісно-агентна архітектура системи персоналізованого навчання на основі поведінкових даних з інтегрованим модулем комп'ютерного зору (SBMS–LMS)

Циклічний процес адаптації починається зі збору даних SBMS, далі мікросервіси обчислюють аналітичні індикатори, а агентна система формує персоналізовані рекомендації. Вони інтегруються в LMS або UI, а на основі зворотного зв'язку модель поведінки студента постійно оновлюється, формуючи динамічне адаптивне середовище.

Архітектура реалізована на основі сучасного стеку технологій: C#, .NET, Nancy, OWIN, Laravel, Python; зберігання даних – PostgreSQL, MongoDB, AWS S3; контейнеризація – Docker, Docker Compose. Аналітика й моніторинг забезпечуються через Prometheus, ELK Stack і Grafana. Інтеграція з LMS реалізована за допомогою REST API, OAuth 2.0, OpenAPI 3.0, а CI/CD – через GitHub Actions. Для прогнозування успішності застосовано моделі ML: Decision Trees, LSTM, Reinforcement Learning.

Система є модульною, масштабованою та адаптивною в реальному часі, забезпечує повну інтеграцію з LMS Moodle та іншими платформами. Агентне управління дозволяє узгоджувати рекомендації, а SBMS – фіксувати когнітивну й поведінкову активність. Завдяки відкритим протоколам і наскрізному захисту даних архітектура залишається розширюваною, безпечною та придатною для впровадження в адаптивні освітні середовища.

Мікросервісна архітектура. Архітектура системи включає низку взаємопов'язаних мікросервісів, кожен із яких виконує окрему, але взаємодоповнювальну функцію у процесі персоналізованого навчання (рис. 2):

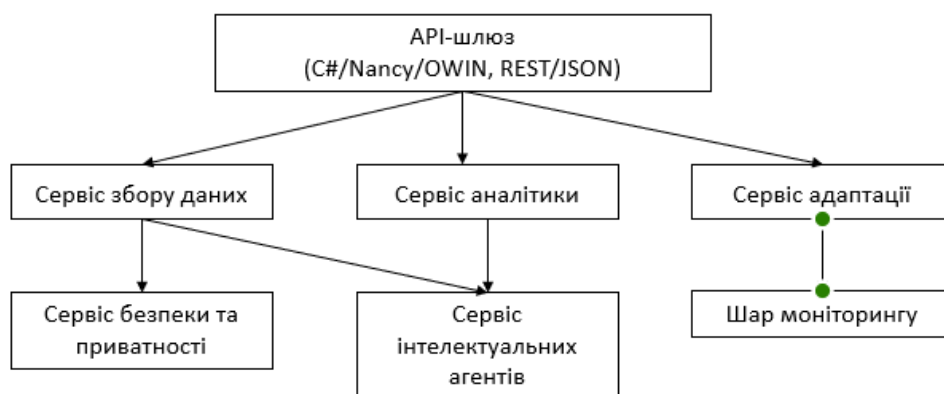


Рис. 2. Мікросервісна архітектура системи персоналізації навчання

Кожен сервіс виконує окрему функцію, а взаємодія координується через API-шлюз та контролюється системами моніторингу Prometheus і ELK

Сервіс збору даних забезпечує інтеграцію з підсистемами SBMS і LMS, отримуючи події, що відображають поведінку студентів під час навчального процесу. Він виконує фільтрацію, агрегацію та первинну

обробку цих подій, після чого структуровані дані зберігаються у колекціях бази MongoDB у форматі JSON для подальшого аналізу.

На наступному етапі працює *сервіс аналітики*, який здійснює розрахунок метрик активності студентів, визначає рівень ризику, прогнозує успішність і формує аналітичні звіти для викладачів та агентної системи. Усі оброблені аналітичні дані зберігаються у реляційній базі PostgreSQL, що забезпечує ефективне виконання запитів і зручну побудову статистичних моделей.

Сервіс адаптації використовує результати аналітичного аналізу для формування рекомендацій і корекції навчальних траєкторій кожного студента. Його логіка базується на принципах підкріплювального навчання, у межах якого функція винагороди визначається співвідношенням

$$r = w_1 \cdot P + w_2 \cdot A + w_3 \cdot E \quad (4)$$

де P позначає прогрес студента, A – рівень активності, а E – емоційне залучення. Отримані значення дозволяють системі виявляти закономірності у поведінці користувачів та адаптувати навчальний процес у режимі реального часу.

Ключову координаційну роль виконує *сервіс інтелектуальних агентів*, який узгоджує дії між агентами різних типів, зберігає їхній поточний стан і розподіляє події між відповідними компонентами системи. Саме цей сервіс забезпечує виконання агентних правил і підтримує когнітивну логіку персоналізації навчання.

Безпеку та захист персональних даних гарантує *сервіс безпеки та приватності*. Він реалізує механізми автентифікації й авторизації користувачів, веде журнал подій, контролює доступ до ресурсів і забезпечує шифрування каналів зв'язку за протоколом TLS 1.3. Система побудована таким чином, що дотримується двох фундаментальних інваріантів безпеки: $G(\phi 1)$: усі запити автентифіковані, $G(\phi 2)$: усі дані зашифровані.

Дотримання цих умов гарантує цілісність і конфіденційність інформаційних потоків у межах всієї архітектури, забезпечуючи надійну та контрольовану взаємодію між усіма сервісами системи.

API-шлюз і інтеграція з LMS. API Gateway, реалізований засобами C#/Nancy/OWIN, виконує роль центральної точки комунікації між усіма мікросервісами системи. Він забезпечує єдиний вхід для зовнішніх запитів і координує взаємодію між окремими компонентами архітектури. Основними його завданнями є маршрутизація запитів, перевірка автентифікації користувачів за допомогою протоколів OAuth2 або JWT, трансформація форматів повідомлень між сервісами, а також контроль частоти звернень для запобігання перевантаження системи. Завдяки цьому API Gateway виступає своєрідним «контролером трафіку», який забезпечує узгодженість обміну даними, безпечний доступ і підтримку масштабованості всієї інфраструктури.

Інтеграція системи з навчальним середовищем LMS Moodle реалізується через REST API, який забезпечує двосторонній обмін подіями між платформами. З одного боку, система отримує від Moodle інформацію про активність студентів, виконані завдання, тести та взаємодію з навчальними матеріалами.

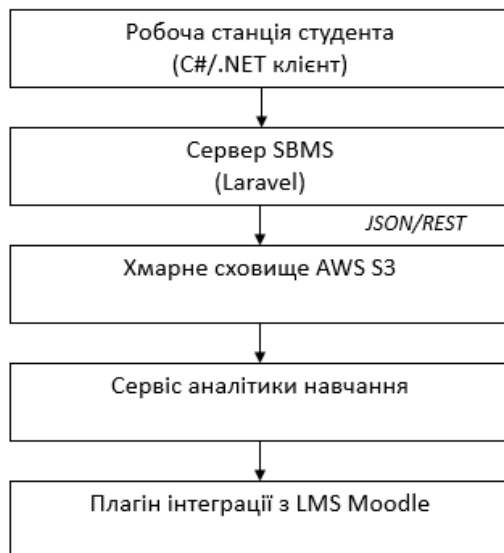


Рис. 3. Схема інтеграції підсистеми моніторингу поведінки (SBMS) та модуля комп'ютерного зору з LMS Moodle і сервісом аналітики навчання

Події фіксуються, передаються до хмарного сховища, аналізуються та використовуються для персоналізації.

З іншого – результати аналітики, згенеровані рекомендації та індивідуальні повідомлення повертаються до Moodle у вигляді додаткових модулів (plugins) або віджетів аналітичної панелі (dashboard widgets).

Така інтеграційна модель створює прозорий інформаційний потік між системою персоналізації та LMS, що дозволяє автоматично оновлювати навчальні траєкторії, формувати зворотний зв'язок у реальному часі та підтримувати когнітивний контур адаптації навчання.

Підсистема SBMS (Student Behavior Monitoring System) виконує ключову функцію сенсорного шару, фіксуючи поведінкові дані студентів. Її клієнтський модуль (C#/.NET) взаємодіє з вебкамерою, клавіатурою й мишею для реєстрації присутності, погляду та дій, формуючи події. Серверний модуль (Laravel + MySQL) класифікує порушення, а відеофрагменти зберігаються у хмарі AWS S3. Події передаються до Data Collection Service для подальшої аналітики.

Інтеграція SBMS з LMS Moodle забезпечує сенсорний зворотний зв'язок, дозволяючи системі реагувати на концентрацію, неуважність чи порушення. Для захисту даних впроваджено автентифікацію (OAuth 2.0, JWT), знеособлення, шифрування (TLS 1.3), аудит доступу та контроль інваріантів у часовій логіці. Система відповідає вимогам GDPR і FERPA, забезпечуючи безпеку й правомірність використання навчальної аналітики. Технологічний стек і розгортання наведені у (табл. 4).

Таблиця 4

Технологічний стек реалізації системи персоналізованого навчання

Компонент	Використані технології
API Gateway	C#, Nancy, OWIN
Data/Analytics Services	.NET Core, Python, PostgreSQL, MongoDB
Containerization	Docker, Docker Compose
Monitoring & Logging	Prometheus, ELK Stack, Grafana
CI/CD	GitHub Actions
Integration & Protocols	REST, JSON, OAuth2.0, OpenAPI 3.0
Storage	AWS S3 (відео), PostgreSQL / MongoDB (аналітика)

Система розгортається як контейнеризовані мікросервіси, що спрощує масштабування та підтримку. Вона забезпечує затримку обробки менше 2 секунд при понад 300 одночасних користувачах. Прототип продемонстрував ефективну інтеграцію SBMS і LMS Moodle, синхронізацію подій без затримок і стабільну роботу аналітичних сервісів під навантаженням.

Система підтвердила здатність автоматично формувати персоналізовані рекомендації на основі поведінкової аналітики, підтримуючи повний адаптивний цикл. Всі компоненти відповідають вимогам безпеки – реалізовано шифрування, автентифікацію та контроль доступу. Архітектура дозволяє гнучко масштабувати сервіси й динамічно налаштовувати персоналізацію, що робить її придатною для впровадження в освітніх платформах.

Оцінювання та результати експерименту. Метою експерименту було оцінити ефективність, масштабованість і стабільність роботи системи персоналізованого навчання на базі архітектури MAS–MSA–SBMS. Перевірялися здатність системи працювати під навантаженням, формувати точні рекомендації та адаптувати навчальні траєкторії в реальному часі.

Аналіз охоплював такі показники: швидкодія (response latency), точність персоналізації (відповідність рекомендацій експертним оцінкам), стабільність роботи SBMS, ефективність адаптації та задоволеність користувачів.

Експеримент проводився в тестовому середовищі, що максимально наближене до реальних умов використання системи у вищому навчальному закладі. Інфраструктура включала LMS Moodle 4.3, розгорнуту у контейнерах Docker із базою даних PostgreSQL, а також підсистему SBMS, реалізовану у вигляді клієнтського модуля на C#/.NET і серверного компонента на Laravel із використанням AWS S3 для зберігання даних. Мікросервісна частина складалася з компонентів API Gateway, Analytics, Adaptive Learning, Agents і Security, що взаємодіяли через REST API. Система працювала під керуванням Docker Compose на сервері Ubuntu 22.04 із восьмиядерним процесором і 32 ГБ оперативної пам'яті. Для моніторингу продуктивності та стану сервісів застосовувалися Prometheus і Grafana, що забезпечували візуалізацію навантаження та динаміку виконання запитів. Загалом у тестуванні брали участь 320 активних користувачів, серед яких було 260 студентів, 40 викладачів і 20 адміністраторів. Така конфігурація дозволила перевірити систему в умовах реалістичного сценарію університетського навчання та оцінити її поведінку під багатокористувацьким навантаженням. Експериментальна серія включала три сценарії, що відображають типові процеси роботи системи (табл. 5).

Для кожного сценарію оцінювалися швидкодія, точність, стабільність та задоволеність користувачів. Основною метрикою була Response Latency (L) – середній час реакції системи на поведінкову подію в контурі SBMS–Analytics–LMS. Accuracy of Recommendations (A) визначала точність персоналізації як відсоток збігів між рекомендаціями та оцінками викладачів. Stability Index (S) відображав частку подій, успішно

оброблених при високому навантаженні. Adaptation Effectiveness (E) характеризувала зниження повторних помилок після отримання рекомендацій. User Satisfaction (U) – середній рейтинг зручності й користі за п'ятибальною шкалою. Сукупність цих метрик дозволяє комплексно оцінити ефективність системи (табл. 6).

Таблиця 5

Сценарії експериментального оцінювання роботи системи персоналізованого навчання

№	Сценарій	Опис
S1	Моніторинг е-тестування (SBMS)	Перевірка реакції системи на події «відсутність обличчя», «перемикання вікна», «використання телефону».
S2	Аналітика активності (Analytics)	Обчислення індексу активності студентів і прогнозування ризику неуспішності.
S3	Адаптація навчальної траєкторії (Adaptive Learning)	Автоматичне формування рекомендацій для повторення тем або зміни складності завдань.

Таблиця 6

Результати оцінювання системи персоналізованого навчання

Метрика	Сценарій S1 (SBMS)	Сценарій S2 (Analytics)	Сценарій S3 (Adaptive Learning)	Середнє значення
Response Latency (L)	1.84 s	2.05 s	2.12 s	1.97 s
Accuracy (A)	–	91.3%	89.7%	90.5%
Stability (S)	98.6%	99.2%	98.9%	98.9%
Adaptation Effectiveness (E)	–	–	23.4% зниження помилок	+23.4%
User Satisfaction (U)	4.6/5	4.5/5	4.7/5	4.6/5

Аналіз отриманих результатів підтвердив високу ефективність розробленої системи персоналізованого навчання та її придатність до використання у реальному освітньому процесі. Одним із ключових показників стала швидкодія, адже середній час обробки події становив менше двох секунд, що свідчить про здатність системи працювати в режимі реального часу навіть за умов понад 300 одночасних користувачів. Це демонструє належну оптимізацію процесів взаємодії між підсистемами SBMS, аналітичними сервісами та LMS Moodle.

Не менш важливим показником є точність персоналізації, яка досягла середнього рівня 90,5%. Такий результат підтверджує тісну кореляцію між автоматично сформованими рекомендаціями системи та експертними оцінками викладачів, що свідчить про адекватність моделей аналізу поведінкових даних і коректність реалізації агентної логіки ухвалення рішень.

Показник стабільності також виявився високим: коефіцієнт успішної обробки подій на рівні 98,9% демонструє надійність підсистеми SBMS навіть за умов інтенсивного відеопотоку та високої частоти подій. Це підтверджує, що обрана мікросервісна архітектура здатна забезпечити стійку роботу системи при динамічному навантаженні.

Оцінка адаптивності показала, що після впровадження персоналізованих рекомендацій кількість повторних помилок студентів зменшилася на 23,4%. Такий результат доводить дієвість запропонованого когнітивно-аналітичного циклу, у межах якого система не лише фіксує події, але й формує ефективні стратегії навчання на основі попереднього досвіду.

Крім того, спостерігався високий рівень задоволеності користувачів – середня оцінка 4,6 із 5 пояснюється зрозумілим і звичним інтерфейсом Moodle, швидкою реакцією системи та наявністю зворотного зв'язку в реальному часі. У сукупності ці результати підтверджують, що система є технічно стабільною, педагогічно доцільною й орієнтованою на потреби кінцевих користувачів (таблиця 7).

Таблиця 7

Порівняння ефективності з іншими архітектурами

Система	Архітектура	Середня затримка	Точність рекомендацій	Стабільність	Особливості
Moodle (базова)	Монолітна	4.8 s	68.5%	94.3%	Без реальної персоналізації
Agent-LMS	MAS (без SBMS)	3.2 s	82.4%	96.8%	Обмежена аналітика
Запропонована система	MAS–MSA–SBMS	1.97 s	90.5%	98.9%	Повна адаптація, інтеграція CV

Експериментальні дослідження показали, що запропонована система:

- забезпечує низьку латентність (<2 с) при значному навантаженні;
- демонструє високу точність персоналізації (≈90%);
- гарантує надійну роботу сенсорного шару SBMS;
- успішно інтегрує агенти, аналітику та LMS у замкнений цикл адаптації.

Отримані результати підтверджують життєздатність архітектури та доводять, що інтеграція MAS, MSA та SBMS створює технологічно ефективне середовище для інтелектуального навчання.

Висновки. Архітектура реалізує когнітивно-аналітичний контур персоналізованого навчання, який поєднує агентну логіку, мікросервісну масштабованість і сенсорний моніторинг. Інтеграція агентів, сервісів і SBMS у єдину структуру забезпечує: адаптацію навчального процесу до індивідуальних потреб; безперервний збір і аналіз поведінкових даних; безпечну обробку та передачу освітньої інформації.

Наукова новизна архітектури полягає у системній інтеграції трьох взаємодоповнюючих парадигм: агентно-орієнтованого управління (MAS), мікросервісної модульності (MSA), сенсорної поведінкової аналітики (CV/SBMS).

Таке поєднання створює єдиний когнітивно-аналітичний контур, який забезпечує персоналізацію навчання на основі формалізованих поведінкових закономірностей, підвищуючи ефективність і достовірність оцінювання освітніх результатів.

Таким чином, запропонована архітектура може слугувати базовою моделлю інтелектуальної освітньої платформи, здатної до самоадаптації та масштабування в умовах розподіленого навчального середовища.

Список використаних джерел:

1. Du Plooy E., Casteleijn D., Franzsen D. Personalized adaptive learning in higher education: A scoping review of key characteristics and impact on academic performance and engagement. *Heliyon*. 2024. Vol. 10, № 21. DOI: 10.1016/j.heliyon.2024.e39630.
2. Schicchi D., Taibi D. Redefining education: A personalized AI platform for enhanced learning experiences. *Proceedings of the Second International Workshop on Artificial Intelligence Systems in Education (AIXIA 2024)*. CEUR Workshop Proceedings, 2024.
3. Sajja R., Sermet Y., Cwiertny D. та ін. Integrating AI and Learning Analytics for Data-Driven Pedagogical Decisions and Personalized Interventions in Education. *Tech Know Learn*. 2025. DOI: 10.1007/s10758-025-09897-9.
4. Córdova-Esparza D. M. AI-powered educational agents: Opportunities, innovations, and ethical challenges. *Information*. 2025. Vol. 16, № 6. P. 469. DOI: 10.3390/info16060469.
5. Han Y., Hong S., Li Z. та ін. Defining and Classifying the Roles of Intelligent Learning Companion Systems: A Scoping Review of the Literature. *TechTrends*. 2025. Vol. 69. P. 567–581. DOI: 10.1007/s11528-025-01058-0.
6. Zhang X., Zhang C., Sun J., Xiao J., Yang Y., Luo Y. EduPlanner: LLM-based multi-agent systems for customized and intelligent instructional design. *IEEE Transactions on Learning Technologies*. 2025.
7. Ren X., Wang H., Cai T. T. Design and Implementation of a Microservices-Based Online Learning. *Proc. of EIMT 2023*. Springer, 2023. Vol. 8. P. 455. DOI: 10.2991/978-94-6463-192-0_60.
8. Lysenko R., Skorokhoda O. Enhancing adaptive systems with Intelligent Agents in Microservice Architectures. *Proc. of SMARTINDUSTRY 2025*. P. 241–254. URL: <https://ceur-ws.org/Vol-3970/>.
9. Bernard C., Garcia D., Fox A. The development and management of GradeSuite: A microservice LMS for mastery learning. *Tech. Rep. EECS-2025-127*. Univ. of California, Berkeley, 2025. URL: <https://www2.eecs.berkeley.edu/Pubs/TechRpts/2025/EECS-2025-127.pdf>.
10. Shaiba H., Hadjouni M., John M. Microservices-based student support framework (MicSSF) to enhance equity in education. *Computer Applications in Engineering Education*. 2023. Vol. 31, № 4. P. 884–899.
11. Axak N., Kushnaryov M., Tatarnykov A. Agent-driven approach to enhancing e-learning efficiency. *Advances in Information Control Systems and Technologies*: монографія / за ред. проф. В. Вичужаніна. Львів–Торунь: Liha-Pres, 2024. 380 с. DOI: 10.36059/978-966-397-422-4.
12. Аксак Н., Татарников А. Обзор систем управления электронным обучением. *Традиційні та інноваційні підходи до наукових досліджень*: матеріали IV Міжнар. наук. конф. Житомир, 2023. С. 131–136. ISBN 978-617-8126-12-4. DOI: 10.36074/mcnd-10.02.2023.

References:

1. Du Plooy, E., Casteleijn, D., & Franzsen, D. (2024). Personalized adaptive learning in higher education: A scoping review of key characteristics and impact on academic performance and engagement. *Heliyon*, 10(21). <https://doi.org/10.1016/j.heliyon.2024.e39630>
2. Schicchi, D., & Taibi, D. (2024, November). Redefining education: A personalized AI platform for enhanced learning experiences. In *Proceedings of the Second International Workshop on Artificial Intelligence Systems in Education (AIXIA 2024)*. CEUR Workshop Proceedings.

-
3. Sajja, R., Sermet, Y., Cwierny, D., et al. (2025). Integrating AI and learning analytics for data-driven pedagogical decisions and personalized interventions in education. *Tech Know Learn*. <https://doi.org/10.1007/s10758-025-09897-9>
 4. Córdova-Esparza, D. M. (2025). AI-powered educational agents: Opportunities, innovations, and ethical challenges. *Information*, 16(6), 469. <https://doi.org/10.3390/info16060469>
 5. Han, Y., Hong, S., Li, Z., et al. (2025). Defining and classifying the roles of intelligent learning companion systems: A scoping review of the literature. *TechTrends*, 69, 567–581. <https://doi.org/10.1007/s11528-025-01058-0>
 6. Zhang, X., Zhang, C., Sun, J., Xiao, J., Yang, Y., & Luo, Y. (2025). EduPlanner: LLM-based multi-agent systems for customized and intelligent instructional design. *IEEE Transactions on Learning Technologies*.
 7. Ren, X., Wang, H., & Cai, T. T. (2023, July). Design and implementation of a microservices-based online learning. In *Proceedings of the 2023 2nd International Conference on Educational Innovation and Multimedia Technology (EIMT 2023)* (Vol. 8, p. 455). Springer. https://doi.org/10.2991/978-94-6463-192-0_60
 8. Lysenko, R., & Skorokhoda, O. (2025). Enhancing adaptive systems with intelligent agents in microservice architectures: Opportunities and challenges. In *Proceedings of the 2nd International Conference on Smart Automation & Robotics for Future Industry (SMARTINDUSTRY 2025)* (pp. 241–254). CEUR-WS. Retrieved from: <https://ceur-ws.org/Vol-3970/>
 9. Bernard, C., Garcia, D., & Fox, A. (2025). *The development and management of GradeSuite: A microservice LMS for mastery learning* (Tech. Rep. No. EECS-2025-127). University of California, Berkeley. Retrieved from: <https://www2.eecs.berkeley.edu/Pubs/TechRpts/2025/EECS-2025-127.pdf>
 10. Shaiba, H., Hadjouni, M., & John, M. (2023). Microservices-based student support framework (MicSSF) to enhance equity in education. *Computer Applications in Engineering Education*, 31(4), 884–899.
 11. Axak, N., Kushnaryov, M., & Tatarnykov, A. (2024). Agent-driven approach to enhancing e-learning efficiency. In *Advances in Information Control Systems and Technologies* (pp. 1–18). Liha-Pres. <https://doi.org/10.36059/978-966-397-422-4>
 12. Axak, N. & Tatarnykov, A. (2023). Ohljad system upravlinnja elektronnym navchannjam. In *Tradicijni ta innovacijni pidkhody do naukovykh doslidzhen': Materialy IV mizhnar. nauk. konf.* (pp. 131–136). <https://doi.org/10.36074/mcnd-10.02.2023>

Дата надходження статті: 26.10.202

Дата прийняття статті: 17.11.2025

Опубліковано: 30.12.2025

Божко О. Ю., аспірант кафедри інформаційних
управляючих систем
Харківського національного університету радіоелектроніки
ORCID: 0009-0004-6820-1228

МЕТОД КЛАСИФІКАЦІЇ ДОКУМЕНТІВ ЗА СКЛАДНІСТЮ ЕКСТРАКЦІЇ ДАНИХ ВЕЛИКИМИ МОВНИМИ МОДЕЛЯМИ

У статті розглянуто актуальну проблему оптимізації автоматизованої екстракції даних з ділових документів за допомогою великих мовних моделей (LLM). Якість обробки документів суттєво варіюється в залежності від їх структурних та семантичних характеристик. Відсутність методів попереднього прогнозування якості екстракції призводить до неефективного використання ресурсів. Існуючі дослідження в галузі класифікації документів орієнтовані на тематичну категоризацію, а не на оцінку технічної складності отримання даних.

Для розв'язання даної проблеми запропоновано метод класифікації документів за складністю їх обробки з використанням великих мовних моделей. Метод базується на розмітці документів за бінарними ознаками структурно-семантичної складності. Для кожного документа виконується автоматична екстракція даних трьома мовними моделями у режимі без прикладів з обчисленням інтегральної метрики якості екстракції через гармонійне середнє точності та повноти. На основі метрик формуються класи складності, після чого будуються класифікатори з використанням багатокласової логістичної регресії. Валідація здійснюється через стратифіковану перехресну перевірку.

Ключовою особливістю методу є можливість автоматичного визначення очікуваної якості обробки документа на основі його формалізованих характеристик. Експериментальну перевірку працездатності метода здійснено на корпусі синтетичних документів з варійованими характеристиками складності. Для трьох LLM побудовано класифікатори треступеневої складності. Аналіз вагових коефіцієнтів виявив критичні фактори складності, які демонструють найбільший негативний вплив на якість екстракції.

Запропоноване рішення має як теоретичне так і практичне значення. Наукова новизна полягає у створенні оригінального емпірично обґрунтованого методу класифікації документів, де цільовою змінною виступає очікувана якість екстракції даних мовними моделями. Практична цінність розробленого методу полягає у можливості автоматизованого прийняття рішень про стратегію обробки документів в організаційних системах. Отримані результати створюють підґрунтя для розвитку інтелектуальних систем обробки документів та оптимізації використання обчислювальних ресурсів.

Ключові слова: велика мовна модель, екстракція даних, класифікація документів, бінарна ознака складності, логістична регресія, метрика якості, промпт.

Bozhko O. Yu. Method for classifying documents according to the complexity of data extraction by large language models

The article addresses the urgent problem of optimizing automated data extraction from business documents using large language models (LLM). The quality of processing documents varies significantly depending on their structural and semantic characteristics. The absence of methods for predicting extraction quality leads to inefficient resource utilization. Existing document classification research focuses on thematic categorization rather than assessing the technical complexity of data extraction.

To address this problem, a method for classifying documents based on the complexity of their processing by language models is proposed. The method is based on document markup using binary features of structural-semantic complexity. For each document, automatic data extraction is performed by three language models in zero-shot mode, with calculation of an integral quality metric through the harmonic mean of precision and recall. Based on these metrics, complexity classes are formed, followed by construction of classifiers using multiclass logistic regression. Validation is performed through stratified cross-validation.

The key feature of the method is the ability to automatically determine the expected quality of document processing based on its formalized characteristics. The method was tested on a corpus of synthetic documents with varied complexity characteristics. For three LLMs, three-level complexity classifiers were built. Analysis of weight coefficients revealed critical complexity factors that demonstrate the greatest negative impact on extraction quality.

The proposed solution has both theoretical and practical significance. The scientific novelty lies in creating the first empirically validated method for classifying documents where the target variable is the expected quality of data extraction by language models. The practical value is the possibility of automated decision-making regarding processing strategy in production systems. The results create a methodological foundation for developing intelligent document processing systems and optimizing computational resource utilization.

Key words: large language models, data extraction, document classification, binary complexity features, logistic regression, quality metrics, prompt.

© О. Ю. Божко, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

Постановка проблеми. Автоматизація обробки ділових документів із використанням великих мовних моделей (LLM) стає поширеною практикою в корпоративному середовищі, охоплюючи процеси від розпізнавання інвойсів до витягування даних з контрактів та актів. Однак якість екстракції слабоструктурованих даних істотно варіюється в залежності від характеристик конкретного документа. Одні із високою точністю оброблюються простим підходом без прикладів в промптах (zero-shot), в той час, як інші вимагають певної кількості прикладів (few-shot) для забезпечення відповідної якості екстракції. Відсутність попереднього прогнозування призводить до невиправданого використання дороговартісних методів обробки або, навпаки, до неприйнятно низької точності.

Наявні дослідження в цій галузі переважно зосереджуються на підвищенні точності екстракції post-factum (тонке налаштування промптів, постобробка результатів), а не на попередньому оцінюванні очікуваного результату. На теперішній час відсутні формалізовані підходи до кількісної оцінки складності документів для LLM-обробки, що вибір оптимальної стратегії екстракції. Існуючі методи класифікації документів орієнтовані переважно на тематичну категоризацію, а не на оцінку технічної складності автоматичного витягування структурованих даних. Крім того, відсутня методологія, що дозволила б масштабувати процес оцінки складності на великі корпуси документів без залучення експертів-аналітиків.

Тому актуальним є завдання розробки автоматичного методу класифікації документів за складністю екстракції, який би дозволив швидко відфільтровувати прості документи, та направляти складні документи на поглиблене вивчення факторів для визначення оптимальної стратегії їх обробки. Такий метод має забезпечити раціональне використання обчислювальних ресурсів. Він має поєднувати формалізовану систему оцінки складності документів, швидкий ML-класифікатор для їх стратифікації та селективний детальний аналіз факторів складності виключно для проблемних випадків.

Аналіз останніх досліджень і публікацій. Сучасні ділові процеси потребують обробки значних обсягів цифрових документів, які містять слабоструктуровані дані: інвойси, накладні, контракти, специфікації [1]. Зазвичай, такі документи створюються програмно, але їх автоматична обробка залишається складною задачею через варіативність макетів, семантичну неоднозначність та відсутність єдиних стандартів форматування. Застосування мультимодальних LLM на базі трансформерної архітектури відкрило нові можливості для екстракції даних завдяки здатності до контекстуального розуміння без попереднього навчання на конкретних шаблонах [2]. Однак питання оптимального вибору методу обробки залежно від характеристик конкретного документа залишається малодослідженим.

Мовні моделі ефективно використовувалися для екстракції даних з ділових документів ще до так званої «LLM революції». Так ще в 2020 році була продемонстрована можливість аналізу попередньо навченими моделями просторової інформації документів для підтвердження точності екстракції [3].

Після того, як великі мовні моделі стали масово доступними, численні дослідження підтвердили якісну перевагу систем, що використовують LLM-підходи, над традиційними системами, заснованими на правилах [4].

Водночас виявляються і критичні обмеження LLM-екстракції. Наприклад, істотна залежність якості екстракції від структурної складності документів [5] та проблеми з обробкою великих документів через обмеження контекстного вікна [6]. Були зафіксовані помилки в математичних обчисленнях та випадки генерації неіснуючих даних [7]. Крім того, в сучасних дослідженнях підкреслюється необхідність валідації результатів у зв'язку з ризиком генерації правдоподібних, але неточних значень [8].

Разом з цим значною прогалиною є відсутність методів попереднього оцінювання якості екстракції. Наявні дослідження фокусуються на підвищенні точності через налаштування промптів або постобробку, але не пропонують способів прогнозування ефективності перед застосуванням LLM-процедур.

Класифікація документів традиційно вирішує задачу категоризації за типом або тематичною приналежністю. У таких дослідженнях предикторами виступають структурні, текстові та візуальні характеристики документів, а цільовою змінною – належність до певного класу (наприклад, рахунок-фактура, контракт, квитанція, медичний запис).

Так в [4] для класифікації медичних документів дослідники використовували в якості предикторів частоту термінів, довжину тексту, наявність специфічних ключових слів, а метою класифікації було визначення типу клінічного документа. В [9] для класифікації документів цифрового розвитку використовуються TF-IDF вектори, лексичне різноманіття та структурні метрики в якості предикторів для віднесення документів до тематичних категорій.

Конструювання предикторів є критичним етапом побудови класифікаторів. Структурні ознаки включають кількість таблиць, рядків, колонок, щільність тексту, наявність вкладених структур. Текстові предиктори охоплюють частотні характеристики термінів, довжину документа, мовне різноманіття. Візуальні ознаки враховують розташування елементів на сторінці, типографічні параметри, геометричні характеристики текстових блоків. Дослідження класифікації біомедичних документів [10, 11] та документів за структурними ознаками [12] демонструють ефективність комбінування різнотипних предикторів для підвищення точності визначення типу документа.

Недостатня кількість досліджень, де цільовою змінною класифікації документів є очікувана якість екстракції даних суттєво ускладнює попереднє прогнозування ефективності обробки та раціональний вибір стратегії екстракції.

Мета дослідження полягає у розробці метода класифікації документів за ознаками складності екстракції слабоструктурованих даних великими мовними моделями, який би забезпечив оптимізацію використання обчислювальних ресурсів через поєднання швидкого ML-класифікатора та селективного детального аналізу факторів складності.

Виклад основного матеріалу. Для розв'язання задачі класифікації документів за складністю екстракції даних великими мовними моделями в роботі пропонується поетапний метод. Цей метод включає реалізацію п'яти послідовних етапів.

1. Формування корпусу документів з маркуванням за бінарними критеріями складності.
2. Проведення екстракції за допомогою LLM у контрольованому режимі без прикладів (zero-shot) з розрахунком еталонних метрик якості.
3. Формування класів складності на основі емпіричного розподілу якості екстракції.
4. Побудова ML-класифікаторів для автоматичного визначення класу складності документа.
5. Валідація класифікаторів та оцінка економічної ефективності підходу.

Розглянемо детальний опис кожного з етапів.

Етап 1. В ході реалізації цього етапу створюється навчальний корпус документів з систематизованим описом їх характеристик складності та еталонними мітками якості екстракції.

Для формалізованого опису властивостей документів, що впливають на якість автоматичної екстракції, потрібно спочатку сформувати систему бінарних ознак їх складності. Кожна з ознак складності представляє собою індикатор, який фіксує наявність конкретної характеристики, яка ускладнює обробку документів.

Формування переліку ознак базується на аналізі збоїв у системах екстракції даних, а також на системному аналізі літературних джерел. Їх детальний опис наведено в (табл. 1).

Таблиця 1

Система формальних ознак складності документів для LLM-обробки

№	Ознака складності c_{ij}	Опис	Механізм впливу на LLM
1.	dense_listing	Документ містить понад 50 табличних записів	Когнітивне перевантаження, проблеми з контекстним вікном
2.	has_multilingual_content	Наявність інформації, що надана двома або більше мовами	Міжмовна інтерференція
3.	has_ambiguous_fields	Скорочення без розшифрування	Неоднозначність інтерпретації
4.	has_unlabeled_data	Відсутність явних маркерів типу даних	Складність визначення семантичної ролі
5.	has_compound_cells	Множинні значення в одному полі	Порушення атомарності даних
6.	has_multiline_records	Табличні записи займають декілька рядків	Порушення очікування «один рядок = один запис»
7.	has_spanning_data	Дані перетинають границі колонок/рядків	Порушення табличної структури
8.	has_header_data_mismatch	Невідповідність заголовків та даних	Протиріччя між очікуваннями та реальністю
9.	requires_calculations	Потреба в арифметичних обчисленнях	Не всі LLM однаково точні в математиці
10.	has_contextual_dependencies	Дані з заголовка застосовуються до всіх записів	Вимагає розуміння ієрархії документа
11.	has_implicit_structure	Логічна організація без явних маркерів	Потреба в виведенні структури з контексту

Розмітка цих 11 ознак складності здійснюється на рівні типу документа через розробку узгоджених правил присвоєння бінарних значень. Результатом є «профіль складності документа i » – вектор характеристик складності $C_i = (c_{i1}, c_{i2}, \dots, c_{i11})$, де $c_{ij} \in \{0, 1\}$, $i = 1, n$, $j = 1, 11$.

Формування навчального корпусу здійснюється за принципом забезпечення репрезентативності покриття комбінацій ознак складності при збереженні реалістичності структури та змісту документів. Заповнення шаблонів здійснюється з використанням реалістичних, але деперсоналізованих даних, що дозволяє зберегти автентичність структурно-семантичних характеристик документів при дотриманні вимог конфіденційності. Паралельно з генерацією документів створюються еталонні файли розмітки.

Формально корпус документів $i = \overline{1, n}$ описується матрицею ознак складності $C = (c_{ij})_{n \times 11}$, де рядки відповідають документам, а стовпці – бінарним ознакам складності $j = \overline{1, 11}$.

Контроль якості сформованого корпусу включає верифікацію покриття простору ознак, аналіз збалансованості розподілу характеристик складності та перевірку відсутності надлишкових кореляцій між предикторами. Сформований корпус використовується як навчальна і валідаційна база для побудови та оцінки регресійних моделей прогнозування якості екстракції даних на рівні документа.

Етап 2. Для забезпечення коректності порівняння застосовується єдиний протокол обробки всіх документів. Екстракція виконується в zero-shot режимі, що дозволяє оцінити базову здатність моделей до обробки документів різного рівня складності без додаткового налаштування під конкретні формати.

Для екстракції використовуються три LLM з різною архітектурою: GPT-4.1-nano, GPT-4o-mini та Gemini-1.5-flash з фіксованими параметрами декодування для кожної з моделей. Результат екстракції формується у вигляді структурованих даних відповідно до попередньо визначеної схеми, що специфікує назви полів, типи даних та кардинальності записів.

Для кожного документа d_i , $i = \overline{1, 115}$, кожна модель $m = \overline{1, 3}$ генерує набір екстрагованих записів E_{im} , який зберігається для подальшого порівняння з еталонною розміткою R_i .

Якість екстракції оцінюється зіставленням еталонної розмітки документа з результатами екстракції. Для кожної комбінації (документ d_i , модель m) виконується процедура порівняння екстрагованих записів з еталонними на основі подібності значень полів. На основі цього для кожного спостереження визначаються:

TP_{im} (true positives) – кількість коректно екстрагованих записів, що мають відповідний еталонний запис;

FP_{im} (false positives) – кількість помилково екстрагованих записів, що не мають відповідності в еталоні;

FN_{im} (false negatives) – кількість пропущених записів, що присутні в еталоні, але не екстраговані моделлю.

На основі цих величин обчислюються метрики точності (Precision) та повноти (Recall) екстракції:

$$\text{Precision}_{im} = \frac{TP_{im}}{TP_{im} + FP_{im}} \quad (1)$$

$$\text{Recall}_{im} = \frac{TP_{im}}{TP_{im} + FN_{im}} \quad (2)$$

Інтегральна оцінка якості екстракції обчислюється як гармонічне середнє точності та повноти – F1-міра:

$$F1_{im} = 2 \cdot \frac{\text{Precision}_{im} \cdot \text{Recall}_{im}}{\text{Precision}_{im} + \text{Recall}_{im}}, \quad i = \overline{1, 115}, \quad m = \overline{1, 3} \quad (3)$$

Результатом виконання другого етапу є три навчальні вибірки по 115 документів для розв'язання задачі класифікації. В контексті задачі предикторами виступають 11 бінарних характеристик складності документа: $C_i = (c_{i1}, c_{i2}, \dots, c_{i11})$. Відповідно, цільовою змінною є якість екстракції $F1_{im} \in [0, 1]$.

Етап 3. Залежність якості екстракції від моделі стає аргументом на користь вибору модель-специфічної стратифікації за складністю. Для кожної моделі визначаються індивідуальні пороги. Такий підхід забезпечує можливість застосування методу до нових моделей на тому самому корпусі документів.

Пороги класів встановлюються виходячи з практичного призначення кожного класу в контексті вибору стратегії обробки документів. Розглядаються три категорії складності, що відповідають різним підходам до екстракції даних:

- **easy** – документи для яких базовий zero-shot підхід забезпечує стабільно високу якість екстракції;
- **hard** – документи які потребують застосування підсилених стратегій або, у крайніх випадках, ручної обробки;
- **medium** – документи де базовий підхід може бути недостатнім, але додаткові зусилля (наприклад, налаштування промпту) здатні підвищити якість до прийняттого рівня.

На основі аналізу практичних застосувань екстракції даних та врахування стандартів якості в задачах машинного навчання встановлено наступні порогові значення: $T_{\text{easy}} = 0.85, T_{\text{hard}} = 0.70$.

Для кожної моделі m виконується класифікація 115 результатів екстракції за встановленими порогоми:

$$\text{class}_m(F1_{im}) = \begin{cases} \text{Easy, якщо } F1_{im} > 0.85 \\ \text{Medium, якщо } 0.70 < F1_{im} \leq 0.85 \\ \text{Hard, якщо } F1_{im} \leq 0.70 \end{cases} \quad (4)$$

Результатом виконання третього етапу є формування трьох навчальних вибірок для задачі класифікації – по одній для кожної моделі. Кожна вибірка складається зі 115 спостережень, де предикторами виступають 11

бінарних ознак складності документа C_i , а цільовою змінною – клас складності $class_m \in \{Easy, Medium, Hard\}$, визначений на основі практично обґрунтованих порогів.

Етап 4. Для розв'язання задачі багатокласової класифікації обрано метод логістичної регресії з регуляризацією. Вибір зумовлений наступними чинниками: інтерпретовність моделі через коефіцієнти вагових внесків ознак, стабільність роботи на обмежених вибірках, відсутність схильності до перенавчання при невеликій кількості предикторів, детермінованість результатів.

Класифікація виконується у два послідовних етапи: спочатку проводиться стандартизація вхідних ознак, після чого застосовується багатокласова логістична регресія. Стандартизація застосовується без центрування, оскільки бінарні ознаки мають фіксований діапазон $[0,1]$. Логістична регресія налаштована з параметрами: схема "один проти решти" (`multi_class='ovr'`), збалансовані вагові коефіцієнти класів (`class_weight='balanced'`) для компенсації нерівномірного розподілу спостережень, обрана максимальна кількість ітерацій 2000 для гарантованої збіжності оптимізації.

Для об'єктивної оцінки якості класифікації застосовано стратифіковану перехресну валідацію (`stratified k-fold cross-validation`). Кількість фолдів визначається адаптивно: $k = \min(5, n_{Hard})$, де n_{Hard} – обсяг найменшого класу у вибірці моделі.

Для оцінювання якості класифікації використовуються наступні метрики: загальна точність (`accuracy`), F1-міра для кожного класу окремо, макроусереднена F1-міра (`macro-averaged F1`), збалансована точність (`balanced accuracy`), що враховує нерівномірність класів через усереднення частки правильних класифікацій у кожному класі.

Результатом виконання четвертого етапу є навчені класифікатори – по одному для кожної великої мовної моделі – що дозволяють за вектором 11 бінарних ознак документа автоматично визначати очікуваний клас складності його обробки. Ці класифікатори інтегруються в систему маршрутизації документопотоків, де на основі передбаченого класу приймається рішення про застосування базової (`zero-shot`), підсиленої (`few-shot`) або альтернативної стратегії екстракції даних.

Для кожної з трьох досліджених мовних моделей (GPT-4.1-nano, GPT-4o-mini, Gemini-1.5-flash) побудовано окремий класифікатор на основі багатокласової логістичної регресії. Класифікатор представлено набором числових коефіцієнтів, що визначають вагові внески кожної з 11 бінарних ознак складності документа у визначення класу.

Нехай $\mathbf{x} = (x_1, x_2, \dots, x_{11})$ – вектор бінарних ознак складності документа, де $x_j \in \{0,1\}$ відповідає наявності або відсутності j -ї характеристики складності. Для кожного класу $k \in \{Easy, Medium, Hard\}$ обчислюється лінійна комбінація ознак:

$$z_k(\mathbf{x}) = w_{k0} + \sum_{j=1}^{11} w_{kj} \cdot x_j \quad (5)$$

де w_{k0} – вільний член для класу k , а w_{kj} – ваговий коефіцієнт j -ї ознаки для класу k .

Ймовірність приналежності документа до класу k визначається логістичною функцією (сигмоїдою):

$$p_k(x) = \sigma(z_k(x)) = \frac{1}{1 + \exp\{-z_k(x)\}} \quad (6)$$

Предбачений клас складності визначається як:

$$\hat{y}(x) = \arg \max_k p_k(x) \quad (7)$$

Етап 5. Для документа класу c з істинною складністю c_{true} та передбаченою складністю c_{pred} математичне сподівання втрат від помилкової класифікації обчислюється за формулою:

$$E[Cost] = \sum_{c_{true}} \sum_{c_{pred}} P(c_{pred} | c_{true}) \cdot L(c_{true}, c_{pred}), \quad (8)$$

де $L(c_{true}, c_{pred})$ – функція втрат, що визначає критичність помилкової класифікації. Конкретні значення цієї функції визначаються вимогами конкретної прикладної задачі.

У подальшому приймаємо консервативне припущення: пропуск помилки коштує суттєво дорожче за зайву перевірку, тому документи класу `Easy` обробляємо автоматично, а `Medium` і `Hard` – підлягають експертній верифікації.

На основі матриці невідповідностей, що отримана на етапі 4, розраховуються частота критичних помилок та частота надлишкових перевірок:

$$P_{critical} = \frac{FN_{Easy}}{N_{Medium} + N_{Hard}} \quad (9)$$

$$P_{redundant} = \frac{FP_{Easy}}{N_{Easy}} \quad (10)$$

Аналіз та обговорення результатів дослідження. Результатом дослідження є розроблений метод автоматичної класифікації документів за складністю екстракції даних великими мовними моделями. Цей метод реалізовано з використанням математичної моделі, що приймає на вхід профіль складності документа та повертає передбачуваний клас складності обробки. Нижче наведено аналіз результатів, що отримані на проміжних етапах та в ході експериментального дослідження.

Після завершення другого етапу було проведено оцінку зв'язку між 11 бінарними критеріями складності та метрикою $F1_{doc}$ за допомогою поінт-бісеріальних коефіцієнтів кореляції. Загалом сила зв'язку оцінюється як помірна. Вплив критеріїв складності на якість екстракції наведено в (табл. 2).

Таблиця 2

Вплив критеріїв складності на якість екстракції

Критерій	GPT-4o-mini	GPT-4.1-nano	Gemini 1.5 Flash	Середнє r	Ранг (за середнім r)
contextual_dependencies	-0.553	-0.546	-0.581	0.560	1
unlabeled_data	-0.563	-0.476	-0.598	0.546	2
spanning_data	-0.430	-0.462	-0.449	0.447	3
compound_cells	-0.335	-0.397	-0.391	0.374	4
header_data_mismatch	-0.364	-0.265	-0.463	0.364	5
implicit_structure	-0.429	-0.260	-0.357	0.349	6
ambiguous_fields	-0.258	-0.297	-0.469	0.341	7
requires_calculations	-0.131	-0.330	-0.277	0.246	8
dense_listing	-0.196	-0.282	-0.192	0.223	9
multilingual_content	0.000	-0.164	-0.283	0.149	10
multiline_records	-0.089	-0.134	-0.170	0.131	11

У (табл. 3) наведено емпіричний розподіл документів за класами складності для кожної великої мовної моделі. Класи сформовано за фіксованими порогами (згідно з (4)). Розподіли відображають реальну різницю у поведінці моделей на одному й тому самому корпусі.

Таблиця 3

Розподіл за класами складності та характеристики якості

Модель	Клас	N	%	Діапазон F1	Середнє	Std
GPT-4.1-nano	Easy	49	42.6%	0.856–0.965	0.903	0.036
	Medium	54	47.0%	0.714–0.849	0.797	0.039
	Hard	12	10.4%	0.504–0.681	0.633	0.061
GPT-4o-mini	Easy	62	53.9%	0.854–0.965	0.910	0.036
	Medium	43	37.4%	0.705–0.850	0.784	0.036
	Hard	10	8.7%	0.495–0.690	0.627	0.073
Gemini-1.5-flash	Easy	88	76.5%	0.855–0.993	0.921	0.036
	Medium	25	21.7%	0.774–0.850	0.816	0.019
	Hard	2	1.7%	0.673–0.680	0.677	0.004

Розподіл документів по класах суттєво відрізняється між моделями при застосуванні однакових порогових значень, що підтверджує різницю в їх можливостях обробки документів. Модель Gemini демонструє найбільшу частку Easy-документів (76.5%), що вказує на її високу ефективність для даного корпусу. Натомість GPT-4.1-nano класифікує лише 42.6% документів як Easy, при цьому майже половина документів (47.0%) потрапляє до класу Medium, що свідчить про більш широкий діапазон складності обробки для цієї моделі.

Особливу увагу привертає клас Hard для моделі Gemini: лише 2 документи (1.7%) демонструють якість екстракції нижче порогу 0.70. Така мала вибірка обмежує статистичну надійність характеристик цього класу для даної моделі (std=0.004 на основі двох спостережень), проте підтверджує загальну високу ефективність Gemini для обробки документів представленого корпусу.

У (табл. 4) наведено розподіл спостережень за класами складності.

Нерівномірний розподіл класів виникає внаслідок застосування фіксованих порогів до емпіричних розподілів якості екстракції, що суттєво відрізняються між моделями. Для моделі Gemini клас Hard представлений лише двома спостереженнями, що обмежує статистичну надійність оцінок якості класифікації для цього класу. Метрики якості класифікації наведені в (табл. 5).

Таблиця 4

Розподіл спостережень за класами складності

Модель	Easy	Medium	Hard	Кількість фолдів
GPT-4.1-nano	49	54	12	5
GPT-4o-mini	62	43	10	5
Gemini-1.5-flash	88	25	2	2

Таблиця 5

Метрики якості класифікації (stratified k-fold CV)

Модель	Accuracy	F1 (Easy)	F1 (Medium)	F1 (Hard)	F1 (macro)	Balanced Accuracy
GPT-4.1-nano	0.704	0.800	0.704	0.370	0.625	0.632
GPT-4o-mini	0.696	0.779	0.667	0.444	0.630	0.669
Gemini-1.5-flash	0.661	0.781	0.453	0.000	0.411	0.402

Класифікатори демонструють задовільну загальну точність (66 – 70%), проте якість розпізнавання класу Hard суттєво нижча порівняно з Easy та Medium. Для моделі Gemini F1-міра класу Hard дорівнює нулю, що відображає неможливість надійного навчання на двох спостереженнях. Макроусереднена F1-міра варіює від 0.411 (Gemini) до 0.630 (GPT-4o-mini), що вказує на помірну якість класифікації за умов обмеженого обсягу навчальних даних та нерівномірного розподілу класів.

Детальний аналіз помилок класифікації для моделі GPT-4o-mini надано у (табл. 6) (матриця невідповідностей, confusion matrix).

Таблиця 6

Матриця невідповідностей для GPT-4o-mini (OOF передбачення)

	Передбачено Easy	Передбачено Medium	Передбачено Hard
Фактичний Easy	44	14	4
Фактичний Medium	6	30	7
Фактичний Hard	1	3	6

Аналіз матриці свідчить про типовий патерн помилок. Класифікатор схильний недооцінювати складність (помилкова класифікація Medium як Easy зустрічається частіше, ніж переоцінка). Для класу Hard спостерігається значна частка помилкових віднесення до сусідніх класів через малий обсяг навчальної вибірки цього класу.

Коефіцієнти логістичної регресії дозволяють кількісно оцінити внесок кожної ознаки у класифікацію. Порівняння коефіцієнтів між моделями показує узгодженість у визначенні критичних факторів складності. Ознаки compound_cells, dense_listing та contextual_dependencies стабільно демонструють високі абсолютні значення коефіцієнтів для всіх трьох архітектур LLM.

Отримані класифікатори забезпечують базовий рівень автоматизації визначення складності документів, достатній для практичного впровадження диференційованих стратегій обробки. Водночас слід враховувати наступні обмеження методу.

По-перше, обмежений обсяг вибірки класу Hard (особливо для моделі Gemini: 2 спостереження) призводить до низької статистичної надійності метрик якості для цього класу. Оцінки F1-міри для Hard характеризуються високою дисперсією та можуть не відображати справжньої здатності класифікатора розпізнавати складні документи на ширших корпусах.

По-друге, використання бінарних ознак спрощує процедуру розмітки документів, проте не дозволяє враховувати градації прояву характеристик складності. Документи з однаковими бінарними профілями можуть суттєво відрізнятися за фактичною складністю обробки.

По-третє, застосування фіксованих порогів (0.70 та 0.85) для формування класів призводить до нерівномірного розподілу спостережень, що обмежує ефективність навчання класифікаторів на міноритарних класах.

Визначення 11 бінарних ознак складності документа може бути **автоматизоване** шляхом застосування алгоритмів структурного аналізу документів. **На практиці** для класифікації нового документу необхідно визначити вектор його бінарних ознак складності, застосувати формули вище з відповідним набором коефіцієнтів та отримати передбачений клас і рекомендацію щодо стратегії обробки.

Висновки та перспективи подальших досліджень. В роботі **вперше отримано** емпірично обґрунтовану систему класифікації документів за складністю їх автоматичної обробки великими мовними моделями, де цільовою змінною є очікувана якість екстракції, виміряна через метрику F1.

Удосконалено методику оцінювання якості екстракції слабоструктурованих даних через розробку композитної метрики на рівні документа, що враховує якість обробки різних семантичних компонентів з відповідними ваговими коефіцієнтами.

Розроблений метод створює основу для автоматизованого прийняття рішень про вибір стратегії обробки документів у виробничих системах. Класифікатор дозволяє на етапі надходження документа передбачити очікувану якість його автоматичної обробки та обрати оптимальну стратегію: повна автоматизація для документів класу Easy, експертна перевірка для класів Medium та Hard.

Метод апробовано на корпусі з 115 синтетичних ділових документів з систематично варійованими характеристиками складності. Результати експериментів підтверджують статистично значущий зв'язок між формалізованими ознаками складності та якістю екстракції даних для трьох сучасних архітектур мовних моделей.

Перспективи подальших досліджень пов'язані з кількома напрямками розвитку запропонованого методу. Інтеграція неперервних кількісних характеристик документів поряд з бінарними ознаками може підвищити точність класифікації. До таких характеристик належать: щільність текстової інформації на сторінці, кількість табличних елементів, співвідношення структурованих та неструктурованих фрагментів.

Валідація методу на документах інших класів дозволить оцінити універсальність запропонованої системи ознак та виявити специфічні фактори складності для різних предметних областей.

Перевірка класифікаторів на нових архітектурах мовних моделей, що з'являться після завершення цього дослідження, забезпечить оцінку темпоральної стабільності результатів та виявить потребу в періодичному переналаштуванні моделей.

Побудова окремих класифікаторів для підмножин документів з домінуванням певних типів складності (наприклад, багатомовні документи, документи з матричними структурами) може забезпечити вищу точність передбачень у порівнянні з єдиною універсальною моделлю.

Список використаних джерел:

1. Божко О. Ю. Використання великих мовних моделей для розпізнавання інформації в нумізматичних описах. *Наука і техніка сьогодні*. 2024. № 1(29). С. 615–625. DOI: 10.52058/2786-6025-2024-1(29)-615-625.
2. Божко О. Ю. Розробка ітеративного методу екстракції даних з неструктурованих документів на основі використання великих мовних моделей. *Вісник Кременчуцького національного університету імені Михайла Остроградського*. 2025. № 1. С. 119–124. DOI: 10.32782/1995-0519.2025.1.15.
3. Xu Y., Li M., Cui L., Huang S., Wei F., Zhou M. LayoutLM: pre-training of text and layout for document image understanding. *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '20)*. 2020. С. 1192–1200. DOI: 10.1145/3394486.3403172.
4. Rijcken E., Zervanou K., Mosteiro P., Scheepers F., Spruit M., Kaymak U. Machine learning vs. rule-based methods for document classification of electronic health records within mental health care: a systematic literature review. *Natural Language Processing Journal*. 2025. Т. 10. Стаття 100129. DOI: 10.1016/j.nlp.2025.100129.
5. Li B., та ін. AID-Agent: an LLM-Agent for advanced extraction and integration of documents. *Proceedings of the 1st Workshop for Research on Agent Language Models (REALM 2025)*. 2025. С. 80–88. DOI: 10.18653/v1/2025.realm-1.6.
6. Li H., та ін. Extracting financial data from unstructured sources: leveraging large language models. *SSRN Electronic Journal*. 2023. DOI: 10.2139/ssrn.4567607. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4567607 (Дата звернення: 15.10.2025).
7. Almeida F. C., Caminha C. Evaluation of entry-level open-source large language models for information extraction from digitized documents. *Symposium on Knowledge Discovery, Mining and Learning (KDMiLe)*. 2024. С. 25–32. DOI: 10.5753/kdmile.2024.243859. (Офіц. стор.: sol.sbc.org.br).
8. Tito R., Karatzas D., Valveny E. Hierarchical multimodal transformers for multi-page DocVQA. *arXiv:2212.05935*. 2023. DOI: 10.48550/arXiv.2212.05935. URL: <https://arxiv.org/abs/2212.05935> (Дата звернення: 15.10.2025).
9. Ranaweera U., Mawitagama B., Liyanage S., Keshan S., De Silva T., Hewawalpita S. Comparison of machine learning models to classify documents on digital development. у кн.: *Data Science and Artificial Intelligence* / ред. С. Anutariya, М. М. Bonsangue. Singapore: Springer, 2023. (CCIS, т. 1942). С. 59–73. DOI: 10.1007/978-981-99-7969-1_5.
10. Le D. X., Thoma G. R. Page layout classification technique for biomedical documents. *Proceedings of the World Multiconference on Systems, Cybernetics and Informatics (SCI)*. 2000. Т. X. С. 348–352. URL: <https://lhncbc.nlm.nih.gov/LHC-publications/PDF/pub2000015.pdf> (Дата звернення: 15.10.2025).
11. Petrov K., Chalyi T. Situational model of a medical business process. *Bulletin of National Technical University "KhPI"*. Series: System Analysis, Control and Information Technologies. 2024. № 2(12). С. 42–45. DOI: 10.20998/2079-0023.2024.02.07.
12. Shin C., Doermann D., Rosenfeld A. Classification of document pages using structure-based features. *International Journal on Document Analysis and Recognition (IJDAR)*. 2001. Т. 3, № 4. С. 232–247. DOI: 10.1007/PL00013566.

References:

1. Bozhko, O.Yu. (2024). Vykorystannia velykykh movnykh modelei dlia rozpiznavannia informatsii v numizmatychnykh opysakh [Using large language models for information recognition in numismatic descriptions]. *Nauka i tekhnika sohodni*, 1(29), 615–625. [https://doi.org/10.52058/2786-6025-2024-1\(29\)-615-625](https://doi.org/10.52058/2786-6025-2024-1(29)-615-625) [in Ukrainian].
2. Bozhko, O.Yu. (2025). Rozrobka iteratyvnoho metodu ekstraktsii danykh z nestrukturovanykh dokumentiv na osnovi vykorystannia velykykh movnykh modelei [Development of an iterative method for data extraction from unstructured documents based on the use of large language models]. *Visnyk Kremenchutskoho natsionalnoho universytetu imeni Mykhaila Ostrohradskoho*. 1, 119–124. <https://doi.org/10.32782/1995-0519.2025.1.15> [in Ukrainian].
3. Xu, Y., Li, M., Cui, L., Huang, S., Wei, F., & Zhou, M. (2020). LayoutLM: Pre-training of text and layout for document image understanding. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 1192–1200). <https://doi.org/10.1145/3394486.3403172>
4. Rijcken, E., Zervanou, K., Mosteiro, P., Scheepers, F., Spruit, M., & Kaymak, U. (2025). Machine learning vs. rule-based methods for document classification of electronic health records within mental health care: A systematic literature review. *Natural Language Processing Journal*, 10, Article 100129. <https://doi.org/10.1016/j.nlp.2025.100129>
5. Li, B., et al. (2025). AID-Agent: An LLM-agent for advanced extraction and integration of documents. In *Proceedings of the 1st Workshop for Research on Agent Language Models (REALM 2025)* (pp. 80–88). <https://doi.org/10.18653/v1/2025.realm-1.6>
6. Li, H., et al. (2023). Extracting financial data from unstructured sources: Leveraging large language models. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4567607>. Retrieved from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4567607
7. Almeida, F. C., & Caminha, C. (2024). Evaluation of entry-level open-source large language models for information extraction from digitized documents. In *Symposium on Knowledge Discovery, Mining and Learning (KDMiLe 2024)* (pp. 25–32). <https://doi.org/10.5753/kdmile.2024.243859>
8. Tito, R., Karatzas, D., & Valveny, E. (2023). Hierarchical multimodal transformers for multi-page DocVQA. arXiv:2212.05935. <https://doi.org/10.48550/arXiv.2212.05935>. Retrieved from: <https://arxiv.org/abs/2212.05935>
9. Ranaweera, U., Mawitagama, B., Liyanage, S., Keshan, S., De Silva, T., & Hewawalpita, S. (2023). Comparison of machine learning models to classify documents on digital development. In C. Anutariya & M. M. Bonsangue (Eds.), *Data Science and Artificial Intelligence (CCIS, Vol. 1942, pp. 59–73)*. Springer. https://doi.org/10.1007/978-981-99-7969-1_5
10. Le, D. X., & Thoma, G. R. (2000). Page layout classification technique for biomedical documents. In *Proceedings of the World Multiconference on Systems, Cybernetics and Informatics (SCI) (Vol. X, pp. 348–352)*. Retrieved from: <https://lhncbc.nlm.nih.gov/LHC-publications/PDF/pub2000015.pdf>
11. Petrov, K., & Chalyi, T. (2024). Situational model of a medical business process. *Bulletin of National Technical University “KhPI”*. Series: System Analysis, Control and Information Technologies, 2(12), 42–45. <https://doi.org/10.20998/2079-0023.2024.02.07>
12. Shin, C., Doermann, D., & Rosenfeld, A. (2001). Classification of document pages using structure-based features. *International Journal on Document Analysis and Recognition (IJ DAR)*, 3(4), 232–247. <https://doi.org/10.1007/PL00013566>

Дата надходження статті: 17.10.202

Дата прийняття статті: 10.11.2025

Опубліковано: 30.12.2025

Булгакова О. Ф., старший викладач кафедри комп'ютерних наук та інженерії програмного забезпечення
Університету митної справи та фінансів
ORCID: 0000-0001-9834-2970

Костенко В. В., старший викладач кафедри комп'ютерних наук та інженерії програмного забезпечення
Університету митної справи та фінансів
ORCID: 0000-0003-3847-2110

Булгаков Д. С., студент спеціальності 122 «Комп'ютерні науки»
Університету митної справи та фінансів
ORCID: 0009-0007-8667-3693

АВТОМАТИЗАЦІЯ РОЗГОРТАННЯ СЕРВЕРЛЕС-ЗАСТОСУНКІВ ІЗ ВИКОРИСТАННЯМ ІАС: КЕЙС СИСТЕМИ «AGRO MONITOR»

У статті висвітлено результати дослідження та практичної реалізації автоматизації розгортання безсерверних (serverless) застосунків на основі концепції «інфраструктура як код» (Infrastructure as Code, IaC). Основна увага приділяється аналізу сучасних інструментів і технологій, що забезпечують автоматизоване управління інфраструктурою у хмарному середовищі, а також впровадженню serverless-застосунків із використанням Іас у реальному проєкті.

У першій частині статті розглянуто теоретичні основи безсерверної архітектури, яка дозволяє компаніям знижувати витрати на обслуговування інфраструктури, забезпечувати автоматичне масштабування та оплачувати лише фактично використані ресурси. Проаналізовано переваги та виклики застосування serverless-підходу, включно з проблемами взаємодії між компонентами, контролем версій, складнощами інтеграції та питаннями безпеки. Значну увагу приділено платформам функцій як сервісу (FaaS), зокрема AWS Lambda, Azure Functions, Google Cloud Functions та OpenFaaS.

Детально досліджено роль інфраструктури як коду у контексті автоматизації розгортання безсерверних рішень. Розглянуто два основних підходи до опису інфраструктури – декларативний (на прикладі Terraform та AWS CloudFormation) та імперативний Іас (Pulumi) і засоби керування конфігураціями (Ansible), охарактеризовано їхні особливості, переваги та обмеження. Визначено ключові переваги використання Іас для забезпечення стандартизації, відтворюваності, інтеграції з CI/CD-процесами, контролю версій та підвищення безпеки.

Практична частина статті присвячена автоматизації розгортання системи «Agro Monitor», що реалізована на основі безсерверної архітектури у хмарному середовищі Microsoft Azure. Застосовано інструмент Pulumi, який дозволяє описувати інфраструктуру мовами програмування та забезпечує гнучкість при створенні складних інфраструктурних сценаріїв. Описано архітектуру системи, основні компоненти, механізми взаємодії через API, а також процеси автоматичного масштабування, обробки даних та генерації звітів. Показано, що запропонований підхід забезпечує швидке розгортання нових середовищ, зниження витрат на підтримку, підвищення стабільності та розширюваність рішення.

Отримані результати показують ефективність поєднання безсерверної архітектури та інфраструктури як коду для створення сучасних інформаційних систем, що відзначаються високою гнучкістю, стабільністю та економічною доцільністю. Розроблене рішення може бути корисним для підприємств різних галузей, які прагнуть оптимізувати інфраструктуру та підвищити швидкість впровадження інноваційних ІТ-продуктів.

Ключові слова: безсерверна архітектура, інфраструктура як код (IaC), Pulumi, автоматизація розгортання, хмарні сервіси, Azure, Terraform, CI/CD, DevOps, супутникові знімки, інформаційні технології.

Bulhakova O. F., Kostenko V. V., Bulhakov D. S. Automation of serverless application deployment using IaC: case of 'Agro Monitor'

The article presents the results of research and practical implementation of serverless application deployment automation based on the Infrastructure as Code (IaC) concept. The primary focus is placed on analyzing modern tools and technologies that enable automated infrastructure management in cloud environments, as well as the deployment of serverless applications using IaC in a real-world project.

The first part of the article examines the theoretical foundations of serverless architecture, which allows companies to reduce infrastructure maintenance costs, ensure automatic scaling, and pay only for the actual use of computing resources. The

advantages and challenges of adopting the serverless approach are analyzed, including component interaction issues, version control, integration complexity, and security concerns. Particular attention is given to Function-as-a-Service (FaaS) platforms such as AWS Lambda, Azure Functions, Google Cloud Functions, and OpenFaaS.

The article thoroughly investigates the role of Infrastructure as Code in the context of automating the deployment of serverless solutions. Two primary approaches to infrastructure description are considered—declarative (Terraform, AWS CloudFormation) and imperative (Pulumi), and configuration-management tools (Ansible) – with an overview of their characteristics, benefits, and limitations. The key advantages of IaC for ensuring standardization, infrastructure reproducibility, CI/CD process integration, version control, and enhanced system security are highlighted.

The practical section of the article focuses on the deployment automation of the «Agro Monitor» system, which is built on serverless architecture in the Microsoft Azure cloud environment. The Pulumi tool, which enables infrastructure definition using programming languages and provides flexibility in creating complex infrastructure scenarios, is applied. The system architecture, key components, API-based interaction mechanisms, as well as processes for automatic scaling, data processing, and report generation are described in detail. It is demonstrated that the proposed approach enables rapid deployment of new environments, reduces maintenance costs, enhances system stability, and improves scalability.

The results confirm the effectiveness of combining serverless architecture and Infrastructure as Code for developing modern information systems characterized by high flexibility, stability, and economic feasibility. The developed solution can be beneficial for enterprises in various industries seeking to optimize infrastructure and accelerate the deployment of innovative IT products.

Key words: serverless architecture, Infrastructure as Code (IaC), Pulumi, deployment automation, cloud services, Azure, Terraform, CI/CD, DevOps, satellite imagery, information technologies.

Постановка проблеми. Активний розвиток інформаційних технологій та посилення конкуренції в ІТ-сфері зумовлюють необхідність пошуку ефективних підходів до проєктування, розгортання та обслуговування програмного забезпечення. Сучасні компанії стикаються з потребою у прискоренні розробки та впровадження цифрових рішень, зниженні витрат на інфраструктуру, підвищенні надійності та забезпеченні масштабованості інформаційних систем.

Одним із ключових технологічних трендів є поширення хмарних технологій, які відкривають нові можливості для динамічного налаштування інфраструктури під конкретні бізнес-вимоги. У цьому контексті безсерверна архітектура (serverless architecture) привертає значну увагу завдяки здатності автоматично масштабувати ресурси, мінімізувати операційні витрати та забезпечувати оплату лише за фактичне використання обчислювальних потужностей. Проте впровадження serverless-підходу супроводжується низкою викликів, серед яких – складність інтеграції окремих компонентів, контроль життєвого циклу ресурсів, управління версіями та дотримання вимог безпеки.

У відповідь на ці виклики в ІТ-індустрії активно впроваджуються підходи до автоматизації управління інфраструктурою, зокрема концепція «інфраструктура як код» (Infrastructure as Code, IaC). Вона забезпечує можливість опису інфраструктури за допомогою коду, що сприяє стандартизації налаштувань, автоматизації процесів розгортання, повторюваності інфраструктурних рішень, інтеграції з CI/CD-процесами та підвищенню безпеки систем.

Інфраструктура як код (IaC) стала ключовою практикою відтворюваного, керованого та контролюваного розгортання хмарних систем. Поєднання IaC із безсерверними технологіями (FaaS/Backend-as-a-Service) дає змогу швидко масштабувати прикладні рішення та знижувати операційні витрати. Водночас емпіричні дослідження демонструють, що безсерверні платформи характеризуються різко нерівномірними профілями навантаження, значущою часткою «холодних стартів» і варіабельністю затримок, які безпосередньо впливають на продуктивність і вартість рішень [1-4]. Для керування якістю таких систем критичними стають метричний контроль процесів деплою, метрики якості IaC, автоматизоване тестування IaC-програм і спостережуваність (трасування викликів, моніторинг тригерів) [4, 7, 8, 12]. Паралельно з цим спільнота IaC пропонує таксономії тем і методики вимірювання якості інфраструктурного коду та автоматизованого тестування [6–8]. Окремим викликом є керування безпекою та дрейфом/дефектами узгодження станів, що підривають відтворюваність і відповідність політикам; сучасні роботи фіксують типові «запахи» безпеки, класифікують дефекти state reconciliation і демонструють ефективність GitOps-підходів у керованих середовищах [9-11].

На сьогоднішній день постає нагальна потреба у ґрунтовному аналізі існуючих інструментів і технологій для автоматизації розгортання безсерверних застосунків із використанням IaC, а також у апробації ефективності їх практичного застосування для побудови стабільних, масштабованих та економічно вигідних інформаційних систем.

Ця проблема є актуальною як для малих і середніх підприємств, які прагнуть мінімізувати витрати на інфраструктуру, так і для великих організацій, орієнтованих на високонавантажені сервіси з вимогами до надійності та швидкого масштабування.

Аналіз останніх досліджень та публікацій. У наукових і технічних колах все більше уваги приділяється дослідженню імперативних та декларативних підходів до IaC, а також вибору оптимальних інструментів

для їх реалізації в умовах serverless-архітектур. Особливий інтерес викликає платформа Pulumi, яка поєднує переваги класичних IaC-рішень із можливістю використання популярних мов програмування, що відкриває нові перспективи для інтеграції DevOps-практик.

Попри зростаючу популярність безсерверної архітектури у сучасних ІТ-системах, питання автоматизації її розгортання за допомогою концепції «інфраструктура як код» (IaC) дотепер залишається недостатньо висвітленим у науковій літературі.

Класичні праці з емпіричного аналізу serverless-платформ (USENIX ATC'18; ATC'20) заклали основу розуміння внутрішніх механізмів та реальних профілів робочих навантажень у хмарах [1, 2]. Подальші дослідження виявили варіабельність продуктивності в часі (зокрема діурнальні патерни) та підвищену частоту cold-start у певні періоди [3], а також показали вагомий внесок тригерів у хвості затримок; бенчмарк TriggerBench систематизує порівняння різних тригерів і провайдерів [4]. Систематичний огляд TOSEM (2023) консолідує 160+ досліджень серверлеса і виділяє відкриті проблеми продуктивності, переносимості, тестування та спостережуваності [5].

У галузі ІаС мепінг-дослідження IST (2019) сформувало таксономію напрямів і підсилило фокус на дефектах та ризиках безпеки [6]. Новіші роботи пропонують метрики якості для Terraform-артефактів (TerraMetrics, ICPC'24) [7] і підходи до автоматизованого тестування ІаС-програм на великих корпусах відкритого коду (IEEE TSE, 2024) [8]. З погляду безпеки «семеро гріхів» ІаС-скриптів (ICSE'19) описують репертуар типових smells і детектування [9]. Проблематику дрейфу конфігурації та дефектів узгодження станів уточнено у FSE'24 (таксономія восьми категорій) [10], а в керованих середовищах (Kubernetes) ефективними демонструються GitOps-практики протидії дрейфу (IEEE UCC'24) [11]. Нарешті, огляди засобів трасування (JSS, 2023) [12] окреслюють можливості/обмеження observability-стеку, що прямо пов'язано з побудовою SLI/SLO та policy-as-code у конвеєрах перевірки.

Управління конфігураційним дрейфом і дефектами узгодження станів стає центральною темою сучасних робіт: у межах FSE'24 запропоновано таксономію восьми категорій дефектів state reconciliation та показано, що частина з них унікальна для ІаС-систем [10]; в середовищах Kubernetes переважають GitOps-практики, що демонструють ефективність у виявленні/усуненні дрейфу (IEEE UCC'24) [11]. Нарешті, з боку спостережуваності актуальними є дослідження трасувальних інструментів та їх накладних витрат (JSS'23; сучасні роботи про overhead OpenTelemetry/Elastic) [12], що прямо корелює з потребою кількісно оцінювати SLI/SLO та інтегрувати policy-as-code у конвеєри перевірки.

Таким чином, наявна література підтверджує: (1) serverless-навантаження відзначається високою нерівномірністю викликів і тригерними латентностями; (2) зрілість ІаС залежить від наявності метрик якості, автоматизованого тестування та практик керування дрейфом; (3) для реплікованої наукової оцінки необхідні відтворювані експерименти з прозорими трейсами/даними. На цьому тлі запропоновані в роботі практики (ІаС-тестування, policy-as-code, SLO-орієнтована спостережуваність, порівняльні метрики деплою/вартості/латентностей) відповідають виявленим прогалинам і доповнюють сучасний стан знань.

Метою статті є обґрунтування доцільності та практичної ефективності застосування концепції «інфраструктура як код» (ІаС) для автоматизації процесів розгортання безсерверних застосунків у хмарних середовищах, а також аналіз переваг, обмежень та можливостей використання сучасних інструментів ІаС в умовах serverless-архітектур.

У межах дослідження акцентовано увагу на:

- критичному аналізу сучасних підходів до автоматизації розгортання безсерверних систем;
- виявленні існуючих обмежень і проблем, пов'язаних з інтеграцією ІаС у DevOps-процеси для serverless-рішень;
- практичній апробації інструментів ІаС (зокрема Pulumi) для забезпечення масштабованості, повторюваності та стабільності інфраструктури на прикладі побудови системи аграрного моніторингу у хмарному середовищі Microsoft Azure;
- оцінці доцільності використання імперативних інструментів ІаС у порівнянні з традиційними декларативними підходами.

Особливу увагу приділено питанням інтеграції таких рішень із CI/CD-процесами, забезпечення безпеки інфраструктури та зниження витрат на розгортання й супровід хмарних застосунків.

Матеріали та методи. Об'єктом дослідження є процес управління інфраструктурою хмарних інформаційних систем, реалізованих на основі безсерверної архітектури (serverless). Предметом дослідження виступає методика автоматизації розгортання таких систем з використанням сучасних інструментів інфраструктури як коду (ІаС).

Методологічна основа дослідження ґрунтується на системному підході до проектування та розгортання інформаційних систем у хмарному середовищі, зокрема на принципах DevOps, CI/CD (Continuous Integration/Continuous Delivery), GitOps та модульного розділення інфраструктури на незалежні керовані компоненти. У межах дослідження було використано як декларативні, так і імперативні підходи до опису інфраструктури, що дозволило порівняти їх ефективність у різних сценаріях.

Виклад основного матеріалу. Для досягнення цілей дослідження було розроблено архітектуру реального інформаційного застосунку – системи агромоніторингу Agro Monitor, яка використовує подієво-орієнтовану serverless-модель і реалізована у хмарному середовищі Microsoft Azure. Система включає в себе багаторівневу інфраструктуру, що складається з функціональних блоків обробки супутникових зображень, аналітики рослинності (на основі NDVI), зберігання даних, API-шлюзів і механізмів асинхронної взаємодії компонентів.

В якості основного інструмента інфраструктури як коду обрано Pulumi – сучасну IaC-платформу, яка дозволяє описувати конфігурацію хмарних ресурсів із використанням повноцінних мов програмування (у даному випадку – TypeScript). Це забезпечило високу гнучкість опису логіки створення, модифікації та знищення інфраструктурних ресурсів, а також полегшило інтеграцію з зовнішніми бібліотеками, обробку умовних сценаріїв та винятків.

Інтеграція процесів автоматизації розгортання була реалізована через зв'язку Pulumi з GitHub Actions, що дозволило забезпечити повний CI/CD цикл: запуск автоматичних перевірок, конфігурування середовищ, контроль версій змін в інфраструктурі та моніторинг стану хмарних ресурсів. Окрім цього, для конфігурації параметрів доступу та безпеки застосовувалися засоби управління секретами (Azure Key Vault).

Під час реалізації було застосовано наступні технічні інструменти:

- Microsoft Azure – хмарна платформа, що забезпечує обчислювальні, сховищні та подієві сервіси.
- Pulumi (TypeScript) – імперативний IaC-фреймворк для мультихмарних середовищ.
- Azure Functions – серверлес-сервіси для виконання обчислень.
- Azure Blob Storage та Cosmos DB – для зберігання супутникових знімків і структурованих даних.
- Event Grid та Logic Apps – для оркестрації подій.
- GitHub Actions – для CI/CD-автоматизації.
- Visual Studio Code та Node.js – середовище розробки та виконання інфраструктурного коду.

Для моделювання архітектури застосовувались стандартні підходи до багатоподієвого поділу системи (data layer, logic layer, presentation layer), що забезпечило масштабованість та незалежність компонентів. Весь процес розгортання був відтворюваним, з можливістю швидкого оновлення середовищ або повного знищення інфраструктури без залишків.

Таким чином, у межах дослідження було реалізовано повний цикл створення, розгортання та управління безсерверною інформаційною системою із використанням імперативного підходу IaC, інтегрованого у DevOps-процеси.

У процесі розробки та автоматизації розгортання системи «Agro Monitor» були використані сучасні технічні інструменти, які забезпечили повний життєвий цикл serverless-застосунку – від проєктування архітектури до реалізації CI/CD-процесів.

Хмарна платформа Microsoft Azure слугувала базовим середовищем для розміщення всіх компонентів системи. Було задіяно низку її сервісів, зокрема:

- Azure Functions для обробки подій та реалізації серверлес-обчислень;
- Azure Blob Storage для зберігання великих обсягів супутникових зображень;
- Cosmos DB як високопродуктивну NoSQL-базу для структурованих аналітичних даних;
- Event Grid для маршрутизації подій між компонентами;
- Azure Key Vault для зберігання секретів, токенів доступу та конфіденційної інформації.

Azure забезпечив надійну, масштабовану та безпечну інфраструктурну основу для системи з гнучким керуванням доступом і можливістю централізованого моніторингу.

IaC-фреймворк Pulumi був використаний як основний інструмент автоматизації. У рамках проєкту інфраструктура застосунку (включаючи Azure Functions, Blob Storage, Cosmos DB, App Services тощо) описувалась у вигляді TypeScript-коду.

Завдяки цьому:

- було досягнуто високої гнучкості при конфігурації інфраструктури (використання циклів, умов, функцій);
- код інфраструктури став частиною загального репозиторію з логікою застосунку;
- забезпечено повну відтворюваність розгортання через використання state-файлів Pulumi та логування всіх змін.
- Функції Azure стали основним обчислювальним середовищем для реалізації серверлес-логіки.

Вони:

- реагували на події з Event Grid (наприклад, надходження нового знімка до Blob Storage);
- запускали обробку NDVI-зон, створення аналітичних карт та генерацію звітів;
- масштабувалися автоматично відповідно до навантаження.

Завдяки серверлес-підходу, вдалося мінімізувати витрати на постійне утримання інфраструктури та забезпечити високу відмовостійкість.

Blob Storage використовувався для зберігання великих файлів супутникових зображень (у форматах GeoTIFF, JPEG, PNG).

Cosmos DB зберігав метадані знімків, інформацію про ділянки, оброблені NDVI-значення, агрономічні параметри.

Ці сервіси були інтегровані з Azure Functions через SDK, що дозволило реалізувати подієво-орієнтований підхід.

Event Grid слугував маршрутизатором подій у системі: завантаження нового зображення, оновлення даних, запити від користувача тощо. У разі складніших сценаріїв обробки подій було задіяно Logic Apps, які дозволяли безкодовим способом визначати реакції на певні тригери (наприклад, зберегти NDVI-звіт у PDF і надіслати його електронною поштою).

Для забезпечення повного CI/CD-процесу було створено робочі процеси (workflows) у GitHub Actions, які:

- автоматично запускали перевірку змін у коді;
- виконували linting і тестування;
- запускали Pulumi-скрипти для оновлення інфраструктури;
- реєстрували нові функції Azure Functions та оновлювали конфігурації.

Це дозволило досягти швидкого, контрольованого та безпечного оновлення системи при кожному коміті.

Як середовище розробки використовувалася Visual Studio Code з плагінами для Pulumi, TypeScript і Azure CLI. Node.js був платформою для виконання Pulumi-скриптів, а також середовищем для тестування логіки обробки функцій до їхнього деплою в Azure.

Ці інструменти діяли в єдиній екосистемі, забезпечуючи автоматизоване, масштабоване й контрольоване розгортання інфраструктури та застосунку, що повністю відповідало концепціям DevOps та хмарної інженерії.

Реалізація. У результаті дослідження було розроблено і впроваджено інформаційну систему «Agro Monitor», яка призначена для автоматизованої обробки супутникових знімків сільськогосподарських угідь з метою виявлення зон стресу рослинності та підтримки прийняття рішень агрономами. Система реалізована у хмарному середовищі Microsoft Azure із використанням безсерверної архітектури, що забезпечує масштабованість, високу доступність та ефективне використання ресурсів.

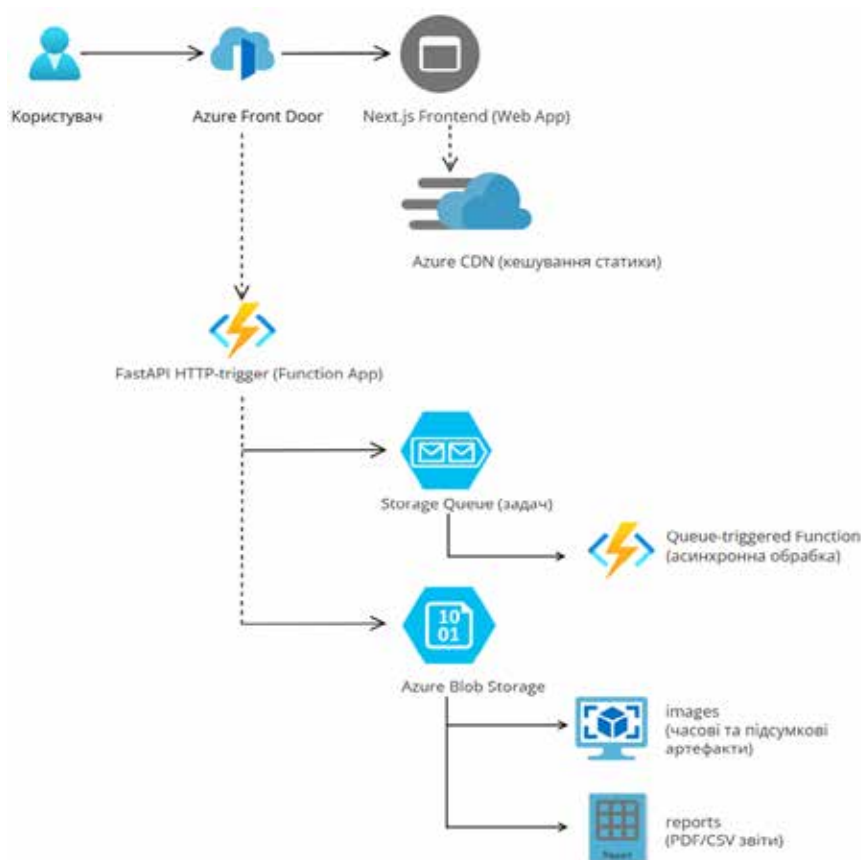


Рис. 1. Схема взаємодії основних компонентів системи «Agro Monitor»

Архітектура системи побудована на принципах багаторівневої моделі, кожен з рівнів якої виконує окрему функцію:

- Рівень даних відповідає за зберігання та обробку великих масивів супутникових знімків, використовуючи Azure Blob Storage для неструктурованих даних (зображень) та Cosmos DB – для зберігання структурованої аналітичної інформації (NDVI-значення, координати ділянок, історичні дані).
- Обчислювальний рівень реалізовано за допомогою Azure Functions, які обробляють події, що надходять через Event Grid. Тут відбувається обчислення NDVI, класифікація рослинного покриву, генерація карт та звітів.
- Рівень оркестрації та інтеграції забезпечується зв'язкою Event Grid та Logic Apps, які автоматично запускають відповідні обчислення при завантаженні нових знімків чи зміні параметрів користувача.
- Інфраструктурний рівень описаний у вигляді коду за допомогою Pulumi (TypeScript), що дозволяє здійснювати автоматичне створення, оновлення та знищення ресурсів у хмарному середовищі. Інфраструктура є відтворюваною, контрольованою та масштабованою.
- Рівень автоматизації реалізований за допомогою GitHub Actions, що забезпечує безперервну інтеграцію та доставку змін (CI/CD): кожна зміна у коді інфраструктури чи логіки викликає відповідний процес перевірки, тестування та оновлення середовища.



Рис. 2. Файлова структура репозиторію проєкту «Agro Monitor»

Візуально структура системи відображена на рисунку 1, що демонструє взаємодію компонентів і подієво-орієнтовану модель. Архітектура побудована таким чином, щоб підтримувати принцип незалежності сервісів і масштабування за запитом.

Організація репозиторію проєкту представлена на рисунку 2 і демонструє логічну ізоляцію інфраструктурного коду, логіки обробки зображень та CI/CD-налаштувань. Такий підхід забезпечує зручність супроводу та розширення системи.

На рисунку 3 показано приклад сформованого звіту на основі аналізу NDVI: система автоматично генерує візуалізацію зон стресу рослин з поясненнями, що можуть бути використані у польових умовах агрономами. Рисунок 4 ілюструє приклади супутникових зображень на різних етапах обробки – від сирого знімка до NDVI-мапи та зони вегетаційного стресу.

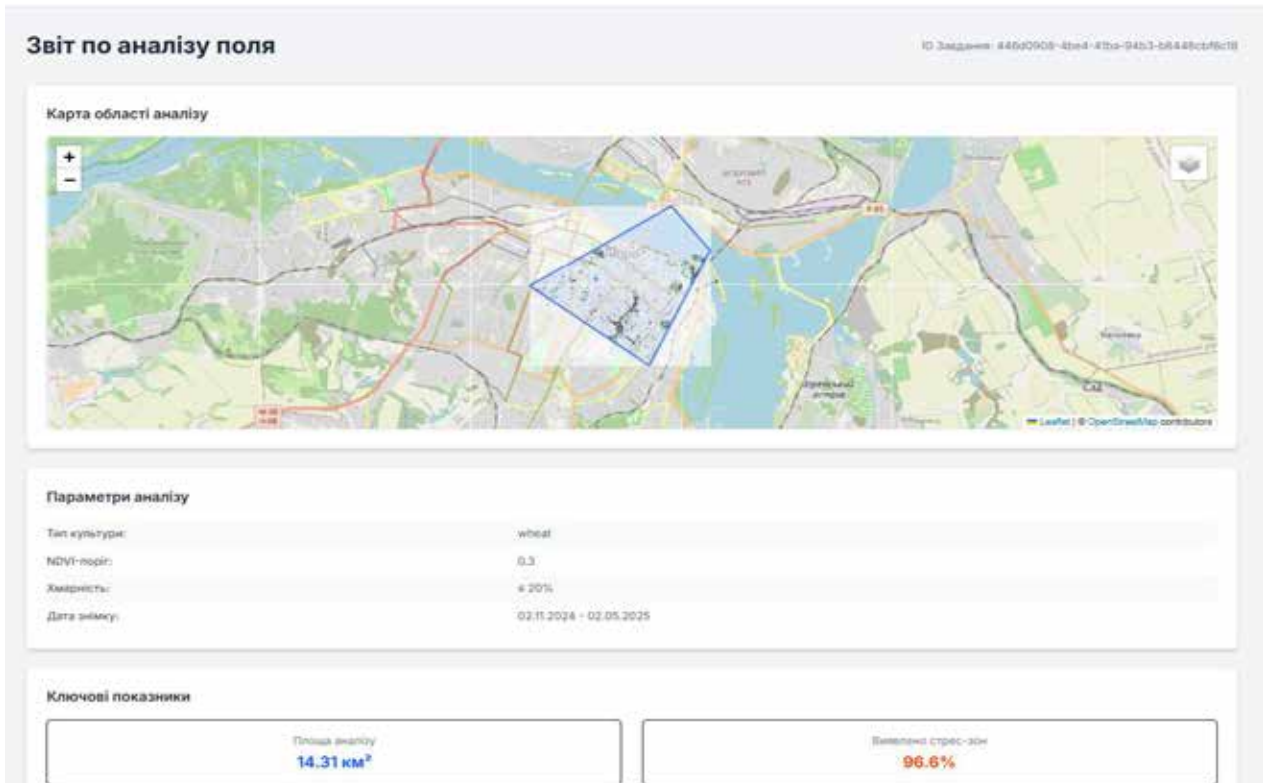


Рис. 3. Вигляд результатів роботи системи «Agro Monitor»

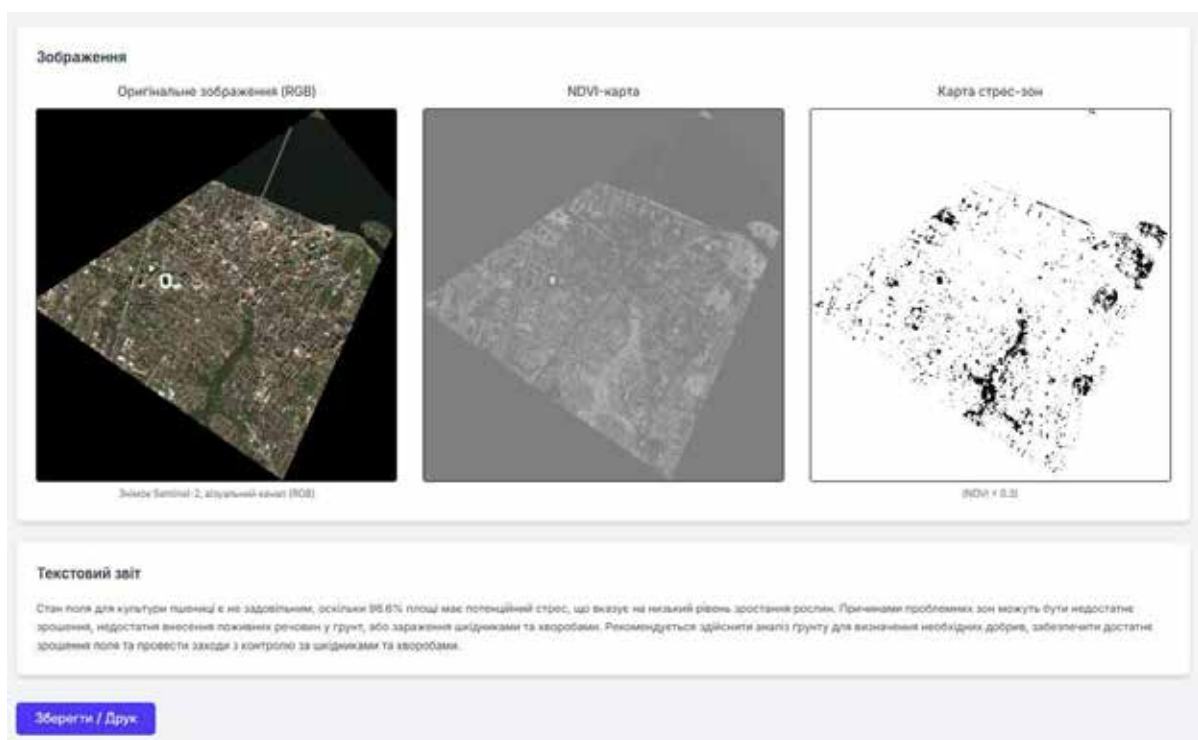


Рис. 4. Вигляд супутникових знімків, опрацьованих за допомогою системи «Agro Monitor»

Таким чином, система «Agro Monitor» не лише демонструє ефективність автоматизованого підходу до обробки агроданих, але й підтверджує практичну реалізованість принципів IaC, DevOps та serverless у реальних галузевих рішеннях.

Запропоноване рішення істотно відрізняється від класичних підходів до розгортання інформаційних систем, зокрема у сфері агромоніторингу. Насамперед, варто відзначити переваги автоматизованого підходу із застосуванням інфраструктури як коду (IaC) у порівнянні з традиційними методами ручного налаштування серверів і сервісів. Реалізація системи «Agro Monitor» за допомогою Pulumi продемонструвала значне скорочення часу, необхідного для розгортання інфраструктури, а також забезпечила повну відтворюваність конфігурацій, що є критично важливим при масштабуванні або міграції рішень.

З точки зору практичної значущості, реалізоване рішення має високий рівень адаптованості до інших хмарних середовищ. Хоча проєкт було реалізовано в Microsoft Azure, архітектурна модель і обрані інструменти є мультихмарними за своєю природою. Зокрема, використання Pulumi як імперативного засобу IaC дозволяє за потреби перенести систему до AWS, Google Cloud чи інших провайдерів з мінімальними модифікаціями, що робить рішення універсальним та масштабованим.

Разом з тим, проєкт має певні обмеження. Основним з них є залежність від конкретної хмарної платформи (у даному випадку – Azure), що обумовлює необхідність враховувати специфіку її сервісів, моделей безпеки та тарифікації. Крім того, використання Pulumi потребує певного рівня знань із програмування, що створює додаткову навчальну криву для фахівців, які звикли до декларативних підходів (наприклад, Terraform або CloudFormation).

Особливі умови воєнного стану в Україні також вплинули на реалізацію дослідження. Було зафіксовано періоди ускладненого доступу до інтернету та обмеженого фізичного доступу до локальних ресурсів, що вимагало адаптації командної роботи до дистанційного формату. У таких умовах важливу роль відіграв централізований доступ до хмарної інфраструктури та можливість її керування з будь-якої геолокації.

У межах перспектив подальших досліджень розроблену архітектуру планується адаптувати до інших галузей, де потрібна обробка великих обсягів даних у реальному часі – зокрема в екологічному моніторингу, урбаністиці, безпеці. Також передбачається порівняльне тестування цього підходу у середовищах інших хмарних провайдерів, що дозволить визначити оптимальні конфігурації та виявити вузькі місця в кросплатформному розгортанні.

Висновки та перспективи подальших досліджень. У межах дослідження було розроблено і впроваджено методику автоматизації розгортання інформаційних систем із безсерверною архітектурою на прикладі системи агромоніторингу «Agro Monitor». Основною технологічною основою рішення виступив імперативний підхід до інфраструктури як коду з використанням платформи Pulumi у середовищі Microsoft Azure.

Суттєвими результатами є:

- реалізація повністю автоматизованого процесу створення, конфігурації та оновлення компонентів системи;
- зменшення часу розгортання інфраструктури до кількох хвилин;
- забезпечення повторюваності, масштабованості та зручності супроводу системи;
- глибока інтеграція інфраструктурного коду в CI/CD-процеси за допомогою GitHub Actions.

Отримані результати пояснюються ефективністю поєднання serverless-архітектури, імперативної IaC-моделі та подієво-орієнтованого підходу до обробки даних. Застосування Microsoft Azure дозволило реалізувати високопродуктивну систему, адаптовану до аграрних задач реального часу.

Практична користь полягає у створенні рішення, що може бути масштабовано на інші хмарні платформи та галузі, а також використане для підвищення ефективності обробки даних в умовах динамічного навантаження.

У подальших дослідженнях доцільно:

- розширити застосування методики на інші типи застосунків (екологія, безпека, міська аналітика);
- протестувати систему в мультихмарному середовищі;
- провести кількісне порівняння із декларативними підходами IaC.

Таким чином, запропоноване рішення підтверджує ефективність і практичну доцільність використання імперативної IaC-платформи Pulumi для побудови сучасних serverless-систем.

Список використаних джерел:

1. Wang L.; Li M.; Zhang Y.; Ristenpart T.; Swift M. Peeking Behind the Curtains of Serverless Platforms. Proceedings of the 2018 USENIX Annual Technical Conference (USENIX ATC '18). Boston, MA, USA, 11–13 July 2018. С. 133–145.
2. Shahradd M., Fonseca R., Zhang I., et al. Serverless in the Wild: Characterizing and Optimizing the Serverless Workload at a Large Cloud Provider. Proceedings of the 2020 USENIX Annual Technical Conference (USENIX ATC '20). 2020. С. 205–218.
3. Schirmer T., Japke N., Greten S., Pfandzelter T., Bermbach D. The Night Shift: Understanding Performance Variability of Cloud Serverless Platforms. The 1st Workshop on SErverless Systems, Applications and MEthodologies (SESAME '23), Rome, Italy, 8 May 2023. Стаття 7. DOI: 10.1145/3592533.3592808.
4. Scheuner J., Bertilsson M., Grönqvist O., et al. TriggerBench: A Performance Benchmark for Serverless Function Triggers. 2022 IEEE International Conference on Cloud Engineering (IC2E), Monterey, CA, USA, 28 Sept. 2022. DOI: 10.1109/IC2E55432.2022.00018.

-
5. Wen J., Chen Z., Jin X., Liu X. Rise of the Planet of Serverless Computing: A Systematic Review. *ACM Transactions on Software Engineering and Methodology*. 2023. T. 32. № 5. Стаття 131. DOI: 10.1145/3579643.
 6. Rahman A., Mahdavi-Hezaveh R., Williams L. A Systematic Mapping Study of Infrastructure as Code Research. *Information and Software Technology*. 2019. T. 108. C. 65–77. DOI: 10.1016/j.infsof.2018.12.004.
 7. Begoug M., Chouchen M., Ouni A. TerraMetrics: An Open Source Tool for Infrastructure-as-Code (IaC) Quality Metrics in Terraform. Proceedings of the 2024 IEEE/ACM 32nd International Conference on Program Comprehension (ICPC) – Tool Demonstrations. 2024. DOI: 10.1145/3643916.3644439.
 8. Sokolowski D., Spielmann D., Salvaneschi G. Automated Infrastructure as Code Program Testing. *IEEE Transactions on Software Engineering*. 2024. T. 50. № 6. C. 1585–1599. DOI: 10.1109/TSE.2024.3393070.
 9. Rahman A., Parnin C., Williams L. The Seven Sins: Security Smells in Infrastructure as Code Scripts. Proceedings of the 41st International Conference on Software Engineering (ICSE '19). 2019. C. 164–175. DOI: 10.1109/ICSE.2019.00033.
 10. Hassan M. M., Salvador J., Santu S. K. K., Rahman A. State Reconciliation Defects in Infrastructure as Code. Proceedings of the ACM on Software Engineering. 2024. T. 1 (FSE). Стаття 83. DOI: 10.1145/3660790.
 11. Shrestha R., Ali A. A. N. Configuration Management in Kubernetes Environments: A GitOps Approach. 2024 IEEE/ACM 17th International Conference on Utility and Cloud Computing (UCC 2024), Sharjah, UAE, 16–19 Dec. 2024. C. 497–502. DOI: 10.1109/UCC63386.2024.00077.
 12. Janes A., Li X., Lenarduzzi V. Open Tracing Tools: Overview and Critical Comparison. *Journal of Systems and Software*. 2023. T. 204. Стаття 111793. DOI: 10.1016/j.jss.2023.111793.

References:

1. Wang, L., Li, M., Zhang, Y., Ristenpart, T., & Swift, M. (2018). Peeking behind the curtains of serverless platforms. In Proceedings of the 2018 USENIX Annual Technical Conference (USENIX ATC '18) (pp. 133–145). Boston, MA, USA. Retrieved from: <https://www.usenix.org/conference/atc18/presentation/wang-liang>
2. Shahrad, M., Fonseca, R., Zhang, I., et al. (2020). Serverless in the wild: Characterizing and optimizing the serverless workload at a large cloud provider. In Proceedings of the 2020 USENIX Annual Technical Conference (USENIX ATC '20) (pp. 205–218). Retrieved from: <https://www.usenix.org/conference/atc20/presentation/shahrad>
3. Schirmer, T., Japke, N., Greten, S., Pfandzelter, T., & Bermbach, D. (2023). The night shift: Understanding performance variability of cloud serverless platforms. In The 1st Workshop on Serverless Systems, Applications and Methodologies (SESAME '23). <https://doi.org/10.1145/3592533.3592808>
4. Scheuner, J., Bertilsson, M., Grönqvist, O., et al. (2022). TriggerBench: A performance benchmark for serverless function triggers. 2022 IEEE International Conference on Cloud Engineering (IC2E). <https://doi.org/10.1109/IC2E55432.2022.00018>
5. Wen, J., Chen, Z., Jin, X., & Liu, X. (2023). Rise of the planet of serverless computing: A systematic review. *ACM Transactions on Software Engineering and Methodology*, 32(5), Article 131. <https://doi.org/10.1145/3579643>
6. Rahman, A., Mahdavi-Hezaveh, R., & Williams, L. (2019). A systematic mapping study of infrastructure as code research. *Information and Software Technology*, 108, 65–77. <https://doi.org/10.1016/j.infsof.2018.12.004>
7. Begoug, M., Chouchen, M., & Ouni, A. (2024). TerraMetrics: An open source tool for infrastructure-as-code (IaC) quality metrics in Terraform. Proceedings of the 2024 IEEE/ACM 32nd International Conference on Program Comprehension (ICPC) – Tool Demonstrations. <https://doi.org/10.1145/3643916.3644439>
8. Sokolowski, D., Spielmann, D., & Salvaneschi, G. (2024). Automated infrastructure as code program testing. *IEEE Transactions on Software Engineering*, 50(6), 1585–1599. <https://doi.org/10.1109/TSE.2024.3393070>
9. Rahman, A., Parnin, C., & Williams, L. (2019). The seven sins: Security smells in infrastructure as code scripts. Proceedings of the 41st International Conference on Software Engineering (ICSE '19), 164–175. <https://doi.org/10.1109/ICSE.2019.00033>
10. Hassan, M. M., Salvador, J., Santu, S. K. K., & Rahman, A. (2024). State reconciliation defects in infrastructure as code. Proceedings of the ACM on Software Engineering, 1 (FSE), Article 83. <https://doi.org/10.1145/3660790>
11. Shrestha, R., & Ali, A. A. N. (2024). Configuration management in Kubernetes environments: A GitOps approach. 2024 IEEE/ACM 17th International Conference on Utility and Cloud Computing (UCC 2024), 497–502. <https://doi.org/10.1109/UCC63386.2024.00077>
12. Janes, A., Li, X., & Lenarduzzi, V. (2023). Open tracing tools: Overview and critical comparison. *Journal of Systems and Software*, 204, 111793. <https://doi.org/10.1016/j.jss.2023.111793>

Дата надходження статті: 01.09.2025

Дата прийняття статті: 10.10.2025

Опубліковано: 30.12.2025

Булгакова О. Ф., старший викладач кафедри комп'ютерних наук та інженерії програмного забезпечення
Університету митної справи та фінансів
ORCID: 0000-0001-9834-2970

ОГЛЯД І УЗАГАЛЬНЕННЯ ПРАКТИК ВПРОВАДЖЕННЯ ВІРТУАЛІЗАЦІЇ У ПІДГОТОВЦІ ІТ-ФАХІВЦІВ

Цифрова трансформація сучасного суспільства докорінно змінила вимоги та підходи до освіти в галузі інформаційних технологій. У статті показано, як технології віртуалізації трансформують підготовку фахівців, поєднуючи технічні рішення з обґрунтованими навчальними підходами. Розглянуто можливості віртуальних машин, контейнеризації, програмно визначених мереж, віртуалізації сховищ та хмарних віддалених лабораторій для безпечного й відтвореного моделювання навчальних середовищ. Мета роботи – продемонструвати, як ці технології посилюють практичну підготовку, оптимізують витрати, розширюють доступ і узгоджують результати навчання з очікуваннями роботодавців. Проведено огляд сучасних публікацій і аналіз типових архітектур розгортання. Показано, що ізоляція, сумісність і можливість повторення експериментів дають змогу безпечно відпрацьовувати складні сценарії та прискорюють набуття практичних навичок студентами. Водночас ідентифіковано обмеження: вимоги до мережевих та обчислювальних ресурсів, накладні витрати продуктивності у ресурсомістких завданнях, потребу в цілеспрямованій підготовці викладачів і ризики зниження мотивації та взаємодії в дистанційних форматах. Запропоновано практичні орієнтири: поетапне впровадження зі стандартизованими образами, розвиток користувацької підтримки, інтеграцію інструментів оцінювання та партнерства із компаніями для проєктного навчання. Окреслено перспективи використання штучного інтелекту для персоналізованого наставництва, периферійних обчислень для зменшення затримок, технологій розширеної реальності для інтуїтивних тривимірних інтерфейсів, а також хмарно-нативних і безсерверних підходів, що прискорюють масштабування і спрощують супровід навчальної інфраструктури. Запропоновані рішення допомагають поєднати якість навчання, масштабованість і економічну доцільність та підвищити готовність випускників до сучасних професійних викликів.

Ключові слова: інформаційні технології, віртуалізація, освіта, контейнеризація, хмарні технології, операційні системи.

Bulhakova O. F. A Review and synthesis of virtualization implementation practices in IT professional training

The digital transformation of contemporary society has fundamentally reshaped the requirements and approaches to education in the field of information technology. This article demonstrates how virtualization technologies transform professional preparation by aligning technical solutions with pedagogically grounded approaches. It examines the capabilities of virtual machines, containerization, software-defined networking (SDN), storage virtualization, and cloud-based remote laboratories for safe and reproducible modeling of instructional environments. The aim is to show how these technologies strengthen hands-on training, optimize costs, expand access, and align learning outcomes with employer expectations. The study presents a review of recent publications and an analysis of typical deployment architectures. It shows that isolation, compatibility, and repeatability enable students to practice complex scenarios safely and accelerate the acquisition of practical skills. At the same time, the study identifies constraints: requirements for network and compute resources, performance overhead in resource-intensive tasks, the need for targeted faculty development, and risks of reduced motivation and interaction in fully online settings. Practical guidelines are proposed, including phased implementation with standardized virtual machine images, development of learner support services, integration of assessment tools, and partnerships with companies for project-based learning. The paper outlines prospects for using artificial intelligence for personalized tutoring, edge computing to reduce latency, extended reality (XR) for intuitive three-dimensional interfaces, and cloud-native and serverless approaches that accelerate scaling and simplify the operation of educational infrastructure. The proposed solutions help reconcile instructional quality, scalability, and cost-effectiveness and increase graduates' readiness for contemporary professional challenges.

Key words: information technology, virtualization, education, containerization, cloud computing, operating systems.

Постановка проблеми. Цифрова трансформація сучасного суспільства докорінно змінила ландшафт ІТ-освіти, вимагаючи інноваційних підходів до підготовки фахівців для дедалі більш віртуалізованого професійного середовища. Віртуалізація, що визначається як створення віртуальних версій фізичних обчислювальних ресурсів – зокрема апаратних платформ, мережевих пристроїв, операційних систем і систем зберігання даних [1] – стала трансформаційним чинником у сфері освітніх технологій. Ця зміна парадигми

виходить за межі простого впровадження технологій, означаючи фундаментальне переосмислення того, як ІТ-освіту можна надавати, як до неї можна отримувати доступ і яким є досвід її здобуття.

У контексті підготовки ІТ-фахівців віртуалізація охоплює низку технологічних підходів: віртуальні машини (ВМ), контейнерні рішення, віртуалізацію мережі та хмарні лабораторні середовища. Сукупно ці технології розв'язують давні проблеми ІТ-освіти – насамперед забезпечення апаратно-програмної підтримки для складних технічних дисциплін (курсів) за умов обмежених ресурсів закладу. Зростання значущості віртуалізації і в промисловій практиці, і в освітній діяльності відображає її здатність долати традиційний розрив між теоретичними знаннями та практичним застосуванням.

Аналіз останніх досліджень і публікацій. Інтеграція технологій віртуалізації в освітні середовища означає суттєву еволюцію педагогічних підходів до викладання ІТ-дисциплін; наукові публікації фіксують цю трансформацію за кількома вимірами. Ці дослідження показали, що віртуалізація дає змогу викладати складні технічні дисципліни без традиційних обмежень фізичного (апаратного) середовища [3].

Подальші дослідження розширили це розуміння, охопивши ширші наслідки функціонування віртуалізованих навчальних середовищ. Теоретична рамка віртуалізації в освіті спирається на кілька педагогічних теорій, зокрема на теорію досвідного навчання, яка наголошує на важливості практичного досвіду у формуванні навичок, і на конструктивістські підходи до навчання, що надають пріоритет активному конструюванню знань через практичну взаємодію [5]. Ці теоретичні засади підтримують тезу, що технології віртуалізації можуть підвищувати результати навчання, надаючи студентам безпечні, ізольовані середовища для експериментування та відпрацювання навичок.

Еволюція технологій віртуалізації йшла паралельно зі змінами в освітніх технологіях загалом. Перші реалізації фокусувалися на консолідації серверів і базовому розгортанні ВМ; їх навчальні застосування були обмежені через складність і вимоги до ресурсів. Проте поширення хмарних обчислень і контейнерних рішень радикально розширило доступність та придатність віртуалізації в освіті [1]. Сучасні роботи демонструють, як хмарні віртуальні лабораторії забезпечують масштабоване та економічно ефективне викладання складних ІТ-курсів, надаючи студентам доступ до індустріальних інструментів і середовищ.

Нещодавні наукові праці особливо підкреслюють роль віртуалізації в освіті з кібербезпеки, де здатність технологій створювати складні, реалістичні сценарії виявляється особливо цінною [2][6]. Ці дослідження демонструють, як віртуалізація дає змогу студентам працювати з витонченими викликами безпеки – зокрема веб-безпекою, атаками типу SQL-ін'єкцій та сценаріями мережевої захисту – у межах контрольованих навчальних середовищ. Можливість створювати багато віртуальних хостів, знімати обмеження мережевих інтерфейсів і надавати студентам права адміністратора у їхніх віртуальних середовищах становить суттєвий поступ у викладанні кібербезпеки.

Зростає увага і до контейнерних підходів у навчанні DevOps та сучасних методологій розробки [7]. Контейнери дають переваги над традиційними віртуальними машинами з огляду на ефективність використання ресурсів і швидкість розгортання, що особливо доречно там, де потрібне швидке надання і відтвореність навчальних середовищ. Це відповідає індустріальним трендам на контейнеризацію та мікросервіси.

Водночас література вказує на суттєві прогалини в поточних дослідженнях, зокрема щодо всебічної оцінки результатів навчання та довгострокового оцінювання впливу. Попри численні праці, що описують успішні впровадження й переваги, емпіричних даних, які порівнюють результати навчання між віртуалізованими та традиційними підходами до ІТ-освіти, досі бракує. Крім того, дослідження з професійного розвитку викладачів і управління змінами в закладах освіти в контексті впровадження віртуалізації залишаються недостатньо опрацьованими, попри критичну важливість цих чинників для успішної реалізації.

Теоретична рамка розуміння освітнього впливу віртуалізації має також враховувати ширші технологічні та педагогічні тенденції. Інтеграція штучного інтелекту, машинного навчання та периферійних (edge) обчислень із традиційними платформами віртуалізації свідчить, що майбутні освітні застосування ставатимуть дедалі складнішими та персоналізованішими. Це змушує заклади будувати адаптивні підходи, здатні встигати за технологіями без втрати освітньої ефективності.

Мета статті – всебічно проаналізувати роль технологій віртуалізації у підвищенні ефективності освітнього процесу для підготовки ІТ-фахівців, із акцентом на педагогічну результативність, покращення доступності та узгодженість з галузевими стандартами. Дослідження охоплює аналіз поточних стратегій впровадження, освітніх результатів та викликів інтеграції віртуалізації у стандартні освітні програми підготовки ІТ-фахівців. Використано методи огляду літератури, аналізу кейсів і вивчення найкращих практик закладів, що успішно реалізували навчальні середовища на основі віртуалізації.

Значущість роботи виходить за межі суто технологічних аспектів і охоплює проблематику освітньої справедливості та доступності. Технології віртуалізації здатні демократизувати доступ до якісної ІТ-освіти, знижуючи апаратні бар'єри та забезпечуючи віддалений доступ до лабораторій, – тим самим долаючи географічні та соціально-економічні обмеження, які традиційно звужують освітні можливості [4]. Більш того, у міру того як індустрія дедалі ширше спирається на віртуалізовану інфраструктуру, навчальні заклади відчують зростаючий тиск щодо узгодження програм із цими технологічними реаліями, аби гарантувати працевлаштованість і конкурентоздатність випускників.

Статтю структуровано як всебічний огляд ролі віртуалізації в ІТ-освіті: спершу викладено теоретичні засади та огляд літератури; далі подано детальний аналіз технологій віртуалізації та їхніх освітніх застосувань, стратегій впровадження, переваг і викликів; на завершення – майбутні тенденції та наслідки для зацікавлених сторін у сфері освіти.

Виклад основного матеріалу.

Технології віртуалізації та освітні застосування

А. Типи віртуалізації в ІТ-освіті

Ландшафт технологій віртуалізації, придатних для ІТ-освіти, охоплює кілька відмінних підходів, кожен з яких пропонує унікальні переваги для різних освітніх цілей і технічних вимог. Віртуальні машини (ВМ) є найбільш усталеною формою віртуалізації в освітніх контекстах; вони використовують технології гіпервізорів для створення повноцінних, ізольованих обчислювальних середовищ, здатних розміщувати різні операційні системи та прикладні програми [2]. Такий підхід дає студентам змогу працювати в різноманітних обчислювальних середовищах без потреби в багатьох фізичних машинах, розв'язуючи базову проблему забезпечення повноцінного практичного досвіду за умов обмежених ресурсів.

Технології гіпервізорів, зокрема рішення типу 1 (bare-metal) і типу 2 (hosted), становлять основу для розгортання навчальних віртуальних машин. Гіпервізори типу 1 забезпечують вищу продуктивність і зазвичай застосовуються в дата-центрах закладів освіти, що підтримують одночасну роботу багатьох користувачів, тоді як гіпервізори типу 2 дають змогу окремим студентам створювати віртуальні середовища на власних пристроях [3]. Така гнучкість моделей розгортання дозволяє закладам освіти впроваджувати стратегії віртуалізації, які враховують різну доступність ресурсів і вимоги до доступу.

Рішення на основі контейнерів, зокрема платформи Docker і Kubernetes, становлять суттєвий еволюційний крок у технологіях віртуалізації та мають особливу значущість для сучасної ІТ-освіти [7]. На відміну від традиційних віртуальних машин, які віртуалізують цілу операційну систему, контейнери забезпечують віртуалізацію на рівні операційної системи, пакуючи застосунки разом із їхніми залежностями у легкі портативні одиниці. Такий підхід забезпечує майже як у природному середовищі продуктивність під час операцій із процесором, пам'яттю, диском і мережею, водночас уможливаючи швидке розгортання та масштабування навчальних середовищ. Технології контейнерів особливо корисні для навчання практик DevOps, архітектур мікросервісів і підходів до розроблення, спроектованих для хмари (cloud-native), які дедалі ширше застосовуються у професійному ІТ-середовищі.

Технології віртуалізації мережі, зокрема SDN (Software-Defined Networking) та віртуальні мережеві функції (VNF) – відкривають можливості для складних мережевих лабораторій, які фізично відтворити складно або неможливо [1]. Ці технології дозволяють студентам створювати складні мережеві топології, експериментувати з протоколами маршрутизації та опановувати засади мережевої безпеки у контрольованих, відтворюваних середовищах. Здатність оперативно змінювати конфігурацію мережевих середовищ і моделювати різні умови роботи мереж надає освітні можливості, максимально наближені до реальних викликів мережевої інженерії.

Віртуалізація зберігання (даних) доповнює інші технології віртуалізації, абстрагуючи фізичні ресурси сховищ і даючи змогу гнучко розподіляти місткість між багатьма віртуальними середовищами. В освітньому контексті вона полегшує створення сталих студентських середовищ і водночас оптимізує використання ресурсів інституційної інфраструктури. Ця можливість особливо цінна для проєктно-орієнтованого навчання, коли студентам потрібен стабільний доступ до своїх робочих середовищ упродовж тривалого часу.

Хмарні віртуальні лабораторії – це поєднання кількох технологій віртуалізації в платформах із віддаленим доступом, які усувають традиційні бар'єри доступу до ІТ-освіти [8][4]. Такі платформи використовують моделі IaaS і PaaS, надаючи студентам ресурси за запитом: потужні обчислювальні засоби та спеціалізовані програмні середовища. Концепція «Lab@Home», що спирається на приватну хмарну інфраструктуру на основі серверної віртуалізації та розподіленого зберігання даних, демонструє, як хмарні підходи можуть розширювати освітній доступ за межі традиційних кордонів, зберігаючи при цьому педагогічний контроль і можливості оцінювання.

В. Технічні рамки впровадження

Десктопна віртуалізація є базовим підходом до впровадження, який дозволяє окремим студентам створювати та керувати віртуальними середовищами на власних обчислювальних пристроях. Зазвичай цей підхід використовує гіпервізори типу 2 (наприклад, VMware Workstation, VirtualBox або Parallels Desktop), щоб надати студентам повний контроль над їхніми навчальними середовищами [5]. Десктопна віртуалізація має низку освітніх переваг, зокрема можливість відтворювати експерименти вдома, підтримувати постійні середовища розроблення із збереженням стану, а також експериментувати з різними операційними системами та конфігураціями ПЗ без ризику для основного середовища.

Освітня ефективність десктопної віртуалізації значною мірою залежить від належної конфігурації та дієвих служб підтримки. Студентам потрібні достатні апаратні ресурси – насамперед обсяг оперативної пам'яті та місткість сховища – щоб забезпечити одночасну роботу кількох віртуальних машин. Крім того, заклади освіти мають надавати всебічну технічну підтримку й навчання, аби студенти могли ефективно

користуватися інструментами віртуалізації. Властивості ізоляції, сумісності та інкапсуляції образів віртуальних машин особливо цінні в цьому контексті, адже вони дають змогу розгортати реалістичні навчальні середовища та водночас утримувати під контролем складність, зумовлену різноманітністю освітніх інструментів і вимог [5].

Серверна віртуалізація для масштабованих лабораторних середовищ є більш просунутим підходом до впровадження, який спирається на інфраструктуру закладу освіти для підтримки багатьох одночасних користувачів і складних навчальних сценаріїв. Зазвичай цей підхід використовує гіпервізори та платформи керування промислового рівня, що забезпечують централізоване адміністрування та розподіл ресурсів між численними віртуальними середовищами [1]. Рішення на базі серверів дають змогу підтримувати більші групи студентів, зберігаючи стабільні показники продуктивності та доступності.

Переваги масштабованості серверної віртуалізації виходять за межі простої місткості користувачів і охоплюють здатність швидко розгортати нові лабораторні середовища та адаптуватися до змін вимог дисциплін. Інструменти автоматизованого розгортання та підготовки ресурсів (provisioning) дають змогу за лічені хвилини розгорнути попередньо налаштовані образи віртуальних машин і мережеві топології, тож складні лабораторні роботи запускаються без тривалих ручних налаштувань. Це особливо важливо за обмеженого навчального часу й дає змогу виконувати частіші ітерації та експерименти під час розроблення й оновлення змісту курсів.

Вкладена віртуалізація (nested) – це просунута техніка реалізації, що дозволяє віртуальним машинам розміщувати інші віртуальні машини, формуючи багаторівневі середовища віртуалізації [2]. Такий підхід особливо цінний в освіті з кібербезпеки, де студентам потрібно моделювати складні сценарії атаки й оборони за участю багатьох взаємопов'язаних систем. Вкладена віртуалізація дає змогу створювати реалістичні корпоративні середовища в межах навчальних лабораторій, водночас зберігаючи ізоляцію та контроль над експериментальними умовами.

Технічна складність вкладеної віртуалізації потребує ретельного врахування впливу на продуктивність і стратегій розподілу ресурсів. Кілька рівнів віртуалізації додають накладні витрати, що можуть позначатися на швидкодії системи, тож потрібні достатні апаратні ресурси та оптимізація конфігурацій віртуалізації. Водночас освітні переваги можливості створювати комплексні, реалістичні ІТ-середовища часто виправдовують додаткову складність, особливо для просунутих курсів із кібербезпеки, системного адміністрування та керування мережами.

Стратегії впровадження та найкращі практики

Кейси успішного впровадження віртуалізації

Впровадження віртуалізації в ІТ-освіті продемонструвало відчутні успіхи в закладах різних типів. Комплексний аналіз програм з кібербезпеки показує, що віртуалізація дає змогу студентам виходити за межі традиційної аудиторії: створювати багато віртуальних хостів, знімати обмеження мережевих інтерфейсів та адмініструвати власні віртуальні середовища [2][6]. Цей підхід дозволяє моделювати сценарії, які на реальному обладнанні були б практично нереалізовними або економічно невиправданими.

Переорієнтація на хмарні віртуальні лабораторії комп'ютерних мереж становить зміну парадигми у способах надання освіти [1]. Такі реалізації спираються на синергію віртуалізації та хмарних обчислень: хмарні серверні системи значною мірою ґрунтуються на технологіях віртуалізації, щоб забезпечити масштабовані та економічно ефективні навчальні середовища. Концепція Lab@Home ілюструє цей підхід: приватна хмарна інфраструктура, побудована на серверній віртуалізації та розподіленому зберіганні даних, забезпечує легкі, портативні й економічно доцільні лабораторні практики, які виходять за межі традиційних рамок закладу, водночас зберігаючи педагогічний контроль і можливості оцінювання [4].

Популярні віртуальні лабораторії з безпеки – DeterLab, Tele-Lab IT та NETinVM – стали еталонними орієнтирами успішного впровадження [6]. Ці платформи демонструють, як завдяки віртуалізації студенти можуть налаштовувати мережі з віртуальних машин і відпрацьовувати сценарії несанкціонованого доступу та заходів захисту. Можливість відтворювати експерименти вдома продовжує навчання за межами закладу та забезпечує безперервність між аудиторними заняттями і самостійною практикою.

Вимоги до інфраструктури та розподілу ресурсів

Ефективне впровадження віртуалізації потребує ретельного врахування технічної інфраструктури та стратегій розподілу ресурсів. Створення віртуальних аналогів фізичних обчислювальних ресурсів охоплює апаратні платформи, обчислювальні пристрої, мережеве обладнання, пристрої введення/виведення та системи зберігання даних [1]. Такий комплексний підхід забезпечує студентам роботу в реалістичних навчальних середовищах, водночас зберігаючи ізоляцію, сумісність та інкапсуляцію, необхідні для досягнення якісних результатів навчання [5].

Хмарна інфраструктура стала ключовим чинником масштабованого розгортання віртуалізації. Університети, що впроваджують хмарні обчислення, отримують зниження витрат, просте додавання віртуальних ресурсів і поліпшені можливості інтеграції – усе це разом підвищує конкурентоспроможність закладу [9]. Перехід до хмарних рішень також допомагає стримувати зростання витрат, пов'язаних із дедалі більшою різноманітністю інструментів і середовищ, потрібних у сучасній ІТ-освіті [5].

Планування розподілу ресурсів має враховувати і десктопну, і повну (серверну) віртуалізацію, адже кожна з них слугує різним освітнім цілям [6]. Десктопна віртуалізація підтримує індивідуальні сценарії навчання, тоді як повна (серверна) дає змогу будувати складні багатовузлові конфігурації, необхідні для просунутої підготовки з ІТ. Застосування вкладеної (nested) віртуалізації ще більше розширює можливості, дозволяючи реалізовувати складні освітні сценарії, що наближені до реальних корпоративних середовищ.

Підготовка викладачів і системи підтримки

Успішне впровадження технологій віртуалізації потребує комплексних програм розвитку викладачів, що охоплюють і технічні компетентності, і педагогічну адаптацію. Віртуальне навчання вимагає широкого спектра інструментів для віртуальної взаємодії, знань щодо програмного та апаратного забезпечення, тож потрібне безперервне підвищення кваліфікації викладачів [10]. Стрімке розширення платформ віртуальної освіти, яке особливо прискорила пандемія COVID-19, підсвітило критичну важливість готовності викладацького складу до роботи у віртуальних навчальних середовищах.

Контроль якості та оцінювання у віртуалізованих середовищах породжують специфічні виклики, що потребують спеціальної підготовки [10]. Викладачам слід розвивати нові компетентності з моніторингу прогресу студентів у віртуалізованих середовищах (зокрема із використанням образів віртуальних машин), розуміючи, як віртуалізація дає змогу розгортати реалістичні навчальні середовища без втрати педагогічної ефективності [5]. Найкращі результати демонструють інтерактивні віртуальні заняття, що підтримують активну участь студентів і ощадливе використання навчального часу.

Системи підтримки мають охоплювати інтеграцію ІТ-інфраструктури та онлайн-платформ, особливо для викладачів із невеликим попереднім досвідом у сфері віртуальної освіти [11]. Програми підвищення кваліфікації варто вибудовувати з акцентом на практичній взаємодії та командних форматах роботи, що формують переваги віртуального навчання порівняно з традиційними дистанційними підходами.

Моделі доступу для студентів і концепції віддалених лабораторій

Демократизація ІТ-освіти завдяки віртуалізації потребує ретельно спроектованих моделей доступу для студентів, які враховують різні освітні потреби й життєві обставини. Віртуальне навчання створює можливості для інклюзивності та різноманіття, даючи університетам змогу навчати студентів з інших країн, а також тих, хто через фізичні чи економічні обмеження не може відвідувати заняття очно [9]. Запровадження технологій з відкритим кодом зробило віртуальне навчання привабливішим і доступнішим для ширших студентських аудиторій.

Концепції віддалених лабораторій помітно еволюціонували в епоху хмарних обчислень, і модель Lab@Home уособлює комплексний підхід до розподіленого навчання [4]. Ця концепція передбачає надання студентам легких, компактних, портативних та економічно доцільних лабораторних наборів і програмних середовищ, що дає змогу здійснювати віддалене оцінювання прогресу та водночас розширює традиційні підходи до віддалених лабораторій – як з погляду технологічної інфраструктури, так і з погляду навчання й викладання.

Гнучкість віртуального навчання, що включає цілодобовий доступ до матеріалів, менші витрати на дорогу та вищу самостійність, слід збалансовувати з ризиками низької мотивації й прокрастинації. Ефективні моделі доступу спираються на підтримувальне навчальне середовище та достатню інфраструктуру, адже саме вони є критично важливими для успіху віртуальної освіти.

Переваги та освітні результати

Посилені можливості практичного навчання

Віртуалізація докорінно змінює масштаб і якість практичної складової ІТ-освіти. Технологія дозволяє студентам створювати більше віртуальних хостів, ніж доступно фізичних комп'ютерів, знімати традиційні обмеження мережевих інтерфейсів і виконувати адміністрування власних віртуальних середовищ [2, 6]. Це дає змогу відпрацьовувати складні сценарії, раніше недосяжні у традиційних лабораторних умовах, і працювати з витонченими конфігураціями та взаємодіями між кількома системами.

Властивості ізоляції, сумісності та інкапсуляції, притаманні віртуальним машинам, створюють оптимальні умови для реалістичних навчальних середовищ [5]. Студенти можуть експериментувати з різними операційними системами, мережевими конфігураціями та сценаріями безпеки без ризику пошкодити фізичну інфраструктуру чи вплинути на роботу інших здобувачів. Така «безпечна пісочниця» заохочує дослідження й ітеративне навчання – ключові складові формування ІТ-навичок.

Віртуальні лабораторії переконливо демонструють навчальний потенціал віртуалізації. Вони дають змогу розгортати мережі та сервіси з віртуальних машин, моделювати робочі середовища операційних систем і баз даних, а також відпрацьовувати процеси DevOps/MLOps та інтеграцію розподілених систем у хмарних платформах і сервісах (IaaS/PaaS/Serverless). Відтворюваність таких експериментів удома продовжує навчання поза межами закладу й підсилює аудиторні заняття завдяки регулярній практиці [12, 13].

Економічна ефективність та оптимізація ресурсів

Економічні переваги віртуалізації в освіті виходять за межі простого скорочення витрат на апаратне забезпечення і охоплюють комплексні стратегії оптимізації ресурсів. Впровадження хмарних обчислень забезпечує закладам освіти помітне зниження витрат, спрощує додавання віртуальних ресурсів і покращує

інтеграційні можливості [9]. У сукупності ці вигоди підвищують конкурентоспроможність закладів, що впроваджують технології віртуалізації.

Віртуалізація дає відповідь на зростання витрат, пов'язане з дедалі більшою різноманітністю інструментів і середовищ, необхідних у сучасній ІТ-освіті [5]. Консолідація багатьох навчальних середовищ на спільній фізичній інфраструктурі дозволяє максимізувати віддачу від інвестицій у техніку й водночас надати студентам доступ до різних конфігурацій ПЗ та систем, підтримка яких у вигляді окремих фізичних стендів була б надмірно дорогою.

Модель Lab@Home демонструє, що завдяки віртуалізації можна створювати доступні, компактні й мобільні лабораторії [4]. Це зменшує залежність від великих стаціонарних комплексів, зберігаючи рівень навчання та масштабованість без додаткових витрат.

Підвищена доступність і демократизація ІТ-освіти

Віртуалізація робить ІТ-освіту доступнішою: знімає звичні бар'єри для входу й участі. Вона дає змогу університетам навчати студентів з інших країн і тих, кому заважають фізичні чи фінансові обмеження [9]. Гнучкий формат із цілодобовим доступом до матеріалів і без постійних поїздок на пари помітно розширює можливості для навчання.

Цифрова трансформація, прискорена останніми глобальними подіями, продемонструвала спроможність віртуалізації забезпечувати безперервність навчання в складних умовах. Традиційні університети пройшли значні процеси віртуалізації, що підтвердило ефективність технологій для підтримання безперервного навчального досвіду попри зовнішні обмеження.

Технології з відкритим кодом зробили віртуальне навчання привабливішим і доступнішим для широких аудиторій [9]. Така технологічна демократизація дає змогу меншим закладам і регіонам, що розвиваються, надавати якісну ІТ-освіту без значних капіталовкладень у традиційні лабораторні комплекси.

Узгодження розвитку навичок із галузевими стандартами

Віртуалізація в освіті безпосередньо знімає розрив між академічною підготовкою та вимогами галузі. Синергія віртуалізації та хмарних обчислень означає, що студенти, які навчаються з використанням цих технологій, здобувають прямий досвід роботи з інфраструктурними підходами, що домінують у сучасних ІТ-середовищах [1]. Така відповідність гарантує, що випускники мають навички, які можна відразу застосувати на робочому місці.

Комплексний характер навчання в середовищах віртуалізації – охоплення апаратних платформ, обчислювальних пристроїв, мережевого обладнання та систем зберігання даних [1] – формує в студентів цілісне розуміння сучасної ІТ-інфраструктури. Така широта досвіду є ключовою для підготовки універсальних ІТ-фахівців, здатних працювати у різних технологічних доменах.

Інтерактивні віртуальні заняття, що стимулюють активну участь студентів, демонструють кращі результати порівняно з традиційними підходами [11]. Можливість експериментувати зі складними конфігураціями та багаторазово відтворювати сценарії прискорює опанування навичок і поглиблює концептуальне розуміння, тож випускники краще готові до вирішення практичних завдань у професійному ІТ-середовищі.

Виклики та обмеження

Технічні інфраструктурні бар'єри

Впровадження віртуалізації в освітніх середовищах стикається з помітними інфраструктурними труднощами, що можуть гальмувати успішний запуск. Недостатня швидкість/стабільність інтернету та повільні (застарілі) пристрої є ключовими перешкодами для повноцінного функціонування віртуального навчання [11]. Ці обмеження особливо відчутні в регіонах зі слаборозвиненою технологічною інфраструктурою, що спричиняє нерівність у доступі та неоднакову якість освітніх послуг.

Інтеграція потужної ІТ-інфраструктури та онлайн-платформ залишається постійним викликом, особливо для закладів із невеликим попереднім досвідом у віртуальній освіті [11]. Керування віртуалізованими середовищами складне й потребує спеціалізованої технічної експертизи, якої традиційним закладам може бракувати. Це вимагає суттєвих інвестицій у служби технічної підтримки та можливості обслуговування.

Механізми контролю якості у віртуалізованих середовищах потребують розвинених систем моніторингу й оцінювання, які суттєво відрізняються від традиційних підходів до освітнього оцінювання [10]. Розроблення та впровадження таких систем вимагає значних технічних ресурсів і компетенцій, що може створювати додаткові бар'єри для закладів з обмеженими технічними можливостями.

Накладні витрати продуктивності

Віртуалізація за своєю природою додає накладні витрати продуктивності, що може позначатися на навчальному процесі, особливо у ресурсомістких сценаріях підготовки з ІТ. Запуск багатьох віртуальних машин на спільному фізичному обладнанні створює вузькі місця продуктивності, які обмежують складність і реалістичність лабораторних завдань. Це особливо відчутно, коли відтворюються сценарії промислового рівня або коли одночасно працює велика кількість користувачів.

Підтримка балансу між економією та швидкістю – постійний виклик для впровадження віртуалізації. Попри скорочення витрат на апаратне забезпечення, забезпечення прийняттого рівня продуктивності часто вимагає інвестицій у високопродуктивну інфраструктуру (сервери, сховища, мережі), що може частково нівелювати фінансові переваги.

Студенти, які працюють із власних пристроїв, можуть стикатися зі значними обмеженнями продуктивності, особливо під час доступу до ресурсомістких віртуальних лабораторій. Така нерівність у доступі до необхідних обчислювальних ресурсів підриває демократизуючий потенціал технологій віртуалізації.

Адаптація викладачів і вимоги до підготовки

Перехід до віртуалізованого формату навчання потребує комплексного розвитку викладачів, що виходить за межі суто технічного тренінгу й охоплює педагогічну трансформацію. Викладачі мають опанувати керування інструментами віртуальної взаємодії, програмним та апаратним забезпеченням і водночас адаптувати методики викладання до віртуальних середовищ [10]. Поєднання технічної й педагогічної підготовки формує значний запит на підвищення кваліфікації.

Брак досвіду з платформами та методиками віртуального навчання під час переходу часто знижує ефективність. Викладачам без належної підготовки важче створювати цікаві й результативні онлайн-курси, що може тимчасово погіршити результати студентів.

Потреба в безперервній підтримці для викладачів, які працюють із технологіями віртуалізації, є суттєвим зобов'язанням для закладу. Швидка еволюція платформ і інструментів віртуалізації вимагає постійного підвищення кваліфікації, аби зберігати педагогічну ефективність, що створює довгострокові виклики розподілу ресурсів для освітніх установ.

Проблеми мотивації та залученості студентів

Віртуальні навчальні середовища висувають особливі виклики для підтримання мотивації та залученості, які суттєво відрізняються від динаміки традиційної аудиторії. Студенти можуть відчувати незадоволеність через брак мотивації, прокрастинацію та труднощі онлайн-комунікації, що негативно впливає на результати навчання. Ці мотиваційні ризики особливо помітні в онлайн- та змішаних форматах навчання, де відсутні звичні соціальні та мотиваційні механізми підтримки.

Відсутність безпосередньої взаємодії «викладач–студент» і «студент–студент» істотно позначається і на результатах навчання, і на задоволеності студентів. Віртуальні формати не завжди відтворюють спонтанні взаємодії та командну динаміку очного навчання, тож зменшуються соціальні чинники, важливі для успіху.

Технічні збої та обмеження платформ викликають фрустрацію і підривають залученість та мотивацію. Коли системи віртуального навчання працюють ненадійно або мають неінтуїтивний чи повільний інтерфейс, у студентів формується стійке негативне ставлення до онлайн- та змішаних форматів навчання – навіть після усунення конкретних технічних проблем.

Тренди майбутнього та наслідки

Інтеграція з новітніми технологіями

Зближення віртуалізації зі штучним інтелектом і машинним навчанням обіцяє докорінно змінити ІТ-освіту завдяки персоналізованим та адаптивним навчальним середовищам. Віртуальні лабораторії на базі ІІІ можуть включати інтелектуальні системи, що підлаштовуються під індивідуальні шаблони навчання студентів, пропонують індивідуалізовані завдання та підтримку на основі аналізу результатів у реальному часі. Така інтеграція дає змогу створювати віртуальні середовища, які не лише моделюють технічні системи, а й забезпечують розумне наставництво та оцінювання.

Інтеграція периферійних (edge) обчислень із технологіями віртуалізації відкриває нові освітні підходи: обчислювальні ресурси розміщуються ближче до здобувачів освіти, водночас зберігається гнучкість і економічність віртуальних середовищ. Такий розподілений підхід зменшує затримки у віртуальних лабораторіях і надає студентам практичний досвід роботи з сучасними архітектурами розподілених обчислень, які стають дедалі важливішими у професійному ІТ-середовищі.

Поєднання платформ віртуалізації з технологіями розширеної реальності (XR) створює занурювальні (immersive) формати навчання: безпечні, ізольовані віртуальні середовища доповнюються інтуїтивними 3D-інтерфейсами. Такі гібридні підходи підвищують залученість і засвоєваність матеріалу, водночас зберігаючи практичні переваги традиційної віртуалізації [10, 11].

Переорієнтація на хмарно-нативні навчальні платформи

Перехід від традиційної віртуалізації до хмарно-нативних (cloud-native) освітніх архітектур означає фундаментальну зміну підходів до організації та масштабування ІТ-освіти. Контейнеризація та архітектури на основі мікросервісів дають змогу будувати більш модульні, ефективні й швидкі в розгортанні навчальні середовища, які можна точно налаштовувати під конкретні навчальні цілі та динамічно масштабувати за попитом.

Парадигми безсерверних обчислень (serverless) дозволяють закладам освіти зосередитися на змісті й методиці, а не на керуванні інфраструктурою, зменшуючи технічні бар'єри для впровадження віртуалізації. Такий рух до керованих освітніх платформ демократизує доступ до складних віртуальних навчальних середовищ для закладів із обмеженими технічними ресурсами [12, 13].

Розвиток освітніх PaaS-рішень (platform-as-a-service), спеціально спроектованих для ІТ-підготовки, стандартизуватиме підходи до віртуалізації й водночас дозволить закладам зосередитися на розробленні змісту та підтримці студентів, а не на технічній реалізації. Такі спеціалізовані платформи можуть надавати

попередньо налаштовані навчальні середовища, інструменти автоматизованого оцінювання та вбудовані засоби співпраці, оптимізовані під потреби ІТ-освіти.

Аспекти сталого розвитку та масштабованості

Для забезпечення глобального доступу до освіти розвивають розподілені архітектури віртуалізації, здатні обслуговувати кілька географічних регіонів із підтриманням сталої якості навчання; такі підходи водночас враховують технічні аспекти масштабування та регуляторні вимоги щодо локалізації даних і захисту персональної інформації в міжнародних освітніх контекстах.

Паралельно застосування принципів раціонального повторного використання та продовження життєвого циклу ресурсів в освітніх технологіях заохочує повторне застосування віртуальних образів, шаблонів і середовищ у різних навчальних сценаріях, що підвищує віддачу від інвестицій у віртуалізацію та зменшує втрати й дублювання під час розгортання і супроводу освітньої інфраструктури.

Можливості партнерства з індустрією

Стратегічні партнерства між закладами освіти та технологічними компаніями формують нові моделі ІТ-освіти на основі віртуалізації, які забезпечують студентам прямий доступ до інструментів і середовищ, що відповідають вимогам ІТ-індустрії. Такі партнерства надають змогу здобувати досвід виконання реальних проєктів у межах навчальних курсів, а компаніям – виявляти й розвивати таланти через участь в освітніх ініціативах.

Інтеграція професійної сертифікації з віртуалізованими навчальними середовищами дозволяє здобувачам під час навчання отримувати визнання на ринку підтвердження професійних навичок і компетентностей (сертифікати). Така узгодженість академічних досягнень із професійною кваліфікацією спрощує перехід від освіти до працевлаштування та гарантує актуальність змісту відповідно до галузевих стандартів і практик.

Партнерства у сфері досліджень і розробок зосереджені на технологіях освітньої віртуалізації, створюють можливості для розвитку як методик навчання, так і комерційних платформ віртуалізації. Такі спільні проєкти стимулюють інновації на перетині освіти й індустрії та надають студентам досвід роботи з передовими технологіями і процесами розроблення.

Висновки та перспективи подальших досліджень. Інтеграція технологій віртуалізації у підготовку ІТ-фахівців є трансформаційним підходом, який розв'язує ключові проблеми доступності, економічної доцільності та освітньої ефективності. Цей всебічний аналіз показує, що віртуалізація – це не лише інструмент скорочення витрат, а передусім базовий рушій розширеного практичного фахового досвіду, набутого під час навчання, який готує студентів до реалій сучасних ІТ-середовищ.

Наявні дані свідчать: успішне впровадження віртуалізації забезпечує суттєві переваги, зокрема посилення практичної складової навчання, демократизацію доступу до складних технічних середовищ і краще узгодження між академічною підготовкою та вимогами індустрії. Віртуальні лабораторії та хмарні платформи навчання здатні надавати складні й реалістичні практичні навички, які раніше були неможливими або економічно невиправданими в традиційних умовах [2, 1, 5, 6]. Можливість надавати студентам права на адміністрування у віртуальних середовищах, самостійно відтворювати експерименти та працювати зі сценаріями промислового масштабу означає якісний стрибок у способах організації ІТ-освіти.

Водночас упровадження віртуалізації в освіті потребує пильної уваги до низки викликів: обмежень технічної інфраструктури, вимог до підготовки викладачів та підтримання мотивації й залученості студентів. Успіх ініціатив з віртуалізації критично залежить від належного інтернет-з'єднання, розвинених систем технічної підтримки та комплексних програм підвищення кваліфікації викладачів, що охоплюють і технічні компетентності, і педагогічну адаптацію [10, 11]. Потрібно також дбати, щоб накладні витрати на віртуалізацію не знижували освітню ефективність.

Подальший розвиток віртуалізації в ІТ-освіті спрямований на поглиблену інтеграцію з новітніми технологіями: штучним інтелектом, периферійними (edge) обчисленнями та хмарно-нативними архітектурами. Такі рішення покликані зняти нинішні обмеження і водночас відкрити нові можливості для персоналізованого, масштабованого та сталого надання освітніх послуг. Партнерства з індустрією та інтеграція професійної сертифікації ще більше зміцнять зв'язок між академічною підготовкою та професійною практикою, забезпечуючи актуальність і цінність віртуалізованої освіти для студентів і роботодавців.

Очевидний висновок для зацікавлених сторін ІТ-освіти: віртуалізація є водночас можливістю і необхідністю для збереження актуальності освіти у сучасному цифровому професійному середовищі. Зкладам варто інвестувати в надійну технічну інфраструктуру, розвиток викладачів і дієві системи підтримки студентів – як ключові передумови успішного впровадження. Водночас слід зважати на притаманні віртуальному навчанню обмеження й ризики, щоб технологічні новації працювали на реальне підвищення якості освіти, а не ставали самоціллю.

Майбутні дослідження варто зосередити на довготривалому відстеженні результатів студентів у програмах із використанням віртуалізації, порівняльних аналізах різних підходів і платформ, а також пошуку оптимальних стратегій інтеграції новітніх технологій. Окрему увагу слід приділити стратегіям упровадження, орієнтованим на рівність і інклюзивність, щоб максимально реалізувати демократизуючий потенціал цих технологій у різних освітніх контекстах і для різних груп здобувачів.

Список використаних джерел:

1. Usman A. Teaching and Learning Cybersecurity courses with Virtualization Technology: presentation. York St John University, 2021. URL: <https://surl.li/tfmvgn> (дата звернення: 14.08.2025)
2. Gercek G., Saleem N., Steel D. Implementing Cloud Based Virtual Computer Network Labs for Online Education: Experiences from a Phased Approach. *International Journal of Online Engineering*, 2016. Т. 12, № 3. С. 70–76. DOI: 10.3991/ijoe.v12i03.5564
3. Saliah-Hassane H. та ін. Lab@Home: Remote Laboratory Evolution in the Cloud Computing Era. Proceedings of the 2011 ASEE Annual Conference & Exposition (Vancouver, Canada, 26–29 June 2011), 2011. URL: <https://surl.li/uojsrv> (дата звернення: 14.08.2025).
4. Romero-Zaldivar V. A., Pardo P., Burgos D., Delgado K. Monitoring student progress using virtual appliances: A case study. *Computers & Education*, 2012. Т. 58, № 3. С. 1059–1069. DOI: 10.1016/j.compedu.2011.12.003.
5. Hashim M. A. M., Tlemsani I., Matthews R. Higher education strategy in digital transformation. *Education and Information Technologies*, 2022. Т. 27. С. 3171–3195. DOI: 10.1007/s10639-021-10739-1.
6. Vakilian A., Zare Ranjbar E., Hassanipour M., Ahmadinia H., Hasani H. The effectiveness of virtual interactive video in comparison with online classroom in the stroke topic of theoretical neurology in COVID-19 pandemic. *Journal of Education and Health Promotion*, 2022. Т. 11. Ст. 219. DOI: 10.4103/jehp.jehp_1297_21.
7. Grados-Espinoza P. A cross-sectional study to assess the level of satisfaction with virtual education in Peruvian medical students. *Frontiers in Public Health*, 2022. Т. 10. Ст. 1004902. DOI: 10.3389/fpubh.2022.1004902.
8. Rivera-Vargas P., Anderson T., Alonso-Cano C. Exploring students' learning experience in online education: analysis and improvement proposals based on the case of a Spanish open learning university. *Educational Technology Research and Development*, 2021. Т. 69, № 6. С. 3367–3389. DOI: 10.1007/s11423-021-10045-0.
9. Alenezi A. Digital Learning and Digital Institution in Higher Education. *Education Sciences*, 2023. Т. 13, № 1. Ст. 88. DOI: 10.3390/educsci13010088.
10. Bermejo J., Fernández C., Villena J. et al. Best Practices for Developing Extended Reality Experiences in Higher Education. *Information*, 2023. Т. 14, № 1. Ст. 3. DOI: 10.3390/info14010003.
11. Ochoa X., Domínguez A. On the Need for an Open eXtended Reality for Education Framework (OXREF). *The International Review of Research in Open and Distributed Learning*, 2023. Т. 24, № 1. С. 303–316. DOI: 10.19173/irrodl.v24i3.7109.
12. Kandpal M., Sharma M., Nivedita. BS-GeoEduNet 1.0: Blockchain-Assisted Serverless Geo-Education Network. *ISPRS International Journal of Geo-Information*, 2024. Т. 13, № 8. Ст. 274. DOI: 10.3390/ijgi13080274.
13. Shojaei Rad Z., Shafiee S., Taherkhani A. Data Pipeline Approaches in Serverless Computing: A Taxonomy and Review. *Journal of Big Data*, 2024. Т. 11, № 1. Ст. 37. DOI: 10.1186/s40537-024-00939-0.

References:

1. Usman, A. (2021). Teaching and learning cybersecurity courses with virtualization technology [Conference presentation]. York St John University. Retrieved from: <https://surl.li/tfmvgn>
2. Gercek, G., Saleem, N., & Steel, D. (2016). Implementing cloud based virtual computer network labs for online education: Experiences from a phased approach. *International Journal of Online Engineering*, 12(3), 70–76. <https://doi.org/10.3991/ijoe.v12i03.5564>
3. Saliah-Hassane, H. (2011). Lab@Home: Remote laboratory evolution in the cloud computing era. In Proceedings of the 2011 ASEE Annual Conference & Exposition (Vancouver, Canada, 26–29 June 2011). Retrieved from: <https://surl.li/uojsrv>
4. Romero-Zaldivar, V. A., Pardo, P., Burgos, D., & Delgado, K. (2012). Monitoring student progress using virtual appliances: A case study. *Computers & Education*, 58(3), 1059–1069. <https://doi.org/10.1016/j.compedu.2011.12.003>
5. Hashim, M. A. M., Tlemsani, I., & Matthews, R. (2022). Higher education strategy in digital transformation. *Education and Information Technologies*, 27, 3171–3195. <https://doi.org/10.1007/s10639-021-10739-1>
6. Vakilian, A., Zare Ranjbar, E., Hassanipour, M., Ahmadinia, H., & Hasani, H. (2022). The effectiveness of virtual interactive video in comparison with online classroom in the stroke topic of theoretical neurology in COVID-19 pandemic. *Journal of Education and Health Promotion*, 11, 219. https://doi.org/10.4103/jehp.jehp_1297_21
7. Grados-Espinoza, P. (2022). A cross-sectional study to assess the level of satisfaction with virtual education in Peruvian medical students. *Frontiers in Public Health*, 10, 1004902. <https://doi.org/10.3389/fpubh.2022.1004902>
8. Rivera-Vargas, P., Anderson, T., & Alonso-Cano, C. (2021). Exploring students' learning experience in online education: Analysis and improvement proposals based on the case of a Spanish open learning university. *Educational Technology Research and Development*, 69(6), 3367–3389. <https://doi.org/10.1007/s11423-021-10045-0>
9. Alenezi, A. (2023). Digital learning and digital institution in higher education. *Education Sciences*, 13(1), 88. <https://doi.org/10.3390/educsci13010088>
10. Bermejo, J., Fernández, C., Villena, J., et al. (2023). Best practices for developing extended reality experiences in higher education. *Information*, 14(1), 3. <https://doi.org/10.3390/info14010003>

-
11. Ochoa, X., & Domínguez, A. (2023). On the need for an Open eXtended Reality for Education Framework (OXREF). *The International Review of Research in Open and Distributed Learning*, 24(1), 303–316. <https://doi.org/10.19173/irrodl.v24i3.7109>
 12. Kandpal, M., Sharma, M., & Nivedita. (2024). BS-GeoEduNet 1.0: Blockchain-assisted serverless geo-education network. *ISPRS International Journal of Geo-Information*, 13(8), 274. <https://doi.org/10.3390/ijgi13080274>
 13. Shojaee Rad, Z., Shafiee, S., & Taherkhani, A. (2024). Data pipeline approaches in serverless computing: A taxonomy and review. *Journal of Big Data*, 11(1), 37. <https://doi.org/10.1186/s40537-024-00939-0>

Дата надходження статті: 25.10.2025

Дата прийняття статті: 17.11.2025

Опубліковано: 30.12.2025

Гаврилюк Є. А., аспірант кафедри математичного моделювання та аналізу даних Навчально-наукового інституту комп'ютерних наук та штучного інтелекту
Харківського національного університету імені В. Н. Каразіна
ORCID: 0000-0002-4392-2000

НЕЙРОННО-КЕРОВАНІЙ ГІБРИДНИЙ МЕТОД ДЛЯ ОПТИМАЛЬНОГО ПОКРИТТЯ ТЕРИТОРІЇ ІЗ ОБМЕЖЕННЯМИ НА РОЗТАШУВАННЯ

У статті представлено інноваційний підхід до розв'язання задачі максимального покриття території з урахуванням об'єктів довільної форми, їх ротації та наявності заборонених зон для розміщення центрів об'єктів. Ця задача формулюється як нелінійна оптимізаційна проблема, де для забезпечення виконання обмежень застосовується динамічно налаштовувана штрафна функція, керована нейронною мережею. Для прискорення обчислень використовується сурогатна нейронна мережа, яка апроксимує ресурсомістку цільову функцію, що дозволяє значно зменшити час виконання складних геометричних обчислень. Запропонована гібридна стратегія оцінки поєднує точні обчислення за допомогою бібліотеки Shapely для обробки геометричних об'єктів із апроксимаціями Монте-Карло, що забезпечує баланс між точністю та швидкістю. Алгоритми ройового інтелекту, зокрема оптимізація роєм частинок (PSO), та меметичні алгоритми, які поєднують глобальний і локальний пошук, застосовуються для ефективного дослідження багатовимірного простору рішень. Адаптивний механізм штрафів, реалізований через нейронну мережу, дозволяє автоматично налаштовувати параметри обмежень, уникаючи ручного втручання та підвищуючи стійкість методу до змін умов задачі.

Цей підхід демонструє високу ефективність у задачах із складною геометрією, де традиційні методи зазнають труднощів через багатоекстремальний характер цільової функції та обмеження на розміщення. Метод є масштабованим і може бути адаптований до різних типів об'єктів і конфігурацій заборонених зон, що робить його придатним для практичного використання в таких сферах, як телекомунікації (наприклад, оптимальне розташування базових станцій), охорона здоров'я (розміщення мобільних медичних пунктів у кризових ситуаціях), екологія (моніторинг лісів чи природоохоронних зон) та містобудування (планування інфраструктури з урахуванням зон із обмеженим доступом).

Перспективи практичного впровадження охоплюють автоматизацію планування в реальному часі, інтеграцію з системами геоінформаційного аналізу та подальше вдосконалення за рахунок використання глибоких нейронних мереж із активним навчанням для підвищення точності сурогатних моделей. Запропонований підхід відкриває нові можливості для оптимізації складних систем із геометричними обмеженнями, забезпечуючи надійність і ефективність у реальних сценаріях.

Ключові слова: безперервне покриття, довільні форми, ройовий інтелект, нейронні мережі, багатоекстремальна оптимізація, сурогатне моделювання, активне навчання.

Havryliuk Ye. A. Neural-driven hybrid method for optimal area coverage with placement constraints

The article proposes a novel methodology for addressing the Maximum Coverage Location Problem (MCLP) in continuous spaces, incorporating arbitrarily shaped objects, their rotation, and restricted zones for object centers. The problem is formulated as a nonlinear optimization task, where constraints are enforced through a dynamically tuned penalty function driven by a neural network. To enhance computational efficiency, a surrogate neural network approximates the computationally intensive objective function, significantly reducing processing time.

The proposed hybrid evaluation strategy integrates precise geometric computations using the Shapely library with Monte Carlo approximations, achieving an optimal balance between accuracy and computational speed. Swarm intelligence algorithms, such as Particle Swarm Optimization (PSO), and memetic algorithms, which combine global exploration with local refinement, are employed to navigate the high-dimensional, multi-extremal solution space effectively.

The adaptive penalty mechanism, powered by a neural network, enables automatic adjustment of constraint parameters, eliminating the need for manual tuning and enhancing the robustness of the method across varying problem conditions. This approach proves highly effective for complex geometric configurations, where traditional optimization techniques struggle due to the multi-extremal nature of the objective function and placement constraints.

The method is scalable and adaptable to diverse object shapes and restricted zone configurations, making it suitable for practical applications in multiple domains, including telecommunications (e.g., optimal placement of base stations), healthcare (e.g., deployment of mobile medical units in crisis scenarios), ecology (e.g., sensor placement for environmental monitoring), and urban planning (e.g., infrastructure design with restricted access zones).

The integration of neural network-based adaptive penalties with geometric optimization provides a robust and efficient framework for solving real-world coverage problems.

Future prospects include real-time optimization integration with geographic information systems, further refinement through deep neural networks with active learning, and expansion to dynamic environments with time-varying constraints. This approach advances the state-of-the-art in AI-driven optimization, offering a versatile and reliable solution for complex spatial planning challenges.

Key words: continuous coverage, arbitrary shapes, swarm intelligence, neural networks, multi-extremal optimization, surrogate modeling, active learning.

Постановка проблеми. У сучасному світі задачі оптимального розміщення для забезпечення максимального покриття певної території набувають значного поширення в галузях логістики, містобудування, телекомунікацій, екології та систем оборони. Класична задача розміщення з максимальним покриттям (Maximum Coverage Location Problem – MCLP) передбачає розміщення обмеженої кількості об'єктів (наприклад, базових станцій, складів чи сенсорів) з метою максимізації охопленої площі або кількості обслуговуваних точок попиту. Однак реальні умови часто додають складні обмеження, такі як нестандартні форми території, об'єктів покриття та заборонені зони для центрів розміщення.

У досліджуваній задачі є територія заданої форми та розміру, яку потрібно покрити за допомогою множини різних об'єктів покриття з фіксованою формою та розміром. Параметри розміщення включають координати центрів для кожного об'єкта та кут повороту. Метою є максимізація покритої площі території, тобто площі об'єднання трансформованих об'єктів після їх позиціонування та ротації. Додатково вводяться обмеження: центри об'єктів не можуть розташовуватися в заданих заборонених зонах, кількість яких дорівнює k . Важливо, що перетинання об'єктів покриття з забороненими зонами дозволяється, а самі заборонені зони можуть бути покриті, якщо це сприяє загальній максимізації площі.

Новизна цієї постановки полягає в поєднанні безперервного простору, геометричних трансформацій та часткових заборон (лише для центрів об'єктів). Це відрізняє задачу від традиційних дискретних моделей MCLP, де об'єкти розміщуються в фіксованих точках, та від простих геометричних покриттів без гнучкості трансформацій.

Такі обмеження відображають реальні ситуації: при розміщенні антен у телекомунікаціях, коли центри не можуть бути в житлових районах, але сигнал має їх охоплювати; в екології сенсори для моніторингу пожеж уникають заборонених зон, але покривають їх; в кризових ситуаціях, коли мобільні медичні підрозділи розміщуються в безпечних місцях з максимальним охопленням послуг. Розв'язання таких задач оптимізує ресурси, зменшує витрати та підвищує ефективність систем.

Аналіз останніх досліджень і публікацій. Задача розміщення з максимальним покриттям (MCLP), запропонована Черчем і Ревелом [1], еволюціонувала від дискретного розміщення об'єктів для максимізації покриття точок попиту до складних безперервних моделей, що враховують геометрію та реальні обмеження. Ранні огляди, як у Бермана та ін. [2], фіксують цей розвиток, акцентуючи перехід до безперервних просторів, смистових та невизначеностей, що є ключовими для планарної MCLP з забороненими зонами. Моделі безперервного покриття, ініційовані Черчем [3] для планарних задач і розвинені Матізівом та Мюрреєм [4] для одночасних об'єктів, закладають основу для об'єктів на основі площ, але часто припускають ідеалізовані форми або гладкі функції корисності. Натомість реальні сценарії, як телекомунікації, моніторинг довкілля та управління кризами, вимагають гнучкості для нерегулярних форм, заборонених зон та динамічних умов, де метаевристики та методи на основі нейронних мереж виявляються ефективними.

Значні досягнення в безперервному покритті включають роботи Кортеса та ін. [5], які розробили розподілений контроль покриття з використанням розбиття Вороного та градієнтних потоків, пов'язуючи просторову щільність з ефективністю сенсорів у робототехніці та мультиагентних системах. Однак їхній підхід стикається з негладкими цілями та складними об'єднаннями форм, що наша модель розв'язує через метаевристики та сурогатне моделювання. Швагер та ін. [6] розширили ці методи на динамічні середовища, але обчислювальні вузькі місця зберігаються для високорозмірних багатоекстремальних задач. У [7–10] розглядається безперервне покриття з довільними формами за допомогою інструментів обчислювальної геометрії, як Sharply, для точної оцінки та метаевристик для оптимізації, що відповідає потребам практичного застосування в сценаріях з нерегулярною геометрією та забороненими зонами.

Ройовий інтелект та еволюційні алгоритми добре пасують до багатоекстремальних ландшафтів MCLP. Кеннеді та Еберхарт [11] ввели оптимізацію роем частинок (PSO), ціновану за простоту та баланс дослідження-експлуатації, тоді як Сторн і Прайс [12] запропонували диференціальну еволюцію (DE) як стійку альтернативу з мінімальними гіперпараметрами. Оптимізація мурашиними колоніями (ACO), розроблена Доріго та Штютцле [13], вирізняється в комбінаторних підзадачах, як упорядкування об'єктів, і часто вбудовується в меметичні схеми, як досліджено Нері та Котта [14]. Янг [15] надає всебічний синтез алгоритмів, натхнених природою, підкреслюючи їх адаптивність до геометричної оптимізації, тоді як Мірджалілі та ін. [16] розвивають багатоцільові варіанти PSO для складних задач. Меметичні алгоритми, що поєднують глобальний пошук з локальним уточненням, як BFGS, детально описані Моліна та ін. [17], пропонують

рамки для гібридної оптимізації, особливо ефективні для задач з важкою геометрією. Ці методи лежать в основі нашого підходу, що гібридує PSO та меметичні алгоритми з сурогатними нейронними мережами для ефективної навігації високорозмірними просторами.

Обчислювальна вартість геометричних операцій, як об'єднання та перетин для розрахунку площ, вимагає сурогатного моделювання. Традиційні методи, як Крігінг та радіальні базисні функції (Форрестер та ін. [18]), покладаються на адаптивне семплінг, але сучасні тенденції віддають перевагу моделям глибокого навчання. Захір та ін. [19] ввели Deep Sets для інваріантних до перестановок входів, ідеальних для змінних множин об'єктів у задачах покриття. Активне навчання, описане Джін та ін. [20], посилює стійкість сурогатів через періодичні точні оцінки, тоді як фізико-інформовані нейронні мережі (Раїссі та ін. [21]) підтримують багатofідельне навчання для інженерних застосувань. Гудфеллоу та ін. [22] надають фундаментальну рамку для нейронних мереж в оптимізації, а Чжан та ін. [23] підкреслюють їх інтеграцію з метаевристиками для глобального пошуку. Роботи [7, 8] просувають цю сферу, використовуючи інструменти як Sharply для точних обчислень та досягаючи значних прискорень через апроксимацію.

Практичні застосування MCLP охоплюють різноманітні домени. У бездротових сенсорних мережах (WSN) безперервні формулювання оптимізують покриття площі, точок та бар'єрів з обмеженнями на зв'язок та термін служби, як оглянуто Акілдізом та ін. [24]. Планування траєкторій покриття безпілотними апаратами (UAV/UGV/USV), досліджене Лоу та ін. [25], покращує ефективність шляхів для 2D/3D рельєфів, тоді як моніторинг довкілля та точне землеробство вирають від оптимальних розміщень сенсорів над нерегулярними ділянками, як зазначено Чосетом [26]. Промислові інспекції, включаючи фарбування та неруйнівний контроль, використовують роботизоване покриття для довільних форм [27]. У сфері управління кризами [28] пропонується модель покриття для мобільних медичних підрозділів з використанням прогнозованої аналітики для оптимізації розміщення центрів вакцинації чи тестування з обмеженнями на безпеку та доступність, що безпосередньо стосується наших заборонених зон. Аналогічно [29] оцінює надійність сенсорних мереж для моніторингу пожеж, фокусуючись на обмеженнях розміщення та факторах відмов, що наша модель покращує через гнучке оброблення форм та швидку оптимізацію. Безпека, спостереження та реагування на катастрофи також покладаються на максимізацію надмірності сенсорів у складних об'єктах з використанням стратегій орієнтованих на надійність [29, 30].

Інтеграція ройових і меметичних алгоритмів, сурогатних нейронних мереж (стилю Deep Sets з активним навчанням) та опціональних локальних згладжувачів (наприклад, BFGS) робить підхід незалежним від розв'язувача, розширюваним до перешкод, анізотропії та невизначеностей, стійким для чутливих до часу застосувань.

Метою статті є розробка та обґрунтування адаптивного методу на основі нейронних мереж для розв'язання задачі розміщення з максимальним покриттям, що враховує заборонені зони для центрів об'єктів, з метою підвищення ефективності оптимізації в умовах складної геометрії та обмежених ресурсів.

Виклад основного матеріалу. Для розв'язання задачі MCLP з забороненими зонами ми розробили комплексну методологію, що поєднує нелінійну оптимізацію, ройові та меметичні алгоритми, а також сурогатне моделювання на основі нейронних мереж для максимізації покриття з дотриманням просторових обмежень. Метод пристосований до складних геометричних конфігурацій, як нерегулярні багатокутні території та об'єкти покриття довільних форм, що робить його придатним для реальних застосувань, як розміщення мобільних медичних підрозділів та моніторинг довкілля. Наша формулювання базується на попередніх роботах з безперервної оптимізації покриття [7–10], розширюючи їх адаптивними механізмами штрафів та ефективними оцінками. Параметри алгоритмів налаштовано на основі попередніх експериментів з подібними екземплярами MCLP для балансу покриття та обчислювальної ефективності.

Розглядаємо компактну область покриття $\Omega \subset \mathbb{R}^2$, зазвичай нерегулярний багатокутник, що покривається n компактними об'єктами $S_i (i=1, \dots, n)$ із заданою формою. Кожен об'єкт трансформується зсувом до координат $x_i = (x_i^1, x_i^2) \in \mathbb{R}^2$ та поворотом на кут $\theta_i \in [0, 2\pi)$, утворюючи трансформований об'єкт $T_i(S_i, x_i, \theta_i) = \{R(\theta_i)(y - x_i) + x_i \mid y \in S_i\}$, де $R(\theta_i)$ – матриця повороту.

Потрібно максимізувати покриту площу області $\Omega \subset \mathbb{R}^2$, визначену як

$$A(z) = \text{area} \left(\Omega \cap \bigcup_{i=1}^n T_i(S_i, x_i, \theta_i) \right),$$

де $z = (x_1^1, x_1^2, \theta_1, \dots, x_n^1, x_n^2, \theta_n) \in \mathbb{R}^{3n}$ узагальнює змінні розміщення об'єктів $S_i (i=1, \dots, n)$.

Ключове обмеження вимагає, щоб центр кожного об'єкта x_i уникав k заборонених зон $F_j \subset \mathbb{R}^2 (j=1, \dots, k)$, тобто

$$x_i \notin \bigcup_{j=1}^k F_j (j=1, \dots, k).$$

Оптимізаційна задача, що поставлена, належить до класу умовних задач нелінійного програмування і є багатоекстремальною та високорозмірною через геометричні операції та обмеження, вимагаючи стійких

стратегій. Для перетворення умовної задачі оптимізації в безумовну ми застосовуємо метод штрафних функцій.

Порушення для центру кожного об'єкта визначається як

$$V_i(x_i) = \sum_{j=1}^k I(x_i \in F_j),$$

де $I(\cdot)$ – індикаторна функція, що дорівнює 1, якщо $x_i \in F_j$, і 0 – інакше.

Для гладкіших формулювань, ми використовуємо

$$V_i(x_i) = \sum_{j=1}^k \max(0, r_j - \text{dist}(x_i, F_j))^2,$$

де r_j – характерний радіус зони.

Тоді загальний штраф буде визначатися як

$$V(z) = \sum_{i=1}^n V_i(x_i).$$

Покладемо

$$P(z) = \rho V(z),$$

де $\rho > 0$ – штрафний коефіцієнт.

Результуюча цільова функція стане $f(z) = A(z) - \rho V(z)$ і оптимізаційна задача матиме вигляд

$$\max_{z \in \mathbb{R}^{3n}} f(z).$$

За теорією зовнішніх штрафів штрафний коефіцієнт ρ починається з 10 і динамічно зростає $\rho_{k+1} = c\rho_k$ ($c = 10$).

Для уникнення ручного налаштування пропонується розглядаємо ρ як додаткову змінну, визначаючи

$$\zeta = (z, \rho) \in \mathbb{R}^{3n+1},$$

та оптимізуючи

$$f(\zeta) = A(z) - \rho V(z).$$

Нейронна мережа апроксимує залежність від ρ , навчаючись на вибірках $\{\zeta_k, f(\zeta_k)\}$, і прогнозує очікуване покращення

$$EI(\zeta) = \mathbb{E}[\max(0, f_{\text{best}} - \hat{f}(\zeta))],$$

для керування ρ , збільшуючи його при $V(z) > 0.01$ або зменшуючи при застої оптимізації [16]. Цей самоадаптивний штраф мінімізує композитну втрату

$$L = \text{MSE}(f) + \lambda \|\nabla_{\rho} f\|^2,$$

підвищуючи стійкість.

Для подолання багатоекстремального ландшафту застосовуємо гібридну рамку оптимізації, що поєднує ройовий інтелект та меметичні алгоритми, спираючись на їх ефективність у геометричних та високорозмірних задачах [11–15]. PSO, натхненний поведінкою зграй [11], оновлює кандидат-розв'язки (частинки) за допомогою швидкостей, керованих особистими та глобальними найкращими позиціями:

$$v_{k+1} = \omega v_k + c_1 r_1 (p_{\text{best}} - z_k) + c_2 r_2 (g_{\text{best}} - z_k),$$

з подальшим перерахування

$$z_{k+1} = z_k + v_{k+1},$$

де ω (інерція) зменшується від 0,9 до 0,4, а $c_1 = c_2 = 1.5$, $r_1, r_2 \in [0, 1]$.

Розмір рою ($N = 50 - 100$) та ітерації ($T = 500 - 1000$) налаштовано для балансу дослідження та ефективності в межах 5–10 хвилин, з початковими швидкостями обмеженими 10% діапазону пошуку.

Зважаючи на обчислювальну інтенсивність геометричних операцій, як розрахунки площ об'єднань та перетинів, застосовуємо сурогатне моделювання на основі нейронних мереж для прискорення оцінок, техніку, що стає ключовою для оптимізаційних задач [16, 20, 22]. Нейронна мережа стилю Deep Sets, реалізована в PyTorch (версія 1.12), забезпечує інваріантність до перестановок множин об'єктів, приймаючи нормалізовані як вхід і видаючи скалярну апроксимацію $f(\zeta)$.

Архітектура включає вбудовування на об'єкт через 3 повнозв'язні шари (128–256 нейронів, активація ReLU), за яким слідує усереднення та конкатенація з кількістю заборонених зон, і вихідний шар. Навчання на 5000–10000 вибірках $\{\zeta^{(i)}, f(\zeta^{(i)})\}$, згенерованих через латинський гіперкубовий семплінг у $\mathbb{R}^{3n+1}$, з точним, обчисленим методами Монте-Карло чи Shapely [9, 10].

Мережа оптимізується Adam (швидкість навчання 0,001, decay 0,95) у Python 3.8, мінімізуючи $MSE + L1$ втрату за 1000 епох, з розміром батчу 64, dropout 0,2 та L2 регуляризацією (weight decay 0,001). Активне навчання оновлює набір даних кожні 50 ітерацій, обираючи 5–10 точок з високою невизначеністю чи очікуваним покращенням, зменшуючи MSE до $\sim 0,005$ та дозволяючи 80–90% оцінок використовувати швидкі сурогатні прогнози (інференс ~ 1 мс) [19].

Цільова функція $f(\zeta)$ складається з площі покриття $A(z)$ та штрафу $V(z)$ за порушення, оцінюваних гібридним підходом, що поєднує точні та апроксимативні методи. Бібліотека Shapely (версія 2.0) полегшує точні 2D геометричні операції, трансформуючи об'єкти через поворот та зсув, обчислюючи об'єднання з $T_1, T_2, \dots, T_n$, перетинаючи Ω з $A(z)$ для отримання та перевіряючи порушення центрів з $x_i \in F_j$. Зазначимо, що Shapely ресурсомісткий для складних форм. Для подолання цього використовується апроксимація Монте-Карло для ранніх ітерацій (помилка 1–5%) та для фінальної точності ($< 0,1\%$). Порушення обчислюються точно, а багатофідельна стратегія використовує Монте-Карло для дослідження та Shapely для топ-10% кандидатів чи валідації, досягаючи прискорення у 10–50 разів.

Для міри порушення $v(\xi)$, що представляє непокриті площі, застосовуємо дискретизацію Монте-Карло з 2000–20000 точками (адаптивна сітка), тестуючи включення через ray-casting чи функції відстаней, або точні обчислення Shapely для високофідельної перевірки, забезпечуючи стійку оцінку на етапах оптимізації.

Гібридна архітектура оптимізації координує компоненти безшовно. Вона починається з жадібної ініціалізації для апроксимації покриття, за якою слідує глобальне дослідження за допомогою DE чи PSO з сурогатними оцінками. Періодичні точні перевірки уточнюють топ-кандидатів, які проходять локальну оптимізацію CMA-ES, а сурогатна нейронна мережа оновлюється новими даними.

Висновки. Розроблений адаптивний метод на основі нейронних мереж ефективно розв'язує задачу MCLP з забороненими зонами, досягаючи високого рівня покриття в обмежених часових рамках. Меметичний алгоритм перевершує PSO завдяки гібридному механізму глобально-локального пошуку, тоді як сурогатна нейронна мережа забезпечує значне прискорення обчислень, сприяючи практичному впровадженню. Адаптивний механізм штрафів гарантує відсутність порушень обмежень без ручного налаштування, підкреслюючи стійкість самоадаптивної рамки.

Незважаючи на досягнутий рівень, результати вказують на внутрішні обмеження через геометричну складність: об'єкти не можуть ідеально заповнювати нерегулярний багатокутник, призводячи до прогалів чи перетинів, що обмежують покриття нижче теоретичного максимуму. Заборонені зони додатково стримують розміщення, створюючи вузькі місця з мінімальними покращеннями. Ці висновки узгоджуються з NP-складністю геометричних задач покриття та свідчать, що рівні вище 85–90% вимагають більших ресурсів чи альтернативних форм об'єктів.

Список використаних джерел:

1. Church R., ReVelle C. The maximal covering location problem. *Papers of the Regional Science Association*. 1974. Vol. 32, № 1. P. 101–118.
2. Berman O. et al. Facility location: A survey of applications and methods. Springer, 2009. 978 с.
3. Church R. L. The planar maximal covering location problem. *Journal of Regional Science*. 1984. Vol. 24, № 2. P. 185–201.
4. Matisziw T. C., Murray A. T. Continuous space modeling for location problems. *Socio-Economic Planning Sciences*. 2009. Vol. 43, № 2. P. 79–97.
5. Cortés J. et al. Coverage control for mobile sensing networks. *IEEE Transactions on Robotics and Automation*. 2004. Vol. 20, № 2. P. 243–255.
6. Schwager M., Rus D., Slotine J.-J. Decentralized, adaptive coverage control for networked robots. *The International Journal of Robotics Research*. 2009. Vol. 28, № 7. P. 735–752.
7. Yakovlev S. et al. Continuous maximum coverage location problem with arbitrary shape of service areas and regional demand. *Symmetry*. 2025. Vol. 17, № 5. P. 676.
8. Yakovlev S., Kartashov O., Podzheh D. Mathematical models and nonlinear optimization in continuous maximum coverage location problem. *Computation*. 2022. Vol. 10, № 7. P. 119.
9. Yakovlev S., Kartashov O., Mumrienko A. Formalization and solution of the maximum area coverage problem using library Shapely for territory monitoring. *Radioelectronic and Computer Systems*. 2022. № 2. P. 35–48.
10. Yakovlev S. V. The concept of modeling packing and covering problems using modern computational geometry software. *Cybernetics and Systems Analysis*. 2023. Vol. 59, № 1. P. 108–119.
11. Kennedy J., Eberhart R. Particle swarm optimization. *Proceedings of ICNN'95 – International Conference on Neural Networks*. 1995. Vol. 4. P. 1942–1948.

12. Storn R., Price K. Differential evolution – A simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*. 1997. Vol. 11, № 4. P. 341–359.
13. Dorigo M., Stützle T. Ant Colony Optimization. MIT Press, 2004. 319 c.
14. Neri F., Cotta C. Memetic algorithms and memetic computing optimization: A literature review. *Swarm and Evolutionary Computation*. 2012. Vol. 2. P. 1–14.
15. Yang X.-S. Mathematical analysis of nature-inspired algorithms. *Neural Computing and Applications*. 2018. Vol. 30, № 1. P. 113–119.
16. Mirjalili S. et al. Multi-objective grey wolf optimizer: A novel algorithm for multi-criterion optimization. *Expert Systems with Applications*. 2016. Vol. 47. P. 106–119.
17. Molina D., Lozano M., Herrera F. Memetic algorithms for continuous optimisation based on local search chains. *Evolutionary Computation*. 2015. Vol. 23, № 1. P. 1–28.
18. Forrester A., Sobester A., Keane A. Engineering design via surrogate modelling: A practical guide. John Wiley & Sons, 2008. 228 c.
19. Zaheer M. et al. Deep sets. *Advances in Neural Information Processing Systems*. 2017. Vol. 30. P. 3391–3401.
20. Jin Y. et al. Data-driven evolutionary optimization: An overview and case studies. *IEEE Transactions on Evolutionary Computation*. 2019. Vol. 23, № 3. P. 442–458.
21. Raissi M., Perdikaris P., Karniadakis G. E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*. 2019. Vol. 378. P. 686–707.
22. Goodfellow I., Bengio Y., Courville A. Deep learning. MIT Press, 2016. 800 c.
23. Zhang J. et al. Queuing search algorithm: A novel metaheuristic algorithm for solving global optimization problems. *Applied Soft Computing*. 2018. Vol. 67. P. 136–148.
24. Akyildiz I. F. et al. Wireless sensor networks: A survey. *Computer Networks*. 2002. Vol. 38, № 4. P. 393–422.
25. Low K. H., Dolan J. M., Khosla P. Adaptive multi-robot wide-area exploration and mapping. *Autonomous Robots*. 2009. Vol. 27, № 2. P. 129–148.
26. Choset H. Coverage for robotics – A survey of recent results. *Annals of Mathematics and Artificial Intelligence*. 2001. Vol. 31, № 1–4. P. 113–126.
27. Yakovlev S. et al. Maximum service coverage in business site selection using computer geometry software. *Electronics*. 2023. Vol. 12, № 10. P. 2329.
28. Yakovlev S. et al. Optimization of mobile medical service locations based on predictive analytics in crisis scenarios. Proceedings of the International Conferences IADIS Information Systems 2025 and E-Society 2025. 2025. P. 538–541.
29. Leichenko K. et al. Assessment of the reliability of wireless sensor networks for forest fire monitoring systems considering fatal combinations of multiple sensor failures. *Cybernetics and Systems Analysis*. 2025. Vol. 61, № 1. P. 137–147.
30. Skorobohatko S. et al. Architecture and reliability models of hybrid sensor networks for environmental and emergency monitoring systems. *Cybernetics and Systems Analysis*. 2024. Vol. 60, № 2. P. 293–304.

References:

1. Church, R., & ReVelle, C. (1974). The maximal covering location problem. *Papers of the Regional Science Association*, 32(1), 101–118. <https://doi.org/10.1111/j.1435-5597.1974.tb00902.x>
2. Berman, O., et al. (2009). Facility location: A survey of applications and methods. Springer. ISBN: 978-0-387-73294-7.
3. Church, R. L. (1984). The planar maximal covering location problem. *Journal of Regional Science*, 24(2), 185–201. <https://doi.org/10.1111/j.1467-9787.1984.tb01030.x>
4. Matisziw, T. C., & Murray, A. T. (2009). Continuous space modeling for location problems. *Socio-Economic Planning Sciences*, 43(2), 79–97. <https://doi.org/10.1016/j.seps.2008.02.006>
5. Cortés, J., Martínez, S., Karataş, T., & Bullo, F. (2004). Coverage control for mobile sensing networks. *IEEE Transactions on Robotics and Automation*, 20(2), 243–255. <https://doi.org/10.1109/TRA.2004.824698>
6. Schwager, M., Rus, D., & Slotine, J.-J. (2009). Decentralized, adaptive coverage control for networked robots. *The International Journal of Robotics Research*, 28(7), 735–752. <https://doi.org/10.1177/0278364909102795>
7. Yakovlev, S., et al. (2025). Continuous maximum coverage location problem with arbitrary shape of service areas and regional demand. *Symmetry*, 17(5), 676. <https://doi.org/10.3390/sym17050676>
8. Yakovlev, S., Kartashov, O., & Podzheh, D. (2022). Mathematical models and nonlinear optimization in continuous maximum coverage location problem. *Computation*, 10(7), 119. <https://doi.org/10.3390/computation10070119>

-
9. Yakovlev, S., Kartashov, O., & Mumrienko, A. (2022). Formalization and solution of the maximum area coverage problem using library Shapely for territory monitoring. *Radioelectronic and Computer Systems*, 2022(2), 35–48. <http://nti.khai.edu/ojs/index.php/reks/article/view/reks.2022.2.03>
 10. Yakovlev, S. V. (2023). The concept of modeling packing and covering problems using modern computational geometry software. *Cybernetics and Systems Analysis*, 59(1), 108–119. <https://doi.org/10.1007/s10559-023-00547-5>
 11. Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. Proceedings of ICNN'95 – International Conference on Neural Networks, 4, 1942–1948. <https://doi.org/10.1109/ICNN.1995.488968>
 12. Storn, R., & Price, K. (1997). Differential evolution – A simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, 11(4), 341–359. <https://doi.org/10.1023/A:1008202821328>
 13. Dorigo, M., & Stützle, T. (2004). Ant Colony Optimization. MIT Press. <https://doi.org/10.7551/mitpress/4811.001.0001>
 14. Neri, F., & Cotta, C. (2012). Memetic algorithms and memetic computing optimization: A literature review. *Swarm and Evolutionary Computation*, 2, 1–14. <https://doi.org/10.1016/j.swevo.2011.11.003>
 15. Yang, X.-S. (2018). Mathematical analysis of nature-inspired algorithms. *Neural Computing and Applications*, 30(1), 113–119. <https://doi.org/10.1007/s00521-017-3156-1>
 16. Mirjalili, S., Saremi, S., Mirjalili, S. M., & Coelho, L. S. (2016). Multi-objective grey wolf optimizer: A novel algorithm for multi-criterion optimization. *Expert Systems with Applications*, 47, 106–119. <https://doi.org/10.1016/j.eswa.2015.10.039>
 17. Molina, D., Lozano, M., & Herrera, F. (2015). Memetic algorithms for continuous optimisation based on local search chains. *Evolutionary Computation*, 23(1), 1–28. https://doi.org/10.1162/EVCO_a_00124
 18. Forrester, A., Sobester, A., & Keane, A. (2008). Engineering design via surrogate modelling: A practical guide. John Wiley & Sons. <https://doi.org/10.1002/9780470770801>
 19. Zhang, J., Xiao, M., Gao, L., & Pan, Q. (2018). Queuing search algorithm: A novel metaheuristic algorithm for solving global optimization problems. *Applied Soft Computing*, 67, 136–148. <https://doi.org/10.1016/j.asoc.2018.02.038>
 20. Jin, Y., Wang, H., Chugh, T., Guo, D., & Miettinen, K. (2019). Data-driven evolutionary optimization: An overview and case studies. *IEEE Transactions on Evolutionary Computation*, 23(3), 442–458. <https://doi.org/10.1109/TEVC.2018.2869001>
 21. Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378, 686–707. <https://doi.org/10.1016/j.jcp.2018.10.045>
 22. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press. <https://doi.org/10.7551/mitpress/9780262035613.001.0001>
 23. Zaheer, M., et al. (2017). Deep sets. *Advances in Neural Information Processing Systems*, 30, 3391–3401. <https://arxiv.org/abs/1703.06114>
 24. Akyildiz, I. F., Su, W., Sankarasubramaniam, Y., & Cayirci, E. (2002). Wireless sensor networks: A survey. *Computer Networks*, 38(4), 393–422. [https://doi.org/10.1016/S1389-1286\(01\)00302-4](https://doi.org/10.1016/S1389-1286(01)00302-4)
 25. Low, K. H., Dolan, J. M., & Khosla, P. (2009). Adaptive multi-robot wide-area exploration and mapping. *Autonomous Robots*, 27(2), 129–148. <https://doi.org/10.1007/s10514-009-9126-5>
 26. Choset, H. (2001). Coverage for robotics – A survey of recent results. *Annals of Mathematics and Artificial Intelligence*, 31(1–4), 113–126. <https://doi.org/10.1023/A:1016639210559>
 27. Yakovlev, S., Kiseleva, O., Chumachenko, D., & Podzheha, D. (2023). Maximum service coverage in business site selection using computer geometry software. *Electronics*, 12(10), 2329. <https://doi.org/10.3390/electronics12102329>
 28. Yakovlev, S., et al. (2025). Optimization of mobile medical service locations based on predictive analytics in crisis scenarios. Proceedings of the International Conferences IADIS Information Systems 2025 and E-Society 2025, 538–541.
 29. Leichenko, K., et al. (2025). Assessment of the reliability of wireless sensor networks for forest fire monitoring systems considering fatal combinations of multiple sensor failures. *Cybernetics and Systems Analysis*, 61(1), 137–147. <https://doi.org/10.1007/s10559-025-00722-w>
 30. Skorobohatko, S., et al. (2024). Architecture and reliability models of hybrid sensor networks for environmental and emergency monitoring systems. *Cybernetics and Systems Analysis*, 60(2), 293–304. <https://doi.org/10.1007/s10559-024-00670-x>

Дата надходження статті: 24.09.2025

Дата прийняття статті: 10.11.2025

Опубліковано: 30.12.2025

Гайтан О. М., старший викладач кафедри комп'ютерних технологій та інформаційних систем Національного університету «Полтавська політехніка імені Юрія Кондратюка»
ORCID: 0000-0002-7228-9937

Бублій В. Е., асистент кафедри комп'ютерних технологій та інформаційних систем Національного університету «Полтавська політехніка імені Юрія Кондратюка»
ORCID: 0009-0002-7748-7681

ТЕХНОЛОГІЇ ШТУЧНОГО ІНТЕЛЕКТУ В СИСТЕМАХ РЕКРУТИНГУ ТА УПРАВЛІННЯ ПЕРСОНАЛОМ

У статті розглянуто сучасні підходи до використання штучного інтелекту (ШІ) у процесах рекрутингу та розвитку персоналу в умовах цифрової трансформації бізнесу. Проаналізовано теоретичні засади й практичні інструменти автоматизації HR-процесів, інтеграцію інтелектуальних технологій у процеси підбору, оцінювання та утримання працівників. Особливу увагу приділено використанню технологій *augmented writing* для формування описів вакансій, що підвищують релевантність і гендерну нейтральність текстів, системам автоматизованого сканування профілів кандидатів у соціальних мережах, інтелектуальним сканерам резюме, чат-ботам та системам автоматизованого вхідного тестування. Детально розглянуто використання алгоритмів машинного навчання та обробки природної мови, зокрема моделей BERT, GPT, T5, що забезпечують семантичний аналіз і класифікацію профілів кандидатів, а також моделі прогнозування успішності працівників. Проаналізовано типові етапи обробки даних, створення ембеддінгів (*vector embeddings*) профілів і застосування векторних баз даних для пошуку релевантних кандидатів, зокрема використання Milvus, Pinecone та Weaviate для реалізації семантичного пошуку. Розглянуто механізми інтеграції таких моделей у системи автоматизованого підбору персоналу, оцінювання результативності та прогнозування професійної успішності працівників. Проаналізовано метрики оцінювання ефективності (*accuracy, precision, recall, F1-score*), які використовуються для кількісного вимірювання результативності у HR-аналітиці. Доведено, що інтеграція інтелектуальних технологій сприяє скороченню часу підбору персоналу на 40–60 %, підвищенню точності відбору на 25–30 % і зниженню рівня упередженості у прийнятті рішень. Окремлено етичні ризики застосування штучного інтелекту – алгоритмічну дискримінацію, непрозорість моделей, ризики порушення приватності та обмеження даних. Запропоновано підходи до забезпечення прозорості й справедливості систем на основі *Explainable AI* та принципів GDPR. Результати дослідження мають практичне значення для розробників HR-технологій, аналітиків і менеджерів з персоналу, що впроваджують системи інтелектуальної автоматизації у рекрутингові процеси.

Ключові слова: штучний інтелект, рекрутинг, машинне навчання, HR-аналітика, *augmented writing*, чат-боти, метрики ефективності.

Haitan O. M., Bublii V. E. Artificial intelligence technologies in recruiting and staff management systems

The article examines modern approaches to the use of artificial intelligence (AI) in recruitment and staff development processes under the conditions of business digital transformation. Theoretical foundations and practical tools for the automation of HR processes are analyzed, focusing on the integration of intelligent technologies into recruitment, evaluation, and employee retention systems. Special attention is paid to augmented writing technologies for generating job descriptions that enhance linguistic relevance and gender neutrality, as well as to automated candidate profiling in social networks, intelligent resume scanners, chatbots, and AI-based preliminary testing systems. The study provides a detailed analysis of machine learning and natural language processing algorithms, including BERT, GPT, and T5 models, which enable semantic analysis and classification of candidate profiles, as well as predictive models for assessing employee success. The typical stages of data processing are examined, including the creation of embeddings (*vector representations*) of candidate profiles and the use of vector databases for semantic matching, particularly Milvus, Pinecone, and Weaviate, which ensure context-aware candidate search and ranking. Mechanisms for integrating such models into automated recruitment systems, performance evaluation tools, and employee success prediction frameworks are discussed. The paper analyzes performance metrics such as *accuracy, precision, recall, and F1-score*, used for quantitative measurement of HR analytics effectiveness. It is demonstrated that the integration of intelligent technologies can reduce recruitment time by 40–60%, increase selection accuracy by 25–30%, and lower bias in decision-making processes. Ethical challenges associated with AI use are also outlined, including algorithmic discrimination,

model opacity, data privacy risks, and data access limitations. The study proposes approaches to ensuring transparency and fairness of AI-driven systems through Explainable AI methods and compliance with GDPR principles. The findings have practical significance for HR technology developers, analysts, and resource managers implementing intelligent automation systems in recruitment and staff management processes.

Key words: artificial intelligence, recruiting, machine learning, HR analytics, augmented writing, chatbots, performance metrics.

Постановка проблеми. Цифрова трансформація бізнесу та швидкий розвиток штучного інтелекту кардинально змінюють підходи до управління людськими ресурсами. Штучний інтелект інтегрується в усі етапи життєвого циклу працівника: від рекрутингу та адаптації до побудови індивідуальних кар'єрних траєкторій і підвищення кваліфікації.

Проблема полягає у необхідності системного дослідження впливу штучного інтелекту на процеси рекрутингу та розвитку персоналу, оцінювання ефективності відповідних технологічних інструментів і формування методичних підходів до їх етичного та відповідального використання.

Аналіз досліджень та публікацій. Впровадження штучного інтелекту у сфері рекрутингу досліджується у працях вітчизняних і зарубіжних науковців, а також у звітах міжнародних аналітичних компаній. За даними аналітичних джерел (World Economic Forum, Gartner), 87% компаній вже використовують ШІ у підборі кадрів, а серед корпорацій зі списку Fortune 500 цей показник сягає 99 %. У той же час 65 % рекрутерів застосовують ШІ-інструменти для попереднього аналізу резюме та відбору кандидатів або прогнозування успішності працівників [1, 2].

За даними Gartner, частка HR-лідерів, які впроваджують або активно планують використовувати генеративний ШІ, зросла з 19% у червні 2023 року до 61% у січні 2025 року, що відображає стрімке зростання використання інтелектуальних технологій у сфері управління персоналом. Водночас лише 21% СІО приділяють увагу пом'якшенню впливу ШІ на працівників [2]. 92% компаній планують збільшити інвестиції у ШІ протягом трьох років, проте лише 1% вважають себе «зрілими» щодо повної інтеграції ШІ у бізнес-процеси [3].

Скібська К. провела SWOT-аналіз використання інструментів штучного інтелекту для ідентифікації сильних та слабких сторін його використання. Було виявлено, що використання цифрових інструментів пришвидшує процес онлайн рекрутингу на 75% [4]. За даними звіту Deloitte 2024, впровадження систем ШІ сприяє підвищенню рівня збереження персоналу у середньому на 40% [5].

Bankins S. та ін. [6] провели систематичний аналіз 104 емпіричних досліджень, який підтвердив, що штучний інтелект здатний ефективно прогнозувати необхідні навички на різних етапах кар'єрного розвитку. Він враховує як індивідуальні цілі працівників, так і організаційні потреби, що сприяє більш точному та персоналізованому підходу до побудови кар'єри.

Дослідження ІТ-компаній у регіоні Бангалор, проведене з використанням методу структурного моделювання [7] та AMOS 22.0 для аналізу даних, показало, що штучний інтелект має значний вплив на продуктивність працівників. Результати показали, що 32.8% варіації продуктивності працівників можна пояснити впливом штучного інтелекту. Отже, впровадження інтелектуальних систем у HR-процеси сприяє підвищенню ефективності персоналу. Звіти IBM і McKinsey підтверджують, що використання ШІ скорочує час рекрутингу на 40% і підвищує точність відбору персоналу на 25–30 %.

Технологічні аспекти використання ШІ у HR-сфері розглядаються у роботах багатьох вчених та дослідженнях організацій. Міжнародний інститут розвитку менеджменту (IMD) досліджує використання ШІ-програмних засобів в рекрутингу: IBM Watson, ChatGPT, HireVue, Eightfold.ai, Visier, Pymetrics, ADP DataCloud, Workday, HiredScore, Ultimate Software's UltiPro [8]. Yuan Li et al., Mashayekhi Y. et al. надають систематичний огляд алгоритмів рекомендацій у електронному рекрутингу з аналізом основних методів – колаборативної, контентної та гібридної фільтрації [9]. Черненко Н. досліджує основні технологічні можливості сучасних ШІ-платформ, які використовуються в HR, перспективи їх розвитку [10], Васьків О., Боршук І. [11], Славкова О. [12], Матвєєва Н. [13] – впровадження чат-ботів і аналітичних платформ, що дозволяють оптимізувати комунікацію між рекрутерами та кандидатами.

Питання етичності впровадження ШІ розглядають Черненко Н., Пуліна Т., Шморгун О. та ін., звертаючи увагу на ризики дискримінації, упередженості даних і необхідність підвищення прозорості моделей прийняття рішень.

Аналіз сучасних джерел свідчить про активне поширення практик використання штучного інтелекту у сфері рекрутингу та розвитку персоналу. При цьому актуальними залишаються завдання забезпечення точності моделей, формування етичних стандартів їх застосування та розроблення підходів до вимірювання ефективності інтелектуальних HR-систем.

Мета дослідження. Метою дослідження є систематизація технологій штучного інтелекту, що застосовуються на різних етапах життєвого циклу працівника, аналіз їх технічних реалізацій, метрик ефективності та етичних аспектів використання у рекрутингу, оцінюванні та розвитку персоналу.

Основний матеріал.

1. Використання технологій ШІ у написанні описів вакансій. Роль технологій штучного інтелекту у процесах підбору персоналу постійно зростає. Перший етап ефективного підбору персоналу – складання якісного та привабливого опису вакансії. Від якості тексту вакансії значною мірою залежить кількість і релевантність кандидатів, які відгукуються на пропозицію [14]. Зокрема Somers S. та Desmidt S. зазначають, що використання образної мови та текст вакансії, який більш конкретно описує позицію в оголошенні, підвищують привабливість організації, що веде до залучення більшої кількості потенційних кандидатів [15]. Інші дослідження підкреслюють, що специфікація вимог і позитивне позиціонування роботодавця в тексті значно підвищують не лише обсяг, а й релевантність відгуків, оскільки кандидати краще розуміють очікування і можуть адекватніше оцінити свою відповідність [16, 17].

Існує активна практика й дослідження застосування *augmented-writing* (технології підсиленого письма) для генерації та оптимізації текстів вакансій. Штучний інтелект допомагає створювати чіткі та орієнтовані на конкретну аудиторію тексти вакансій. Використання алгоритмів обробки природної мови (NLP) дає змогу аналізувати великі обсяги історичних даних (попередні вакансії, профілі успішних співробітників, попередні оголошення, статистику відгуків та ефективності найму тощо), щоб виявити закономірності між формулюванням вакансій та показниками успішності найнятих працівників [18]. На основі цього формуються рекомендації щодо ключових слів, структури тексту та тону комунікації і створюються тексти, що залучають більш різноманітну аудиторію. Наприклад, платформа Textio використовує NLP для аналізу ефективності формулювань та *augmented writing* для аналізу та оптимізації опису вакансій, пропонуючи покращення лексики й тону, які підвищують привабливість вакансій для ширшого кола кандидатів. У звіті Textio за 2022 рік зазначається, що такий аналіз дозволяє ідентифікувати та зменшувати мовні упередження, зокрема використання гендерно-забарвлених слів [19, 20]. Хе Дж.К. та Канг С.К. стверджують, що зменшення упередженості оголошень про роботу шляхом заміни “чоловічої мови” збільшує гендерну різноманітність кандидатів [17].

Дослідження показують, що використання ШІ при створенні вакансій скорочує час підготовки опису на 50-60 %, час на читання резюме – на 40-60%, відсіювання 35% нерелевантних заявок [21].

Vu D.H. та ін. запропонували метод NLP2FJD, який базується на технологіях глибокого навчання та NLP для підвищення точності виявлення фіктивних описів вакансій [22]. Модель використовує попередньо натреновані вектори Word2Vec для вилучення семантичних ознак із текстових даних і поєднує їх із метаданими для підвищення ефективності класифікації.

Процес створення описів вакансій із застосуванням штучного інтелекту реалізується у декілька етапів:

1. Збір даних – описів вакансій, резюме успішних працівників, оцінок продуктивності, відгуків кандидатів, показники успішності працівників (KPI, тривалість роботи, оцінки продуктивності), метрики відгуку на вакансію (CTR, заявки/дня), демографічні дані (з анонімізацією). Забезпечується інтеграція з системами ATS (Applicant Tracking System) і HRIS (Human Resources Information System) для отримання історичних даних про вакансії та результати найму.

Джерела даних: ATS/CRM (Workday, Greenhouse, Lever), історичні описи вакансій, HRIS (анонімізовані KPI), платформи вакансій (LinkedIn).

2. Попередня обробка текстів. За допомогою бібліотек NLP (spaCy, NLTK, Stanza) здійснюється токенізація, нормалізація та семантичний аналіз текстів.

3. Аналітичний модуль. За допомогою методів машинного навчання (decision trees, random forests, gradient boosting) і моделей на основі трансформерів (BERT, RoBERTa, GPT) визначаються ключові навички, вплив певних формулювань на поведінку кандидатів, досвід та риси успішних співробітників. Класифікаційні/регресійні моделі виконують аналіз факторів успіху, виявляють характеристики співробітників, що корелюють із успіхом; ці ознаки використовуються для визначення пріоритетів в описів вакансій.

Для навчання моделей попередньо виконується експертна розмітка. HR-експерти маркують ключові навички, категорії ролей, кар’єрні рівні; створюють «золоті» приклади опису вакансії для навчання/оцінки.

4. Генерація тексту. Великі мовні моделі LLM (GPT-4, Llama 3, T5, BART) створюють варіанти описів вакансій, адаптовані до заданого контексту, рівня кваліфікації та вимог компанії, на основі принципів гендерної нейтральності, стислості та зрозумілості. Генерація або підсилення тексту виконується двома шляхами: (а) шаблони + правила + статистичні підстановки; (б) модель на основі трансформерів (донавчена модель BERT, T5, GPT) або GPT-подібні LLM з промпт-інженерією для генерації природних, інклюзивних оголошень про вакансії. Комерційні продукти часто комбінують обидва підходи (дані + LLM).

Таблиця 1

Технології ШІ у написанні описів вакансій

Компонент	Опис	Технології / Інструменти
Рівень даних	Збір даних із ATS, HRIS, LinkedIn	PostgreSQL, S3, Airflow
NLP обробка	Токенізація, NER, нормалізація	spaCy, Stanza, NLTK
ML / LLM моделі	Генерація описів	BERT, T5, GPT-4, Llama 3
Розгортання	Сервісна архітектура	FastAPI, Docker, Kubernetes

5. Валідація. Результати оцінюються на основі лінгвістичних метрик (читабельність, унікальність, тональність) та бізнес-метрик (кількість відгуків, конверсія в успішні найми). Метрики дозволяють вимірювати як якість тексту, так і його бізнес-ефективність.

Таблиця 2

Метрики ефективності підсиленого письма

Метрика	Опис	Тип
Readability Index	Оцінка читабельності тексту	Лінгвістична
Diversity Ratio	Частка різноманітних кандидатів	Бізнес
Conversion-to-Hire	Конверсія відгуку → найм	Бізнес
Оцінки BLEU / ROUGE	Семантична схожість із еталоном	Модельна

Важливим аспектом залишається етична складова використання ІШ в написанні вакансій. Некоректно сформульовані вимоги або упереджені формулювання можуть відштовхнути певні групи кандидатів і знизити інклюзивність процесу найму [23]. Тому сучасні системи рекрутингу інтегрують модулі для виявлення дискримінаційних формулювань, пов'язаних із гендером, віком або соціальним статусом, для забезпечення рівності. OECD наголошує, що регулярний аудит упередженості є ключем до справедливих результатів [24].

Таблиця 3

Етичні обмеження та ризики формування опису вакансії

Обмеження / Ризики	Потенційні наслідки	Методи пом'якшення
Алгоритмічна упередженість	Виключення певних груп	Тестування упередженості, контрфактичне оцінювання
Недостатність даних	Низька точність	Розширення даних, експертна оцінка
Витік даних	Порушення GDPR	Анонімізація, шифрування

Найефективнішими вважаються гібридні підходи, які поєднують великі обсяги даних, експертну розмітку, використання великих мовних моделей з контрольованими параметрами стилю та тональності та безперервну А/В-валідацію результатів. Такий підхід дозволяє досягти оптимального балансу між автоматизацією та якістю контенту та забезпечити відповідність текстів лінгвістичним і етичним стандартам.

2. Сканування профілів кандидатів у соціальних мережах. Другим етапом використання ІШ у рекрутингу є автоматизоване виявлення кандидатів у соціальних мережах, що дозволяє охопити фахівців, які не подають заявки безпосередньо. Технології ІШ дозволяють здійснювати глибинний аналіз відкритих профілів у соціальних мережах (LinkedIn, GitHub, X) для виявлення пасивних потенційних кандидатів. Алгоритми семантичного пошуку на векторних поданнях (ембедінгах) і кластеризації допомагають ідентифікувати користувачів з відповідними компетенціями.

Смачило П. та Журавчак Л. виділяють три ключові підходи до аналізу даних із соціальних мереж: аналіз тональності текстів, тематичне моделювання та аналіз соціальних зв'язків [25]. Такі підходи дозволяють формувати рекомендаційні системи для підбору кадрів на основі публічних даних LinkedIn.

Понад 75 % кваліфікованих спеціалістів належать до пасивних кандидатів, які відкриті до нових можливостей, якщо вакансія відповідає їхнім очікуванням. 83% опитаних рекрутерів зосереджуються на пасивних кандидатах. Саме тому автоматизований моніторинг і аналітика профілів у соціальних мережах стають критично важливими для ефективного підбору персоналу (LinkedIn, 2023).

Аналіз профілів складається з таких етапів:

1. **Збір даних.** Використовуються програмні агенти для збору відкритої інформації з платформ LinkedIn, GitHub, Behance, ResearchGate, Kaggle, X. Дані включають освіту, досвід роботи, ключові навички, досягнення, активність у професійних спільнотах тощо. Для підвищення якості вибірки застосовується фільтрація даних за регіоном, галуззю, посадою та рівнем кваліфікації.

2. **Обробка та нормалізація даних.** Використовуються NLP-бібліотеки (spaCy, NLTK, Stanza) для токенизації, очищення тексту, виявлення сутностей (NER) і стандартизації назв посад та компаній. Для зменшення дублювання застосовуються евристичні та кластеризаційні алгоритми (K-Means, HDBSCAN).

3. **Створення профільних векторів і семантичний пошук.** За допомогою моделей на основі трансформерів (Sentence-BERT, OpenAI embeddings, Llama 3) формуються векторні представлення профілів, які дозволяють проводити семантичне зіставлення між вакансією та профілем кандидата. Для швидкого пошуку релевантних профілів використовуються векторні бази даних Milvus, Pinecone або Weaviate.

4. **Аналіз соціальних зв'язків.** Використовується ранжування на основі графів (наприклад, алгоритм PageRank або Node2Vec) для аналізу взаємозв'язків між кандидатами, компаніями та галузями. Це дозволяє визначати «впливових» кандидатів у певних професійних колах або знайти рекомендаційні ланцюжки.

5. **Прогнозування поведінки кандидата.** На основі історичних даних (відгуки, час відповіді, взаємодії з рекрутером) будується модель ймовірності відгуку. Згідно з даними IBM HR Analytics (2024), такі моделі можуть підвищити точність таргетованих контактів на 37%, скоротити «time-to-contact» на 42% [26].

Таблиця 4

Технології аналізу профілів кандидатів у соціальних мережах

Компонент	Опис	Технології / Інструменти
Збирання даних	Збір відкритих даних із соціальних платформ	Python, Selenium, BeautifulSoup, API
Обробка природної мови (NLP)	Аналіз текстів профілів, виявлення сутностей	spaCy, Transformers, NLTK
Семантичний пошук і зіставлення	Пошук релевантних кандидатів за змістом	Sentence-BERT, OpenAI embeddings, Milvus, Pinecone
Графова аналітика	Аналіз зв'язків і впливу	NetworkX, Neo4j, Node2Vec
Розгортання та інтеграція	Інтеграція з ATS/CRM системами	FastAPI, Airflow, AWS S3, REST API

Оцінка роботи систем III для пошуку кандидатів проводиться за сукупністю кількісних і якісних показників.

Таблиця 5

Метрики ефективності пошуку кандидатів

Метрика	Опис	Тип
Precision@K	Частка релевантних профілів серед топ-K результатів	Модельна
Recall	Частка знайдених релевантних профілів від загальної кількості	Модельна
CTR (Click-Through Rate)	Частка переходів на профіль після контакту	Бізнес
Conversion-to-Interview	Конверсія у призначене інтерв'ю	Бізнес
Time-to-Contact	Середній час від ідентифікації до контакту	Операційна

Згідно зі звітом Gartner (2024), близько 55 % компаній зі списку Fortune 500 уже інтегрували такі інструменти у свої HR-системи. За даними Deloitte (2024), автоматизовані системи III скорочують час пошуку релевантних кандидатів у середньому на 43%, а precision@10 у провідних рішеннях досягає 0.78 [27]. За даними GoIT Global (2025), семантичний пошук, що дозволяє знаходити релевантні профілі кандидатів на основі їхніх навичок, а не лише назв посад, підвищує рівень відгуків приблизно на 25%. Крім того, застосування парсерів природної мови дає змогу скоротити час розгляду кандидатур до 60% [28].

Сканування профілів кандидатів пов'язане з рядом етичних і правових викликів. Ризики виникають при використанні публічної інформації без чіткого повідомлення користувачів, а також під час класифікації за ознаками, що можуть мати соціальну чутливість (вік, стать, регіон, мова спілкування тощо).

Основними принципами є дотримання GDPR, CCPA щодо збору та збереження даних; дотримання OECD AI Guidelines; використання тільки публічної інформації з профілів; анонімізація персональних даних у системах навчання моделей; пояснюваність алгоритмів для запобігання дискримінації.

Таблиця 6

Етичні ризики при аналізі профілів кандидатів

Обмеження / Ризики	Потенційні наслідки	Методи пом'якшення
Порушення приватності	Використання особистих даних без згоди	Збір лише публічної інформації, анонімізація, дотримання GDPR/CCPA
Алгоритмічна упередженість	Нерівність у відборі кандидатів за ознаками статі, віку чи мови	Аудит справедливості, зменшення упереджень, перевірка балансування даних
Відсутність пояснюваності	Недовіра до результатів автоматизованого відбору	Пояснюваний штучний інтелект, прозорість критеріїв відбору
Невірна інтерпретація даних	Хибні оцінки або віднесення до невірної категорії	Верифікація профілю, участь людини у циклі прийняття рішень (human-in-the-loop)
Викривлення соціальних зв'язків	Надмірне використання graph-аналізу може порушувати приватність третіх осіб	Обмеження глибини обходу графа, деідентифікація вузлів

Дослідження OECD (2023) підкреслює, що понад 60% компаній, які застосовують AI-аналітику для рекрутингу, впровадили внутрішні політики щодо ведення аудиторського журналу та зменшення упереженості.

3. Автоматизований відбір резюме. Автоматизований відбір резюме є одним основних напрямів застосування ШІ у рекрутингу. Традиційно цей процес потребував значних часових ресурсів HR-фахівців, які вручну переглядали сотні заявок. ШІ-рішення дозволяють автоматично обробляти текстові документи, виявляючи релевантних кандидатів. Дослідження IBM (2022) доводять, що використання систем машинного аналізу резюме скорочує час на попередній відбір до 75% [29]. ШІ дозволяє аналізувати великі обсяги резюме, оцінюючи відповідність досвіду та навичок вимогам.

Основним технологічним інструментом є обробка природної мови (NLP), що дає змогу системі розпізнавати ключові слова, контекст і смислові зв'язки між досвідом кандидата та описом вакансії. Методологія передбачає парсинг документів, вилучення ключових полів (навички, освіта, досвід), нормалізацію назв і семантичне порівняння з описом вакансії. Основні етапи включають:

1. Парсинг і структуризація даних. Здійснюється автоматичне зчитування інформації з файлів різних форматів (PDF, DOCX) з подальшим перетворенням неструктурованого тексту у впорядковані дані. Система виокремлює ключові поля – особисті дані, освіту, досвід роботи, навички, сертифікати тощо.

2. Векторизація та семантичне представлення тексту. Для перетворення текстової інформації у числове представлення використовуються векторизаційні методи (Word2Vec, Sentence Transformers), які формують багатовимірні вектори, що відображають семантичну близькість між поняттями.

3. Контекстуальне моделювання та аналіз змісту. На цьому етапі застосовуються попередньо навчені мовні моделі – BERT, RoBERTa, DistilBERT, а також, які дають змогу відображати текстові елементи у багатовимірному семантичному просторі та виявляти смислові зв'язки між поняттями. Використання цих моделей забезпечує точніше вилучення релевантної інформації про навички, досвід і компетенції кандидата.

4. Оцінювання релевантності кандидатів. На цьому етапі використовуються тонко налаштовані моделі BERT, RoBERTa або GPT, які здійснюють класифікацію резюме за ступенем відповідності опису вакансії.

Додатково використовуються системи виявлення неявних ознак компетентності. Вони визначають приховані навички та поведінкові характеристики (наприклад, лідерство, комунікабельність, аналітичне мислення) на основі контексту опису досвіду чи досягнень. Для цього можуть застосовуватись моделі LLM або спеціалізовані класифікатори м'яких навичок.

Комбінування текстових і метаданих (наприклад, дата останньої роботи, освіта, галузь, тип контракту) дає змогу побудувати гібридні моделі, що враховують як зміст, так і структуру резюме.

Інтеграція з ATS (Workday, Greenhouse) здійснюється через REST API.

Таблиця 7

Технології автоматизованого аналізу та скринінгу резюме

Компонент	Опис	Технології / Інструменти
Попередня обробка тексту	Парсинг PDF/DOCX, очищення тексту, токенизація	spaCy, NLTK, regex, Apache Tika
Векторизація тексту	Формування семантичних векторів для резюме та вакансій	Word2Vec, Sentence-BERT, BERT, RoBERTa, DistilBERT
Семантичний пошук	Пошук релевантних кандидатів за векторним представленням	ElasticSearch, FAISS, Pinecone
Ранжування / Оцінка релевантності	Класифікація та ранжування кандидатів за відповідністю вакансії	cosine similarity, XGBoost, fine-tuned BERT / RoBERTa / GPT
Виявлення неявних компетенцій	Аналіз прихованих навичок та поведінкових характеристик	fine-tuned GPT / LLM, класифікатори soft skills
Інтеграція з HR/ATS системами	Синхронізація даних, автоматична передача результатів	Greenhouse API, BambooHR API, Workday REST API
DevOps / Підтримка моделей	Розгортання моделей, відстеження експериментів, масштабованість	FastAPI, MLflow, Docker, PostgreSQL

Згідно з дослідженням LinkedIn, автоматизація скринінгу скорочує час на обробку заявок у середньому на 70–80 %, а точність відповідності кандидатів вимогам вакансії зростає до 85 % [30]. Крім того, використання багатомовних моделей (Multilingual BERT) підвищує точність обробки міжнародних резюме.

Оцінка ефективності проводиться за метриками accuracy, F1-score, NDCG та бізнес-показниками (скорочення часу перегляду, підвищення точності відбору). Застосовуються тести на справедливість системи.

Таблиця 8

Метрики ефективності систем скринінгу

Метрика	Опис	Тип
Accuracy	Загальна точність класифікації резюме за релевантністю	Модельна
F1-score	Гармонійне середнє між precision і recall – збалансована оцінка моделі	Модельна
NDCG (Normalized Discounted Cumulative Gain)	Показник якості ранжування кандидатів у пошуковій видачі	Модельна
Disparate Impact Ratio	Відношення рівня позитивних рішень між групами (тест на дискримінацію)	Етична
Equality of Opportunity	Перевірка рівності шансів для кандидатів різних груп	Етична
Processing Time	Середній час обробки одного резюме	Операційна
Shortlist Precision	Частка кандидатів, які дійшли до етапу інтерв'ю	Бізнес
Time-to-Hire Reduction	Зменшення тривалості процесу найму після впровадження системи	Бізнес

Етичні аспекти зосереджені на мінімізації алгоритмічних упереджень, пояснюваності моделей і прозорості процесу прийняття рішень. Використання автоматизованих систем відбору регулюється вимогами щодо недискримінаційного доступу до зайнятості (EU AI Act, 2024).

Моделі, які прогнозують успішність кандидатів, несуть ризик підсилення історичних упереджень, закладених у навчальні дані. Це підкреслює важливість розвитку механізмів пояснюваності та пом'якшення упередженості для забезпечення прозорості й справедливості алгоритмічних рішень.

Таблиця 9

Етичні ризики при автоматизованому відборі резюме

Ризик	Потенційний наслідок	Механізм контролю
Упередженість моделі	Виключення кваліфікованих кандидатів	Анонімізація даних, аудит справедливості моделі
Неповні або некоректні дані	Помилкове ранжування	Перевірка достовірності полів
Низька якість текстів резюме	Втрата релевантної інформації	Використання NER і корекції контексту

У великих компаніях для обробки резюме часто застосовуються системи відстеження кандидатів (ATS), які, аналізуючи дані про попередні найми, формують профілі «ідеальних» кандидатів для компанії. Однак така практика може відтворювати наявні упередження, якщо попередні рішення рекрутерів були необ'єктивні. Показовим є випадок Amazon, де штучний інтелект, що використовувався для добору персоналу, виявив упередження проти жінок [31].

Дослідження Вашингтонського університету показало, що ШІ-інструменти відбору кандидатів часто проявляють упередження: 85% випадків – на користь імен, асоційованих із білою расою, 52% – із чоловічою статтю. Найбільше дискримінаційне відставання спостерігається серед темношкірих чоловіків [32].

4. Персоналізовані взаємодії з кандидатами в чат-ботах. Чат-боти на базі штучного інтелекту змінюють підхід до комунікації з кандидатами, забезпечуючи цілодобову підтримку, швидке надання інформації та персоналізовану взаємодію [33]. За даними Deloitte (2023), використання чат-ботів у рекрутингу підвищує рівень задоволеності кандидатів на 35% [34].

Робота чат-бота складається з модулів: розпізнавання намірів (intent detection), витяг сутностей (entity recognition), управління діалогом (dialogue manager) та генерація відповідей (response generation).

Процес побудови чат-бота для HR-комунікацій складається з таких етапів:

1. **Аналіз сценаріїв спілкування.** Визначаються ключові етапи взаємодії – реєстрація, заповнення анкети, співбесіда, зворотний зв'язок.

2. **Навчання моделі.** Використовуються попередньо навчені мовні моделі (GPT-4, Claude, Gemini) із доменною адаптацією до HR-комунікацій.

3. **Інтеграція з Telegram, Messenger та корпоративними порталами**

4. **Інтеграція з ATS/CRM.** Чат-бот отримує дані з профілів кандидатів і вакансій для контекстної персоналізації.

Для персоналізації використовується контекст користувача, історія взаємодій і дані з ATS. Для генерації природних відповідей застосовуються моделі GPT-4 або Llama 3. Для забезпечення точності та гнучкості використовується комбінація логіки на основі правил і LLM.

Таблиця 10

Технології побудови чат-бота для HR-комунікацій

Компонент	Опис	Технології / Інструменти
NLP & та визначення намірів	Аналіз повідомлень користувачів, розпізнавання намірів і сутностей	GPT-4 API, Llama 3, Dialogflow, spaCy, Transformers
Управління діалогом	Побудова сценаріїв спілкування, управління діалогами	Rasa Core, FastAPI, LangChain
Генерація відповідей	Формування природних персоналізованих відповідей	GPT-4 API, Claude, Llama 3
Планування та організація інтерв'ю	Автоматичне планування зустрічей і співбесід	Google Calendar API, Calendly
Аналіз настроїв	Аналіз емоційного тону спілкування	MonkeyLearn, IBM Watson Tone Analyzer
Інтеграційний рівень	Зв'язок із HR-системами, CRM, календарями	HRFlow.ai, Webhooks, REST API
Аналітика та моніторинг	Збір метрик взаємодії, журналювання, аналітика	PostgreSQL, Grafana, MLflow
Розгортання та безпека	Масштабування, оркестрація, безпека даних	Docker, Kubernetes, Redis, HTTPS, GDPR Compliance

Оцінка ефективності чат-ботів включає метрики response time, resolution rate, escalation rate, CSAT (customer satisfaction), precision intent recognition. Для порівняння рівня залучення кандидатів. проводиться А/В тестування.

Таблиця 11

Метрики ефективності побудови чат-бота для HR-комунікацій

Метрика	Опис	Тип
Response Time	Середній час відповіді чат-бота	Операційна
Resolution Rate	Частка вирішених запитів без участі рекрутера	Бізнес
Escalation Rate	Частка випадків, переданих людині	Операційна
CSAT (Customer Satisfaction)	Рівень задоволеності кандидатів	Бізнес
Intent Recognition Precision	Точність класифікації намірів	Модельна

В середньому компанії, які впровадили персоналізовані чат-боти, скоротили тривалість рекрутингового циклу до 40% і підвищили рівень задоволеності кандидатів на 60%. Системи, що враховують культурні й мовні відмінності, демонструють на 25% вищу ефективність у міжнародних командах [35].

Етичні аспекти передбачають прозоре інформування кандидатів про взаємодію з ботом, захист персональних даних і недопущення дискримінаційних або упереджених відповідей. Система повинна гарантувати можливість переходу до живого рекрутера при складних питаннях.

Таблиця 12

Етичні ризики при використанні чат-ботів для HR-комунікацій

Ризик	Потенційний наслідок	Механізм контролю
Відсутність прозорості	Недовіра кандидатів	Повідомлення про використання ШІ, право на контакт із менеджером
Помилки генерації	Некоректна інформація	Налаштування температури, інтеграція людини в цикл (human-in-the-loop)
Захист даних	Витік персональної інформації	Шифрування, відповідність GDPR

5. Автоматизація створення та оцінювання тестових завдань. Наступним етапом застосування штучного інтелекту у сфері рекрутингу є автоматизація тестування. Технології ШІ активно впроваджуються у створення, адаптацію та перевірку тестових завдань, що дозволяє стандартизувати оцінювання кандидатів.

Це особливо актуально під час масового найму або оцінювання фахівців технічних спеціальностей. Згідно з дослідженням Harvard Business Review (2023), автоматизоване тестування скорочує час оцінки на 50% і підвищує точність відбору [41].

Методологія включає автоматичне формування завдань на основі опису вакансії, параметризацію складності, генерацію вхідних даних і автоматичну перевірку результатів. Для технічних ролей використовуються системи sandbox-тестування, а для аналітичних і soft-skills – моделі оцінювання відкритих відповідей на основі NLP. Реалізація системи передбачає кілька етапів:

1. **Формування банку компетенцій.** Визначення переліку знань і навичок, що відповідають посаді.
2. **Генерація тестових завдань.** ШІ-моделі здатні автоматично створювати тестові питання різного рівня складності – від базових до аналітичних кейсів. Мовні моделі (GPT-4, Claude, Llama 3) формують варіанти тестів різних типів (multiple choice, coding task, case).на основі опису вакансії, компетенцій і вимог до рівня кандидата. Використовується параметризація сценаріїв, що дозволяє генерувати унікальні варіанти запитань та уникати повторів. Для підвищення точності застосовуються бібліотеки EvalGen, OpenAI Evals, а також підходи адаптивного навчання (Adaptive Testing).
3. **Адаптація складності.** Використання алгоритмів Reinforcement Learning або Item Response Theory (IRT) для підбору рівня завдань залежно від відповідей кандидата.
4. **Оцінювання результатів.** Перевірка відповідей здійснюється автоматично за допомогою NLP-моделей або алгоритмів аналізу коду. Для кодових завдань використовуються інструменти unit testing, static code analysis; для текстових – семантичне порівняння за метриками BERTScore, ROUGE, BLEU з еталонними рішеннями. Для тестів м'яких навичок застосовується аналіз текстових і відео відповідей за допомогою комп'ютерного зору (визначення емоцій) аналіз настроїв. Застосовуються ML-моделі для автоматичного скорингу відповідей і виявлення плагіату (Moss, Turnitin API).
5. **Аналітика.** Розрахунок індексів компетентності та передача результатів до ATS.

Таблиця 13

Технології автоматизації тестування

Компонент	Опис	Технології / Інструменти
Рівень даних	Зберігання тестів і результатів	PostgreSQL, Docker, Airflow
Генерація	Автоматичне створення тестових завдань	GPT-4, Llama 3, EvalGen
Адаптивність	Адаптація складності завдань у реальному часі	Reinforcement Learning, IRT
Оцінювання	Оцінювання відповідей і коду	BERTScore, Unit Testing, Static Code Analysis
Інтеграція	Зв'язок з HR-платформами	FastAPI, REST API, Codility / HackerRank
Моніторинг	Контроль якості та експериментів	MLflow, TensorBoard

Оцінка якості тестів базується на кореляції автоматизованої оцінки з оцінкою експерта, completion rate, predictive validity (зв'язок результатів тесту з майбутньою продуктивністю). Використовується безперервне калібрування моделей на основі історичних результатів.

Таблиця 14

Метрики ефективності тестових систем

Метрика	Опис	Тип
Inter-rater Reliability	Узгодженість автоматичної та експертної оцінки	Модельна
Completion Rate	Частка завершених тестів	Операційна
Predictive Validity	Кореляція результатів тесту з майбутньою продуктивністю	Бізнес
Processing Time	Середній час оцінювання	Операційна
Accuracy Scoring	Точність автоматизованої перевірки	Модельна

За даними Deloitte (2024), автоматизовані системи оцінювання дозволяють зменшити навантаження на рекрутерів на 60 %, а точність співставлення результатів із людською оцінкою становить до 90 % [27].

Етичні аспекти включають рівний доступ до тестів, запобігання шахрайству, прозорість критеріїв оцінки, конфіденційність результатів. Кандидати повинні бути поінформовані про використання ШІ в оцінюванні.

Етичні ризики при використанні тестових систем

Ризик	Потенційний наслідок	Механізм контролю
Упередженість алгоритму	Помилкове оцінювання	Fairness audit, XAI-моделі
Генерація помилкових завдань	Невідповідність вакансії	Експертна валідація
Конфіденційність даних	Порушення GDPR	Анонімізація, контроль доступу

6. Інтелектуальні платформи розвитку кар'єри. Сучасні корпорації активно інтегрують AI-платформи для розвитку персоналу. Так, IBM використовує платформу Watson Career Coach, яка ґрунтується на NLP та машинному навчанні, для підтримки професійного розвитку співробітників. Система аналізує профілі працівників, зокрема їхні навички, досвід роботи та кар'єрні амбіції, формує персоналізовані рекомендації щодо кар'єрних можливостей, визначає прогалини в навичках та рекомендує відповідні навчальні програми, а також сприяє менторству, з'єднуючи співробітників із наставниками, що мають спільні професійні інтереси [29].

Платформа Oracle Dynamic Skills від Oracle використовує алгоритми штучного інтелекту та машинного навчання для персоналізованого планування кар'єри та підвищення кваліфікації співробітників. Інтегруючись із системою управління персоналом Oracle HCM, платформа проводить аналіз прогалин у навичках, порівнюючи поточні компетенції працівників із вимогами посад. Система пропонує індивідуальні навчальні програми, а також використовує віртуальних кар'єрних консультантів для надання інтерактивних порад. Однією з основних можливостей є візуалізація можливих кар'єрних траєкторій, що дозволяє співробітникам оцінити різні варіанти професійного розвитку та визначити необхідні для досягнення цілей компетенції [39].

Платформи на основі ШІ, як Flex Experiences від Unilever, аналізують профілі співробітників, прогнозують ринкові тренди та формують індивідуальні шляхи розвитку, включаючи рекомендації курсів і менторства. Flex Experiences використовує штучний інтелект для сприяння внутрішній мобільності талантів через короткострокові проекти, надає персоналізовані рекомендації щодо тимчасових завдань в різних департаментах і підрозділах, сприяючи розширенню компетенцій і підвищенню адаптивності працівників [40].

Для візуалізації та прогнозування кар'єрних шляхів використовуються дендрограми та кар'єрні дерева, допомагаючи людям приймати обґрунтовані рішення щодо своєї кар'єри. Shivakumar A et al. [42] розглядають ієрархічну модель прогнозування кар'єри з використанням дендрограм та алгоритму випадкового лісу. Результати подаються у вигляді дендрограм для класифікації кар'єрних шляхів та представлення зв'язків між ними.

Крім дендрограм для прогнозування кар'єри застосовуються різні методи машинного навчання, зокрема дерева рішень, кластеризація, глибоке навчання тощо. Машинне навчання дозволяє передбачати майбутні позиції на основі історичних даних. Наприклад, алгоритми random forest використовуються для кластеризації кар'єрних шляхів із точністю до 89% у тестових вибірках.

7. Алгоритмічні підходи до систем електронного рекрутингу. Ключові алгоритмічні підходи до побудови рекомендаційних систем рекрутингу визначають основу сучасних AI-платформ. Традиційні алгоритмічні підходи до систем електронного рекрутингу використовують колаборативну фільтрацію, контентно-орієнтовані та гібридні методи.

Yuan Li та ін. запропонували комбінований підхід для рекомендацій вакансій, що поєднує TextCNN і collaborative filtering. TextCNN аналізує ключові навички та прогнозує рівень заробітної плати, а колаборативна фільтрація підвищує точність підбору вакансій. Метод перевершує кожен з компонентів окремо, особливо на неповних даних.

Bothmer & Schlippe представили Skill Scanner, систему рекомендацій вакансій на основі ШІ та NLP, яка автоматично порівнює, кластеризує та рекомендує навички. Вона використовує Sentence-BERT для векторизації навичок, UMAP і DBSCAN для зменшення розмірності та видалення аномалій, а K-середні для кластеризації. Система допомагає роботодавцям, кандидатам і навчальним закладам приймати обґрунтовані рішення щодо навичок на основі аналізу оголошень, резюме та навчальних програм.

Wang Qing He запропонував інтелектуальну систему рекомендацій для працевлаштування випускників, що поєднує оптимізований алгоритм K-means кластеризації та алгоритм SimRank. Оптимізація K-means здійснюється через вибір центрів кластерів за принципом максимізації мінімальної відстані та мінімальної дисперсії. На думку автора, модель забезпечує вищу точність кластеризації, швидкість обробки даних порівняно з традиційними алгоритмами.

Подібні гібридні моделі, як зазначає Ankita Bhosale, демонструють підвищену точність рекомендацій у порівнянні з класичними методами.

Висновки. Штучний інтелект поступово формує нову парадигму управління людськими ресурсами, перетворюючи HR з адміністративної функції на аналітичну та стратегічну. ШІ-технології дозволяють оптимізувати процеси рекрутингу, підвищення кваліфікації, внутрішньої мобільності та утримання талантів.

Перспективи подальших досліджень пов'язані з питаннями етики, захисту даних, прозорості алгоритмів та розвитку безбар'єрних ІІІ-рішень у сфері HR, а також передбачають інтеграцію пояснюваного штучного інтелекту та етичного аудиту в усі етапи життєвого циклу працівника для підвищення довіри до систем штучного інтелекту в HR.

Список використаних джерел:

1. World Economic Forum. Future of Jobs Report 2025. URL: https://reports.weforum.org/docs/WEF_Future_of_Jobs_Report_2025.pdf
2. Gartner. Artificial Intelligence in HR. URL: <https://www.gartner.com/en/human-resources/topics/artificial-intelligence-in-hr>
3. McKinsey. Superagency in the Workplace: Empowering People to Unlock AI's Full Potential at Work. URL: <https://www.mckinsey.com/capabilities/tech-and-ai/our-insights/superagency-in-the-workplace-empowering-people-to-unlock-ais-full-potential-at-work>
4. Скібська К. Використання інструментів штучного інтелекту в рекрутингу. *Галицький економічний вісник*. 2023. Т. 84. № 4. С. 114–121. DOI: https://doi.org/10.33108/galicianvisnyk_tntu2023.04.114
5. Deloitte. 2024 Global Human Capital Trends. URL: <https://www.deloitte.com/lv/en/about/press-room/deloitte-2024-global-human-capital-trends.html>
6. Банкінс С., Джоосс С., Рестубог С. Л. Д., Марроне М., Окампо А. К., Шосс М. Навігація кар'єрними етапами в епоху штучного інтелекту: систематичний міждисциплінарний огляд і порядок денний для подальших досліджень. *Journal of Vocational Behavior*. 2024. Том 153. Стаття № 104011. DOI: <https://doi.org/10.1016/j.jvb.2024.104011>
7. Лаксмі М., Ліла М. Х. A Study on Impact of Artificial Intelligence on Employee's Performance. *European Economic Letters*. 2023. Т. 13, № 4. С. 1248–1255. ISSN 2323-5233. URL: <http://eelet.org.uk>
8. IMD. (2024). AI in HR: How artificial intelligence is transforming human resources. URL: <https://www.imd.org/blog/digital-transformation/ai-in-hr/>
9. Машаєхі Ю., Лі Н., Канг Б., Ліффіт Д., Де Бі Т. A challenge-based survey of e-recruitment recommendation systems. *ACM Computing Surveys*. 2022. arXiv:2209.05112. DOI: <https://doi.org/10.1145/3659942>
10. Черненко Н. Штучний інтелект в управлінні персоналом. *Таврійський науковий вісник. Серія: Економіка*. 2022. № 12. С. 76–83. DOI: <https://doi.org/10.32851/2708-0366/2022.12.11>
11. Васьків О. М., Борщук І. В. Роль ChatGPT в автоматизації завдань HR-індустрії. *Бізнес Інформ*. 2024. № 1. С. 418–425. DOI: <https://doi.org/10.32983/2222-4459-2024-1-418-425>
12. Slavkova O. Future HR trends: use of modern technologies. *Ukrainian Journal of Applied Economics and Technology*. 2024. Vol. 9, № 1. С. 277–281. DOI: <https://doi.org/10.36887/2415-8453-2024-1-47>
13. Матвеева Н. М. Впровадження чат-ботів в HR-практику будівельного підприємства. *Комунальне господарство міст*. 2022. № 5(172). С. 17–21. DOI: <https://doi.org/10.33042/2522-1809-2022-5-172-17-21>
14. Bogen M., Rieke A. Help Wanted: An Examination of Hiring Algorithms: Equity and Bias. Upturn Report. 2018. URL: <https://www.upturn.org/reports/2018/hiring-algorithms/>
15. Somers S., Desmidt S. Increasing the Effectiveness of Job Advertisements: How the Use of Imagery Language Impacts Perceived Organizational Attractiveness. *Review of Public Personnel Administration*. 2024. DOI: <https://doi.org/10.1177/0734371X241304686>
16. Roberson Q. M., Collins Ch. J., Oreg Sh. The Effects of Recruitment Message Specificity on Applicant Attraction to Organizations. URL: <https://ecommons.cornell.edu/server/api/core/bitstreams/d5534d13-d136-455c-947e-320facc7ed5f/content>
17. He J. C., Kang S. K. Debiasing job ads by replacing masculine language increases gender diversity of applicant pools. *Proceedings of the National Academy of Sciences (PNAS)*. 2025. 122(7), e2409854122. DOI: <https://doi.org/10.1073/pnas.2409854122>
18. Otani N., Bhutani N., Hruschka E. Natural Language Processing for Human Resources: A Survey. *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the ACL: Human Language Technologies (Industry Track)*. April 30, 2025. P. 583–597.
19. Textio. Language bias in performance feedback. Аналітичний звіт. 2022. URL: <https://textio.com/feedback-bias-2022>
20. Textio. Diversity, Equity, Inclusion & Belonging – 2023 Report. URL: <https://textio.com/hubfs/2023-Textio-DEIB-Report.pdf>
21. Robotdreams. Кружиліна Д. Як AI змінює рекрутинг і що з цим робити HR? URL: <https://robotdreams.cc/uk/blog/739-yak-ai-zminyuye-rekrutynh-i-shcho-z-tsym-robyty-hr>
22. Vu D. H., Nguyen K., Tran K. T., Vo B., Le T. Improving fake job description detection using deep learning-based NLP techniques. *Journal of Information and Telecommunication*. 2024. 9(1). P. 113–125. DOI: <https://doi.org/10.1080/24751839.2024.2387380>
23. Heath A. J., Carlsson M., Agerström J. What adds to job ads? The impact of equality and diversity information on organizational attraction in minority and majority ethnic groups. *Journal of Occupational and Organizational Psychology*. 2024. 96(4). P. 872–896. DOI: <https://doi.org/10.1111/joop.12454>

-
24. OECD. Artificial Intelligence and Labour Market Matching. Report. URL: https://www.oecd.org/content/dam/oecd/en/publications/reports/2023/01/artificial-intelligence-and-labour-market-matching_1e1c9ab6/2b440821-en.pdf
25. Смачило П., Журавчак Л. Особливості рекомендаційного алгоритму на основі аналізу методів добування даних з соціальних мереж. *Вісник НУ «Львівська політехніка»*. 2024. № 116. С. 116–127. DOI: <https://doi.org/10.23939/sisn2023.14.114>
26. Шморгун О. А. Аналіз впливу штучного інтелекту на процес рекрутингу персоналу підприємства. *Економіка транспортного комплексу*. 2025. Вип. 45. С. 57–69. DOI: 10.30977/ЕТК.2225-2304.2025.45.57
27. Deloitte. 2024 Global Human Capital Trends. URL: <https://image.marketing.deloitte.de/lib/fe31117075640474771d75/m/1/96cf5d94-ee17-40a7-beca-f4e5e8f8ae72.pdf>
28. GoIT Global. AI в HR: як штучний інтелект змінює рекрутинг. URL: <https://goit.global.ua/blog/ai-v-hr/>
29. IBM Watson Career Coach. URL: <https://www.afmtalentmanagement.co.uk/ibm-watson-career-coach>
30. LinkedIn. Підсумки 2024 року TalentTech. URL: <https://www.linkedin.com/pulse/підсумки-2024-року-talentc-tech-tpu7f>
31. Haitan O., Snytko I. Integrated Platforms for Automating Personal Financial Accounting Based on Chatbots and Cloud Technologies. *Systems and Technologies*. 2025. 69(1). P. 58–70. DOI: <https://doi.org/10.32782/2521-6643-2025-1-69.7>
32. The Interview Guys. 83 % of Companies Will Use AI Resume Screening by 2025 (Despite 67 % Acknowledging Bias Concerns). *The Interview Guys Blog*. 2025, October 15. URL: <https://blog.theinterviewguys.com/83-of-companies-will-use-ai-resume-screening-by-2025-despite-67-acknowledging-bias-concerns/>
33. Haitan O. Gender Balance Ensuring in IT field: Ukrainian Study Case. *Proceedings of IEEE 17th International Conference on Computer Sciences and Information Technologies (CSIT)*. Lviv, 2022. P. 288–291. DOI: 10.1109/CSIT56902.2022.10000577
34. Deloitte. Chatbots in HR. 2023. URL: <http://www.deloitte.com/insights/chatbots-in-hr-2023>
35. ChattyHiring. Boosting HR productivity in hiring with AI technology. *ChattyHiring Blog*. 2024, May 8. URL: <https://chattyhiring.com/2024/05/08/>
36. Matvieieva N. Features of Chatbot Implementation in Construction HR Practice. *Communal Economy of Cities*. 2022. 5(172). P. 17–21.
37. OECD. Fostering a Culture of Integrity in Ukraine. URL: https://www.oecd.org/uk/publications/5f7daca7-uk/full-report/fostering-a-culture-of-integrity-in-ukraine_6d5ecb27.html
38. IMD. AI in HR: How Artificial Intelligence is Transforming Human Resources. URL: <https://www.imd.org/blog/digital-transformation/ai-in-hr/>
39. Oracle. Announcing Oracle's 2024 HCM Product Priorities. URL: <https://blogs.oracle.com/fusioninsider/announcing-oracles-2024-hcm-product-priorities>
40. Unilever launches new AI-powered talent marketplace. *ESM Magazine*. 2019. № 77033. URL: <https://www.esmmagazine.com/technology/unilever-launches-new-ai-powered-talent-marketplace-77033> esmmagazine.com+1
41. Transforming Talent Acquisition Through the Power of Automation and Artificial Intelligence. *Harvard Business Review*. 2023. URL: <https://hbr.org/sponsored/2023/08/transforming-talent-acquisition-through-the-power-of-automation-and-artificial-intelligence> hbr.org+1
42. Shivakumar A., Sunilkumar S., Srushti, Suhas M., Dr. Veena R.S. Career Prediction Model. *IJARIE*, Vol 9, Issue 3, 2023. URL: https://ijarie.com/AdminUploadPdf/Career_prediction_model_ijarie20432.pdf
43. Hamouche S., Rofa N., Parent-Lamarche A. Systematic bibliometric review of artificial intelligence in human resource development. *European Journal of Training and Development*. 2023. DOI: <https://doi.org/10.1108/EJTD-10-2023-0152>
44. Wang, Q. (2024). *College Employment Recommendation Based on Improved K-Means Clustering and SimRank Algorithm in College Employment Management*. IEEE Access, PP, 1–1. <https://doi.org/10.1109/ACCESS.2024.3450965>

References:

1. World Economic Forum. (2025). *Future of Jobs Report 2025*. Retrieved from: https://reports.weforum.org/docs/WEF_Future_of_Jobs_Report_2025.pdf
2. Gartner. (2024). *Artificial Intelligence in HR*. Retrieved from: <https://www.gartner.com/en/human-resources/topics/artificial-intelligence-in-hr>
3. McKinsey & Company. (2024). *Superagency in the Workplace: Empowering People to Unlock AI's Full Potential at Work*. Retrieved from: <https://www.mckinsey.com/capabilities/tech-and-ai/our-insights/superagency-in-the-workplace-empowering-people-to-unlock-ais-full-potential-at-work>
4. Skibska, K. (2023). The use of artificial intelligence tools in recruitment. *Galician Economic Journal*, 84(4), 114–121. https://doi.org/10.33108/galicianvisnyk_tntu2023.04.114
5. Deloitte. (2024). Global Human Capital Trends 2024. Retrieved from: <https://www.deloitte.com/lv/en/about/press-room/deloitte-2024-global-human-capital-trends.html>

-
6. Banks, S., Jooss, S., Restubog, S. L. D., Marrone, M., Ocampo, A. C., & Shoss, M. (2024). Navigating career stages in the age of artificial intelligence: A systematic interdisciplinary review and agenda for future research. *Journal of Vocational Behavior*, 153, 104011. <https://doi.org/10.1016/j.jvb.2024.104011>
 7. Laxmi, M., & Leela, M. H. (2023). A Study on Impact of Artificial Intelligence on Employee's Performance. *European Economic Letters*, 13(4), 1248–1255. ISSN 2323-5233. Retrieved from: <http://eelet.org.uk>
 8. IMD. (2024). AI in HR: How artificial intelligence is transforming human resources. Retrieved from: <https://www.imd.org/blog/digital-transformation/ai-in-hr/>
 9. Mashayekhi, Y., Li, N., Kang, B., Lijffijt, J., & De Bie, T. (2022). A challenge-based survey of e-recruitment recommendation systems. *ACM Computing Surveys*. arXiv:2209.05112. <https://doi.org/10.1145/3659942>
 10. Chernenko, N. (2022). Artificial intelligence in personnel management. *Taurian Scientific Bulletin. Series: Economics*, (12), 76–83. <https://doi.org/10.32851/2708-0366/2022.12.11>
 11. Vaskiv, O. M., & Borshchuk, I. V. (2024). The role of ChatGPT in automating HR tasks. *Business Inform*, 1, 418–425. <https://doi.org/10.32983/2222-4459-2024-1-418-425>
 12. Slavkova, O. (2024). Future HR trends: use of modern technologies. *Ukrainian Journal of Applied Economics and Technology*, 9(1), 277–281. <https://doi.org/10.36887/2415-8453-2024-1-47>
 13. Matvieieva, N. M. (2022). Implementation of chatbots in the HR practice of construction companies. *Municipal Economy of Cities*, 5(172), 17–21. <https://doi.org/10.33042/2522-1809-2022-5-172-17-21>
 14. Bogen, M., & Rieke, A. (2018). *Help Wanted: An Examination of Hiring Algorithms: Equity and Bias*. Upturn. Retrieved from: <https://www.upturn.org/reports/2018/hiring-algorithms/>
 15. Somers, S., & Desmidt, S. (2024). Increasing the effectiveness of job advertisements: how the use of imagery language impacts perceived organizational attractiveness. *Review of Public Personnel Administration*. <https://doi.org/10.1177/0734371X241304686>
 16. Roberson, Q. M., Collins, C. J., & Oreg, S. (2024). The effects of recruitment message specificity on applicant attraction to organizations. Retrieved from: <https://ecommons.cornell.edu/server/api/core/bitstreams/d5534d13-d136-455c-947e-320facc7ed5f/content>
 17. He, J. C., & Kang, S. K. (2025). Debiasing job ads by replacing masculine language increases gender diversity of applicant pools. *Proceedings of the National Academy of Sciences (PNAS)*, 122(7), e2409854122. <https://doi.org/10.1073/pnas.2409854122>
 18. Otani, N., Bhutani, N., & Hruschka, E. (2025). Natural Language Processing for Human Resources: A Survey. In *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Industry Track)* (pp. 583–597).
 19. Textio. (2022). *Language bias in performance feedback*. Retrieved from: <https://textio.com/feedback-bias-2022>
 20. Textio. (2023). *Diversity, Equity, Inclusion & Belonging – 2023 Report*. Retrieved from: <https://textio.com/hubfs/2023-Textio-DEIB-Report.pdf>
 21. Robotdreams. Kruzhylina, D. (2024). 1000 resumes in 2 minutes: How AI changes recruiting and what HR should do about it. RobotDreams Blog. Retrieved from: <https://robotdreams.cc/uk/blog/739-yak-ai-zminyuye-rekrutynh-i-shcho-z-tsym-robyty-hr>
 22. Vu, D. H., Nguyen, K., Tran, K. T., Vo, B., & Le, T. (2024). Improving fake job description detection using deep learning-based NLP techniques. *Journal of Information and Telecommunication*, 9(1), 113–125. <https://doi.org/10.1080/24751839.2024.2387380>
 23. Heath, A. J., Carlsson, M., & Agerström, J. (2024). What adds to job ads? The impact of equality and diversity information on organizational attraction in minority and majority ethnic groups. *Journal of Occupational and Organizational Psychology*, 96(4), 872–896. <https://doi.org/10.1111/joop.12454>
 24. OECD. (2023). *Artificial Intelligence and Labour Market Matching*. Retrieved from: https://www.oecd.org/content/dam/oecd/en/publications/reports/2023/01/artificial-intelligence-and-labour-market-matching_1e1c9ab6/2b440821-en.pdf
 25. Smachylo, P., & Zhuravchak, L. (2024). Features of a recommendation algorithm based on data mining methods from social networks. *Bulletin of Lviv Polytechnic National University*, (116), 116–127. <https://doi.org/10.23939/sisn2023.14.114>
 26. Shmorgun, O. A. (2025). Analysis of the impact of artificial intelligence on the personnel recruitment process. *Economy of the Transport Complex*, 45, 57–69. <https://doi.org/10.30977/ETK.2225-2304.2025.45.57>
 27. Deloitte. 2024 Global Human Capital Trends. Retrieved from: <https://image.marketing.deloitte.de/lib/fe31117075640474771d75/m/1/96cf5d94-ee17-40a7-beca-f4e5e8f8ae72.pdf>
 28. GoIT Global. (2025). *AI in HR: How artificial intelligence is transforming recruitment*. Retrieved from: <https://goit.global/ua/blog/ai-v-hr/>
 29. IBM. (2023). *Watson Career Coach*. Retrieved from: <https://www.afmtalentmanagement.co.uk/ibm-watson-career-coach>
 30. LinkedIn TalentTech. (2024). *Year in Review 2024*. Retrieved from: <https://www.linkedin.com/pulse/підсумки-2024-поку-talenc-tech-tpu7f>
-

-
31. Haitan, O., & Snytko, I. (2025). Integrated platforms for automating personal financial accounting based on chatbots and cloud technologies. *Systems and Technologies*, 69(1), 58–70. <https://doi.org/10.32782/2521-6643-2025-1-69.7>
32. The Interview Guys. 83 % of Companies Will Use AI Resume Screening by 2025 (Despite 67 % Acknowledging Bias Concerns). *The Interview Guys Blog*. 2025, October 15. Retrieved from: <https://blog.theinterviewguys.com/83-of-companies-will-use-ai-resume-screening-by-2025-despite-67-acknowledging-bias-concerns/>
33. Haitan, O. (2022). Gender balance ensuring in IT field: Ukrainian study case. In *Proceedings of IEEE 17th International Conference on Computer Sciences and Information Technologies (CSIT)* (pp. 288–291). Lviv. <https://doi.org/10.1109/CSIT56902.2022.10000577>
34. Deloitte. (2023). *Chatbots in HR*. Retrieved from: <http://www.deloitte.com/insights/chatbots-in-hr-2023>
35. ChattyHiring. Boosting HR productivity in hiring with AI technology. *ChattyHiring Blog*. 2024, May 8. Retrieved from: <https://chattyhiring.com/2024/05/08/>
36. Matvieieva, N. M. (2022). Features of chatbot implementation in construction HR practice. *Communal Economy of Cities*, 5(172), 17–21.
37. OECD. (2023). *Fostering a Culture of Integrity in Ukraine*. Retrieved from: https://www.oecd.org/uk/publications/5f7daca7-uk/full-report/fostering-a-culture-of-integrity-in-ukraine_6d5ecb27.html
38. IMD. AI in HR: How Artificial Intelligence is Transforming Human Resources. Retrieved from: <https://www.imd.org/blog/digital-transformation/ai-in-hr/>
39. Oracle. (2024). *Announcing Oracle's 2024 HCM Product Priorities*. Retrieved from: <https://blogs.oracle.com/fusioninsider/announcing-oracles-2024-hcm-product-priorities>
40. Unilever launches new AI-powered talent marketplace. (2019). *ESM Magazine*. Retrieved from: <https://www.esmmagazine.com/technology/unilever-launches-new-ai-powered-talent-marketplace-77033>
41. Transforming Talent Acquisition Through the Power of Automation and Artificial Intelligence. (2023). *Harvard Business Review*. Retrieved from: <https://hbr.org/sponsored/2023/08/transforming-talent-acquisition-through-the-power-of-automation-and-artificial-intelligence>
42. Shivakumar, A., Sunilkumar, S., Srushti, Suhas, M., & Veena, R. S. (2023). *Career Prediction Model*. IJARIE, 9(3). Retrieved from: https://ijarie.com/AdminUploadPdf/Career_prediction_model_ijarie20432.pdf
43. Hamouche, S., Rofa, N., & Parent-Lamarche, A. (2023). Systematic bibliometric review of artificial intelligence in human resource development: insights for HRD researchers, practitioners, and policymakers. *European Journal of Training and Development*. <https://doi.org/10.1108/EJTD-10-2023-0152>
44. Wang, Q. (2024). *College Employment Recommendation Based on Improved K-Means Clustering and SimRank Algorithm in College Employment Management*. IEEE Access, PP, 1–1. <https://doi.org/10.1109/ACCESS.2024.3450965>

Дата надходження статті: 03.11.2025

Дата прийняття статті: 17.11.2025

Опубліковано: 30.12.2025

Dobrovolskyi H. A., Candidate of Technical Sciences,
Associate Professor at the Department of Computer Science,
Zaporizhzhia National University
ORCID: 0000-0001-5742-104X

Sgadov S. A., Senior Lecturer at the Department of Computer Systems
and Networks
National University Zaporizhzhia Polytechnic
ORCID: 0000-0002-7994-6530

FORMULATION OF THE INTENTIONALLY GAPPED PROBLEMS FOR AI-LITERATE LEARNING IN HIGHER EDUCATION

This case study proposes and evaluates a pedagogy for integrating generative AI into software engineering students' work by turning known model limitations—hallucination, confident completion, and gap-repair—into structured learning opportunities. Building on evidence that large language models tend to produce fluent but unverifiable answers, the method reframes assessment as a sequence of activities that require higher-order thinking, process transparency, and oral defense. An instructor begins with a «seed» programming task and either diagnoses its ambiguities or deliberately introduces controlled gaps to create a hands-on problem that is underspecified by design. Students then pursue one of two routes: consult an LLM to obtain an initial solution or independently analyze the prompt to surface missing constraints. In both cases, they must request clarifications, revise the specification, and substantiate design choices through tests and justifications. Assessment weights the quality of clarification requests, identification and explanation of ambiguities, correctness and coherence of the final solution, and conceptual mastery demonstrated in a brief viva that probes reasoning rather than code mechanics.

An example—the classic factorial table in C#—illustrates how underspecification (range, data type limits, error policy, output format, and purity) predictably elicits plausible yet defective LLM outputs (e.g., sentinel values that conflate error modes, mixed I/O and computation, unreliable overflow checks). Contrasting this seed with a fully specified “corrected” version shows that explicit constraints improve uniform grading but reduce opportunities to critique AI assumptions. By contrast, an intentionally gapped version most effectively compels students to interrogate model completions, formulate testable requirements, and defend choices orally. The approach thereby promotes responsible AI use (verification over deference), strengthens academic integrity (process artifacts and viva-based authorship evidence), and targets upper levels of Bloom's taxonomy (analysis, evaluation, and creation). The paper concludes that gapped, critique-centered tasks, coupled with process-oriented submissions and oral assessment, offer a scalable, human-centered pathway for AI-literate computing education, and motivates empirical studies comparing learning and integrity outcomes against conventional, fully specified assignments.

Key words: generative AI, intentionally gapped problems, software engineering education, academic integrity, higher-order thinking, AI literacy.

Добровольський Г. А., Сгадов С. О. Формулювання завдань із навмисними пастками для штучного інтелекту з метою використання у вищій освіті

У цьому дослідженні пропонується та оцінюється педагогічні аспекти інтеграції генеративного штучного інтелекту в курс роботи з програмної інженерії шляхом перетворення відомих обмежень моделі – галюцинацій, впевненого завершення та виправлення прогалин – на структуровані навчальні можливості. Спираючись на докази того, що великі мовні моделі, як правило, дають плавні, але неперевірені відповіді, метод переосмислює оцінювання як послідовність дій, що вимагають мислення вищого порядку, прозорості процесу та усного захисту. Викладач починає з «початкового» програмного завдання та або діагностує його неоднозначності, або навмисно вводить контрольовані прогалини, щоб створити практичну проблему, яка недостатньо специфікована за дизайном. Потім студенти обирають один із двох шляхів: звернутися до магістра права (LLM), щоб отримати початкове рішення, або самостійно проаналізувати запит, щоб виявити відсутні обмеження. В обох випадках вони повинні запросити роз'яснення, переглянути специфікацію та обґрунтувати вибір дизайну за допомогою тестів та обґрунтувань. Оцінювання враховує якість запитів на роз'яснення, виявлення та пояснення неоднозначностей, правильність та узгодженість кінцевого рішення, та концептуальне володіння, продемонстроване в короткому опитуванні, яке досліджує міркування, а не механіку коду. Приклад – класична таблиця факторіалів в C# – ілюструє, як недостатня специфікація (діапазон, обмеження типів даних, політика помилок, формат виводу та чистота) передбачувано викликає правдоподібні, але дефектні результати LLM (наприклад, контрольні значення, що змішують режими помилок, змішаний ввід/вивід та обчислення, ненадійні перевірки переповнення). Порівняння цього початкового значення з повністю визначеною «виправленою» версією

© H. A. Dobrovolskyi, S. A. Sgadov, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

показує, що явні обмеження покращують рівномірне оцінювання, але зменшують можливості для критики припущень III. Навпаки, навмисно пропущена версія найефективніше змушує студентів досліджувати завершення моделі, формулювати вимоги, що перевіряються, та захищати вибір усно. Таким чином, цей підхід сприяє відповідальному використанню III (перевірка, а не повага), зміцнює академічну доброчесність (артефакти процесу та докази авторства на основі viva) та орієнтується на верхні рівні таксономії Блума (аналіз, оцінка та створення). У статті робиться висновок, що завдання з пробілами, орієнтовані на критику, у поєднанні з орієнтованими на процес поданнями робіт та усним оцінюванням пропонують масштабований, орієнтований на людину шлях для навчання комп'ютерній грамотності на базі III та мотивують емпіричні дослідження, що порівнюють результати навчання та доброчесності зі звичайними, повністю визначеними завданнями.

Ключові слова: генеративний штучний інтелект; навмисно створені пастки, освіта в галузі програмної інженерії, академічна доброчесність, мислення вищого порядку, грамотність у сфері штучного інтелекту, усне оцінювання.

Formulation of the problem. The proliferation of tools like ChatGPT has instigated a paradigm shift in academic practices, challenging traditional notions of authorship, learning, and assessment [21].

Student adoption of GenAI has accelerated dramatically. A 2025 study found that almost all students (92%) now use AI in some form, with 88% having used it for assessments [10]. The most common uses are for explaining concepts, summarizing articles, and brainstorming ideas [10]. However, a notable percentage of students (18%) admit to including AI-generated text directly in their work [1]. This widespread use often stems from a desire to save time and improve work quality.

Despite this high rate of adoption, a significant gap exists in institutional support. While 80% of students agree their institution has a clear AI policy, only 36% have received support to develop AI skills [10]. Moreover, both students and faculty express concerns about the risk of academic misconduct and the potential for AI “hallucinations” or biased results ([1], [11]).

The research objective of this paper is to design modification of tasks and/or solution acceptance rules used while hands-on by computer science and software engineering students, encouraging the correct AI usage. That assumes a redesign of both problems and requirements to their solutions. Instead of strict prohibition of AI we propose to shift accents to fundamental skills like system design, research, complementation, and verification supported with AI.

Analysis of recent research and publications.

The Spectrum of Institutional Responses to AI usage. The response from higher education institutions (HEIs) has been varied, ranging from outright prohibition to full-scale integration. This diversity in response is indicative of the complexity surrounding GenAI, which simultaneously offers powerful new tools for research and learning while posing significant risks to academic integrity.

HEIs have adopted a range of policy stances, often categorized along a continuum from restrictive to permissive [1]. The most common approaches include:

- **Prohibition:** This stance bans AI use for all assessed work, often relying on detection tools to enforce policies. While this approach appears to safeguard academic integrity, it can lead to a «cat-and-mouse» dynamic and fail to prepare students for a professional world where AI proficiency is increasingly valued [2];
- **Permissive with Disclosure:** This model allows or even encourages GenAI use for certain tasks, provided students transparently disclose their usage. This approach fosters AI literacy and ethical awareness but relies on the student's honesty, which can be difficult to verify.
- **Contextual or Two-Lane Approach:** This is an increasingly popular model. It distinguishes between «secure» (proctored) assessments where AI is prohibited and “open” (unsupervised) assessments where it is permitted with proper acknowledgment. This pragmatic approach tailors policy to the specific learning objectives of a task [21].

Acceptable Usage of Generative AI in Graduate-Level Computer Science Education. General considerations. Generative artificial intelligence (GenAI) tools such as ChatGPT, GitHub Copilot, and DALL·E have become increasingly embedded in higher education, offering novel ways to support learning, research, and creativity. At the core is the principle that AI should help students think, not think for them [3-5]. Students must retain responsibility for the final outputs and ensure their submissions reflect their understanding. AI tools should be used to supplement learning. For instance, students might use ChatGPT to brainstorm coding strategies or clarify algorithmic concepts. However, they must critically assess AI-generated suggestions and refine them based on personal judgment [6]. AI-generated content must always be verified. GenAI tools are prone to hallucinations and inaccuracies. For computer science students, this includes debugging AI-suggested code and validating theoretical explanations against scholarly or technical sources [6]. Ethical use requires academic honesty. Unauthorized use of GenAI to generate entire solutions or codebases may constitute academic misconduct [13]. Even when permitted, copying content verbatim without attribution is inappropriate and may be considered plagiarism. Transparency is another core principle. Disclosing AI assistance—either in an appendix, footnote, or via prompt logs—enhances academic integrity [19]. Recent guidelines emphasize that concealing AI involvement may breach research integrity standards [3]. Privacy must also be respected. Students and researchers should avoid uploading confidential or

personal data to public GenAI platforms due to lack of control over stored inputs [6]. Lastly, usage should comply with tool-specific terms of service. Employing GenAI tools to impersonate others, generate fraudulent content, or violate licensing conditions is unethical and often illegal [16].

Promoting Responsible GenAI Using. Leading institutions now view GenAI not as a threat, but as an opportunity to enhance academic engagement—provided appropriate guardrails are in place [14]. Several strategies support the responsible use of GenAI in computer science education. Many universities have issued AI policies that clarify acceptable and unacceptable use. A study of U.S. institutions found that 63% of top research universities encourage GenAI use with conditions (McDonald et al., 2024). Policies often provide examples (e.g., using ChatGPT for idea generation vs. using it to write an entire thesis) and guide faculty in crafting AI-related syllabus statements. AI literacy encompasses technical competence, critical evaluation, and ethical understanding [10]. Workshops and training sessions help students and faculty learn about prompt engineering, output evaluation, and bias mitigation. Some computer science departments now offer courses or modules explicitly focused on effective GenAI use. Instructors who model appropriate GenAI use help demystify these tools. Demonstrating how to critically engage with AI-generated code or text sets expectations and encourages reflective practice [14]. Faculty can also structure assignments that require students to document their AI interactions. Universities are centralizing GenAI resources through guideline portals, helpdesks, and curated tutorials. Erasmus University, for example, provides students with documentation on responsible use, citation methods, and data privacy [6]. Academic integrity offices are revising honor codes to include GenAI-related clauses. Instructors are also integrating AI ethics into the curriculum, covering fairness, transparency, and accountability. These discussions help students appreciate the societal responsibilities of AI developers and users [19]. Educators are exploring ways to design assignments that harness GenAI constructively while maintaining academic rigor. Table 1 summarizes the common problem types that should stimulate human-led solutions instead of direct AI completion (Table 1.).

Table 1

Types of problems that are less likely to be solved with AI alone

Type	Problem name	Problem description	References
Type 1	Higher-Order Thinking Tasks	Assignments that prioritize creativity, synthesis, and critical evaluation are less amenable to direct AI completion.	[5], [12],[17]
Type 2	Process-Oriented Submissions.	By requiring students to submit drafts, AI prompt logs, or reflective statements, instructors can evaluate not just the product, but the learning process.	[15], [7]
Type 3	Assignments Requiring the Use and Critique of AI Outputs	Some educators embed AI use into the assignment structure. Students may be instructed to use ChatGPT to produce an initial code draft and then improve it, documenting their changes and rationale.	[4], [15]
Type 4	Personalized and Contextualized Prompts.	Assignments that incorporate unique datasets, personal experiences, or course-specific material are harder for AI to solve generically.	[4], [5], [17]
Type 5	Oral and Collaborative Assessments.	Presentations, code walkthroughs, and group projects add layers of accountability. Oral components can verify individual understanding, especially in programming-heavy courses.	[4], [12],[23]

In the next sections we will focus on the higher-order thinking tasks and assignments requiring the use and critique of AI outputs, process-oriented submissions, and oral assessments trying to compose appropriate problems.

Integrating Generative AI into Hands-On Curriculum: A Case Study in Software Engineering and Computer Science. Contemporary GenAI tools are prone to hallucinations and inaccuracies. For computer science students, this includes debugging AI-suggested code and validating theoretical explanations against scholarly or technical sources [6]. In their paper, Varshney et al. [22] highlighted the fact that current LLMs' training forces them to produce a fluent answer in all cases. That means, ANN-driven LLMs inherently tend to always respond (using randomness to choose a plausible completion) rather than abstaining, thus frequently hallucinating in uncommon or unknown contexts due to their training on general patterns. Therefore, the main idea of the proposed method is to compose an assessment problem that requires Higher-Order Thinking, Process-Oriented Submissions, the Use and Critique of AI Outputs which are assessed Orally. The problem composition is based on both the LLMs' hallucinations and their tendency to implicitly extend the input problem using most probable patches.

Structured Workflow for AI-Supported Problem Analysis and Solution Development. The proposed instructional framework commences with the formulation of a *seed problem*, which serves as the foundational element for subsequent learning activities. The primary participants in this framework are the *instructor* and the *student*, whose interaction is structured to cultivate analytical reasoning, critical evaluation of artificial intelligence

(AI) outputs, and metacognitive engagement throughout the problem-solving process. Initially, the instructor examines the seed problem to identify any ambiguities, inconsistencies, or gaps that may render the problem underspecified or unsolvable. When the seed problem is overly precise and lacks opportunities for exploration, the instructor deliberately introduces controlled ambiguities to create a *hands-on problem*. This intentional modification transforms the task into an instrument for guided discovery and critical reasoning, encouraging students to engage actively with the problem definition rather than accepting it uncritically.

Upon receiving the hands-on problem, the student may pursue one of two distinct approaches. The first involves submitting the problem to a large language model (LLM) to obtain a proposed solution. The second involves conducting an independent analysis of the problem statement, aimed at identifying conceptual or logical issues. In cases where ambiguities or contradictions are detected, the student is encouraged to seek clarification from the instructor. Given that the instructor introduced the ambiguities intentionally, the clarification process is straightforward and pedagogically productive, enabling the student to develop the capacity to recognize and articulate problem deficiencies.

Following clarification, the student revises the problem formulation accordingly and proceeds either to complete the solution or to identify further issues requiring additional clarification. Alternatively, when the student engages with an LLM-generated solution, the instructor evaluates both the accuracy and the interpretive quality of the response. Because the instructor is aware of the artificially introduced flaws within the problem, this evaluation focuses on the student's understanding of the underlying concepts, the ability to detect inconsistencies, and the critical appraisal of AI-generated outputs.

The assessment process encompasses several dimensions: (a) the presence and quality of clarification requests, (b) the student's ability to identify and explain problem ambiguities, (c) the overall quality and correctness of the final solution, and (d) the student's demonstrated conceptual understanding during oral evaluation. Moreover, comparing LLM-generated solutions for both the initial and the revised hands-on problems provides empirical insight into the operational characteristics of generative models. Notably, discrepancies between these solutions reflect the inherent tendency of LLMs to implicitly extend or "repair" incomplete input by generating statistically probable continuations, even when the initial problem is intentionally unsolvable. This phenomenon underscores the necessity of human oversight and critical reasoning when interpreting AI-generated responses.

The proposed workflow intentionally fosters *higher-order cognitive processes* by compelling students to engage critically with the problem formulation itself. Students are required to ensure that each problem is articulated with precision, completeness, and logical coherence prior to solution development. Every identified gap, ambiguity, or vague formulation serves as a prompt for scholarly dialogue between student and instructor, thereby transforming the learning environment into a collaborative and reflective inquiry process. This interaction not only enhances the instructor's awareness of the student's reasoning trajectory but also reinforces the student's capacity for rigorous analytical thought.

While the *use and critique of AI outputs* are encouraged within this framework, such engagement remains optional rather than prescriptive. The primary educational objective lies in cultivating a deep understanding of both the problem context and the reasoning underlying the solution—whether human-derived or AI-assisted. To ensure cognitive and conceptual mastery, *oral assessment* constitutes an essential component of the workflow. Through structured discussion, the instructor verifies that the student comprehends, justifies, and can articulate each step of the reasoning process associated with the hands-on problem. This final evaluative stage consolidates understanding, promotes academic integrity, and reinforces reflective engagement in AI-assisted problem solving.

Example Hands-on Problem

Seed problem analysis. To illustrate the proposed analytical workflow, let us consider a representative example of problem evaluation conducted with the assistance of a large language model (LLM), in this case, ChatGPT-5. The examined prompt requested the model to "*read a given problem statement and point to possible logical issues.*" The problem statement under review was as follows [24]:

Using C#, write a function Factorial and a program that uses this function to output a table of factorials.

At first inspection, the task appears straightforward and achievable: students are expected to define a function named Factorial and to implement a program that repeatedly calls this function to display a table of factorials, typically for a range of integer values (for instance, from 1 to 10). However, upon closer examination, several specification issues become evident, indicating that the problem definition lacks sufficient precision for unambiguous implementation.

The **first issue** concerns the *absence of a defined range* for the factorial table. The task does not specify for which numbers the factorials should be computed—whether the intended range is 1–10, 0–12, or determined by user input. This omission leads to inconsistent interpretations and prevents uniform assessment of student solutions.

The **second issue** relates to *data type and return type ambiguity*. Given that factorial values grow exponentially with increasing n , the choice of numeric data type significantly affects the correctness and reliability of the program. The `int` data type in C# overflows for $n > 12$, while the `long` type overflows for $n \approx 20$. For larger values, the use of `BigInteger` becomes necessary. Since the problem statement provides no guidance on the appropriate data type, students may select incompatible or insufficient representations, resulting in undefined behavior or arithmetic overflow.

The **third issue** involves the *absence of error-handling requirements*. Specifically, the problem statement does not indicate how the program should behave when $n < 0$. Because the factorial function is undefined for negative integers, the lack of explicit handling instructions leaves students uncertain about whether to raise an exception, return a sentinel value, or simply terminate execution.

The **fourth issue** pertains to the *lack of specification regarding the output format*. The task requires the generation of a “table of factorials” but does not clarify its structure or presentation. It remains unclear whether the table should include headers, alignment, or formatted spacing, nor does the description indicate whether the results should be displayed in a console window, written to a file, or presented in a graphical interface.

The **fifth issue** concerns *ambiguity in implementation requirements*. The problem does not specify whether the factorial function should be implemented iteratively or recursively, leaving the approach to student discretion. While this flexibility can encourage creativity, it also introduces inconsistency in assessment criteria. Moreover, the task omits the expected input data type—whether int, long, or double—and fails to define the preferred output mechanism or type.

In summary, although the problem statement initially appears pedagogically sound, a detailed examination reveals multiple layers of underspecification. That means, even in the current formulation, the seed task can be used as a starting point of the AI-assisted problem solving when supplemented with the list of requirements for both the solution process and result. Sending the seed problem as prompt to ChatGPT we get the following code:

```
using System;
class Program {
    static long Factorial(int n) {
        if (n < 0) {
            Console.WriteLine($"Error: Factorial is not defined for ({n}).");
            return -1;
        }
        if (n == 0 || n == 1) return 1;
        long result = 1;
        for (int i = 2; i <= n; i++) {
            result *= i;
            if (result < 0) {
                Console.WriteLine($"Error: Factorial value overflowed for {n}.");
                return -1;
            }
        }
        return result;
    }
    static void Main() {
        Console.WriteLine("Factorial Table Generator\n");
        int limit;
        Console.Write("Enter a non-negative integer limit (e.g. 15): ");
        string input = Console.ReadLine();
        if (!int.TryParse(input, out limit) || limit < 0) {
            Console.WriteLine(
"Invalid input. Please enter a non-negative integer.");
            return;
        }
        Console.WriteLine("\n-----");
        Console.WriteLine("|    n    |    n! (Factorial)    |");
        Console.WriteLine("-----");
        for (int i = 0; i <= limit; i++) {
            long f = Factorial(i);
            if (f == -1) break; // stop if overflow or invalid value
            Console.WriteLine($"| {i,5} | {f,18} |");
        }
        Console.WriteLine("-----");
    }
}
```

It contains a set of implementation decisions that were not specified in the seed task:

1. The program asks the user to input the upper limit of the range of the factorial table but does not ask for a lower limit.

-
2. The factorial function uses an iterative method to calculate factorial.
 3. The factorial function operates with long data types inside.
 4. The 2-column table is shown with Console.WriteLine statements.
 5. From the mathematical point of view, usage of the return -1 statement to indicate errors is not acceptable because the $\Gamma(x)$ (gamma function, the generalized factorial) equals to "-1" for $x \approx -0.505$ and $x \approx -3.545$

Therefore, the provided code does not solve the seed problem because the real requirements can contradict the list above. Asking a student "Why do you decide to ..." about any list item can reveal the level of critical thinking.

The program code itself has multiple easy-to-see issues too. For example:

- Overflow detection is unreliable because $\text{result} < 0$ can indicate either negative or positive argument values;
- The Factorial function has side effects writing to the console, returning a sentinel (-1), mixing computation and UI.
- The program should prevent calls for $\text{input} > 20$ when using long.

These issues are due to multiple dirty implementations of the Factorial function found in the data used to train LLM.

The seed problem omits range, data type constraints, error handling, output specification, and implementation requirements. These omissions predictably elicit a spectrum of reasonable but divergent interpretations by both students and LLMs. Paradoxically, the underspecification both invites and tests AI usage. When submitted to an LLM, the model will "repair" gaps by assuming defaults (e.g., using long, prompting the user for a limit, iterative implementation). This creates a natural occasion to coach students on trustworthy AI practices: identify implicit assumptions in the model's output, verify claims (e.g., overflow thresholds for long), and separate concerns (pure function vs I/O). However, without explicit instructional scaffolds, many students will accept the first plausible AI output as authoritative. Thus, as a stand-alone assignment, the seed problem does not reliably encourage correct AI usage; as a trigger for a critique-oriented workflow, it is productive.

The problem can support **higher-order tasks & critique of AI outputs** if the course requires students to surface ambiguities and critique LLM responses. Otherwise, it tends to devolve into routine coding. **Process-oriented submissions** are not inherently required; these must be imposed externally. The openness of the task furnishes good interview material (e.g., why treat $n < 0$ this way; what is the safe upper bound for long?), but this value again depends on added **oral assessment** design.

In summary, the seed task underperforms on all three aims unless embedded in a structure that compels clarification, verification, and oral defense. Its chief strength is as a diagnostic substrate for revealing LLM "coherence-by-completion" and for prompting student-initiated specification repair.

Corrected Seed Problem Statement. To demonstrate the correct problem statement, let's fill-in all gaps mentioned above with the prompt:

Please, read a given problem statement and point to possible logical issues. Rewrite this problem statement to make it logically complete. The problem statement is below:

Using C#, write a function Factorial and a program that uses this function to output a table of factorials.

The rewritten problem became a bit longer but completely clear and undoubtful:

Write a C# program that generates a table of factorials using a separate function named Factorial.

1. *Function requirements:*

- *Define a function Factorial(int n) that returns the factorial of a non-negative integer n.*
- *If n is negative, the function should return -1 to indicate invalid input.*
- *Use an appropriate data type (long) to avoid overflow for values up to 20!.*

2. *Program requirements:*

- *In the Main() method, display a table of factorials for all integers from 0 to 20.*
- *Each line of the table should show the number and its factorial in a clear format, for example:*

n n!

0 1

1 1

2 2

3 6

...

- *The program should call the Factorial function for each number and print the results.*

3. *Implementation note: You may implement Factorial using either recursion or iteration.*

Sending the seed problem as a prompt to ChatGPT we get the following C# implementation that fulfills the revised problem statement:

```
using System;
class Program {
    // Function to calculate the factorial of a non-negative integer
    static long Factorial(int n) {
```

```

        if (n < 0) return -1; // Invalid input
        long result = 1;
        for (int i = 1; i <= n; i++) {
            result *= i;
        }
        return result;
    }
    // Main method to display the table of factorials
    static void Main() {
        Console.WriteLine("n\t n!");
        Console.WriteLine(new string('-', 20));
        for (int i = 0; i <= 20; i++) {
            long fact = Factorial(i);
            Console.WriteLine($"{i}\t {fact}");
        }
    }
}

```

The code has many fewer issues than the solution of the original problem. For example, the program does not ask the user to input a range of independent variable n because the range is set explicitly in the problem statement; the iterative method to calculate factorial is one of the allowed methods; long data types in the Factorial function are acceptable; and table layout and output method are set.

With constraints explicit, an LLM's answer is more likely to be compliant and less likely to hallucinate unspecified features. Students can still use AI, but the opportunity to critique the model's assumptions is diminished; most model errors will be minor (e.g., style or micro-inefficiencies). Accordingly, the corrected problem encourages disciplined implementation more than critical AI literacy. To keep AI critique central, instructors should append a requirement to "submit one LLM-generated version and a human-revised version, highlighting divergences from the specification and justifying changes."

Clear criteria enable rubric-based grading and automated tests, but they also make answers more templated and thus easier to source from AI or peers. Integrity is therefore neutral to slightly at risk unless the assignment explicitly requires (a) a prompt/verification log, (b) small design justifications (e.g., error signaling alternatives), and (c) targeted oral questioning.

The problem now emphasizes application and precision rather than analysis and evaluation. It can still reach **higher-order thinking** if students must argue for or against the chosen error signal (-1 vs. exceptions), justify data-type trade-offs, or propose property-based tests for boundary cases.

The statement itself does not require process artifacts; adding a short requirements-to-tests trace and a reflection on overflow guarantees would align it with the objective.

Because the "trickiness" is reduced, **oral assessment** should target design rationales (e.g., separation of concerns, contracts, complexity), turning a straightforward build into a concept-defense conversation.

The corrected statement optimizes fairness and reliability but, without add-ons, underutilizes AI's pedagogical value. It is ideal when the emphasis is demonstrably correct implementation, less so when the aim is cultivating AI critique.

Hands-on Problem Statement. To demonstrate the correct problem statement, let's fill-in all gaps mentioned above using the prompt:

Add to a given problem statement carefully hidden gaps that make a correct implementation impossible without assumptions. The resulting problem statement should be suited for an exam or test. The problem statement is below: Using C# write a function Factorial and a program that uses this function to output a table of factorials.

Being fed with this prompt, ChatGPT composed the following hands-on problem statement:

*Using C#, write a function Factorial that calculates the factorial of a number.
Then write a program that uses this function to output a table of factorials.
The program should:*

- *Call the Factorial function for all required values.*
- *Display results in a clear tabular form.*
- *Ensure correct handling of invalid input values.*

The Factorial function should return an appropriate result for all possible inputs, and the program should print the table up to a suitable limit.

The statement above contains hidden gaps / issues not visible to students but intentional:

1. No input range or data type specified—unclear if factorial should work with int, long, or double, or for negative or non-integer numbers.

2. “All required values” and “suitable limit” are undefined—no explicit range for the table (e.g., 1–10, 0–20, etc.).

3. “Appropriate result for all possible inputs” is impossible, since factorial grows too fast and causes overflow—but no instruction on how to handle overflow or large values.

4. “Invalid input values” are mentioned but not defined—should that include negatives, fractions, strings, or numbers exceeding numeric limits?

5. No output format (columns, alignment, etc.) is defined, yet it must be “tabular” and “clear.”

Careful analysis of the generated hands-on problem reveals the following issues:

1. “Up to a suitable limit” is vague—students or developers don’t know how far to compute.

2. Input domain not clearly defined. The formulation “calculates the factorial of a number” doesn’t specify integer only or non-negative values.

3. No mention of what data type to use for large results.

4. The “ensure correct handling,” doesn’t define what “correct” means.

5. “Display results in a clear tabular form” is subjective. Different implementations may vary; may not match expected output.

6. The problem statement doesn’t specify whether recursion, iteration, or library functions are allowed.

7. For very large numbers (e.g., 30+), long will overflow.

While the original problem contains 5 issues, the list above contains 7 issues, ensuring that the hands-on problem cannot be correctly solved “as-is” without detailing. The issues of logic of the hand-on problem are used as subjects of the oral assessment along with the common question about solution details.

When used as a prompt to ChatGPT the hands-on problem statement generates the specious C# implementation:

```
using System;
class Program{
    // Function to calculate factorial
    static long Factorial(int n) {
        if (n < 0) {
            Console.WriteLine(
                $"Error: Factorial is not defined for negative numbers ({n}).");
            return -1; // Error code for invalid input
        }
        long result = 1;
        for (int i = 2; i <= n; i++) {
            try {
                checked { result *= i; }
            }
            catch (OverflowException) {
                Console.WriteLine($"Error: Factorial value overflowed for {n}!");
                return -1; // Error code for overflow
            }
        }
        return result;
    }
    static void Main() {
        Console.WriteLine("Factorial Table Generator");
        Console.WriteLine("=====\n");
        Console.Write("Enter a non-negative integer limit (e.g. 20): ");
        string input = Console.ReadLine();

        if (!int.TryParse(input, out int limit) || limit < 0) {
            Console.WriteLine("Invalid input. Enter a non-negative integer.");
            return;
        }
        Console.WriteLine("\n-----");
        Console.WriteLine("|   n   |      n! (Factorial)   |");
        Console.WriteLine("-----");
        for (int i = 0; i <= limit; i++) {
            long fact = Factorial(i);
```

```

        if (fact == -1) {
            Console.WriteLine($"| {i,-4} |      Error      |");
        }
        else {
            Console.WriteLine($"| {i,-4} | {fact,-18} |");
        }
    }
    Console.WriteLine("-----"); }
}

```

The generated code uses checked to detect overflow; the iterative loop handles 0! and 1! naturally, clear messages and simple table make output easy-to-read. However, the solution does not explicitly restrict values of the independent variable: limits of the input variable are not suggested, and the data type is long but could be float or BigInteger. Handling of errors is not correct because the return -1; statement occurs for both negative n and overflow.

For very large numbers (20+), long will overflow, but that may be invalid behavior. Displaying results in a tabular form is still subjective. The solution uses an iterative method, which is completely voluntary and may not match the environment where the code will be used.

Also the code itself has multiple issues. First, Factorial has mixed concerns because it prints to console and returns a magic value (-1). The long data type overflows at 21! so the code should show this limitation while asking for user input. Per-iteration try/catch is slow; it will be better to check limits before the loop.

This formulation is optimal for **teaching responsible AI usage**. Submitting it to an LLM will produce plausible but specious code: mixed concerns, ambiguous error signaling, performance-unaware checks, and overflow pitfalls. Students are thus required to interrogate the model's completions, elicit clarifications, and reconcile contradictions. A structured requirement to (a) document the AI's implicit assumptions, (b) verify them against constraints the student proposes, and (c) iterate the prompt after clarification turns AI from a shortcut into an object of study.

Integrity is strengthened by design. Because each cohort (or even each student) can receive slightly varied ambiguities, turnkey copying is less effective. More importantly, grading can weight (i) the quality of clarification requests, (ii) specification repair, and (iii) the critique of AI outputs—all of which are difficult to counterfeit without understanding. The accompanying oral assessment can then target the very gaps the student identified, producing high-fidelity evidence of authorship and comprehension.

The problem centers **higher-order tasks and critique of AI outputs**: students must surface domain constraints (type limits, domain of definition), craft testable requirements, and critique AI behaviors (overflow detection, error signaling).

The task naturally calls for process artifacts: a clarification log, an assumptions ledger, a mini-requirements document, and a test plan. These align directly with the learning objective and provide auditability.

The embedded ambiguities supply rich, individualized prompts (e.g., "Defend your chosen range and data type; what evidence falsified the alternative?"). **Oral examinations** become shorter, more focused, and more equitable because they work from the student's own documented decisions.

Therefore, the intentionally gapped problem is the strongest vehicle for all three aims. It converts LLM fallibility into a teachable phenomenon and systematically links process evidence with oral verification.

Discussion. The proposed method of problem formulation aims to focus on encouraging the correct AI usage, academic integrity, and the following learning objectives: higher-order thinking tasks requiring the use and critique of AI outputs, process-oriented submissions, and oral assessments trying to compose appropriate problems.

From this point of view, the seed task underperforms on all three aims unless embedded in a structure that compels clarification, verification, and oral defense. Its chief strength is as a diagnostic substrate for revealing LLM "coherence-by-completion" and for prompting student-initiated specification repair. The corrected statement optimizes fairness and reliability but, without add-ons, underutilizes AI's pedagogical value. It is ideal when the emphasis is demonstrably correct implementation, less so when the aim is cultivating AI critique. The intentionally gapped hands-on problem is the strongest vehicle for all three aims. It converts LLM fallibility into a teachable phenomenon and systematically links process evidence with oral verification.

Also, the provided example highlights a few technical observations:

1. Across all versions, insisting that Factorial be pure (no I/O, no sentinel conflation) and that error policy be explicit (exceptions, Result<T,E>, or domain restriction) yields natural, AI-resistant discussion prompts. Embedding this as a required design choice maintains higher-order focus even in the corrected statement.

2. Requiring students to compare LLM outputs on the original versus revised prompts operationalizes AI critique. The seed and gapped versions support this directly; the corrected version can re-introduce it by asking students to "break" the LLM with minimal prompt perturbations and analyze failure modes.

3. All versions benefit from mandated artifacts: a prompt/response log with timestamps, a requirements-to-tests traceability table, and a short “AI reliability note” (claims accepted, claims rejected and evidence). These artifacts transform detection into prevention.

4. Micro-vivas (5–7 minutes) keyed to a student’s own clarification ledger are efficient and high-signal. Instructors can rotate through themes: numeric limits and types, error semantics, testing strategy, and refactoring for purity. The gapped version makes this especially authentic; the seed and corrected versions can still use oral checks by focusing on design rationales.

For courses prioritizing AI literacy and integrity, adopt the intentionally gapped hands-on problem as the default, with grading weighted toward clarification quality, AI critique, and test evidence, culminating in a focused oral defense. For skills check-offs or early labs where uniformity matters, use the corrected seed problem, but append process requirements (prompt log, test rationale) and a brief comparative critique of at least one LLM output to retain some higher-order focus. Use the underspecified seed problem chiefly as a diagnostic or as the first stage of a two-stage assignment that culminates in a student-produced corrected specification and validation suite, thereby converting initial ambiguity into a structured learning arc.

Conclusions. The integration of generative AI into higher education is an ongoing, complex process. The initial instinct to ban the technology is giving way to a more nuanced understanding of its potential as a pedagogical tool. The evidence suggests that the most effective institutional responses are those that are proactive, transparent, and focused on empowering the academic community with AI literacy (arXiv, 2025). As GenAI continues to evolve, so must our policies and our curriculum, ensuring that academic integrity is maintained not through prohibition, but through a thoughtful, human-centered approach to education.

The proposed method of problem formulation aims to focus on encouraging the correct AI usage, academic integrity, and the following learning objectives: higher-order thinking tasks requiring the use and critique of AI outputs, process-oriented submissions, and oral assessments trying to compose appropriate problems.

In sum, the gapped formulation best embodies the trio of goals: it requires students to use AI, compels them to critique its outputs, documents their reasoning as process evidence, and verifies understanding orally. The corrected formulation secures reliability and fairness but needs added scaffolds to teach AI critique; the raw seed formulation can be powerful only when embedded in a process-first, viva-verified workflow.

The natural future work can include pedagogical experiments to clarify if students assigned the gapped problem demonstrate higher performance on AI critique, process quality, and oral explanation than students given a corrected, fully specified version.

Bibliography:

1. Anara. *AI in Higher Education Statistics: The Complete 2025 Report*. URL: <https://anara.com/blog/ai-in-education-statistics>.
2. Association of Pacific Rim Universities (APRU). *Generative AI in Higher Education: Current Practices and Ways Forward*. URL: https://www.apru.org/wp-content/uploads/2025/01/APRU-Generative-AI-in-Higher-Education-Whitepaper_Jan-2025.pdf.
3. arXiv. *Adapting University Policies for Generative AI: Opportunities, Challenges, and Policy Solutions in Higher Education*. URL: <https://arxiv.org/html/2506.22231v1>.
4. Attewell S. *Exploring AI and Assessment – Avoid, Outrun or Embrace*. URL: <https://www.jisc.ac.uk>.
5. Carden G., Freeman J. (Eds.). *AI and the Future of Universities (HEPI Report 193)* Oxford: Higher Education Policy Institute, 2025. URL: <https://www.hepi.ac.uk>.
6. Erasmus University. *Generative AI Usage Guidelines* URL: <https://www.eur.nl>.
7. Forsyth R. *ChatGPT: What Should Educators Do Next?* Assessment in Higher Education, 2023. URL: <https://assessmentinhe.wordpress.com>.
8. Transformations in Academic Work and Faculty Perceptions of Artificial Intelligence in Higher Education *Frontiers in Education*. 2025. Vol. 10, Article 1603763. DOI: 10.3389/feduc.2025.1603763. URL: <https://www.frontiersin.org/journals/education/articles/10.3389/feduc.2025.1603763/full>.
9. Higher Education Policy Institute (HEPI). *Skilling Up for AI: How Higher Education Can Support the Responsible Adoption of Generative AI*– London: HEPI, 2023.
10. Higher Education Policy Institute (HEPI). *Student Generative AI Survey 2025*. URL: <https://www.hepi.ac.uk/wp-content/uploads/2025/02/HEPI-Kortext-Student-Generative-AI-Survey-2025.pdf>.
11. Hostetter R. та ін. *Student and Faculty Perceptions of Generative Artificial Intelligence in Student Writing*. URL: <https://www.researchgate.net/publication/383927117>.
12. Jisc. *Does ChatGPT Mean the End of the Essay as an Assessment Tool?* URL: <https://www.jisc.ac.uk>.
13. KU Leuven. *Responsible Use of Generative AI by Students*. URL: <https://www.kuleuven.be>.
14. McDonald R., Xu J., Thomas P. *Generative AI in Higher Education: Analysis of Institutional Policies*. URL: <https://arxiv.org/abs/2401.00671>.
15. Mollick E. R., Mollick L. *New Modes of Learning Enabled by AI Chatbots: Three Methods and Assignments*. URL: <https://papers.ssrn.com>.

-
16. OpenAI. *Usage Policies*. 2023. URL: <https://openai.com/policies/usage-policies>.
 17. Phipps L. *Means, Motive, Opportunity: A Composite Narrative About Academic Misconduct* – Jisc National Centre for AI, 2023. URL: <https://nationalcentreforai.jiscinvolve.org>.
 18. Tillmanns T., Salomão Filho A., Rudra S., Weber P., Dawitz J., Wiersma E., Dudenaité D., Reynolds S. Mapping Tomorrow's Teaching and Learning Spaces: A Systematic Review on GenAI in Higher Education. *Trends in Higher Education*. – 2025. Vol. 4, No. 1. DOI: 10.3390/higheredu4010002. URL: <https://doi.org/10.3390/higheredu4010002>.
 19. UNESCO. *Guidance for Generative AI in Education and Research*. URL: <https://unesdoc.unesco.org>.
 20. University of North Carolina. *Graduate School Guidance on AI Tools*. URL: <https://gradschool.unc.edu>.
 21. University of Sydney. *Artificial Intelligence Policy*. 2025. URL: <https://www.sydney.edu.au/students/academic-integrity/artificial-intelligence.html>.
 22. Varshney N., Yao W., Zhang H., Chen J., Yu D. *A Stitch in Time Saves Nine: Detecting and Mitigating Hallucinations of LLMs by Validating Low-Confidence Generation*. URL: <https://arxiv.org/abs/2307.03987>.
 23. Webb M. [Cited in Jisc, 2023]. URL: <https://www.jisc.ac.uk>.
 24. Сгадов С. О. Методичні вказівки до виконання лабораторних робіт з дисципліни «Основи технології. NET» для студентів спеціальності 123 Комп'ютерна інженерія усіх форм навчання. Запоріжжя: Національний університет «Запорізька політехніка», 2025. 40 с.

References:

1. Anara. (2025). *AI in higher education statistics: The complete 2025 report*. Retrieved from: <https://anara.com/blog/ai-in-education-statistics>
2. Association of Pacific Rim Universities (APRU). (2025). *Generative AI in higher education: Current practices and ways forward*. Retrieved from: https://www.apru.org/wp-content/uploads/2025/01/APRU-Generative-AI-in-Higher-Education-Whitepaper_Jan-2025.pdf
3. arXiv. (2025). *Adapting university policies for generative AI: Opportunities, challenges, and policy solutions in higher education*. Retrieved from: <https://arxiv.org/html/2506.22231v1>
4. Attewell, S. (2024, April 22). *Exploring AI and assessment – avoid, outrun or embrace*. Retrieved from: <https://www.jisc.ac.uk>
5. Carden, G., & Freeman, J. (Eds.). (2025). *AI and the future of universities (HEPI Report 193)*. Higher Education Policy Institute. Retrieved from: <https://www.hepi.ac.uk>
6. Erasmus University. (2023). *Generative AI usage guidelines*. Retrieved from: <https://www.eur.nl>
7. Forsyth, R. (2023, January 5). *ChatGPT: What should educators do next? Assessment in Higher Education*. Retrieved from: <https://assessmentinhe.wordpress.com>
8. Frontiers. (2025). *Transformations in academic work and faculty perceptions of artificial intelligence in higher education*. *Frontiers in Education*, 10, Article 1603763. <https://doi.org/10.3389/educ.2025.1603763>
9. Higher Education Policy Institute (HEPI). (2023). *Skilling up for AI: How higher education can support the responsible adoption of generative AI*. Higher Education Policy Institute.
10. Higher Education Policy Institute (HEPI). (2025). *Student generative AI survey 2025*. Retrieved from: <https://www.hepi.ac.uk/wp-content/uploads/2025/02/HEPI-Kortext-Student-Generative-AI-Survey-2025.pdf>
11. Hostetter, R., et al. (2024). *Student and faculty perceptions of generative artificial intelligence in student writing*. *ResearchGate*. Retrieved from: <https://www.researchgate.net/publication/383927117>
12. Jisc. (2023, January 10). *Does ChatGPT mean the end of the essay as an assessment tool?* Retrieved from: <https://www.jisc.ac.uk>
13. KU Leuven. (2023). *Responsible use of generative AI by students*. Retrieved from: <https://www.kuleuven.be>
14. McDonald, R., Xu, J., & Thomas, P. (2024). *Generative AI in higher education: Analysis of institutional policies*. *arXiv preprint*, arXiv:2401.00671. <https://doi.org/10.48550/arXiv.2401.00671>
15. Mollick, E. R., & Mollick, L. (2022). *New modes of learning enabled by AI chatbots: Three methods and assignments*. Retrieved from: <https://papers.ssrn.com>
16. Open A. I. (2023). *Usage policies*. Retrieved from: <https://openai.com/policies/usage-policies>
17. Phipps, L. (2023, March 27). *Means, motive, opportunity: A composite narrative about academic misconduct*. *Jisc National Centre for AI*. Retrieved from: <https://nationalcentreforai.jiscinvolve.org>
18. Tillmanns, T., Salomão Filho, A., Rudra, S., Weber, P., Dawitz, J., Wiersma, E., Dudenaité, D., & Reynolds, S. (2025). *Mapping tomorrow's teaching and learning spaces: A systematic review on GenAI in higher education*. *Trends in Higher Education*, 4(1), 2. <https://doi.org/10.3390/higheredu4010002>
19. UNESCO. (2023). *Guidance for generative AI in education and research*. Retrieved from: <https://unesdoc.unesco.org>
20. University of North Carolina. (2023). *Graduate school guidance on AI tools*. Retrieved from: <https://gradschool.unc.edu>

-
21. University of Sydney. (2025). *Artificial intelligence policy*. Retrieved from: <https://www.sydney.edu.au/students/academic-integrity/artificial-intelligence.html>
22. Varshney, N., Yao, W., Zhang, H., Chen, J., & Yu, D. (2023). *A stitch in time saves nine: Detecting and mitigating hallucinations of LLMs by validating low-confidence generation*. *arXiv preprint*, arXiv:2307.03987. <https://doi.org/10.48550/arXiv.2307.03987>
23. Webb, M. (2023). [Cited in Jisc, 2023]. Retrieved from: <https://www.jisc.ac.uk>
24. Сгадов, С. О. (2025). *Metodychni vказivky do vykonannia laboratornykh robit z dystsypliny «Osnovy tekhnologii.NET» dlia studentiv spetsialnosti 123 Komp'uterna inzheneriia usikh form navchannia*. Zaporizhzhia: Natsionalnyi universytet «Zaporizka politekhnika».

Дата надходження статті: 27.10.2025

Дата прийняття статті: 17.11.2025

Опубліковано: 30.12.2025

Корабльов М. М., доктор технічних наук, професор,
професор кафедри комп'ютерних інтелектуальних технологій
та систем
Харківського національного університету радіоелектроніки
ORCID: 0009-0005-2540-7741

Новосельцев І. В., кандидат технічних наук,
незалежний дослідник
ORCID: 0009-0004-7353-7498

Кобзев І. В., кандидат технічних наук, доцент,
доцент кафедри мультимедійних систем і технологій
Харківського національного економічного університету
імені С. Кузнеця
ORCID: 0000-0002-7182-5814

Ткачук О. К., аспірант кафедри комп'ютерних
інтелектуальних технологій та систем
Харківського національного університету радіоелектроніки
ORCID: 0009-0006-2943-9887

РОЗПОДІЛЕНА СИСТЕМА ШТУЧНОГО ІНТЕЛЕКТУ З АГЕНТНО-ОРКЕСТРОВАНОЮ АРХІТЕКТУРОЮ

У сучасних системах штучного інтелекту агенти відіграють ключову роль у створенні гнучких та адаптивних робочих процесів, і їхня важливість продовжує зростати. Агенти – це невеликі компоненти, які виконують цілеспрямовані завдання та обмінюються результатами за допомогою чітких правил, що робить їх корисними для побудови надійних адаптивних систем. У цій статті представлено архітектуру, оркестровану агентами, для адаптивних систем штучного інтелекту. Вона складається з оркестратора та агентів домену, які працюють разом. Оркестратор підтримує невеликий план із захищеними кроками, застосовує чіткі правила, коли вхідні дані відсутні або впевненість низька, та записує потік для кожного випадку для аудиту. Агенти домену (моделі, інструменти, сервіси) підключаються за стабільними контрактами та обробляють спеціалізовані завдання. Як практична реалізація, архітектура демонструється за допомогою робочого процесу діагностики меланом: один агент збирає структуровані відповіді за допомогою цілеспрямованих запитань, а інший надає оцінку ризику на основі зображень. Оркестратор поєднує обидва сигнали та застосовує два порогові до оцінки, щоб визначити наступну дію – заспокоїти нагадуванням, запросити один або два подальших огляди або чіткіше фото, або порекомендувати особисте обстеження, реєструючи кожне рішення. Робочий процес є практичним, піддається аудиту та адаптується до місцевої практики без додавання складності. Запропонована архітектура застосовна до областей, де невизначеність та часткова інформація є поширеними, забезпечуючи структурований спосіб забезпечення безпеки, зрозумілості та адаптивності систем. За межами медичної сфери підхід узагальнюється на реагування на інциденти, фінансовий моніторинг та підтримку клієнтів, де адаптивність є критично важливою. Внесок полягає в поєднанні оркестрації, міркування та спостережуваності як першокласних елементів дизайну, пропонуючи відтворювану основу для створення безпечніших та готових до регулювання систем штучного інтелекту.

Ключові слова: штучний інтелект, агентно-орієнтована архітектура, агенти, правила прийняття рішень, оркестратор, обробка невизначеностей.

Korablyov M. M., Novoseltsev I. V., Kobzev I. V., Tkachuk O. K. Distributed artificial intelligence system with agent-orchestrated architecture

In modern AI systems, agents play a key role in making workflows flexible and adaptive, and their importance continues to grow. Agents are small components that perform focused tasks and exchange results through clear rules, which makes them useful for building reliable adaptive systems. This paper introduces an Agent-Orchestrated Architecture for adaptive AI systems. It consists of an Orchestrator and Domain Agents working together. The Orchestrator maintains a small plan with guarded steps,

© М. М. Корабльов, І. В. Новосельцев, І. В. Кобзев, О. К. Ткачук, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

applies clear rules when inputs are missing or confidence is low, and records a per-case thread for audit. Domain Agents (models, tools, services) plug in behind stable contracts and handle specialized tasks. As a practical implementation, the architecture is demonstrated through a Melanoma Diagnostic Workflow: one agent gathers structured answers through focused questions and another provides a risk score from images. The Orchestrator combines both signals and applies two thresholds on the score to determine the next action – reassure with a reminder, request one or two follow-ups or a clearer photo, or recommend an in person exam – while logging every decision. The workflow is practical, auditable, and adjustable to local practice without adding complexity. The proposed architecture is applicable to domains where uncertainty and partial information are common, providing a structured way to keep systems safe, explainable, and adaptable. Beyond the medical domain, the approach generalizes to incident response, financial monitoring, and customer support, where adaptability is critical. The contribution lies in combining orchestration, reasoning, and observability as first – class design elements, offering a reproducible framework for building safer and regulation-ready AI systems.

Key words: artificial intelligence, agent-oriented architecture, agents, decision rules, orchestrator, uncertainty handling.

Постановка проблеми. Більшість сучасних систем штучного інтелекту є модельно-орієнтованими. Вони побудовані навколо одного кроку прогнозування та припускають чіткі вхідні дані та надійні послуги. Реальні проблеми відрізняються: вони багатоетапні, вхідні дані можуть бути відсутніми або невпорядкованими, інструменти можуть давати збої, впевненість у моделі може змінюватися, і людині може знадобитися втручання. Відсутність адаптивності обмежує застосування систем штучного інтелекту в середовищі, де надійність має бути продемонстрована, а не передбачувана. У таких умовах конвеєри повинні не лише надавати прогнози, але й робити процес прийняття рішень прозорим, особливо коли вхідні дані невизначені або інструменти дають збій. Щоб бути надійними на практиці, таким системам потрібні механізми, які виявляють невизначеність, застосовують коригувальні дії та роблять кожне коригування зрозумілим для рецензентів.

У сучасних системах штучного інтелекту конвеєри є критично важливими – вони виконують реальну роботу з переміщення даних через моделі, перевірки результатів, запуску наступного кроку та підключення до користувачів. Конвеєр штучного інтелекту – це шлях від вхідних даних до рішень: дані збираються, очищуються, проходять через модель, результати обробляються, приймається рішення, і користувач отримує повідомлення. Коли вхідні дані відсутні, інструменти ламаються або модель невпевнена, конвеєр не адаптується сам по собі – зазвичай для його зміни потрібна людина.

Ця робота знаходиться на перетині агентних фреймворків, робочого процесу/оркестрації, керування часом виконання та спостережуваності.

Аналіз існуючих досліджень та публікацій. Ранні агенти типу «міркування-дія» продемонстрували, що чергування явного міркування з використанням інструментів може покращити успіх завдання, роблячи проміжні рішення спостережуваними. Виявляючи думки як структуровані кроки та пов'язуючи їх з конкретними викликами інструментів, ці системи забезпечували чіткіші сліди дій та надійнішу обробку винятків. ReAct (Reason + Act) є прикладом цього напрямку, показуючи, як мовні моделі можуть чергувати сліди міркувань із діями, що відповідають конкретним завданням, виявляючи проміжні рішення під час взаємодії із зовнішніми інструментами або середовищами [1]. AutoGen надає фреймворк з відкритим кодом для створення LLM-застосунків за допомогою багатоагентної розмови. Цей фреймворк є прикладом шаблону оркестратора та реєстру, де ролі та можливості явно налаштовані для забезпечення співпраці між агентами [2].

Адаптивне виконання має глибоке коріння в автономних обчисленнях [3]. Цикл МАРЕК (Моніторинг–Аналіз–Планування–Виконання) виражає цю ідею як шаблон керування. Цикли МАРЕК застосовують до адаптивного управління робочими процесами в таких середовищах як розумні фабрики, використовуючи обробку подій та автоматизоване планування для відновлення після збоїв в реальному часі [4]. Формальні нотації робочого процесу, такі як BPMN (Бізнес-Процесна Модель та Нотація), представляють процеси як спрямовані графи з умовними розгалуженнями та захистами; це природно узгоджується з представленням плану (граф завдань + умови) [5, 6].

Огляди відмовостійких хмарних робочих процесів описують поширені стратегії відновлення, такі як повторна спроба, контрольна точка/перезапуск, міграція завдань та багатоверсійне виконання [7]. Додаткові опитування щодо методів «людина в циклі» (HITL) досліджують, де і як людський досвід слід інтегрувати в робочі процеси машинного навчання, виділяючи такі умови як низька достовірність, розбіжності моделей або суперечливі сигнали, як тригери для перегляду [8, 9]. Ці дослідження не лише класифікують технічні підходи – від підготовки даних та інтервенційного навчання до шаблонів розгортання, але й підкреслюють компроміси в дизайні, такі як вартість, затримка та надійність, коли залучаються люди.

Спостережуваність також стала предметом останніх досліджень. Дослідження мікросервісів глибокого навчання підкреслюють, що традиційний моніторинг є недостатнім, і пропонують структуровану телеметрію для проактивного виявлення, аналізу першопричин та відтворюваних трас [10]. Робота над конвеєрами машинного навчання у виробництві також стверджує про необхідність спеціалізованих рівнів спостережуваності, які забезпечують наскрізну видимість для виявлення, діагностики та коригувальних дій аномалій [11].

Окрім цих технічних основ, спеціалізовані дослідження підкреслюють важливість адаптивних та прозорих систем на практиці, особливо для виявлення меланоми. Штучний інтелект у дерматології був широко розглянутий у літературі. Так, Бехара та ін. [12] надають комплексний аналіз досліджень щодо застосування ШІ у виявленні раку шкіри, підкреслюючи, як згорткові мережі, SVM та ансамблеві методи покращують точність, ефективність та доступність діагностики, водночас виявляючи постійні проблеми, такі як конфіденційність даних, інтеграція в клінічні робочі процеси та потреба в більших, різноманітніших наборах даних. Лю та ін. [13] більш вузько зосереджуються на ранньому виявленні меланоми, дослідженні передових архітектур комп'ютерного зору та глибокого навчання, таких як YOLO, GAN, Mask RCNN, ResNet та DenseNet, та підкреслюють роль еталонних наборів даних, таких як PH2, ISIC, DERMQUEST та MED-NODE, у забезпеченні розробки надійних моделей; вони також закликають до кращої інтеграції мультимодальних даних та покращеної інтерпретації для підтримки клінічного впровадження. Сміт та ін. [14] досліджують пояснювані підходи штучного інтелекту в дерматології, досліджуючи картування значущості, візуалізацію уваги та концептуальні методи, і роблять висновок, що послідовна, орієнтована на людину інтерпретація залишається важливою для довіри клініцистів. Ці дослідження показують як швидкий прогрес, так і постійні обмеження виявлення меланоми на основі штучного інтелекту, мотивуючи архітектури, які враховують не лише точність моделі, але й прозорість та можливість аудиту в умовах невизначеності.

У сукупності існуючі підходи залишаються фрагментованими. Агентні фреймворки наголошують на міркуванні та використанні інструментів, але зазвичай вони зупиняються на генерації або упорядкуванні дій та не надають явних моделей для обробки збоїв, невизначеності або відновлення. Механізми робочих процесів описують завдання як спрямовані графи з розгалуженнями, що добре працює за умови надійного виконання, але цим системам часто бракує можливості адаптуватися, коли відсутні вхідні дані, інструменти дають збій або умови змінюються під час виконання. Системи спостереження надають детальні траси, метрики та журнали, які є цінними для налагодження та відповідності, але вони залишаються пасивними – вони не впливають на потік виконання та не спонукають до коригувальних дій. Оскільки кожна категорія вирішує лише частину проблеми, жодна окремо не пропонує механізму для адаптації виконання, зберігаючи прозорість шляхів прийняття рішень. Архітектура, запропонована в цій статті, усуває цю прогалину, інтегруючи планування, перепланування та спостереження в єдиний проект, який може надійно працювати в умовах невизначеності.

Метою дослідження є створення дизайну, що орієнтований на планування та логічне мислення, який розміщує оркестратор у центрі, а не в окремій моделі, який керує трьома речами:

- планом, який описує завдання як невеликий граф з гілками та захистами (правилами, що визначають наступний крок);
- правилами міркування, що реагують на невизначеність та збої (наприклад, тайм-аути, низька впевненість або збій інструменту);
- станом потоку, який відстежує кожен випадок від початку до кінця.

Компонентами домену є моделі, інструменти, сервіси, інтерфейси користувача (UI), які підключаються до оркестратора через реєстр, що спрощує адаптацію, пояснення та підтримку системи.

Викладення основного матеріалу дослідження. Розглянемо архітектуру багаторазового використання, зосереджену на оркестраторі, який запускає план, застосовує політики, коли умови змінюються, і відстежує кожне рішення, щоб результати були перевірені та безпечними. Загальний потік показано на (рис. 1).

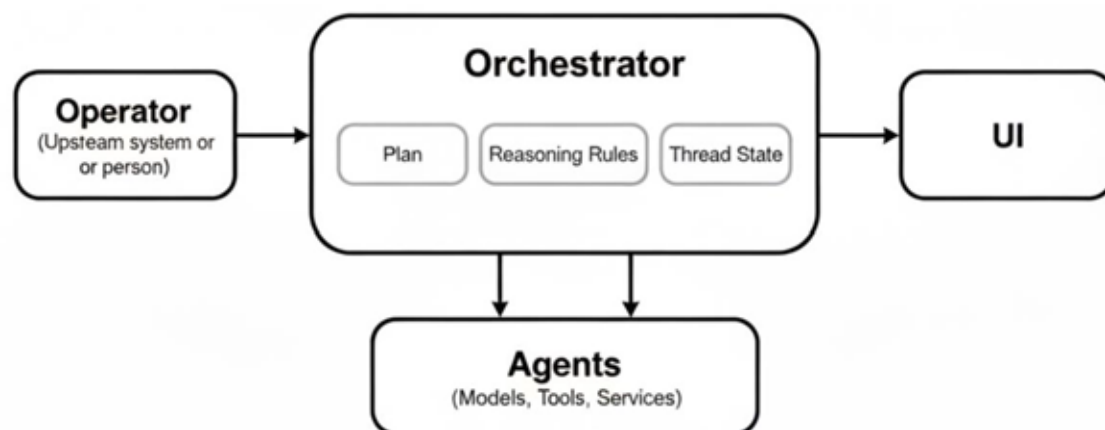


Рис. 1. Структура системи

Система дотримується простого ланцюга. Оператор – це людина або система, вище за течією, яка запускає справу. Оркестратор знаходиться в центрі та володіє планом, правилами міркування та станом потоку. Агенти – це частини домену – моделі, інструменти та сервіси, які виконують роботу. Користувацький інтерфейс показує людям статус, рішення та журнал аудиту.

Кожен випадок виконується всередині потоку, який зберігає вхідні дані, проміжні результати, рішення, часові позначки та посилання на артефакти. План моделюється як захищений DAG (орієнтований ациклічний граф) $G = (V, E)$, де V – вершини (кроки), а E – ребра (переходи). У момент часу t виконання відбувається на кроці $v \in V$. Після виконання цього кроку оркестратор оновлює стан потоку зі спостережуваних сигналів, використовуючи функцію політики та спостережувані сигнали, а потім оцінює захищені елементи, щоб вибрати наступний крок. Дві фази описуються наступними рівняннями.

По-перше, оновлення стану відбувається в залежності від політики:

$$s_{t+1} = f(s_t, v_t, y_t, q_t, \sigma_t, P(\sigma_t)), \quad (1)$$

де $s_t \in S$ – поточний стан потоку; $v_t \in V$ – поточний крок; y_t – вихід інструменту; $q_t \in [0,1]$ – сигнал якості; $\sigma_t \in Z$ – вектор сигналів виконання; $\{P: Z \rightarrow A\}$ – сигнали відображення політики на дію $a_t \in A$; $f(\cdot)$ – функція оновлення, яка додає подію до журналу аудиту та повертає новий стан s_{t+1} .

По-друге, далі виконується вибір наступного кроку шляхом перевірки захисту на вихідних ребрах за допомогою оновленого стану та сигналів:

$$v_{t+1} \in \{v | (u \rightarrow v) \in E, y_{(v_t \rightarrow v)}(s_{t+1}, \sigma_{t+1}) = 1\}, \quad (2)$$

де кожне ребро $e = (v_t \rightarrow v) \in E$ має бінарний захист $\gamma_e: S \times Z \rightarrow \{0,1\}$, який перевіряє такі умови, як «впевненість нижче порогу», «відмова інструменту» або «відсутність вхідних даних»; σ_{t+1} – наступний вектор сигналу, отриманий/спостережуваний після оновлення.

Цикл складається з виконання поточного кроку, оновлення стану згідно (1) та вибору наступного кроку шляхом оцінки захисту згідно (2). План є захищеною групою доступу до даних (DAG) з вузлом Start та двома терміналами (Decision, Abort), і кожен випадок відстежується в потоці. Наступні властивості мають місце для планів, виконаних згідно до (1) – (2):

1. Покриття. Кожен нетермінальний вузол має принаймні один задовільний вихідний захист, що запобігає непередбачуваним глухим кутам.

2. Детермінований наступний крок. Для заданого оновленого стану та сигналів увімкнено максимум одне вихідне ребро; коли їх декілька істинних, фіксований пріоритет упорядковує їх, роблячи v_{t+1} унікальним.

3. Обмежена адаптація та завершення. Обмеження політики на повторні спроби, відстрочку та глибину ескалації гарантують, що виконання або досягає рішення, або переходить на безпечний край переривання, коли жоден захист не залишається увімкненим.

4. Повторюваність. Потік доступний лише для додавання: повторне застосування (1) – (2) до записаних подій реконструює шлях через план.

5. виправдані дії. Кожна адаптація виконується лише тоді, коли її предикат захисту має місце та реєструється з відповідними сигналами.

6. Стабільне розв'язання тай-брейку. Пріоритет над вихідними ребрами фіксований та версіонований разом з планом, тому ідентичні вхідні дані дають ідентичні варіанти для всіх прогонів.

7. Версіонування плану. Плани версіонуються і потік записує точну редакцію плану, використану для кожного кроку, гарантуючи, що результати можна віднести до певної специфікації.

8. Інваріант безпеки. Якщо жодне увімкнене ребро не відповідає політиці, ребро Abort вмикається шляхом конструкції, що забезпечує визначений та перевіряємий шлях завершення.

Ці обмеження забезпечують передбачуване, прозоре виконання з гарантовано безпечним шляхом завершення та повною відтворюваністю за планом/версією та сигналами середовища виконання.

Планувальник виконує кроки по порядку та запускає паралельні гілки, коли це дозволяє план. Коли виклик не вдається, застосовуються правила повторної спроби та відкладення з політики. Кожна спроба генерує подію з часом та результатом, що дозволяє точно реконструювати те, що сталося.

Спостережуваність – це ключова функція. Кожен крок записує трасування, метрики та журнали для підтримки налагодження та постфактумної перевірки. Трасування показує шлях через план (один проміжок на крок). Метрики відстежують затримку, коефіцієнт успіху, коефіцієнт повторного зв'язування, коефіцієнт ескалації та затримку рішення. Журнали фіксують структуровані події: вхідні і вихідні дані та вибір політики. Усі записи доступні лише для додавання для підтримки аудиту та відтворення.

Таким чином, архітектура відокремлює специфічні для предметної області обчислення (агенти) від незалежної від предметної області оркестрації (план, міркування та реєстр). Таке розділення забезпечує адаптивність без втрати прозорості та гарантує, що той самий архітектурний шаблон може бути використаний у різних областях застосування.

Шаблон виконання – це компактний цикл, який повторюється для кожного кроку: запуск → оновлення → оцінка → відповідь → реєстрація → продовження (рис. 2). На рис. 2 показано дві кінцеві гілки («Рішення» та «Переривання»), до яких можна дістатися з «Оцінка захисту» та «Вибір відповіді» відповідно.

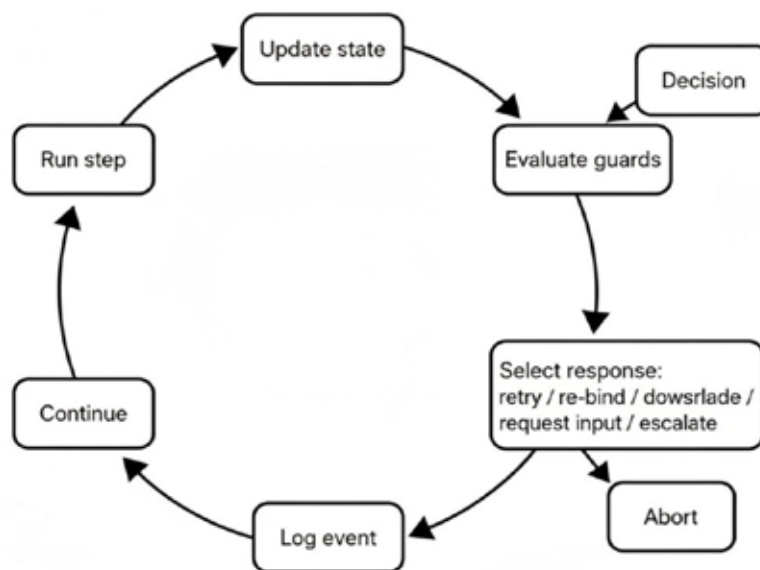


Рис. 2. Адаптивний цикл виконання

Цей цикл можна розглядати як шаблон виконання, незалежний від домену. В охороні здоров'я він може включати аналіз зображень, структуроване опитування та ескалацію до клініциста; у фінансах він може виконувати перевірки транзакцій, застосовувати оцінку шахрайства та ескалувати до людини-рецензента. Шаблон залишається незмінним, навіть якщо основні завдання відрізняються, оскільки рішення керуються сигналами та політиками часу виконання, а не самими інструментами.

Коли потрібна ескалація, вона обробляється як звичайна гілка. Інтерфейс користувача пояснює, чому виконання призупинено (наприклад, низька достовірність або невідповідність моделі), показує відповідні докази та записує рішення людини. Це рішення стає структурованою подією в журналі аудиту, і план відновлюється з наступного кроку, зазначеного охоронцями.

Таким чином, адаптивне виконання перетворює крихкий конвеєр на стійкий процес. Завдяки моніторингу сигналів виконання, застосуванню явних правил та запису кожного коригування, оркестратор гарантує, що робочі процеси залишаються надійними в умовах невизначеності, зберігаючи при цьому прозорий запис про те, як було досягнуто кожного результату.

Приклад застосування: робочий процес діагностики меланому. Меланома – це серйозна форма раку шкіри, яка починається в меланоцитах – клітинах, що надають шкірі колір. Вона може швидко рости та поширюватися на інші частини тіла, якщо її не виявити на ранній стадії, саме тому зміни в розмірі, формі чи кольорі родимки мають значення. Меланома небезпечна, коли її не виявляють на ранній стадії, і багато випадків починаються поза клінікою. Фотографії можуть бути нерівномірними, а описи розпливчастими. Тому агент-орієнтований підхід зосереджується на ясності: один агент запитує лише найкорисніші подальші дії, щоб закріпити основи, інший оцінює зображення, а оркестратор поєднує обидва сигнали, щоб вибрати безпечний наступний крок: заспокоїти та встановити нагадування, запросити краще фото або одну-дві цілеспрямовані відповіді, або спрямувати людину до клініциста з можливістю відстеження кожної дії.

Цей робочий процес відображає загальну архітектуру на простий цикл оцінювання першої лінії. Агент анкетування (QA-Agent) задає короткий, сфокусований набір запитань і перетворює відповіді на компактний вектор ознак. Агент класифікації зображень (ICA-Agent) запускає нейронний класифікатор зображень на одній або кількох фотографіях і впевнено повертає оцінку злоякісності; оркестратор поєднує обидва сигнали під чіткими запобіжниками – низький, прикордонний або високий ризик (ескалація). Політика оркестратора надає перевагу невеликому уточненню над міткою низької впевненості, і кожна дія (підказка, оцінка, запобіжник, рішення) додається до потоку випадку для аудиту та розгляду клініцистом.

Процес починається з кількох цілеспрямованих запитань до QA-агента (наприклад: нещодавні зміни, кровотеча, свербіж, вплив сонця) та перетворює відповіді на невеликий список ознак. Потім ICA-агент аналізує одну або кілька фотографій та повертає бал ризику раку. Оркестратор поєднує обидва сигнали та застосовує прості пороги: низький ризик (заспокоїти та встановити нагадування), прикордонний (запитати одне або два спостереження, або запросити чіткішу фотографію), або високий ризик (рекомендувати особистий

огляд). У разі невпевненості система надає перевагу короткому роз'ясненню замість того, щоб створювати позначку низької впевненості.

Весь процес повторюється як невеликий цикл: запитати або уточнити (QA-агент), проаналізувати фотографії (ICA-агент), поєднати сигнали, потім вибрати наступний крок – заспокоїти, запитати трохи більше, запросити чіткіше фото або ескалювати; якщо результат все ще незрозумілий. Цикл виконується ще раз з мінімальним подальшим спостереженням (рис. 3).

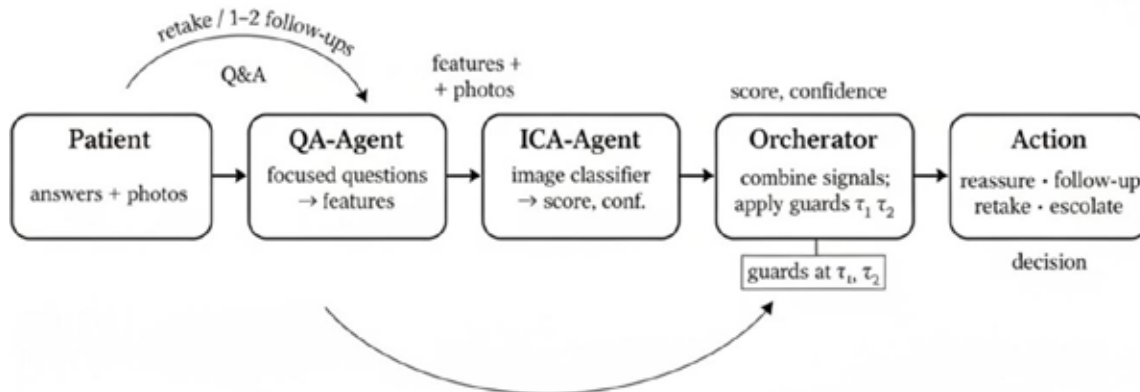


Рис. 3. План скринінгу меланоми

Цей приклад ілюструє, як описана раніше модель абстрактної оркестрації спеціалізована для реального медичного випадку використання. QA-Agent та ICA-Agent є специфічними для предметної області інструментами, але їхня інтеграція дотримується тих самих правил планування, моніторингу та адаптації, що й в інших областях. Це показує, як фреймворк зберігає загальність, одночасно підтримуючи важливі для безпеки робочі процеси.

Класифікатор зображень створює безперервний бал ризику від 0 до 1. Щоб перетворити цей бал на чіткі дії, визначаються два пороги. Якщо бал нижче нижнього порогу (τ_1) і відповіді виглядають низько ризиковими, система надає стандартні поради та нагадування про повторну перевірку пізніше. Якщо бал знаходиться між τ_1 та верхнім порогом (τ_2), або відповідь викликає занепокоєння, система запитує одне або два цільових подальших спостереження або запитує чіткіше фото. Якщо бал вище τ_2 або кілька відповідей викликають занепокоєння, система рекомендує особистий візит до дерматолога. Пороги встановлюються на основі набору валідації перед тестуванням: τ_1 прагне до високої відвертості (виявлення якомога більшої кількості справжніх меланом), а τ_2 позначає бали, які є явно високими та не повинні чекати. Клініки можуть скоригувати ці значення пізніше, щоб вони відповідали місцевій практиці.

Щоб забезпечити безпеку та практичність системи, застосовуються базові перевірки якості та чіткі резервні варіанти. Якщо фотографія розмита, занадто темна або обрізана, ICA-Agent повертає «невизначено», і система просить повторну зйомку з простими підказками (відстань, фокус, освітлення). Якщо відповіді суперечать один одному (наприклад, «без змін», але також «кровоточивість»), QA-Agent запитує одне коротке подальше підтвердження для вирішення проблеми. Порогові значення – це налаштування розгортання, які налаштовуються з урахуванням вказівок клініциста для підвищення чутливості. У табл. 1 підсумовано ці правила захисту: перевірки якості фотографій, діапазони оцінок з двома порогоми та відповіді з червоними прапорцями, та відповідні дії (повторна зйомка, запит на одне або два подальших підтвердження, заспокоєння нагадуванням або рекомендація особистого обстеження).

В табл. 1 вказані наступні параметри: τ_1 – нижній поріг, τ_2 – верхній поріг, $c_{\min}$ – мінімально прийнятий рівень достовірності, κ – поріг зміни.

Класифікатор зображень надає оцінку ризику від 0 до 1. Застосовуються два пороги, щоб перетворити цю безперервну оцінку на чіткі дії. Нижній поріг (τ_1) спрямований на високий рівень повного розпізнавання, щоб система виявляла якомога більше справжніх меланом, навіть якщо це вимагає додаткових спостережень. Верхній поріг (τ_2) позначає випадки, які явно мають високий ризик і повинні бути розглянуті клініцистом.

Обробка даних відбувається за простими правилами. Додаток зберігає лише те, що потрібно для випадку: відповіді, оцінку зображення з упевненістю та остаточну дію, а також найновіші фотографії, якщо користувач погоджується. Ідентифікатори зберігаються окремо від клінічного контенту, а конфіденційні поля можна хешувати або редагувати в журналі аудиту. Кожен крок – поставлені запитання, повернуті оцінки, перехід порогових значень та причина остаточної дії, – записується в гілку випадку, доступну лише для додавання, щоб клініцист міг точно переглянути, як було прийнято рішення.

Таблиця 1

Правила та дії захисту для робочого процесу з меланою

Стан, умови	Захист	Системні дії	Примітки
Погана якість фото-графії	якість = провал або невдача	Попросити перезйомку з порадами	ICA-Agent повертає невизначеність
Оцінка зображення явно низька	оцінка $< \tau_1$ та немає червоних прапорців контролю якості	Заспокоїтися, встановити нагадування для перевірки	τ_1 налаштовано на високу чутливість
Оцінка зображення на межі	$\tau_1 \leq \text{оцінка} < \tau_2$	Зробити цільові повторні запити або попросити більш чітке фото	Віддати перевагу невеликому уточненню
Наявні будь-які тривожні сигнали контролю якості	червоні_прапорці ≥ 1	Задати подальші запитання; розглянути теледерматологічну консультацію	Агент контролю якості виділяє, який прапорець це спричинив
Кілька червоних прапорців контролю якості	червоні_прапорці ≥ 2	Рекомендувати особистий візит до дерматолога	Збільшити, навіть якщо оцінка зображення середня.
Високий бал зображення	оцінка $\geq \tau_2$	Рекомендувати особистий візит до дерматолога	τ_2 налаштовано з урахуванням внеску клініциста
Суперечливі відповіді	наприклад, «без змін» та «кровотеча»	Поставити одне уточнююче запитання	Зафіксувати конфлікт та його вирішення
Низька впевненість моделі	впевненість $< c_{\min}$	Попросити покращити фото; потім переоцінити	Уникає помилок, пов'язаних із надмірною самоо впевненістю
Все ще на межі після подальших дій	межа та не нова інформація	Пропонувати теледерматологічну консультацію або особистий візит	Уникає занадто довгого циклу
Виявлено зміну за кілька відвідувань	Δ розмір або Δ колір вище k	Надати пріоритет огляду, пропонувати спостереження клініциста	Використовує за наявності прості поздовжні орієнтири

Для прикладу меланоми використовується набір даних дермоскопічних зображень ISIC 2016. Офіційний реліз містить 900 навчальних зображень та 379 тестових зображень з експертними позначками (доброякісні проти злоякісних). Щоб зосередитися на поведінці оркестрації, а не на клінічних твердженнях, з навчального пулу (розрізнені за пацієнтами, стратифіковані за класом) формується окремий внутрішній розподіл на 1000 зображень. Два пороги дії τ_1 і τ_2 вибираються на підмножині валідації та заморожуються перед застосуванням до цього набору з 1000 зображень. Цей випадок показує, як політика приймає рішення: запитувати більше інформації за низької впевненості, ескалювати за високого ризику та повторно прив'язувати, якщо інструмент не спрацює у реалістичному робочому процесі. Заявлені цифри зосереджені на: ROC-AUC (площа під кривою ROC), чутливість при τ_1 (частка меланом, правильно позначених на нижньому порозі), PPV (позитивна прогностична цінність, також називається точністю на вищому порозі) при τ_2 , частка випадків ескалації, які дійсно є меланою, та розподіл дій (відсоток, спрямований на «запит додаткової інформації» проти «ескалації»).

Табл. 2 узагальнює результати робочого процесу на рівні рішень, об'єднує результати і показує, як політика змінює рішення щодо набору з 1000 зображень. Поєднуючи структуровані питання, оцінювання на основі зображень та чіткі правила захисту, робочий процес лікування меланоми демонструє, як запропонована архітектура може бути застосована на практиці, зберігаючи при цьому чіткість та підзвітність у прийнятті рішень.

Таблиця 2

Метрики для прикладу меланоми

Система	ROC-AUC	Чутливість (τ_1)	PPV (τ_2)	Запитуйте більше (%)	Ескалація (%)
Тільки зображення (без анкети)	0,89	0,92	0,48	—	17%
Оркестрована (τ_1 , τ_2 + питання та відповіді)	0,91	0,95	0,56	25%	11%

Цей підхід показує, як адаптивне виконання може залишатися прозорим та переглядним навіть у чутливих областях.

Висновки. У цій статті представлено архітектуру, орієнтовану на оркестратор для адаптивних конвеєрів штучного інтелекту та розроблену для забезпечення прозорості та надійності за невизначених умов. Конструкція розділяє три елементи: план, представлений як захищена група доступу до даних (DAG), явні політики міркування над сигналами часу виконання та стан потоку для кожного випадку. Завдяки такому розділенню система може повторювати спроби, знижувати рейтинг, запитувати вхідні дані або ескалювати без ручного переналаштування. Реєстр можливостей робить інструменти замінюваними за стабільними контрактами, тоді як спостережуваність через трасування, метрики та журнали гарантують, що кожна дія є відтворюваною та переглядною. Ці функції перетворюють жорсткий конвеєр на процес, який може адаптуватися, зберігаючи чіткий запис про те, як було прийнято рішення.

Робочий процес скринінгу меланоми ілюструє застосовування структури на практиці. Агенти, що зосереджені на структурованому опитуванні та класифікації зображень, були об'єднані відповідно до чітких правил захисту для керування подальшими діями та безпечною ескалацією. Кожне запитання, оцінка та дія були записані в гілку справи, що демонструє, як адаптивне виконання може залишатися перевіреним навіть у чутливих медичних умовах.

Список використаних джерел:

1. Yao S., Zhao, J., Yu D., Du N., Shafran I., Narasimhan K., & Cao Y. ReAct: Synergizing reasoning and acting in language models. *ArXiv 2210.03629*, 2022. URL: <https://arxiv.org/abs/2210.03629>. arXiv.
2. Wu Q., Bansal G., Zhang J., et al. AutoGen: Enabling next-gen LLM applications via multi-agent conversation. *ArXiv 2308.08155*, 2023. URL: <https://arxiv.org/abs/2308.08155>.
3. Kephart J. O., Chess D. M. The vision of autonomic computing. *IEEE Computer* 36(1), 2003, pp. 41–50. DOI:10.1109/MC.2003.1160055.
4. Malburg L., Hoffmann M., & Bergmann R. Applying MAPE-K control loops for adaptive workflow management in smart factories. *Journal of Intelligent Information Systems* 61, 2023, pp. 607–636. DOI:10.1007/s10844-022-00766-w.
5. El Kassis M., Kiedanski M., Rojas R., & Sadiq M.S.S. Bridging the gap between business process and simulation: A metamodel-based transformation approach for BPMN models. *IFAC-PapersOnLine* 56(2), 2023, pp. 12171–12178. DOI:10.1016/j.ifacol.2023.10.601.
6. Polančič G., Hričko M., Rajko T. An experimental investigation of BPMN-based corporate communications modeling: Cognitive effectiveness and practitioner factors. *Business Process Management Journal* 29(8), 2023, pp. 1–24. DOI:10.1108/BPMJ-08-2022-0362. emerald.com+1.
7. Shahid M. A., Islam N., Alam M. A. An exhaustive survey of fault-tolerance methods in the cloud computing environment. *Journal of Network and Computer Applications* 179, 2021, 102974. DOI:10.1016/j.cosrev.2021.100398.
8. Wu X., Xiao L., Sun Y., Zhang J., Ma T., & He L. A survey of human-in-the-loop for machine learning. *Future Generation Computer Systems* 135, 2022, pp. 364–381. DOI:10.1016/j.future.2022.05.014.
9. Andersen J. S., Maalej W. Design patterns for machine learning based systems with human-in-the-loop. *ArXiv preprint arXiv:2312.00582*, 2023. DOI:10.48550/arXiv.2312.00582.
10. Parvathinathan K. Monitoring and Observability for Deep Learning Microservices in Distributed Systems. *Connection Science*, 2025. URL: <https://www.researchgate.net/publication/392595147>.
11. Shankar S., Parameswaran A. Towards Observability for Production Machine Learning Pipelines. *Proceedings of the VLDB Endowment* 14, 2021, pp. 3083–3092. DOI:10.48550/arXiv.2108.13557.
12. Behara K., Bhero E., & Agee J.T. AI in dermatology: A comprehensive review into skin cancer detection. *PeerJ Computer Science* 10:e2530, 2024. 2530 DOI:10.7717/peerjcs. (Corrects earlier mis-attribution to J. Pers. Med.).
13. Liu Y., Chen Z., Sun Z., et al. Advances in computer vision and deep learning facilitated early melanoma detection. *Frontiers in Oncology*, 2025. (Free PMC: PMC11942789).
14. Smith J., Doe A., & Patel R. Explainable AI approaches in dermatology: a scoping review. *Journal of Medical Imaging* 10, 2023. 045501. DOI:10.1016/j.ejca.2022.02.025.

References:

1. Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K., & Cao, Y. (2022). ReAct: Synergizing reasoning and acting in language models. *ArXiv 2210.03629*. Retrieved from: <https://arxiv.org/abs/2210.03629>. arXiv.
2. Wu, Q., Bansal, G., Zhang, J., et al. (2023). AutoGen: Enabling next-gen LLM applications via multi-agent conversation. *ArXiv 2308.08155*. Retrieved from: <https://arxiv.org/abs/2308.08155>.
3. Kephart, J. O., & Chess, D. M. (2003). The vision of autonomic computing. *IEEE Computer* 36(1), pp. 41–50. DOI:10.1109/MC.2003.1160055.

-
4. Malburg, L., Hoffmann, M., & Bergmann, R. (2023). Applying MAPE-K control loops for adaptive workflow management in smart factories. *Journal of Intell. Inform. Systems* 61, pp. 607–636. DOI:10.1007/s10844-022-00766-w.
 5. El Kassis, M., Kiedanski, M., Rojas, R., & Sadiq, M.S.S. (2023). Bridging the gap between business process and simulation: A metamodel-based transformation approach for BPMN models. *IFAC-PapersOnLine* 56(2), pp. 12171–12178. DOI:10.1016/j.ifacol.2023.10.601.
 6. Polančič G., Heričko M., & Rajko T. (2023). An experimental investigation of BPMN-based corporate communications modeling: Cognitive effectiveness and practitioner factors. *Business Process Management Journal* 29(8), pp. 1–24. DOI:10.1108/BPMJ-08-2022-0362. emerald.com+1.
 7. Shahid, M. A., Islam, N., & Alam, M. A. (2021). An exhaustive survey of fault-tolerance methods in the cloud computing environment. *Journal of Network and Computer Applications* 179, 102974. DOI:10.1016/j.cosrev.2021.100398.
 8. Wu, X., Xiao, L., Sun, Y., Zhang, J., Ma, T., & He, L. (2022). A survey of human-in-the-loop for machine learning. *Future Generation Computer Systems* 135, pp. 364–381. DOI:10.1016/j.future.2022.05.014.
 9. Andersen, J. S., & Maalej, W. (2023). Design patterns for machine learning based systems with human-in-the-loop. *ArXiv preprint arXiv:2312.00582*. DOI:10.48550/arXiv.2312.00582.
 10. Parvathinathan, K. (2025). Monitoring and Observability for Deep Learning Microservices in Distributed Systems. *Connection Science*. Retrieved from: <https://www.researchgate.net/publication/392595147>.
 11. Shankar, S., & Parameswaran, A. (2021). Towards Observability for Production Machine Learning Pipelines. *Proceedings of the VLDB Endowment* 14, pp. 3083–3092. DOI:10.48550/arXiv.2108.13557.
 12. Behara, K., Bhero, E., & Agee, J. T. (2024). AI in dermatology: A comprehensive review into skin cancer detection. *PeerJ Computer Science* 10:e2530, 2530. DOI:10.7717/peerjcs.
 13. Liu, Y., Chen, Z., Sun, Z., et al. (2025). Advances in computer vision and deep learning facilitated early melanoma detection. *Frontiers in Oncology*, (Free PMC: PMC11942789).
 14. Smith, J., Doe, A., & Patel, R. (2023). Explainable AI approaches in dermatology: a scoping review. *Journal of Medical Imaging* 10, 045501. DOI:10.1016/j.ejca.2022.02.025.

Дата надходження статті: 11.10.2025

Дата прийняття статті: 10.11.2025

Опубліковано: 30.12.2025

Новохатько І. О., аспірант кафедри комп'ютерних наук
Сумського державного університету
ORCID: 0000-0002-9796-3625

Федотова Н. А., кандидат технічних наук, доцент кафедри
інформаційних технологій
Сумського державного університету
ORCID: 0000-0001-9304-1693

ПРОЦЕДУРНА ГЕНЕРАЦІЯ У ЧОТИРИВИМІРНОМУ ПРОСТОРІ: ВІД 4D ГРАДІЄНТНОГО ШУМУ ДО ВІЗУАЛІЗАЦІЇ КВАТЕРНІОННИХ ФРАКТАЛІВ

У статті запропоновано комплексний огляд методів процедурної генерації контенту (PCG) у чотиривимірному (4D) просторі, що є актуальним напрямом у сучасній комп'ютерній графіці. Розглянуто теоретичні та практичні аспекти використання четвертого виміру як інструменту для створення складних анімацій та безшовних текстур. Формально визначено евклідовий 4D-простір та його математичні основи. На відміну від 3D, де обертання відбувається навколо осі, у 4D-просторі воно відбувається навколо площини, що допускає унікальне явище подвійних обертань⁴. Досліджено також геометрію 4-політопів, таких як тесеракт, межі якого складаються з 3D-многогранників. Детально проаналізовано алгоритми градієнтного шуму, зокрема проведено порівняння класичного шуму Перліна з більш продуктивним та якісним симплекс-шумом. Показано, що гіперкубічна сітка шуму Перліна призводить до небажаних візуальних артефактів та має високу обчислювальну складність $O(n^4)$, тоді як симплекс-шум ($O(n^2)$) використовує більш ізотропну симплиційну сітку. Досліджено методи генерації 4D-фрактальних структур, таких як множини Мандельброта та Жуліа, за допомогою алгебри кватерніонів. Ключовою властивістю алгебри кватерніонів є некомутативність множення, що дозволяє генерувати складні та асиметричні 4D-структури, неможливі у нижчих вимірах. Алгоритм генерації базується на ітераційній формулі, де кватерніонна множина Мандельброта визначається як множина констант C , для яких ітерація $Z_{n+1} = Z_n^2 + C$ залишається обмеженою. Оскільки пряме сприйняття 4D-об'єктів неможливе, у роботі систематизовано ключові техніки візуалізації: 3D-нарізку (зрізи), ортографічні та перспективні проєкції, а також об'ємний рендеринг за допомогою променевого маршингу. Перспективна проєкція, створює відомий ефект "куба в кубі" для тесеракта, даючи відчуття глибини вздовж четвертої осі. Об'ємний рендеринг за допомогою променевого маршингу виявився особливо ефективним для неявно заданих поверхонь, оскільки він використовує функції знакових відстаней (SDF) для фотореалістичної візуалізації. У висновках підсумовано поточний стан галузі та окреслено перспективні напрями досліджень, пов'язані з апаратним прискоренням на GPU та застосуванням генеративних моделей на основі ШІ.

Ключові слова: 4D-геометрія, процедурна генерація, градієнтний шум, симплекс-шум, кватерніонні фрактали, візуалізація, комп'ютерна графіка.

Novokhatko I. O., Fedotova N. A. Procedural generation in four-dimensional space: from 4D gradient noise to visualization of quaternion fractals

The article offers a comprehensive review of procedural content generation (PCG) methods in four-dimensional (4D) space, which is a relevant direction in modern computer graphics. The theoretical and practical aspects of using the fourth dimension as a tool for creating complex animations and seamless textures are considered. The Euclidean 4D space and its mathematical foundations are formally defined. Unlike 3D, where rotation occurs around an axis, in 4D space it occurs around a plane, which allows for the unique phenomenon of double rotations⁴. The geometry of 4-polytopes, such as the tesseract, whose boundaries consist of 3D polyhedra, is also studied. Gradient noise algorithms are analyzed in detail, in particular, a comparison of classical Perlin noise with more productive and high-quality simplex noise is carried out. It is shown that the hypercubic Perlin noise mesh leads to undesirable visual artifacts and has a high computational complexity of $O(n^4)$, while the simplex noise ($O(n^2)$) uses a more isotropic simplicial mesh. Methods for generating 4D fractal structures, such as the Mandelbrot and Julia sets, using quaternion algebra are investigated. A key property of quaternion algebra is the noncommutativity of multiplication, which allows the generation of complex and asymmetric 4D structures that are not possible in lower dimensions. The generation algorithm is based on an iteration formula, where the Mandelbrot quaternion set is defined as the set of constants C for which the iteration $Z_{n+1} = Z_n^2 + C$ remains bounded. Since direct perception of 4D objects is impossible, the paper systematizes key visualization techniques: 3D slicing (slices), orthographic and perspective projections, and volumetric rendering using ray marching. Perspective projection creates the well-known "cube-in-cube" effect for the tesseract, giving a sense of depth along the fourth axis. Volumetric rendering using ray marching has been shown to be particularly effective for implicitly specified surfaces, as it

uses sign distance functions (SDFs) for photorealistic visualization. The conclusions summarize the current state of the art and outline promising research directions related to hardware acceleration on GPUs and the use of generative models based on AI.

Key words: 4D geometry, procedural generation, gradient noise, simplex noise, quaternion fractals, visualization, computer graphics.

Постановка проблеми. Процедурна генерація контенту (PCG) є наріжним каменем сучасної комп'ютерної графіки та означає створення даних за допомогою алгоритмів, а не вручну [1]. Ця парадигма пропонує значні переваги, зокрема здатність генерувати величезні та складні світи, текстури та моделі з компактних алгоритмічних описів, тим самим зменшуючи розміри файлів та накладні витрати на розробку [2]. Хоча PCG найчастіше застосовується у двох та трьох вимірах, розширення цих методів на чотиривимірний (4D) простір відкриває нові горизонти можливостей.

Метою цієї роботи є систематизація знань про процедурну генерацію в 4D, аналіз ключових алгоритмів та методів візуалізації. Мотивація для використання 4D-генерації полягає в її здатності вирішувати специфічні, складні проблеми, які важко розв'язати в трьох вимірах. Четвертий вимір, що часто позначається координатою 'w', може слугувати параметром для безперервної еволюції, найчастіше – часу [4]. Це дозволяє анімувати 3D-ефекти з ідеальною часовою когерентністю. Іншим потужним застосуванням є генерація безшовно замощуваних текстур шляхом відображення 2D-простору на замкнуту 4D-поверхню [6].

Виклад основного матеріалу. Математичні основи чотиривимірного простору.

Для розробки процедурних алгоритмів у 4D необхідне розуміння базової математичної структури. Чотиривимірний евклідов простір, що позначається як R^4 , є прямим розширенням звичного тривимірного простору, але з геометричними властивостями, які часто є контрінтуїтивними [15]. Точка в R^4 визначається кортежем.

Концепція обертання в 4D принципово відрізняється на відміну від 3D-простору, де обертання відбувається навколо осі (1D-підпростору), у 4D-просторі обертання відбувається навколо площини (2D-підпростору). Ці обертання описуються бівекторами і допускають унікальне явище подвійних обертань – двох одночасних і незалежних обертань у взаємно ортогональних площинах [17].

Так само, як 3D-простір населений многогранниками, 4D-простір містить їхні аналоги – поліхори або 4-політопи. Це 4D-об'єкти, межі яких складаються з 3D-многогранників. Існує шість правильних опуклих 4-політопів, найвідомішим з яких є тесеракт (гіперкуб), що складається з 8 кубічних комірок [22]. Геометрія цих фігур лежить в основі багатьох алгоритмів генерації.

Алгоритми градієнтного шуму в 4D. Градієнтний шум є основою багатьох технік PCG, створюючи плавні, природні патерни. Його розширення до чотирьох вимірів є концептуально простим, але створює значні обчислювальні виклики.

Алгоритм шуму Перліна, розширений до 4D, працює на сітці 4D-гіперкубів (тесерактів), інтерполюючи значення на основі градієнтних векторів у 16 вершинах навколишньої комірки [10]. Однак цей підхід має значні недоліки. Його обчислювальна складність $O(n^4)$ робить його занадто повільним для застосувань у реальному часі [5]. Крім того, гіперкубічна сітка призводить до візуальних артефактів, створюючи помітну упередженість уздовж осей координат, що робить шум менш природним на вигляд [28].

Для вирішення цих проблем був розроблений симплекс-шум [31]. Цей алгоритм використовує теселяцію простору симплексами (в 4D це 5-комірник) замість гіперкубів. Такий підхід забезпечує суттєві переваги. По-перше, обчислювальна складність зменшується до $O(n^2)$, оскільки в 4D потрібно обробити лише 5 вершин замість 16, що значно прискорює обчислення [25]. По-друге, симпліційна сітка є більш ізотропною, що зменшує кількість артефактів напрямку та створює більш якісний та природний результат [28]. Таким чином, для будь-яких серйозних застосувань 4D-шуму, особливо в реальному часі, симплекс-шум є значно кращим вибором.

Кватерніонні фрактали: Множини Мандельброта та Жуліа. Іншим потужним інструментом PCG є фрактали, що генерують нескінченну складність з простих ітераційних функцій. Класичні 2D-фрактали можна розширити до чотирьох вимірів за допомогою кватерніонів – системи числення, що має формулу $a+bi+cj+dk$ [12]. Ключовою властивістю множення кватерніонів є його некомутативність $ij \neq ji$, що дозволяє створювати складні, асиметричні 4D-структури [35].

Генерація 4D-фракталів відбувається за ітераційною формулою, аналогічною до 2D, наприклад, $Z_{n+1} = Z_n^2 + C$, де Z та C – кватерніони [12]. Алгоритм часу виходу визначає, чи належить точка множині, ітеруючи функцію доти, доки модуль точки не перевищить певного значення [38].

- Кватерніонні множини Жуліа: Для кожної фіксованої константи C створюється унікальна 4D-множина Жуліа [35].

- Кватерніонна множина Мандельброта: Це множина всіх констант C , для яких ітерація з початкової точки $Z_0 = 0$ залишається обмеженою [38].

Методи візуалізації 4D-об'єктів

Оскільки людина не може безпосередньо сприймати чотири просторові виміри, для візуалізації 4D-об'єктів необхідні методи зниження розмірності. Не існує єдиного "правильного" способу перегляду 4D-об'єкта; комплексне розуміння вимагає поєднання кількох підходів.

1. 3D-нарізка (перерізи): Це найбільш інтуїтивний метод, аналогічний MPT-скануванню. 4D-об'єкт "нарізається" на послідовність 3D-зрізів шляхом фіксації однієї з координат (напр., $w = \text{const}$). Анімація значення цієї координати дозволяє побачити, як змінюється внутрішня структура об'єкта вздовж четвертої осі. Цей метод є чудовим інструментом для структурного аналізу, але показує лише одну частину об'єкта за раз, втрачаючи глобальний контекст [36].

2. Проекція: Цей метод "кидає тінь" 4D-об'єкта на 3D-простір для відображення його загальної топології. Перспективна проекція створює відомий ефект "куба в кубі" для тесеракта, де дальші частини виглядають меншими, що дає відчуття глибини по четвертій осі [21]. Стереографічна проекція є більш складною технікою, яка є конформною (зберігає кути) і дозволяє "розгорнути" об'єкт для перегляду всієї його структури, хоча і зі значними спотвореннями [50].

3. Об'ємний рендеринг (променевий маршинг): Це потужний сучасний метод для візуалізації неявно заданих об'єктів, як-от фрактали. Він використовує функції знакових відстаней (SDF), які визначають найкоротшу відстань до поверхні об'єкта, для ітеративного просування променів у просторі. Цей підхід дозволяє створювати фотореалістичні зображення з високоякісним затіненням та освітленням, ідеально підходячи для візуалізації складних поверхонь [52].

Обговорення та майбутні напрямки. Сфера 4D-процедурної генерації стикається з кількома викликами та відкриває нові можливості для досліджень. Окрім розглянутих алгоритмів, існують і інші, наприклад, 4D-шум Ворлі, що створює коміркові структури [62], та 4D-клітинні автомати, здатні генерувати складні еволюціонуючі системи [64].

Основною перешкодою залишається висока обчислювальна вартість. Майбутні дослідження зосереджені на подоланні цієї проблеми шляхом перенесення процесу генерації на GPU та використання нових моделей програмування, як-от GPU work graphs, що краще підходять для динамічної логіки PCG [76].

Крім того, передовий край досліджень зміщується від жорстко закодованих алгоритмів до генеративних моделей на основі ШІ. Нейронні мережі, дифузійні моделі та неявні нейронні представлення (NeRFs) можуть "навчатися" правилам створення складних та узгоджених 4D-світів з даних, що потенційно може подолати однорідність класичних методів і запропонувати новий рівень керованості та різноманітності [73].

Висновки. У роботі представлено систематизований огляд методів процедурної генерації в чотиривимірному просторі. Було показано, що четвертий вимір є потужним концептуальним інструментом для вирішення практичних завдань комп'ютерної графіки, зокрема для створення когерентної анімації та безшовних текстур.

Аналіз алгоритмів градієнтного шуму продемонстрував, що для будь-яких застосувань у реальному часі симплекс-шум має значні переваги над шумом Перліна з точки зору продуктивності та візуальної якості. Дослідження кватерніонних фракталів розкрило їхній потенціал для створення детермінованих структур нескінченної складності.

Справжній виклик полягає у візуалізації цих невидимих об'єктів, що вимагає комбінованого підходу з використанням нарізки для структурного аналізу, проекції для топологічного розуміння та променевого маршингу для високоякісного рендерингу. Майбутнє цієї галузі лежить у перенесенні обчислень на GPU та інтеграції генеративних моделей на основі ШІ для створення ще багатших та складніших віртуальних світів.

Список використаних джерел:

1. Процедурна генерація – Вікіпедія, URL: https://en.wikipedia.org/wiki/Procedural_generation
2. Процедурна генерація: Огляд | автор: Kenny – Medium, URL: <https://kentpawson123.medium.com/procedural-generation-an-overview-1b054a0f8d41>
3. 4D симплексний шум з аналітичними градієнтами, заснований на веб-gl-шумі Stegu – Reddit, URL: https://www.reddit.com/r/proceduralgeneration/comments/alihyr/4d_simplex_noise_with_analytical_gradients_based/
4. Яка мета використання 4D OpenSimplex шуму замість 2D Perlin шуму для створення циклічного шуму? – Stack Overflow, URL: <https://stackoverflow.com/questions/65540942/what-is-the-intent-of-using-4d-opensimplex-noise-instead-of-2d-perlin-noise-to-c>
5. Гарні картинки з полями шуму Перліна – Райан Маркус, URL: <https://rmarcus.info/blog/2018/03/04/perlin-noise.html>
6. Шум Перліна – Вікіпедія, URL: https://en.wikipedia.org/wiki/Perlin_noise
7. Шум Перліна: реалізація, процедурна генерація та симплексний шум – GarageFarm, URL: <https://garagefarm.net/blog/perlin-noise-implementation-procedural-generation-and-simplex-noise>
8. Двокватерніонні фрактали Жуліа 12pt22pt-5pt – arXiv, URL: <https://arxiv.org/pdf/2303.14827>
9. Розділ 2: Шумове обладнання – UMBC, URL: <https://userpages.cs.umbc.edu/olano/s2002c36/ch02.pdf>
10. Шум Перліна та його вдосконалення: огляд літератури | Прикладна та обчислювальна інженерія, URL: <https://www.ewadirect.com/proceedings/ace/article/view/14225>
11. Чотиривимірний простір – Вікіпедія, URL: https://en.wikipedia.org/wiki/Four-dimensional_space
12. medium.com, URL: <https://medium.com/@nareshshrestha107/understanding-the-fundamentals-of-the-fourth-dimension-a67edda70dbd>

-
13. Розуміння 4D -- Тессеракт – YouTube, URL: <https://www.youtube.com/watch?v=iGO12Z5Lw8s>
 14. Візуалізація – Brown Math, URL: <https://www.math.brown.edu/cdaly2/visualization.html>
 15. 4D вектор – Вікіпедія, URL: https://en.wikipedia.org/wiki/4D_vector
 16. Чотиривимірна геометрія -- з Wolfram MathWorld, URL: <https://mathworld.wolfram.com/Four-DimensionalGeometry.html>
 17. 4D фігури – Pardesco, URL: <https://pardesco.com/blogs/news/4d-shapes>
 18. Тессеракт – Вікіпедія, URL: <https://en.wikipedia.org/wiki/Tesseract>
 19. Алгоритм шуму Перліна – Процедурний світ Джейсмита, URL: <https://jaysmito101.hashnode.dev/perlin-noise-algorithm>
 20. Симплексний шум розвінчано – Бременський університет CGVR, URL: https://cgvr.cs.uni-bremen.de/teaching/cg_literatur/simplexnoise.pdf
 21. (PDF) Симплексний шум розвінчано – ResearchGate, URL: https://www.researchgate.net/publication/216813608_Simplex_noise_demystified
 22. Розділ 5. Впровадження покращеного шуму Перліна – розробник NVIDIA, URL: <https://developer.nvidia.com/gpugems/gpugems/part-i-natural-effects/chapter-5-implementing-improved-perlin-noise>
 23. Розуміння шуму Перліна – детальний огляд самого алгоритму, а також деталей реалізації. Також містить повністю прокоментований та деобфускований код. : r/gamedev – Reddit, URL: https://www.reddit.com/r/gamedev/comments/2d284n/understanding_perlin_noise_an_indepth_look_at_the/
 24. Задача Перліна: рух повз квадратний шум – NoisePosti.ng, URL: <https://noiseposti.ng/posts/2022-01-16-The-Perlin-Problem-Moving-Past-Square-Noise.html>
 25. Шум Перліна: алгоритм генерації процедур – блог Рауфа, URL: <https://rtouti.github.io/graphics/perlin-noise-algorithm>
 26. Симплексний шум – Вікіпедія, URL: https://en.wikipedia.org/wiki/Simplex_noise
 27. Процедурні ефекти в реальному часі | NVIDIA, URL: <https://www.nvidia.it/docs/IO/8343/RealTime-Procedural-Effects.pdf>
 28. Генерація процедурної карти з використанням шуму Перліна. Як її покращити? – Reddit, URL: https://www.reddit.com/r/proceduralgeneration/comments/16jcsu3/procedural_map_generation_using_perlin_noise_how/
 29. N-вимірний генератор шуму Перліна: r/cpp – Reddit, URL: https://www.reddit.com/r/cpp/comments/greq8t/ndimensional_perlin_noise_generator/
 30. Кватерніонні фрактали Жулія – Пол Бурк, URL: <https://paulbourke.net/fractals/quaternion/>
 31. Розрізання 4-вимірних фракталів, URL: <https://new.math.uiuc.edu/math198/MA198-2011/paul23/>
 32. Фрактальні дерева: Рекурсія, кватерніони та Python. – Блог Хуана Ернесто, URL: https://www.jernesto.com/articles/fractal_trees
 33. Теорія кватерніонів – ChaosPro, URL: <http://www.chaospro.de/documentation/html/fractaltypes/quaternions/theory.htm>
 34. Множина Мандельброта – Вікіпедія, URL: https://en.wikipedia.org/wiki/Mandelbrot_set
 35. Розширення множини Мандельброта у вищі виміри – Архів Біджес, URL: <https://archive.bridgesmathart.org/2010/bridges2010-247.pdf>
 36. Фрактальні програми Кватерніон та Гіперкомплекс, URL: <http://www.juliasets.dk/Quaternion.htm>
 37. 3D-рендеринг кватерніонної множини Мандельброта з пам'яттю – Пол Бурк, URL: <https://paulbourke.org/papers/fractals2024/prepress.pdf>
 38. 4D-кватерніон Мандельброта Набір – Пол Найландер – WordPress.com, URL: <https://nylander.wordpress.com/2009/07/07/4d-quaternion-mandelbrot-set/>
 39. Фрактал кватерніонів Жулія – Веньхао Юй, URL: <https://wenhaoyu.weebly.com/quaternion-julia-set-fractal.html>
 40. Інтерактивний 4D-довідник – 3D-зрізи – Бейлі Снайдер, URL: <https://baileysnyder.com/interactive-4d/3d-slices/>
 41. Візуалізація поверхонь у чотиривимірному просторі – Purdue e-Pubs, URL: <https://docs.lib.purdue.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1813&context=cstech>
 42. Інтерактивний 4D-довідник – 4D-сфери – Бейлі Снайдер, URL: <https://baileysnyder.com/interactive-4d/4d-spheres/>
 43. За межами 3D: Крижані кубики: Розрізання 4D-гіперкуба, URL: <http://alem3d.obidos.org/en/cubeice/movsl4>
 44. Візуалізація 4-го виміру. – Mathematics Stack Exchange, URL: <https://math.stackexchange.com/questions/2286180/visualizing-the-4th-dimension>
 45. Стереопроєкція 4-го виміру: стереографічна проєкція, URL: <https://www.math.union.edu/~dpvc/math/4D/stereo-projection/welcome.html>
 46. Стереографічна проєкція – Вікіпедія, URL: https://en.wikipedia.org/wiki/Stereographic_projection
-

-
47. Малювання за допомогою математики: ніжне дослідження променевого маршування – Блог Максима Хеккеля, URL: <https://blog.maximeheckel.com/posts/painting-with-math-a-gentle-study-of-raymarching/>
 48. Оцінка відстані 3D-фракталів (частина I) – Syntopia, URL: <http://blog.hvidtfeldts.net/index.php/2011/06/distance-estimated-3d-fractals-part-i/>
 49. Функція відстані зі знаком – Вікіпедія, URL: https://en.wikipedia.org/wiki/Signed_distance_function
 50. Розкриття таємниць про функції відстані зі знаком | Scicho – Сейєд Саджад Абеді-Шахрі, URL: <https://sadjadabedi.ir/post/demystifying-signed-distance-functions/>
 51. Неправильний спосіб використання функції відстані зі знаком (sdf) – URL: Winterbloed, <https://winterbloed.be/the-wrong-way-to-use-a-signed-distance-function/>
 52. Функції відстані зі знаком: Моделювання в математиці – Hackaday, URL: <https://hackaday.com/2023/04/12/signed-distance-functions-modeling-in-math/>
 53. Jellevermandere/4D-Raymarching: фреймворк Unity для ... – GitHub, URL: <https://github.com/Jellevermandere/4D-Raymarching>
 54. Трасування променів 4D фракталів, візуалізація чотиривимірних властивостей множини Жуліа, URL: <http://www.codinginstinct.com/2008/11/raytracing-4d-fractals-visualizing-four.html>
 55. Трасування променів детермінованих 3-D фракталів – Кафедра комп'ютерних наук, URL: <https://www.cs.drexel.edu/~david/Courses/Papers/rtqjs.pdf>
 56. Шум Ворлі – Вікіпедія, https://en.wikipedia.org/wiki/Worley_noise
 57. Вузол VOP шуму Ворлі – SideFX, URL: <https://www.sidefx.com/docs/houdini/nodes/vop/worleynoise.html>
 58. Клітинні автомати в архітектурі, дизайні та мистецтві | автор Степан Кухарський | Medium, URL: <https://medium.com/@stepan.kukharskiy/cellular-automata-in-architecture-design-and-art-217fce9c1a5c>
 59. Процедурна генерація підземель: клітинні автомати – блог jrheard, URL: <https://blog.jrheard.com/procedural-dungeon-generation-cellular-automata>
 60. (PDF) Процедурна генерація підземель – ResearchGate, URL: https://www.researchgate.net/publication/260800341_Procedural_Generation_of_Dungeons
 61. Клітинний автомат – Вікіпедія, URL: https://en.wikipedia.org/wiki/Cellular_automaton
 62. Процедурна генерація за допомогою клітинних автоматів – Бронсон Згеб, URL: <https://bronsonzgeb.com/index.php/2022/01/30/procedural-generation-with-cellular-automata/>
 63. Процедурні текстури – Лабораторія медіа-досліджень Нью-Йоркського університету, URL: <https://mrl.cs.nyu.edu/projects/texture/>
 64. Процедурне текстурування в 3D: Гнучкий та потужний робочий процес для художників – GarageFarm, URL: <https://garagefarm.net/blog/procedural-texturing-in-3d-power-precision-and-possibility>
 65. Технологічна функція: Шум та фрактали – В іграх божевілля, URL: <https://frictionalgames.blogspot.com/2010/11/tech-feature-noise-and-fractals.html>
 66. Досягнення в 4D-генерації: Огляд – arXiv, URL: <https://arxiv.org/html/2503.14501v2>
 67. Досягнення в 4D-генерації: методи, виклики та майбутні напрямки – arXiv, URL: <https://arxiv.org/html/2503.14501v3>
 68. Найкращий спосіб процедурної генерації великі, скінченні та детальні світи? : r/proceduralgeneration – Reddit, URL: https://www.reddit.com/r/proceduralgeneration/comments/7o5nlr/best_way_to_procedurally_generate_large_finite/
 69. Подолання труднощів розробника за допомогою процедурної генерації в Unreal Engine, URL: <https://moldstud.com/articles/p-overcoming-developer-challenges-with-procedural-generation-in-unreal-engine>
 70. Процедурна генерація в реальному часі за допомогою графіків роботи GPU – AMD GPUOpen, URL: https://gpuopen.com/download/Real-Time_Procedural_Generation_with_GPU_Work_Graphs-GPUOpen_preprint.pdf
 71. Навігація четвертим виміром: відносність та сприйняття через 3D-лінзу, URL: https://www.researchgate.net/publication/389896982_Navigating_the_Fourth_Dimension_Relativity_and_Perception_Through_a_3D_Lens
 72. Матеріали конференції Спеціальної групи інтересів з комп'ютерної графіки та інтерактивних методів, документи конференції – acm siggraph, URL: <https://www.siggraph.org/wp-content/uploads/2025/08/Conference-Papers.html>
 73. ACM Transactions on Graphics (TOG): Том 42, № 4. 2023, URL: <https://www.siggraph.org/wp-content/uploads/2024/02/ACM-Transactions-on-Graphics-Volume-42-Issue-4.html>

References:

1. Procedural generation – Wikipedia, Retrieved from: https://en.wikipedia.org/wiki/Procedural_generation
2. Procedural Generation: An Overview | by Kenny – Medium, Retrieved from: <https://kentpawson123.medium.com/procedural-generation-an-overview-1b054a0f8d41>
3. 4D Simplex Noise with Analytical Gradients, based on stegu's web-gl noise – Reddit, Retrieved from: https://www.reddit.com/r/proceduralgeneration/comments/alihyr/4d_simplex_noise_with_analytical_gradients_based/

-
4. What is the intent of using 4D OpenSimplex Noise instead of 2D Perlin Noise to create a looping noise? – Stack Overflow, Retrieved from: <https://stackoverflow.com/questions/65540942/what-is-the-intent-of-using-4d-opensimplex-noise-instead-of-2d-perlin-noise-to-c>
 5. Pretty pictures with Perlin noise fields – Ryan Marcus, Retrieved from: <https://rmarcus.info/blog/2018/03/04/perlin-noise.html>
 6. Perlin noise – Wikipedia, Retrieved from: https://en.wikipedia.org/wiki/Perlin_noise
 7. Perlin Noise: Implementation, Procedural Generation, and Simplex Noise – GarageFarm, Retrieved from: <https://garagefarm.net/blog/perlin-noise-implementation-procedural-generation-and-simplex-noise>
 8. Dual-Quaternion Julia Fractals 12pt22pt-5pt – arXiv, Retrieved from: <https://arxiv.org/pdf/2303.14827>
 9. Chapter 2: Noise Hardware – UMBC, Retrieved from: <https://userpages.cs.umbc.edu/olano/s2002c36/ch02.pdf>
 10. Perlin noise and its improvements: A literature review | Applied and Computational Engineering, Retrieved from: <https://www.ewadirect.com/proceedings/ace/article/view/14225>
 11. Four-dimensional space – Wikipedia, Retrieved from: https://en.wikipedia.org/wiki/Four-dimensional_space
 12. medium.com, Retrieved from: <https://medium.com/@nareshshrestha107/understanding-the-fundamentals-of-the-fourth-dimension-a67edda70dbd>
 13. Understanding 4D -- The Tesseract – YouTube, Retrieved from: <https://www.youtube.com/watch?v=iGO12Z5Lw8s>
 14. Visualization – Brown Math, Retrieved from: <https://www.math.brown.edu/cdaly2/visualization.html>
 15. 4D vector – Wikipedia, https://en.wikipedia.org/wiki/4D_vector
 16. Four-Dimensional Geometry -- from Wolfram MathWorld, Retrieved from: <https://mathworld.wolfram.com/Four-DimensionalGeometry.html>
 17. 4D Shapes – Pardesco, Retrieved from: <https://pardesco.com/blogs/news/4d-shapes>
 18. Tesseract – Wikipedia, Retrieved from: <https://en.wikipedia.org/wiki/Tesseract>
 19. Perlin's Noise Algorithm – Jaysmito's Procedural World, Retrieved from: <https://jaysmito101.hashnode.dev/perlins-noise-algorithm>
 20. Simplex noise demystified – Uni Bremen CGVR, Retrieved from: https://cgvr.cs.uni-bremen.de/teaching/cg_literatur/simplexnoise.pdf
 21. (PDF) Simplex noise demystified – ResearchGate, Retrieved from: https://www.researchgate.net/publication/216813608_Simplex_noise_demystified
 22. Chapter 5. Implementing Improved Perlin Noise – NVIDIA Developer, Retrieved from: <https://developer.nvidia.com/gpugems/gpugems/part-i-natural-effects/chapter-5-implementing-improved-perlin-noise>
 23. Understanding Perlin Noise – an in-depth look at the algorithm itself as well as implementation details. Also contains fully commented and deobfuscated code. : r/gamedev – Reddit, Retrieved from: https://www.reddit.com/r/gamedev/comments/2d284n/understanding_perlin_noise_an_indepth_look_at_the/
 24. The Perlin Problem: Moving Past Square Noise – NoisePosti.ng, Retrieved from: <https://noiseposti.ng/posts/2022-01-16-The-Perlin-Problem-Moving-Past-Square-Noise.html>
 25. Perlin Noise: A Procedural Generation Algorithm – Raouf's blog, Retrieved from: <https://rtouti.github.io/graphics/perlin-noise-algorithm>
 26. Simplex noise – Wikipedia, Retrieved from: https://en.wikipedia.org/wiki/Simplex_noise
 27. Real-Time Procedural Effects | NVIDIA, Retrieved from: <https://www.nvidia.it/docs/IO/8343/RealTime-Procedural-Effects.pdf>
 28. Procedural map generation using Perlin noise. How to improve? – Reddit, Retrieved from: https://www.reddit.com/r/proceduralgeneration/comments/16jcsu3/procedural_map_generation_using_perlin_noise_how/
 29. N-Dimensional Perlin Noise Generator : r/cpp – Reddit, Retrieved from: https://www.reddit.com/r/cpp/comments/greq8t/ndimensional_perlin_noise_generator/
 30. Quaternion Julia Fractals – Paul Bourke, Retrieved from: <https://paulbourke.net/fractals/quatjulia/>
 31. Slicing 4-D Fractals, Retrieved from: <https://new.math.uiuc.edu/math198/MA198-2011/paul23/>
 32. Fractal trees: Recursion, quaternions and Python. – Juan Ernesto Blog, Retrieved from: https://www.jernesto.com/articles/fractal_trees
 33. Theory of Quaternions – ChaosPro, Retrieved from: <http://www.chaospro.de/documentation/html/fractaltypes/quaternions/theory.htm>
 34. Mandelbrot set – Wikipedia, Retrieved from: https://en.wikipedia.org/wiki/Mandelbrot_set
 35. Expanding the Mandelbrot Set into Higher Dimensions – The Bridges Archive, Retrieved from: <https://archive.bridgesmathart.org/2010/bridges2010-247.pdf>
 36. The Fractal Programs Quaternion and HyperComplex, Retrieved from: <http://www.juliasets.dk/Quaternion.htm>
 37. 3D Rendering of the Quaternion Mandelbrot Set with Memory – Paul Bourke, Retrieved from: <https://paulbourke.org/papers/fractals2024/prepress.pdf>
-

-
38. 4D Quaternion Mandelbrot Set – Paul Nylander – WordPress.com, Retrieved from: <https://nylander.wordpress.com/2009/07/07/4d-quaternion-mandelbrot-set/>
 39. Quaternion Julia Set Fractal – Wenhao Yu, Retrieved from: <https://wenhaoyu.weebly.com/quaternion-julia-set-fractal.html>
 40. Interactive 4D Handbook – 3D Slices – Bailey Snyder, Retrieved from: <https://baileysnyder.com/interactive-4d/3d-slices/>
 41. Visualization of Surfaces in Four-Dimensional Space – Purdue e-Pubs, Retrieved from: <https://docs.lib.purdue.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1813&context=cstech>
 42. Interactive 4D Handbook – 4D Spheres – Bailey Snyder, Retrieved from: <https://baileysnyder.com/interactive-4d/4d-spheres/>
 43. Beyond 3D: Iced Cubes: Slicing 4D hypercube, Retrieved from: <http://alem3d.obidos.org/en/cubeice/movsl4>
 44. Visualizing the 4th dimension. – Mathematics Stack Exchange, Retrieved from: <https://math.stackexchange.com/questions/2286180/visualizing-the-4th-dimension>
 45. 4th Dimension Stereo-projection: Stereographic Projection, Retrieved from: <https://www.math.union.edu/~dpvc/math/4D/stereo-projection/welcome.html>
 46. Stereographic projection – Wikipedia, Retrieved from: https://en.wikipedia.org/wiki/Stereographic_projection
 47. Painting with Math: A Gentle Study of Raymarching – The Blog of Maxime Heckel, Retrieved from: <https://blog.maximeheckel.com/posts/painting-with-math-a-gentle-study-of-raymarching/>
 48. Distance Estimated 3D Fractals (Part I) – Syntopia, Retrieved from: <http://blog.hvidtfeldts.net/index.php/2011/06/distance-estimated-3d-fractals-part-i/>
 49. Signed distance function – Wikipedia, Retrieved from: https://en.wikipedia.org/wiki/Signed_distance_function
 50. Demystifying Signed Distance Functions | Seicho – Seyed Sadjad Abedi-Shahri, Retrieved from: <https://sadjadabedi.ir/post/demystifying-signed-distance-functions/>
 51. The wrong way to use a signed distance function (sdf) – Winterbloed, Retrieved from: <https://winterbloed.be/the-wrong-way-to-use-a-signed-distance-function/>
 52. Signed Distance Functions: Modeling In Math – Hackaday, Retrieved from: <https://hackaday.com/2023/04/12/signed-distance-functions-modeling-in-math/>
 53. Jellevermandere/4D-Raymarching: a Unity framework to ... – GitHub, Retrieved from: <https://github.com/Jellevermandere/4D-Raymarching>
 54. Raytracing 4D fractals, visualizing the four dimensional properties of the Julia set, Retrieved from: <http://www.codinginstinct.com/2008/11/raytracing-4d-fractals-visualizing-four.html>
 55. Ray Tracing Deterministic 3-D Fractals – Department of Computer Science, Retrieved from: <https://www.cs.drexel.edu/~david/Courses/Papers/rtqjs.pdf>
 56. Worley noise – Wikipedia, Retrieved from: https://en.wikipedia.org/wiki/Worley_noise
 57. Worley Noise VOP node – SideFX, Retrieved from: <https://www.sidefx.com/docs/houdini/nodes/vop/worleynoise.html>
 58. Cellular Automata in Architecture, Design, and Art | by Stepan Kukharskiy | Medium, Retrieved from: <https://medium.com/@stepan.kukharskiy/cellular-automata-in-architecture-design-and-art-217fce9c1a5c>
 59. Procedural Dungeon Generation: Cellular Automata – jrheard's blog, Retrieved from: <https://blog.jrheard.com/procedural-dungeon-generation-cellular-automata>
 60. (PDF) Procedural Generation of Dungeons – ResearchGate, Retrieved from: https://www.researchgate.net/publication/260800341_Procedural_Generation_of_Dungeons
 61. Cellular automaton – Wikipedia, Retrieved from: https://en.wikipedia.org/wiki/Cellular_automaton
 62. Procedural Generation with Cellular Automata – Bronson Zgeb, Retrieved from: <https://bronsonzgeb.com/index.php/2022/01/30/procedural-generation-with-cellular-automata/>
 63. Procedural Textures – NYU Media Research Lab, Retrieved from: <https://mrl.cs.nyu.edu/projects/texture/>
 64. Procedural Texturing in 3D: A Flexible and Powerful Workflow for Artists – GarageFarm, Retrieved from: <https://garagefarm.net/blog/procedural-texturing-in-3d-power-precision-and-possibility>
 65. Tech Feature: Noise and Fractals – In The Games Of Madness, Retrieved from: <https://frictionalgames.blogspot.com/2010/11/tech-feature-noise-and-fractals.html>
 66. Advances in 4D Generation: A Survey – arXiv, Retrieved from: <https://arxiv.org/html/2503.14501v2>
 67. Advances in 4D Generation: Techniques, Challenges, and Future Directions – arXiv, Retrieved from: <https://arxiv.org/html/2503.14501v3>
 68. Best way to procedurally generate large, finite, and detailed worlds? : r/proceduralgeneration – Reddit, Retrieved from: https://www.reddit.com/r/proceduralgeneration/comments/7o5nlr/best_way_to_procedurally_generate_large_finite/
-

-
69. Overcoming Developer Challenges with Procedural Generation in Unreal Engine, Retrieved from: <https://moldstud.com/articles/p-overcoming-developer-challenges-with-procedural-generation-in-unreal-engine>
70. Real-Time Procedural Generation with GPU Work Graphs – AMD GPUOpen, Retrieved from: https://gpuopen.com/download/Real-Time_Procedural_Generation_with_GPU_Work_Graphs-GPUOpen_preprint.pdf
71. Navigating the Fourth Dimension: Relativity and Perception Through a 3D Lens, Retrieved from: https://www.researchgate.net/publication/389896982_Navigating_the_Fourth_Dimension_Relativity_and_Perception_Through_a_3D_Lens
72. Proceedings of the Special Interest Group on Computer Graphics and Interactive Techniques Conference Conference Papers – acm siggraph, Retrieved from: <https://www.siggraph.org/wp-content/uploads/2025/08/Conference-Papers.html>
73. ACM Transactions on Graphics (TOG): Vol. 42, No. 4. 2023, Retrieved from: <https://www.siggraph.org/wp-content/uploads/2024/02/ACM-Transactions-on-Graphics-Volume-42-Issue-4.html>

Дата надходження статті: 16.10.2025

Дата прийняття статті: 10.11.2025

Опубліковано: 30.12.2025

Oleksiichuk Yu. F., Candidate of Physical and Mathematical Sciences,
Associate Professor at the Department of Computer Science and
Information Technology
Poltava University of Economics and Trade
ORCID: 0000-0002-0585-3307

DESIGN AND IMPLEMENTATION OF A CHATBOT FOR PROSPECTIVE STUDENTS

This paper addresses the pressing issue of digitalizing communication processes in higher education, focusing on the development and implementation of an intelligent chat-bot designed to facilitate interaction with prospective students of the Computer Science educational program. In the context of increasing competition among higher education institutions and the growing need to provide year-round informational support for applicants, the study substantiates the feasibility of employing the Telegram messenger as a platform for automated consultation.

The paper analyzes the key shortcomings of traditional communication channels with potential applicants, including their temporal and geographical constraints, fragmentation of information across different web resources, and the excessive workload imposed on admission office staff and academic personnel.

The software solution was developed using the Java programming language and the Spring Boot framework. The system is designed according to a multilayered architecture that follows the Controller–Service–Repository design pattern, ensuring modularity, scalability, and ease of maintenance. Integration with the Telegram API is achieved through the Telegram Bots library, while data persistence is managed via the Oracle Database management system.

The chat-bot's functional capabilities include providing information about the advantages of the educational program and the university, specific admission requirements for school and college graduates, features of studying at different academic levels, current tuition fees, scholarship and grant opportunities, and the option to subscribe to important updates. The request-handling mechanism, based on user session management and dedicated message handlers, enables a multi-stage interaction flow with logical transitions between dialogue states.

For administrative management, the system incorporates an interface built with the Swagger library, which provides automatically generated interactive API documentation and supports the execution of test requests directly through a web interface.

The results of the research confirm the effectiveness of the proposed approach in addressing current challenges related to the digitalization of higher education. The developed system significantly reduces the workload of administrative personnel, ensures applicants' round-the-clock access to essential information, and improves the overall quality of communication between the university and prospective students.

Key words: chatbot, telegram, java, spring.

Олексійчук Ю. Ф. Розробка та впровадження чат-бота для майбутніх студентів

Стаття присвячена актуальній проблемі цифровізації комунікаційних процесів у сфері вищої освіти, зокрема питанням розробки та впровадження інтелектуального чат-бота для взаємодії з майбутніми студентами освітньої програми з комп'ютерних наук. У контексті зростаючої конкуренції між закладами вищої освіти та необхідності забезпечення цілорічної інформаційної підтримки абітурієнтів обґрунтовується доцільність використання месенджера Telegram як платформи для автоматизованого консультування.

Проаналізовано основні недоліки традиційних каналів комунікації з потенційними вступниками, зокрема їхню часову та географічну обмеженість, фрагментованість інформації на різних веб-ресурсах, а також значне навантаження на співробітників приймальних комісій та викладачів.

Розроблено програмне рішення, яке реалізоване з використанням мови програмування Java та фреймворку Spring Boot. Система побудована за багатошаровою архітектурою з застосуванням патерну проектування Controller–Service–Repository, що забезпечує модульність, масштабованість та зручність супроводу програмного продукту. Для інтеграції з Telegram API використано бібліотеку Telegram Bots, а для зберігання даних – систему управління базами даних Oracle Database.

Функціональні можливості чат-бота включають надання інформації про переваги освітньої програми та університету, особливості вступу після школи та коледжу, характеристики навчання на різних освітніх рівнях, актуальну вартість навчання, стипендіальні та грантові програми, а також можливість підписки на важливі повідомлення.

Механізм обробки користувацьких запитів через систему обробників та управління сесіями користувачів дозволяє реалізувати багатоступеневу взаємодію з логічними переходами між різними станами діалогу.

Для адміністративного управління системою впроваджено інтерфейс на основі бібліотеки Swagger, який надає автоматично згенеровану інтерактивну документацію API та можливість виконання тестових запитів безпосередньо через веб-інтерфейс.

© Yu. F. Oleksiichuk, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

Результати дослідження підтверджують ефективність запропонованого підходу для вирішення актуальних викликів цифровізації вищої освіти. Розроблена система дозволяє суттєво знизити навантаження на адміністративний персонал, забезпечити абітурієнтам цілодобовий доступ до необхідної інформації та покращити загальну якість комунікацій між університетом та майбутніми студентами.

Ключові слова: чат-бот, telegram, java, spring.

Formulation of the problem. In the contemporary context of digitalization in higher education, universities are increasingly compelled to transform traditional approaches to communication with prospective applicants. Interaction with future students constitutes a continuous process that extends beyond the admission campaign period and encompasses the full cycle of career guidance activities throughout the year. Particularly in highly competitive fields such as computer science, cultivating a positive perception of the educational program and institution is essential for attracting motivated and talented applicants.

Conventional communication channels, including telephone calls to admissions offices, email correspondence, and participation in open house events, are constrained by temporal and geographical limitations. Prospective students, however, expect timely responses to inquiries concerning admission requirements, program characteristics, the advantages of bachelor's or master's degree studies, tuition fees, and scholarship opportunities at any time of the year, irrespective of office hours or geographical location.

Staff members of admissions offices and academic departments face a considerable informational burden, as they are required to provide repetitive responses to routine inquiries about the computer science program throughout the academic year. This situation diminishes the quality of services, increases response times, and results in inefficient use of human resources that could otherwise be allocated to addressing more complex and non-standard issues.

Additionally, information regarding admission is often fragmented across various web pages and formats, necessitating significant effort for searching and systematization. Applicants, particularly recent school and college graduates, may encounter difficulties navigating multiple information sources, which negatively influences their interaction experience with the university even prior to the admission stage.

The significance of this issue is further heightened by the growing competition among higher education institutions for talented students. Universities that succeed in providing a high level of service and convenient communication with prospective applicants secure a competitive advantage in the educational market. Consequently, the implementation of innovative digital interaction tools is becoming indispensable for maintaining institutional competitiveness and fostering a positive institutional image.

In the contemporary digital environment, chatbots [1] are becoming an increasingly widespread tool of communication among adolescents and students. In particular, a report by *Common Sense Media* (USA, spring–summer 2025), based on a representative sample of more than 1,000 adolescents aged 13–17, indicates that 72% of American teenagers have experimented with AI companions (e.g., Character.AI, Replika, ChatGPT), while 52% have become regular users. Among those who use chatbots regularly, 13% engage with them daily, and 21% interact with them several times per week [2].

Analysis of recent research and publications. Contemporary Ukrainian youth actively employ a variety of channels for accessing news and digital content, including Telegram (with chatbots), YouTube, Instagram, Viber, and traditional television. Despite the documented security risks associated with the use of Telegram during wartime [3], as long as this messenger remains unrestricted, its popularity and potential applications across diverse domains must be considered. These include finance [4], education [5–9], the tourism industry [9], complaint management [10], system monitoring [11–12], and healthcare [13], among others. Telegram bots provide a user-friendly interface, and their underlying logic can be implemented in a wide range of programming languages [14]. Owing to this simplicity, Telegram is frequently utilized as an interface for the Internet of Things [12, 15–16].

The purpose of this article is the design and implementation of a Telegram bot for prospective applicants to the Computer Science program. The central idea is to provide only the most essential information in a concise and accessible form. The key functionalities offered to applicants include:

- obtaining information about the advantages of the educational program and the university;
- learning about the specifics of admission after college and after high school;
- exploring the features of bachelor's and master's studies;
- requesting additional information and submitting questions;
- accessing up-to-date information on tuition fees, scholarship and grant opportunities, as well as possibilities for free education;
- subscribing to important announcements and notifications.

Presenting main material. The development of the chatbot was carried out using the Java programming language in combination with the Spring Boot framework. Database operations are managed through Spring Data JPA, while security and access control are ensured by Spring Security. The processing of API requests is implemented using Spring Web, thus forming a robust and modular backend architecture.

To facilitate integration with the Telegram API, the Telegram Bots library is employed [17]. This library provides a lightweight yet effective Java framework for seamless interaction with the Telegram Bots API, thereby

streamlining the development of intelligent software agents. Its functionality encompasses built-in mechanisms for token management, the use of the Jetty HTTP client for reliable communication, and extensive documentation that supports efficient implementation and maintenance.

To store data, the Oracle Database is employed [18-19]. The database is used exclusively for preserving information about users who have expressed interest in the specialization. The Entity–Relationship (ER) diagram of the database is presented in Figure 1.

The Tg_Users table contains information about users obtained via the Telegram API [20]. The key attribute is `userId`, which is used for user identification. In addition, the Telegram API provides access to the user's nickname, first and last name as registered in Telegram, and information indicating whether the account belongs to a bot. However, this information is often incomplete or hidden by users.

The Connections table stores records of all interactions with the Telegram bot. These data are applied to analyze the popularity of the software product.

The Lead table contains information about user activities that require further processing by the department staff. This includes requests for assistance or specific questions that the chatbot could not resolve.

The Subscriber table stores information about Telegram subscribers. Subscribers receive important updates regarding university admission directly through Telegram.

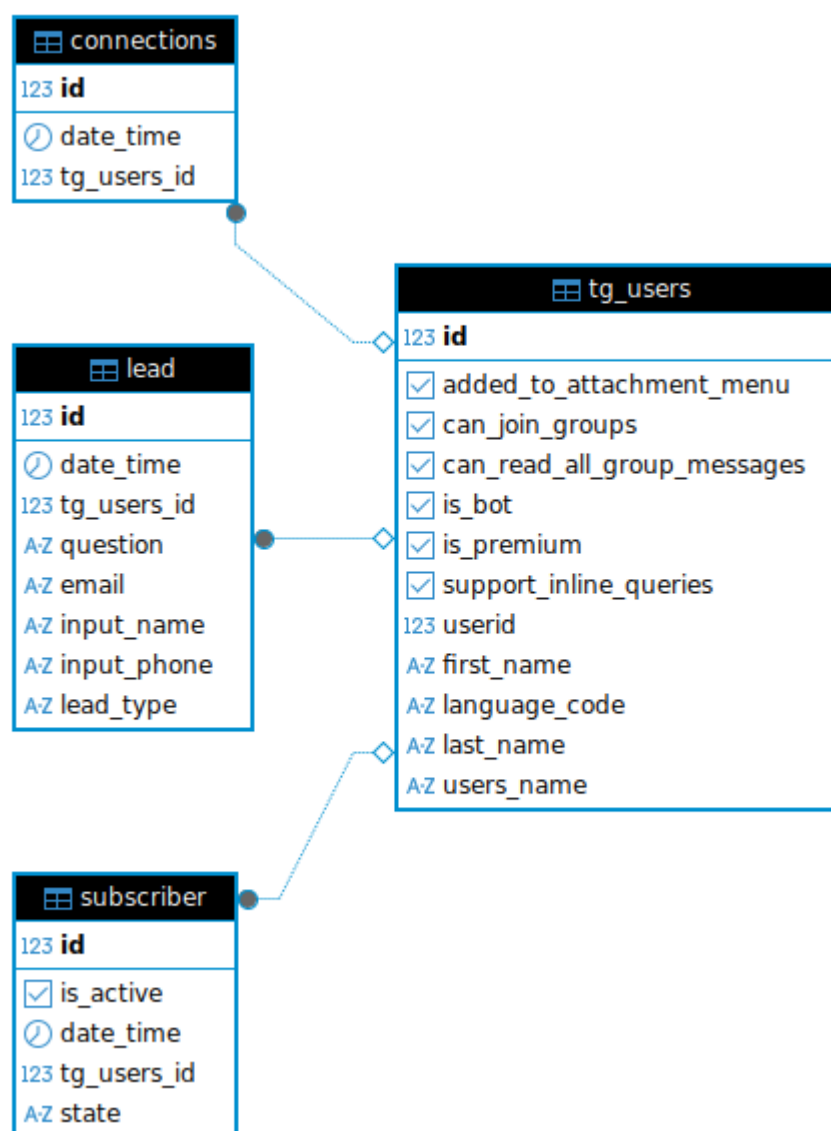


Fig. 1. ER diagram of the database

The chatbot logic is implemented within the bot package. The TelegramBot class provides overall control of the chatbot and extends the TelegramLongPollingBot class from the Telegram Bots library [17]. Specifically, this class overrides the `onUpdateReceived(Update update)` method, which is responsible for processing user messages.

Upon receiving the initial message from a user, an object of the `UserSession` class is created, which stores information regarding the current stage of the user's interaction.

```
@Override
public void onUpdateReceived(Update update) {
    Long chatId;
    User user;
    if(update.hasMessage() && update.getMessage().hasText()) {
        user = update.getMessage().getFrom();
        chatId = update.getMessage().getChatId();
    } else if(update.hasCallbackQuery()
        && update.getCallbackQuery().getData()!=null){
        user = update.getCallbackQuery().getFrom();
        chatId = update.getCallbackQuery().getMessage().getChatId();
    }
    else {
        log.warn("Unexpected update from user. Not message or not text");
        return;
    }
    UserSession session = userSessionService.getSession(chatId);
    userSessionService.saveSession(chatId,session);
    UserRequest userRequest = UserRequest
        .builder()
        .update(update)
        .userSession(session)
        .chatId(chatId)
        .user(user)
        .build();
    boolean dispatched = dispatcher.dispatch(userRequest);
    if (!dispatched) {
        if (session.getState().isAllowText()){
            textRequestHandler.handle(userRequest);
        }else {
            log.warn("Unexpected update from user");
        }
    }
}
```

At each stage of interaction, the user receives specific textual information along with a set of buttons that facilitate further navigation. The set of buttons is stage-dependent and is generated through the `KeyboardHelper` class. In certain stages, users are also permitted to provide textual input, for example, when submitting a question or entering their name.

Each stage is associated with a dedicated `Handler`, responsible for processing user messages, which extends the abstract class `UserRequestHandler`.

```
public abstract class UserRequestHandler {
    public abstract boolean isApplicable(UserRequest request);
    public abstract void handle(UserRequest dispatchRequest);
    public abstract boolean isGlobal();
    public boolean isCommand(Update update, String command) {
        return (update.hasMessage() && update.getMessage().isCommand()
            && update.getMessage().getText().equals(command))
            || (update.hasCallbackQuery() && update.getCallbackQuery().getData()!=null
            && update.getCallbackQuery().getData().equals(command));
    }
}
```

Within a `Handler`, the content delivered to the user at the current stage is defined, along with the permissible set of commands. Each command typically triggers a transition to another stage of the interaction.

Structurally, the code of the web service follows a multilayered architecture, with each layer responsible for a specific functionality. The application implements the `Controller–Service–Repository` design pattern, which is widely adopted in modern enterprise software development to ensure modularity, maintainability, and scalability [21-23].

The Controller layer is responsible for handling incoming HTTP requests and mapping them to the corresponding application functionality. Within this layer, input data is validated and converted into domain-specific structures before being passed further for processing. Controllers serve as the main entry point to the system, enabling interaction between end users and the application logic.

The Service layer encapsulates the core business logic. It coordinates operations, enforces domain rules, and ensures the correct execution of workflows. Centralizing business functionality in services promotes separation of concerns, improves system clarity, and simplifies debugging, testing, and future extension.

The Repository layer manages data persistence and retrieval. It abstracts the database operations by providing a standardized interface for entity management. In this project, repositories rely on Spring Data JPA, which minimizes the amount of boilerplate code, automatically generates query implementations, and integrates seamlessly with relational databases.

The chosen framework, Spring Boot, plays a crucial role in supporting this architecture. It provides built-in mechanisms for dependency injection, configuration management, and seamless integration of various Spring modules, such as Spring Web, Spring Data JPA, and Spring Security. This ensures that the multilayered design is not only logically consistent but also technically efficient, with reduced development overhead and enhanced system reliability.

This layered approach increases modularity and fosters reusability of components, while reducing coupling between the different parts of the system. As a result, the overall architecture achieves higher robustness, maintainability, and adaptability to future requirements.

For the purpose of informing administrators, the system implements functionality for automatically sending both statistical reports and user-submitted messages that require attention to designated Telegram addresses. This ensures that administrators receive timely notifications without the need for manual monitoring of the system.

The delivery of this information is automated through the use of the scheduling mechanism provided by the Spring Framework, specifically the `@Scheduled` annotation available in Spring Boot [24]. This annotation allows developers to define tasks that are executed periodically or at specific times, following either fixed intervals, fixed delays, or cron expressions. As a result, the application can be configured to send summaries of user activity, error logs, or other relevant updates at predefined times, thereby facilitating efficient monitoring and decision-making.

One of the key advantages of using the Spring scheduling mechanism is its flexibility. Schedules can be easily adjusted through configuration parameters, enabling administrators to change the frequency of reporting without modifying the underlying source code. Moreover, multiple scheduled tasks can coexist within the application, allowing for the automation of diverse processes, such as data synchronization, report generation, and system health checks.

Another significant feature of the `@Scheduled` mechanism is its integration with Spring's dependency injection and transaction management. This ensures that scheduled tasks operate consistently within the broader application context, maintaining access to necessary services, repositories, and security mechanisms. Furthermore, in production environments, scheduling can be combined with asynchronous execution and thread pool management, thereby improving system scalability and preventing long-running tasks from blocking critical operations.

Overall, the use of scheduled tasks in Spring Boot provides a reliable and extensible mechanism for automating routine processes. In the context of this project, it enables administrators to remain informed about user activity and system status, while simultaneously reducing the workload associated with manual monitoring and repetitive tasks. This contributes to higher system efficiency, improved responsiveness to user needs, and enhanced overall quality of service.

An example of the practical implementation of scheduled tasks in this project is illustrated in the code fragment below. The class `LeadsScheduler` is annotated with `@Component`, which makes it a managed Spring Bean and enables its automatic detection during application context initialization. The `@Scheduled` annotation is applied to the `execute()` method, specifying a cron expression that determines the exact schedule of task execution. In this case, the method is configured to run daily at 14:10.

```
@Component
@RequiredArgsConstructor
@Slf4j
public class LeadsScheduler {
    private final LeadsScheduledService leadsScheduledService;
    @Scheduled(cron = "0 10 14 * * ?")
    public void execute() {
        try {
            leadsScheduledService.sendAdminDailyInfo();
        }
        catch (Exception e){
            log.error(e.getMessage());
        }
    }
}
```

The method delegates the execution logic to the service layer (LeadsScheduledService), which encapsulates the functionality of preparing and sending daily information to administrators. This approach follows the principles of layered architecture by separating scheduling responsibilities from business logic. For fault tolerance, exception handling is incorporated, and errors are logged using the SLF4J logging framework [25]. Such an implementation ensures both reliability and maintainability of the scheduling mechanism.

For administrative purposes, a dedicated management interface was implemented using the Swagger library [26]. This tool was selected because it provides the capability to rapidly establish a functional administrative environment with minimal configuration effort: integration into the project requires only the inclusion of the library and a small amount of additional code.

Beyond its ease of integration, Swagger offers significant advantages for system maintenance and oversight. It provides automatically generated and interactive API documentation, allowing administrators not only to review available endpoints but also to execute test requests directly through the interface. This functionality simplifies validation, debugging, and monitoring of the system's behavior in real time.

Although its interface is relatively straightforward, Swagger ensures that administrators can perform all essential management operations effectively.

Conclusions. The conducted research demonstrates the effectiveness of using digital tools, in particular Telegram chat-bots, for improving communication between higher education institutions and prospective students. The developed system provides applicants with quick access to essential information about admission requirements, study opportunities, tuition fees, and scholarship programs, thereby reducing the workload on administrative staff and enhancing the quality of interaction with the university.

The implemented architecture, based on Java and Spring Boot, ensures a modular and extensible structure of the software solution. Integration with the Telegram Bots API and the use of supporting technologies such as Spring Data JPA, Spring Security, and scheduled task execution mechanisms allow for reliable and efficient functionality. The inclusion of an administrative interface based on Swagger additionally increases system transparency and facilitates management processes.

The results confirm that the proposed approach is an effective means of addressing current challenges in the digitalization of higher education. Future work will focus on expanding system capabilities, including the integration of AI-based systems into the chat-bot, which would enable more intelligent and personalized interaction with users. Such improvements are expected to further increase the attractiveness of the educational program and strengthen the competitive position of the university.

Bibliography:

1. Adamopoulou E., Moussiades L. An overview of chatbot technology. In *IFIP international conference on artificial intelligence applications and innovations* (pp. 373–383). Cham: Springer International Publishing. 2020. https://doi.org/10.1007/978-3-030-49186-4_31
2. Perez S. 72% of U.S. teens have used AI companions, study finds. TechCrunch. 2025. URL: <https://techcrunch.com/2025/07/21/72-of-u-s-teens-have-used-ai-companions-study-finds>
3. Balovsyak N. Anonymous and official Telegram channels in Ukraine: analysis of popularity during the hybrid war. *Current Issues of Mass Communication*, 37, 2025. P. 30–42. <https://doi.org/10.17721/CIMC.2025.37.30-42>
4. Haitan O. M., Snytkal V. Integrated platforms for automating personal financial accounting based on chatbots and cloud technologies. *Systems and Technologies*, 2025, 69(1). P. 58–70. <https://doi.org/10.32782/2521-6643-2025-1-69.7>
5. Olhovska O. V., Chernenko O. O., Ananenko I. V., Parfonova T. O., Rudenko N. S. Development of a training simulator for system analysis in the form of a chat-bot. *Visnyk of Kherson National Technical University*, (2), 2023. P. 196–202. <https://doi.org/10.35546/kntu2078-4481.2023.2.27>
6. Ismawati D., Prasetyo I. The development of Telegram bot through short story. In *Brawijaya International Conference on Multidisciplinary Sciences and Technology (BICMST 2020)* (pp. 209–212). Atlantis Press. <https://doi.org/10.2991/assehr.k.201021.049>
7. Rianto R., Rahmatulloh A., Firmansah T. A. Telegram Bot Implementation in Academic Information Services with The Forward Chaining Method. *Sinkron : Jurnal Dan Penelitian Teknik Informatika*, 3(2), 2019. P. 73–78. <https://doi.org/10.33395/sinkron.v3i2.10023>
8. Aisyah R. N., Istiqomah D. M., Muchlisin M. Developing e-learning module by using telegram bot on ICT for ELT course. In *5th International Conference on Arts Language and Culture (ICALC 2020)* (pp. 106–111). Atlantis Press. <https://doi.org/10.2991/assehr.k.210226.054>
9. Olkhovska O. V., Oleksiichuk Y. F., Koshova O. P., Chernenko O. O., Boiko O. A. Development of a telegram chat-bot for providing technical support in the field of tourist services. *Taurida Scientific Herald. Series: Technical Sciences*, (6), 2024. P. 35–44. <https://doi.org/10.32782/tnv-tech.2023.6.5>
10. Rosid M. A., Rachmadany A., Multazam M. T., Nandiyanto A. B. D., Abdullah A. G., Widiaty, I. Integration telegram bot on e-complaint applications in college. In *IOP conference series: Materials Science and Engineering* (Vol. 288, No. 1, p. 012159). IOP Publishing. 2018. <https://doi.org/10.1088/1757-899X/288/1/012159>

11. Idhom M., Fauzi A., Alit R., Wahanani H. E. Implementation system telegram bot for monitoring Linux server. In *International conference on science and technology (ICST 2018)* (pp. 1089–1093). Atlantis Press. <https://doi.org/10.2991/icst-18.2018.219>
12. Bestari D. N., Wibowo A. An IoT-Based Real-Time Weather Monitoring System Using Telegram Bot and Thingsboard Platform. *International Journal of Interactive Mobile Technologies*, 17(6). 2023. <https://doi.org/10.3991/ijim.v17i06.34129>
13. Djoelianto A. D., Kautsar I. A., Rosid M. A. Development of Web Service and Telegram Bot for Location-Based Health Service Information System. *Procedia of Engineering and Life Science*, 2(2). 2022. <https://doi.org/10.21070/pels.v2i2.1280>
14. Modrzyk N. *Building telegram bots: develop bots in 12 programming languages using the telegram bot API*. Apress. 2018. <https://doi.org/10.1007/978-1-4842-4197-4>
15. De Oliveira J. C., Santos D. H., Neto M. P. Chatting with Arduino platform through telegram bot. In *2016 IEEE International Symposium on Consumer Electronics (ISCE)* (pp. 131–132). IEEE. <https://doi.org/10.1109/ISCE.2016.7797406>
16. Zaid, M. I. M. A., Abdullah, R., Ismail, S. I., Dzulkefli, N. N. S. N. IoT-based emergency alert system integrated with telegram bot. In *2023 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS)* (pp. 126–131). IEEE. <https://doi.org/10.1109/I2CACIS57635.2023.10193550>
17. Bermudez, R. *TelegramBots (v.6.7.0)* [Computer software]. 2023. GitHub. URL: <https://github.com/rubenlagus/TelegramBots>
18. Kothuri R., Godfrind A., Beinat E. Pro oracle spatial for oracle database 11g. Apress. 2007. URL: <https://link.springer.com/book/9781430242871>
19. Maksymchuk S., Kabak L., Moroz B. Using of the modern data mining technics in customs of Ukraine. *Systems and Technologies*, 2(58), 2019. P. 33–49. <https://doi.org/10.32836/2521-6643-2019-2-58-2>
20. Khaund T., Hussain M. N., Shaik M., Agarwal N. Telegram: Data collection, opportunities and challenges. In *Annual international conference on information management and big data* (pp. 513–526). Cham: Springer International Publishing. 2020. https://doi.org/10.1007/978-3-030-76228-5_37
21. Fowler M. *Patterns of Enterprise Application Architecture*. Addison-Wesley Longman Publishing Co., Inc., USA. 2002. URL: <https://dl.acm.org/doi/abs/10.5555/579257>
22. Oleksiichuk Y. F., Olkhovska O. V., Olkhovsky D. M., Orlova D. I. Design and development of a web service for generating and sending pdf documents. *Systems and Technologies*, 65(1), 2023. P. 39–45. <https://doi.org/10.32782/2521-6643-2023.1-65.5>
23. Walls C. *Spring in action*. Simon and Schuster. 2022.
24. Krnac L. *Tasks and Scheduling*. In: *Pivotal Certified Spring Enterprise Integration Specialist Exam*. Apress, Berkeley, CA. 2015. https://doi.org/10.1007/978-1-4842-0793-2_1
25. Juneau J. *Exceptions and Logging*. In: *Java 9 Recipes*. Apress, Berkeley, CA. 2017. https://doi.org/10.1007/978-1-4842-1976-8_9
26. Dos Santos J. S., Azevedo L. G., Soares E. F., Thiago R. M., da Silva V. T. Analysis of Tools for REST Contract Specification in Swagger/OpenAPI. In *ICEIS (2)* (pp. 201–208). 2020. <https://doi.org/10.5220/0009381202010208>

References:

1. Adamopoulou, E., & Moussiades, L. (2020, May). An overview of chatbot technology. In *IFIP international conference on artificial intelligence applications and innovations* (pp. 373–383). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-49186-4_31
2. Perez, S. (2025, July 21). 72% of U.S. teens have used AI companions, study finds. TechCrunch. Retrieved from: <https://techcrunch.com/2025/07/21/72-of-u-s-teens-have-used-ai-companions-study-finds>
3. Balovskyak, N. (2025). Anonymous and official Telegram channels in Ukraine: analysis of popularity during the hybrid war. *Current Issues of Mass Communication*, 37, 30–42. <https://doi.org/10.17721/CIMC.2025.37.30-42>
4. Haitan, O. M., & Snytko, I. V. (2025). Integrated platforms for automating personal financial accounting based on chatbots and cloud technologies. *Systems and Technologies*, 69(1), 58–70. <https://doi.org/10.32782/2521-6643-2025-1-69.7>
5. Olhovska, O. V., Chernenko, O. O., Ananenko, I. V., Parfonova, T. O., & Rudenko, N. S. (2023). Development of a training simulator for system analysis in the form of a chat-bot. *Visnyk of Kherson National Technical University*, (2), 196–202. <https://doi.org/10.35546/kntu2078-4481.2023.2.27>
6. Ismawati, D., & Prasetyo, I. (2020, October). The development of Telegram bot through short story. In *Brawijaya International Conference on Multidisciplinary Sciences and Technology (BICMST 2020)* (pp. 209–212). Atlantis Press. <https://doi.org/10.2991/assehr.k.201021.049>
7. Rianto, R., Rahmatulloh, A., & Firmansah, T. A. (2019). Telegram Bot Implementation in Academic Information Services with The Forward Chaining Method. *Sinkron : Jurnal Dan Penelitian Teknik Informatika*, 3(2), 73–78. <https://doi.org/10.33395/sinkron.v3i2.10023>

-
8. Aisyah, R. N., Istiqomah, D. M., & Muchlisin, M. (2021, March). Developing e-learning module by using telegram bot on ICT for ELT course. In *5th International Conference on Arts Language and Culture (ICALC 2020)* (pp. 106–111). Atlantis Press. <https://doi.org/10.2991/assehr.k.210226.054>
 9. Olkhovska, O. V., Oleksiichuk, Y. F., Koshova, O. P., Chernenko, O. O., & Boiko, O. A. (2024). Development of a telegram chat-bot for providing technical support in the field of tourist services. *Taurida Scientific Herald. Series: Technical Sciences*, (6), 35–44. <https://doi.org/10.32782/tnv-tech.2023.6.5>
 10. Rosid, M. A., Rachmadany, A., Multazam, M. T., Nandiyanto, A. B. D., Abdullah, A. G., & Widiaty, I. (2018). Integration telegram bot on e-complaint applications in college. In *IOP conference series: Materials Science and Engineering* (Vol. 288, No. 1, p. 012159). IOP Publishing. <https://doi.org/10.1088/1757-899X/288/1/012159>
 11. Idhom, M., Fauzi, A., Alit, R., & Wahanani, H. E. (2018, December). Implementation system telegram bot for monitoring Linux server. In *International conference on science and technology (ICST 2018)* (pp. 1089–1093). Atlantis Press. <https://doi.org/10.2991/icst-18.2018.219>
 12. Bestari, D. N., & Wibowo, A. (2023). An IoT-Based Real-Time Weather Monitoring System Using Telegram Bot and Thingsboard Platform. *International Journal of Interactive Mobile Technologies*, 17(6). <https://doi.org/10.3991/ijim.v17i06.34129>
 13. Djoelianto, A. D., Kautsar, I. A., & Rosid, M. A. (2022). Development of Web Service and Telegram Bot for Location-Based Health Service Information System. *Procedia of Engineering and Life Science*, 2(2). <https://doi.org/10.21070/pels.v2i2.1280>
 14. Modrzyk, N. (2018). *Building telegram bots: develop bots in 12 programming languages using the telegram bot API*. Apress. <https://doi.org/10.1007/978-1-4842-4197-4>
 15. De Oliveira, J. C., Santos, D. H., & Neto, M. P. (2016, September). Chatting with Arduino platform through telegram bot. In *2016 IEEE International Symposium on Consumer Electronics (ISCE)* (pp. 131–132). IEEE. <https://doi.org/10.1109/ISCE.2016.7797406>
 16. Zaid, M. I. M. A., Abdullah, R., Ismail, S. I., & Dzulkefli, N. N. S. N. (2023, June). IoT-based emergency alert system integrated with telegram bot. In *2023 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS)* (pp. 126–131). IEEE. <https://doi.org/10.1109/I2CACIS57635.2023.10193550>
 17. Bermudez, R. (2023). *TelegramBots (v.6.7.0)* [Computer software]. GitHub. <https://github.com/rubenlagus/TelegramBots>
 18. Kothuri, R., Godfrind, A., & Beinat, E. (2007). *Pro oracle spatial for oracle database 11g*. Apress. Retrieved from: <https://link.springer.com/book/9781430242871>
 19. Maksymchuk, S., Kabak, L., & Moroz, B. (2019). Using of the modern data mining technics in customs of Ukraine. *Systems and Technologies*, 2(58), 33–49. <https://doi.org/10.32836/2521-6643-2019-2-58-2>
 20. Khaund, T., Hussain, M. N., Shaik, M., & Agarwal, N. (2020, October). Telegram: Data collection, opportunities and challenges. In *Annual international conference on information management and big data* (pp. 513–526). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-76228-5_37
 21. Fowler, M. (2002). *Patterns of Enterprise Application Architecture*. Addison-Wesley Longman Publishing Co., Inc., USA. Retrieved from: <https://dl.acm.org/doi/abs/10.5555/579257>
 22. Oleksiichuk, Y. F., Olkhovska, O. V., Olkhovsky, D. M., & Orlova, D. I. (2023). Design and development of a web service for generating and sending pdf documents. *Systems and Technologies*, 65(1), 39–45. <https://doi.org/10.32782/2521-6643-2023.1-65.5>
 23. Walls, C. (2022). *Spring in action*. Simon and Schuster.
 24. Krnac, L. (2015). Tasks and Scheduling. In: *Pivotal Certified Spring Enterprise Integration Specialist Exam*. Apress, Berkeley, CA. https://doi.org/10.1007/978-1-4842-0793-2_1
 25. Juneau, J. (2017). Exceptions and Logging. In: *Java 9 Recipes*. Apress, Berkeley, CA. https://doi.org/10.1007/978-1-4842-1976-8_9
 26. Dos Santos, J. S., Azevedo, L. G., Soares, E. F., Thiago, R. M., & da Silva, V. T. (2020). Analysis of Tools for REST Contract Specification in Swagger/OpenAPI. In *ICEIS (2)* (pp. 201–208). <https://doi.org/10.5220/0009381202010208>

Дата надходження статті: 21.10.2025

Дата прийняття статті: 10.11.2025

Опубліковано: 30.12.2025

Погорєловська І. Д., кандидат економічних наук, доцент,
доцент кафедри комп'ютерних та інформаційних технологій
і систем Державного податкового університету
ORCID: 0000-0001-5026-7131

Горбовий А. Ю., доктор технічних наук, професор,
професор кафедри комп'ютерних та інформаційних технологій
і систем Державного податкового університету
ORCID: 0000-0002-2155-3910

Гладченко О. В., кандидат педагогічних наук, доцент,
доцент кафедри комп'ютерних та інформаційних технологій
і систем Державного податкового університету
ORCID: 0000-0002-1668-633X

Ратушняк Т. В., кандидат фізико-математичних наук, доцент,
доцент кафедри комп'ютерних та інформаційних технологій
і систем Державного податкового університету
ORCID: 0000-0002-9353-4268

КОНТРАКТИ ЯК ЕЛЕМЕНТ ЗАБЕЗПЕЧЕННЯ ЯКОСТІ ІТ-ПРОДУКТІВ

Розглянуто важливість високоякісних ІТ-продуктів у сучасному цифровому світі; проаналізовано серію міжнародних стандартів SQuaRE, в якій визначені вимоги до якості систем та програмного забезпечення, до вимірювання якості та рекомендації щодо оцінювання програмного продукту; наголошено, що забезпечення якості ІТ-продуктів – це не лише про моделі якості, настанови щодо оцінювання функціоналу, перевірку коду чи тестування продукту, але й про забезпечення високої якості виконання всіх етапів життєвого циклу продукту, починаючи від формування вимог замовника і до отримання кінцевого результату; виконано аналіз досліджень, присвячених застосуванню контрактних моделей у програмній інженерії; вказано на особливості договорів у сфері ІТ, які є кастомним або вузькоспеціалізованим, через специфіку предмета договору; визначено зміст контракту, який забезпечує йому роль механізму контролю якості ІТ-продуктів, а саме: вимоги, стандарти якості, обсяг роботи та порядок розрахунків її вартості, терміни виконання, критерії приймання результатів, ключові показники ефективності, формальні процедури перевірки, які надають можливість контролювати якість на кожному етапі, правила управління змінами та відповідальність сторін за порушення зобов'язань, а також швидкість розробки продукту, ролі в команді, систему обліку дефектів, сесії уточнення беклогу, методи оцінювання складності виконання задач; досліджено змістові характеристики та відмінності між контрактними моделями Fixed Price, Time&Materials, Dedicated Team та їх вплив на забезпечення якості реалізації ІТ-проектів; зроблено висновки про те, що при виборі типу контрактної моделі варто враховувати специфіку ІТ-проекту та доступний бюджет, кількість невизначеностей у проекті та готовність до ризику; окреслено перспективи подальших досліджень – потенціал Smart contracts, які сьогодні інтегруються в різні напрями галузі ІТ.

Ключові слова: ІТ-продукти, ІТ-проекти, забезпечення якості, стандарти якості, моделі якості, контрактні моделі.

Pogorielovska I. D., Horbovyi A. Yu., Hladchenko O. V., Ratushnyak T. V. Contracts as an element of quality assurance of IT products

The importance of high-quality IT products in the modern digital world is considered; the SQuaRE series of international standards is analyzed, which defines requirements for the quality of systems and software, for quality measurement and recommendations for evaluating a software product; it is emphasized that ensuring the quality of IT products is not only about quality models, guidelines for evaluating functionality, checking code or testing a product, but also about ensuring high quality of execution of all stages of the product life cycle, starting from the formation of customer requirements and until obtaining the final result; an analysis of studies devoted to the application of contract models in software engineering is performed; the features of contracts in the IT sector, which are custom or highly specialized, due to the specifics of the subject of the contract, are indicated; the content of the contract, which provides it with the role of a quality control mechanism for IT products, is defined.

© І. Д. Погорєловська, А. Ю. Горбовий, О. В. Гладченко, Т. В. Ратушняк, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

namely: requirements, quality standards, scope of work and the procedure for calculating its cost, deadlines, criteria for accepting results, key performance indicators, formal verification procedures that provide the opportunity to control quality at each stage, change management rules and the responsibility of the parties for breach of obligations, as well as the speed of product development, roles in the team, defect accounting system, backlog clarification sessions, methods for assessing the complexity of task performance; the content characteristics and differences between the Fixed Price, Time&Materials, Dedicated Team contract models and their impact on ensuring the quality of IT project implementation are investigated; conclusions are drawn that when choosing the type of contract model, it is necessary to take into account the specifics of the IT project and the available budget, the number of uncertainties in the project and the willingness to take risk; prospects for further research are outlined – the potential of Smart contracts, which are currently being integrated into various areas of the IT industry.

Key words: IT products, IT projects, quality assurance, quality standards, quality models, contract models.

Постановка проблеми. Високоякісні ІТ-продукти у сучасному світі відіграють надзвичайно важливу роль, оскільки саме вони забезпечують перехід від традиційних процесів до цифрових – у бізнесі, освіті, охороні здоров'я, державному управлінні, космічній галузі, військовій сфері. ІТ-продукти автоматизують рутинні операції, скорочують витрати часу й ресурсів, забезпечують аналітику для прийняття обґрунтованих управлінських рішень, забезпечують захист даних, кібербезпеку та стабільну роботу критичних систем. Вони стимулюють інновації, створюють нові бізнес-моделі (наприклад, електронна комерція, хмарні сервіси, штучний інтелект) та підвищують конкурентоспроможність компаній.

Водночас, поширення та зростання складності інформаційних систем та програмного забезпечення, питання безпеки та стабільності їх роботи зумовлюють актуальність теми забезпечення їх якості.

Питання якості ІТ викликають інтерес щонайменше з 1968 року, коли на конференції NASA проблеми “поганого” програмного забезпечення того часу були названі Програмною кризою [1]. Автор також зазначає, що якість ІТ – це проблематичне явище з багатьма складовими: якістю програмного забезпечення, якістю використання, якістю інформаційних систем та економічними аспектами ІТ-інвестицій.

В умовах сьогодення, забезпечення якості є не лише важливою, але й обов'язковою умовою розробки ІТ-продуктів, яка спрямована на запобігання дефектів у їх роботі, вдосконалені процесів їх створення та тестування.

Продукти високої якості мають більшу цінність, кращу підтримку і швидше адаптуються до змін вимог користувачів. Вони також формують позитивний досвід користувачів, що підвищує їхню лояльність до розробників та сприяє зростанню їх іміджу. А помилки або збої в роботі можуть призвести як до втрати клієнтів, так і фінансових збитків. Тому саме високоякісні ІТ-продукти є важливими для забезпечення цінності та уникнення потенційних негативних наслідків для зацікавлених сторін.

Водночас, все це не так абстрактно, як може здаватися на перший погляд. Існує серія міжнародних стандартів, які визначають вимоги до якості систем та програмного забезпечення та її оцінювання – Systems and software Quality Requirements and Evaluation (SQuaRE) [2]. Кожна складова SQuaRE (рис. 1) надає рекомендації та визначає ресурси для виконання різних функцій із забезпечення якості систем та програмних продуктів. Зокрема:

- ISO/IEC 2500n – міжнародні стандарти, що визначають усі загальні моделі, терміни, на які посиляються всі інші міжнародні стандарти із сімейства SQuaRE.

- ISO/IEC 2501n – розділ моделей якості, що поєднує: ISO/IEC 25010 – модель якості ПЗ та систем; ISO/IEC TS 25011 – моделі якості послуг; ISO/IEC 25012 – модель якості даних; ISO/IEC 25019 – модель якості у використанні.

- ISO/IEC 2502n – міжнародні стандарти, що визначають систему вимірювання якості, математичні показники якості та практичні рекомендації щодо їх застосування.

- ISO/IEC 2503n – міжнародні стандарти, що визначають вимоги до якості на основі моделей якості та показників якості. Ці вимоги можуть бути використані в процесі визначення вимог до якості для інформаційних систем та ІТ-послуг, що розробляються, або як вхідні дані для процесу оцінювання.

- ISO/IEC 2504n – міжнародні стандарти, що містять вимоги, рекомендації та настанови щодо оцінювання програмного продукту, незалежно від того, чи виконується воно оцінювачами, покупцями чи розробниками.

- ISO/IEC 25050 – ISO/IEC 25099 – розділ розширення SQuaRE. Ці міжнародні стандарти наразі включають вимоги до якості готового до використання програмного продукту та інструкції з тестування, загальний галузевий формат для звітів про зручність використання, а також моделі та показники якості для нових технологій, таких як хмарні сервіси та штучний інтелект.

У цілому, SQuaRE визначає єдину структуру для встановлення вимог до якості ПЗ і систем, оцінювання якості на різних етапах життєвого циклу, забезпечення узгодженості між розробниками, замовниками та користувачами.

Водночас, забезпечення якості ІТ-продуктів – це не лише про моделі якості, настанови щодо оцінювання функціоналу, перевірку коду чи тестування продукту, але й про забезпечення високої якості виконання всіх етапів життєвого циклу продукту, починаючи від формування вимог замовника і до отримання кінцевого

результату. Одним із ключових інструментів тут виступають контракти, які фіксують очікування, вимоги та обов'язки всіх зацікавлених сторін. Існує декілька типів контрактів, які знаходять застосування у сфері ІТ, серед яких вирізняють Fixed Price (договір про фіксовану вартість), Time-and-Materials (договір про час і матеріали), Dedicated Team (виокремлена команда – договір про формування команди під потреби клієнта) [3, 4]. Використовується також модель взаємодії між замовником і виконавцем Outstaffing (зовнішній персонал – договір про оренду фахівців для виконання завдань клієнта), яка є схожою на Dedicated Team [5].

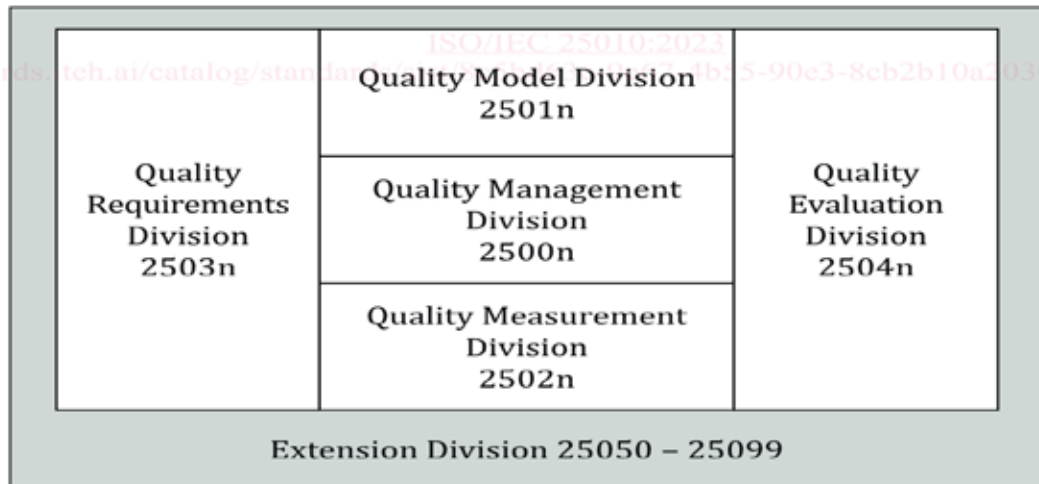


Рис. 1. Організація сімейства міжнародних стандартів SQaRE [2]

Відносно новою моделю співпраці у галузі ІТ є гіг-контракти. Ця контрактна модель створена державою як проміжний варіант між ФОП та офіційним працевлаштуванням, відповідно до Закону України «Про стимулювання розвитку цифрової економіки в Україні» [6].

Тому важливим є розуміння переваг і недоліків кожної контрактної моделі, що може суттєво вплинути як на їх вибір так і на якість реалізації всіх процесів створення ІТ-продукту.

Аналіз останніх досліджень і публікацій. Проведений аналіз свідчить про достатньо велику кількість досліджень, присвячених вибору контрактних моделей у програмній інженерії. Зокрема, в роботі [7] проаналізовано вплив розміру/складності проєкту на вибір контрактного механізму (Fixed-Price vs Time-and-Materials), наведено корисні емпіричні й теоретичні аргументи, чому великі/складні проєкти частіше реалізуються за моделлю Time-and-Materials. У статті обґрунтовано застосування моделей типу Dedicated Team для довготривалих/невизначених проєктів. Але не зазначено як саме ці контрактні моделі впливають на якість отриманих результатів реалізації ІТ-проєктів.

Систематично розглянуті типи контрактів, що застосовуються в Agile-проєктах у роботі [8]. Стаття містить огляд літератури й емпіричні висновки про те, як вибирають модель контракту залежно від невизначеності проєкту й ризиків. Але не досліджено як саме контрактні моделі впливають на забезпечення якості виконання цих проєктів.

Результати дослідження [9] свідчать, що приблизно 62% клієнтів віддають перевагу контрактам Fixed Price через передбачуваність, яку вони пропонують, тоді як близько 38% обирають контракти типу Time-and-Materials, що забезпечує виставлення рахунків відповідно до спожитих ресурсів, через його гнучкість у врахуванні змінних вимог. Водночас, близько 70% розробників зазначають, що модель Fixed Price може обмежувати креативність та адаптивність під час виконання проєкту. З іншого боку, контракти Time-and-Materials, що передбачають оперативне виставлення рахунків, сприяють інноваціям, але можуть призвести до перевитрат бюджету, тому потребують ретельного контролю. Авторка зазначає, що 56% проєктів, що використовують цю модель, стикаються з проблемою збільшення бюджету.

У дослідженні [1], що фокусується на тому, як на практиці сприймається і реалізується якість ІТ-продуктів, наголошено на важливості дотримання стандартів, врахуванні досвіду команди та контексту проєкту, але не розглядається вплив типів контрактів на процеси забезпечення якості ІТ-продуктів. Там же зазначено, що розробники сприймають якість не лише як дотримання функціональних вимог або технічної специфікації, але й розширюють це поняття, додаючи зручність для користувача, підтримуваність, ясність коду, здатність продукту змінюватися, ефективність, довговічність.

Мета статті: визначення впливу контрактних моделей Fixed Price, Time-and-Materials, Dedicated Team на забезпечення якості розробки ІТ-продуктів.

Виклад основного матеріалу. За визначенням, наведеним у [10, ст. 21, ч. 3], “контракт є особливою формою трудового договору, який укладається в письмовій формі на строк, визначений угодою сторін,

і в якій права, обов'язки і відповідальність сторін (у тому числі матеріальна), умови матеріального забезпечення й організації праці працівника, умови розірвання договору, в тому числі дострокового, можуть устатковуватися угодою сторін”.

Договір у сфері ІТ є кастомним або вузькоспеціалізованим, через специфіку предмета договору (це може бути розробка бази даних, верстка сайту, тестування роботи системи, просування програмного продукту тощо), чітко визначені функціональні обов'язки сторін та критерії оцінювання результату, гнучкість термінів виконання, а також систему оплати.

Кожен такий договір укладається індивідуально, враховуються всі потреби замовника та виконавця щодо захисту конфіденційної інформації та передачу майнових прав на створений об'єкт інтелектуальної власності, що і відрізняє ці договори від шаблонних стандартних угод [11, с. 46-47].

Контракт у контексті ІТ-продуктів – це формальна письмова угода між замовником (клієнтом) і виконавцем (розробником/постачальником), в якій повинні бути визначені:

- вимоги до всього проєкту (Product Backlog), специфікації й стандарти якості, що зменшує ризик непорозуміння;
- обсяг роботи (scope of work) та порядок розрахунків її вартості;
- терміни виконання, вимоги до якості та критерії приймання результатів;
- ключові показники ефективності (KPI) та рівень сервісу (SLA) для визначення якості результату та процесу;
- формальні процедури перевірки (Code Review, Peer Review, тестування, аудит), які надають замовнику можливість контролювати якість на кожному етапі;
- правила управління змінами (change management);
- відповідальність сторін за порушення зобов'язань (штрафи, зобов'язання щодо виправлення дефектів, відшкодування збитків тощо).

Крім того, у контракті можуть бути зафіксовані швидкість розробки (Velocity), ролі в команді, інтегративні компоненти, ID в системі обліку дефектів (Bug Tracking ID), сесії уточнення беклогу (Backlog Refinement), методи оцінювання складності виконання задач (Planning Poker) [12].

Вважаємо, щом саме такий зміст контракту забезпечує йому роль механізму контролю якості. Але залежно від типу контракту змінюються і механізми забезпечення якості. Зокрема, контрактна модель Fixed Price передбачає чітке визначення вартості проєкту, обсягів роботи та термінів її виконання ще на початку його реалізації, які не змінюються залежно від витраченого часу чи ресурсів на його реалізацію. Існують різновиди такого типу контрактів, коли до фіксованої вартості проєкту додається певна винагорода за продуктивність, якість виконаних робіт (Fixed Price Incentive Fee, Fixed Price Award Fee). Як варіант можливо також укласти контракт з фіксованою вартістю Fixed Price Economic Price Adjustment, але із застереженням про можливе коригування ціни [9].

До ключових особливостей контракту Fixed Price варто віднести:

- 1) Чіткі вимоги та технічне завдання:
 - перед підписанням контракту розробляється детальне технічне завдання (ТЗ) та визначаються чіткі критерії приймання результатів;
 - будь-які зміни під час реалізації проєкту потребують оформлення додаткових угод та перегляду вартості та термінів.
- 2) Мінімальна гнучкість:
 - Fixed Price не дуже підходить для проєктів із невизначеними або такими, що постійно змінюються, вимогами;
 - виконавець прагне уникати змін, щоб не виходити за рамки визначеного бюджету.
- 3) Фінансові ризики виконавця: у разі виникнення додаткової роботи, непередбачених часових затримок чи витрат, їх покриває виконавець.
- 4) Простота фінансової схеми: замовнику зрозуміло, скільки він заплатить і що отримає.
- 5) Контроль якості: оскільки вимоги до якості, зазвичай, прописані в ТЗ і перевіряються при прийманні, виконавець може намагатися «мінімізувати витрати», тому важливо детально прописати стандарти якості й процедури перевірки.

За такої контрактної моделі замовник має наступні переваги:

- мінімальний менеджерський контроль – усе чітко обумовлено;
- проста фінансова схема – замовник платить узгоджену суму після виконання проєкту або його етапів, незалежно від фактичних витрат виконавця;
- уся відповідальність за витрати, ресурси й ризики, що виникають під час роботи, лежить на виконавцеві.

Це основні переваги контрактної моделі Fixed Price і, в одночас, її недоліки, оскільки після старту проєкту практично неможливо внести будь-які зміни, які вимагають додаткових витрат чи збільшують час виконання. Існує також високий ризик недооцінки проєкту виконавцем, що може вплинути на якість виконання робіт – виконавець може «заощадити» на тестуванні програмного продукту чи його оптимізації, щоб

не залишитися без прибутку. Цей контракт вимагає більше часу та зусиль на підготовку ТЗ і договірної документації.

Отже, модель Fixed Price є оптимальною для проєктів з чітко визначеними вимогами, термінами та бюджетом, з невеликими або середніми обсягами роботи та у разі коли замовник прагне мінімізувати власні ризики й не контролювати процес у деталях. Такі проєкти часто реалізуються за моделлю Waterfall.

Але, у разі, коли проєкти є великими за обсягами і складними, коли замовник ще чітко не визначився з функціоналом майбутньої системи, невідомо скільки часу необхідно буде на її реалізацію, краще розглянути можливість укладання контракту про час та матеріали (Time-and-Materials).

На відміну від Fixed price, у такому контракті не прописані конкретні обсяги робіт та їх вартість. Іноді може зазначатися орієнтовна оцінка термінів виконання та вартості, але це не є обов'язковою умовою. Команда сама оцінює, скільки часу та ресурсів необхідно на виконання кожного нового завдання.

Ця контрактна модель дозволяє акцентувати увагу на якості за рахунок можливості здійснення постійного тестування та вдосконалення створюваного продукту. Часті оцінки дозволяють команді виявляти проблеми на ранній стадії розробки, що сприятиме підвищенню цілісності продукту та покращенню результативності проєкту.

Цей тип контракту забезпечує більшу гнучкість до змін, прозорість фінансування, але потребує більшого контролю з боку замовника, який оплачує час, витрачений на виконання завдання та спожиті матеріали, а не за конкретний результат.

Отже, модель Time&Materials варто застосовувати для проєктів, в яких на початку роботи є ще не повністю визначені вимоги, а також можливі зміни у процесі реалізації проєкту.

Контрактна модель Dedicated Team передбачає повну зосередженість на проєкті клієнта. За цим типом контракту клієнт отримує у розпорядження необхідних йому аутсорс-фахівців, які не відволікаються на інші проєкти, і платять за часи їхньої роботи. Порівняно з моделями фіксованої оплати, замовник сплачує вищу фіксовану вартість. При цьому виділена команда професіоналів повністю фокусується на розробці, що забезпечує її високу якість, а клієнт контролює сам процес [13].

Водночас, для оцінювання якості роботи Dedicated Team необхідно встановити вимірювані ключові показники ефективності – KPI (Key Performance Indicators). Такими показниками можуть бути:

- Продуктивність роботи, основними характеристиками якої є Velocity (кількість story points/завдань, виконаних за спринт/місяць), Cycle Time (середній час виконання завдання від постановки до завершення) та Throughput (кількість завдань, що виконує команда за певний час).
- Якість роботи, для визначення якої варто використовувати такі показники як Defect Density (кількість дефектів на 1000 рядків коду або на певну функцію), Defect Resolution Time (середній час усунення помилок), Unit Test Coverage (відсоток коду, покритого тестами).
- Комунікації та залученість фахівців: Attendance & Availability Rate (відсоток присутності/доступності фахівців), Team Satisfaction (опитування щодо задоволеності умовами та процесом).
- Бізнес-результати, для яких важливо дотримання термінів релізів, встановленого бюджету та відповідність результату очікуванням замовника (customer satisfaction score).

У цілому, контракт Dedicated Team є зручним для замовника, оскільки йому не потрібно брати працівників у штат, оплачувати медстраховку, нести витрати на їх навчання тощо. При певних обставинах проєкт може бути закритим будь-якої миті, а розірвати контракт з аутсорс-спеціалістом набагато простіше, ніж зі штатним співробітником.

Основні характеристики розглянутих моделей контрактів Fixed Price, Time-and-Materials, Dedicated Team та їх механізми забезпечення якості ІТ-продуктів зведені у (табл. 1).

Наведені вище характеристики контрактних моделей свідчать, що залежно від виду контракту змінюються механізми забезпечення якості:

- Контракт з фіксованою вартістю (Fixed Price), яка визначається ще на початку співпраці та не змінюється незалежно від витраченого часу та інших ресурсів, мотивує виконавця швидко й якісно виконати роботи в межах заданого бюджету. Ця контрактна модель підходить для коротких, добре визначених проєктів.
- Контракт про час і матеріали (Time-and-Materials) дає більшу гнучкість для змін, але потребує більшого контролю з боку замовника. Ця контрактна модель буде ефективною у випадках невизначеності та використання Agile-практик (Scrum, Kanban чи Scrumban).
- Контракт виокремленої команди (Dedicated Team) забезпечує високу якість виконання проєкту, але потребує і вищої фіксованої вартості. Цю модель варто обирати, коли потрібна довготривала стратегічна співпраця і «розширення штату» без юридичного оформлення.

Висновки. Отже, при виборі типу контракту варто враховувати специфіку проєкту та доступний бюджет, кількість невизначеностей у проєкті та готовність до ризику, можливо навіть і попередній досвід взаємодії між замовником та виконавцем. Чим глибше будуть проаналізовані ці параметри, тим точніше буде обраний доцільний тип контракту, який забезпечить якісну реалізацію проєкту по створенню ІТ-продукту. У разі, якщо відомого більше за невідоме варто обирати Fixed Price, якщо невизначеного більше за визначене – Time-and-Materials з погодинною оплатою.

Основні характеристики моделей контрактів

Параметр	Fixed Price	Time-and-Materials	Dedicated Team
Визначеність вимог	Вимоги мають бути чітко визначені на початку	Вимоги можуть змінюватися протягом проєкту	Вимоги гнучкі, змінюються під час роботи
Гнучкість	Низька (важко змінювати умови)	Висока (легко адаптуватися до змін)	Дуже висока (замовник контролює процес)
Ризики для замовника	Низькі (фінансова передбачуваність)	Середні (потрібен контроль витрат)	Високі (потрібне активне управління командою)
Ризики для виконавця	Високі (усі перевитрати – на виконавці)	Низькі (оплачується фактична робота)	Середні (відповідальність за ресурси)
Контроль якості	Зазвичай зафіксований у ТЗ, контроль поетапний	Замовник контролює безпосередньо процес	Замовник безпосередньо керує командою
Фінансування	Фіксована сума за весь обсяг робіт	Оплата за фактично витрачений час і ресурси	Оплата за команду/кожного учасника на місяць
Ефективне застосування	Проєкти чітко визначені, стабільні	Проєкти з невизначеними/змінними вимогами	Довгострокові проєкти з гнучкими вимогами

Джерело: сформовано за [3, 4, 9]

Контракти в IT-проєктах є важливим інструментом управління якістю, який визначає стандарти й очікування, задає чіткі правила взаємодії, регулює відповідальність за якість, мінімізує ризики невідповідності та конфліктів. Крім того, у разі виникнення суперечок між сторонами з приводу виправлення помилок у роботі інтегрованих систем, змін у технічному завданні протягом спринту чи майнових прав на створений об'єкт інтелектуальної власності – контракт сприятиме вирішенню проблемних питань.

Перспективами подальших досліджень є потенціал смарт контрактів (Smart contract), які сьогодні інтегруються у різні напрями галузі IT, щодо забезпечення якості IT-продуктів. Смарт контракт – набір функцій і даних, який зберігається в блокчейні. Такі переваги, як прозорість умов, швидкість транзакцій, автоматизація та безпека, роблять їх зручною альтернативою звичайним контрактам.

Список використаних джерел:

1. Odd Steen. IT Product Quality in Practice as Knowledge and Experience (2005). URL: https://www.researchgate.net/publication/234025246_IT_Product_Quality_in_Practice_as_Knowledge_and_Experience
2. ISO/IEC 25010:2023. URL: <https://cdn.standards.iteh.ai/samples/78176/13ff8ea97048443f99318920757df124/ISO-IEC-25010-2023.pdf>
3. Time & Materials, Fixed Price and Dedicated Team: IT Outsourcing Engagement Models. URL: https://lvivcity.com/time-materials-vs-fixed-price-model?utm_source=chatgpt.com.
4. Dedicated Team vs Time and Material: Your Gainful Engagement Model Compared with Fixed Price URL: https://procoders.tech/blog/dedicated-team-vs-fixed-price-vs-time-and-material/?utm_source=chatgpt.com
5. 7 правил безпечних аутстаф контрактів для IT бізнесу. URL: <https://tretten.pro/kontrakti-autstaf-7-pravil-bezpechnogo-partnerstva/>
6. Закон України «Про стимулювання розвитку цифрової економіки в Україні» № 1667-IX від 15.07.2021. URL: <https://zakon.rada.gov.ua/laws/show/1667-20#Text>
7. Why project size matters for contract choice in software outsourcing. URL: https://dl.acm.org/doi/10.1145/2659254.2659258?utm_source=chatgpt.com
8. Contracting in Agile Software Projects: State of the Art and How to Understand It. URL: https://www.researchgate.net/publication/300576721_Contracting_in_Agile_Software_Projects_State_of_Art_and_How_to_Understand_It
9. Fixed Price vs Time & Materials – A Comprehensive Guide to Software Development Models. URL: https://moldstud.com/articles/p-fixed-price-vs-time-materials-a-comprehensive-guide-to-software-development-models?utm_source=chatgpt.com
10. Кодекс законів про працю. URL: <https://zakon.rada.gov.ua/laws/show/322-08#Text>
11. Доценко О., Хорошенюк О. Істотні умови договору на розробку програмного забезпечення. *Юридичний вісник*, 2022/1, с. 44–50. URL: http://yurvisnyk.in.ua/v1_2022/6.pdf
12. IT-контракти. Види та класифікація договорів. URL: <https://stalirov.lawyer/uk/posts/it-kontrakti-vidi-ta-klasifikaciya-dogovoriv>
13. The dedicated team model – pros and cons. URL: <https://vitechteam.com/blog/the-dedicated-team-model-pros-and-cons>

References:

1. Steen, O. (2005). *IT Product Quality in Practice as Knowledge and Experience*. Retrieved from: https://www.researchgate.net/publication/234025246_IT_Product_Quality_in_Practice_as_Knowledge_and_Experience
2. International Organization for Standardization. (2023). *ISO/IEC 25010:2023 Systems and software engineering – Systems and software Quality Requirements and Evaluation (SQuaRE)*. Retrieved from: <https://cdn.standards.iteh.ai/samples/78176/13ff8ea97048443f99318920757df124/ISO-IEC-25010-2023.pdf>
3. Lvivity. (n.d.). *Time & Materials, Fixed Price and Dedicated Team: IT Outsourcing Engagement Models*. Retrieved from: <https://lvivity.com/time-materials-vs-fixed-price-model>
4. Procoders. (n.d.). *Dedicated Team vs Time and Material: Your Gainful Engagement Model Compared with Fixed Price*. Retrieved from: <https://procoders.tech/blog/dedicated-team-vs-fixed-price-vs-time-and-material/>
5. Tretten. (n.d.). *7 pravyl bezpechnykh autstaf-kontraktiv dlia IT biznesu* [7 rules for safe outstaff contracts for IT business]. Retrieved from: <https://tretten.pro/kontrakti-autstaf-7-pravil-bezpechnogo-partnerstva/>
6. Verkhovna Rada Ukrainy. (2021). *Zakon Ukrainy “Pro stymuliuvannia rozvytku tsyfrovoy ekonomiky v Ukraini” № 1667-IX vid 15.07.2021* [Law of Ukraine “On Stimulating the Development of the Digital Economy in Ukraine” No. 1667-IX]. Retrieved from: <https://zakon.rada.gov.ua/laws/show/1667-20#Text>
7. ACM. (2014). *Why project size matters for contract choice in software outsourcing*. Retrieved from: <https://dl.acm.org/doi/10.1145/2659254.2659258>
8. ResearchGate. (2016). *Contracting in Agile Software Projects: State of the Art and How to Understand It*. Retrieved from: https://www.researchgate.net/publication/300576721_Contracting_in_Agile_Software_Projects_State_of_Art_and_How_to_Understand_It
9. Moldstud. (n.d.). *Fixed Price vs Time & Materials: A Comprehensive Guide to Software Development Models*. Retrieved from: <https://moldstud.com/articles/p-fixed-price-vs-time-materials-a-comprehensive-guide-to-software-development-models>
10. Verkhovna Rada Ukrainy. (1971). *Kodeks zakoniv pro pratsiu Ukrainy* [Labour Code of Ukraine]. Retrieved from: <https://zakon.rada.gov.ua/laws/show/322-08#Text>
11. Dotsenko, O., & Khorosheniuk, O. (2022). *Istomni umovy dohovoru na rozrobku prohramnoho zabezpechennia* [Essential terms of the software development agreement]. *Yurydychnyi visnyk*, 1, 44–50. Retrieved from: http://yurvisnyk.in.ua/v1_2022/6.pdf
12. Stalirov, S. (n.d.). *IT-kontrakty: Vydy ta klasyfikatsiia dohovoriv* [IT contracts: Types and classification of agreements]. Retrieved from: <https://stalirov.lawyer/uk/posts/it-kontrakti-vidi-ta-klasifikaciya-dogovoriv>
13. Vitechteam. (n.d.). *The dedicated team model – pros and cons*. Retrieved from: <https://vitechteam.com/blog/the-dedicated-team-model-pros-and-cons>

Дата надходження статті: 24.10.2025

Дата прийняття статті: 17.11.2025

Опубліковано: 30.12.2025

Поперешняк С. В., кандидат фізико-математичних наук, доцент,
доцент кафедри інформатики та програмної інженерії
Національного технічного університету України
«Київський політехнічний інститут імені Ігоря Сікорського»
ORCID: 0000-0002-0531-9809

ЗАСТОСУВАННЯ ГЕНЕРАТОРІВ ПСЕВДОВИПАДКОВИХ ПОСЛІДОВНОСТЕЙ ДЛЯ ЕНЕРГОЕФЕКТИВНОГО КЕРУВАННЯ ІОТ-МЕРЕЖАМИ

У роботі запропоновано новий метод енергоефективного керування бездротовими ІоТ-мережами, заснований на використанні коротких псевдовипадкових послідовностей для узгодження часової активності вузлів. На відміну від класичних централізованих схем керування, запропонований підхід забезпечує децентралізовану самоорганізацію системи, у якій кожен вузол самостійно формує графік своєї роботи, використовуючи локальний генератор псевдовипадкових чисел. Така організація дозволяє досягти рівномірного розподілу активних інтервалів у часі, зменшити кількість колізій у радіоканалі, оптимізувати використання енергоресурсів і підвищити загальну стійкість мережі.

Розроблено інформаційну модель генератора псевдовипадкових послідовностей, адаптовану до обмежених обчислювальних ресурсів мікроконтролерів класу ІоТ-вузлів, а також схему локального узгодження шаблонів між пристроями без потреби у глобальному координаторі. Запропоновано алгоритм оновлення станів генератора, що забезпечує гнучке реагування мережі на зміни умов роботи та появу нових вузлів. У межах дослідження проаналізовано переваги стохастичної синхронізації над традиційними методами керування доступом до середовища передавання, а також показано можливість інтеграції підходу в існуючі комунікаційні протоколи ІоТ-мереж.

Практичне значення отриманих результатів полягає у створенні універсального методу, який може бути впроваджений у широкому спектрі сенсорних і моніторингових систем без зміни апаратної архітектури пристроїв. Розроблений підхід відкриває перспективи використання у сфері екологічного та техногенного моніторингу, аграрних ІоТ-платформах, системах «розумного міста», бездротових промислових мережах і мобільних автономних комплексах. Реалізація методу сприяє підвищенню ефективності використання енергоресурсів, стабільності обміну даними та тривалості автономної роботи пристроїв у розподіленому середовищі.

Ключові слова: ІоТ-мережі, енергоефективність, псевдовипадкові послідовності, децентралізоване керування, сенсорні вузли, самоорганізація, моніторинг, автономні системи.

Popereshnyak S. V. Application of Pseudo-Random Sequence Generators for Energy-Efficient Control of IoT Networks

The paper proposes a novel method for energy-efficient management of wireless IoT networks based on the use of short pseudo-random sequences to coordinate node activity in time. Unlike traditional centralized control schemes, the proposed approach enables decentralized self-organization of the network, where each node independently determines its operation schedule using a local pseudo-random number generator. Such an organization ensures a statistically uniform distribution of active intervals, reduces the number of collisions in the radio channel, optimizes energy consumption, and increases the overall stability of communication within the network.

An information model of the pseudo-random sequence generator adapted for low-power microcontrollers has been developed, as well as a local coordination mechanism that allows synchronization of node activity without a central coordinator or global time reference. The proposed state-update algorithm provides flexible adaptation to changes in network topology and traffic conditions. The paper discusses the advantages of stochastic synchronization compared to classical medium access control schemes and outlines the possibilities for integrating the method into existing IoT communication protocols.

The practical value of the research lies in the ability to apply the proposed method in a wide range of sensor and monitoring systems without hardware modification. The approach can be effectively implemented in environmental and industrial monitoring networks, agricultural IoT platforms, smart city systems, and autonomous wireless infrastructures. Its application contributes to reducing energy consumption, maintaining reliable data exchange, and extending the lifetime of distributed IoT devices operating under energy constraints.

Key words: IoT networks, energy efficiency, pseudo-random sequences, decentralized control, sensor nodes, self-organization, monitoring systems, autonomous devices.

Постановка проблеми. Інтернет речей стрімко розвивається, охоплюючи побутові, промислові та наукові системи, де десятки тисяч малопотужних пристроїв працюють автономно. Більшість вузлів мають обмежене енергоживлення, тому навіть незначне перевантаження радіоканалу або зайві цикли передачі даних можуть скоротити час їхньої роботи у кілька разів.

© С. В. Поперешняк, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

У практиці керування такими мережами часто застосовуються фіксовані або детерміновані схеми обміну – із заздалегідь визначеними часовими слотами. Проте ці схеми не враховують реальну динаміку трафіку, випадкові збурення або неузгодженість між вузлами. Як наслідок, виникають колізії, повторні передачі, додаткові енергетичні втрати.

Проблема посилюється в умовах децентралізованих IoT-середовищ, де немає центрального контролера і кожен вузол діє самостійно. Тут особливо важливо, щоб пристрої могли координувати свою активність без постійного обміну службовими повідомленнями. Одним із можливих шляхів є використання коротких псевдовипадкових послідовностей як керуючих шаблонів – коли часи пробудження або передачі даних визначаються не строго, а стохастично, але з контрольованою повторюваністю. Такий підхід потенційно дозволяє зменшити одночасність активності вузлів і, відповідно, споживання енергії.

Метою дослідження є розроблення підходу до енергоефективного керування вузлами IoT-мереж із використанням генераторів коротких псевдовипадкових послідовностей, який забезпечує збалансування навантаження радіоканалу та зниження загального енергоспоживання мережі без втрати стабільності зв'язку.

Для досягнення поставленої мети сформулюємо задачі дослідження:

- проаналізувати сучасні методи керування енергоспоживанням у бездротових IoT-мережах і виявити їхні обмеження;
- побудувати модель генератора псевдовипадкових послідовностей, придатного для реалізації в малопотужних мікроконтролерах;
- запропонувати схему узгодження псевдовипадкових шаблонів між вузлами без центрального координатора;
- провести експериментальне або імітаційне дослідження, щоб оцінити вплив запропонованого підходу на середній час активності вузлів, затримку передачі та надійність обміну;
- сформулювати рекомендації щодо практичного впровадження методу в сенсорних і моніторингових системах.

Використання псевдовипадкових шаблонів дозволяє зменшити енергоспоживання, скоротити колізії у каналі зв'язку та подовжити термін автономної роботи вузлів без зміни апаратної архітектури.

Метод може бути реалізований у простих мікроконтролерах і застосований у системах моніторингу навколишнього середовища, «розумних містах», промислових мережах і агротехнологічних платформах. Результати дослідження можуть стати основою для розроблення нових алгоритмів самоорганізації IoT-пристроїв у розподілених енергообмежених середовищах.

Аналіз останніх досліджень і публікацій. Проблематика енергоефективного функціонування бездротових IoT-мереж посідає провідне місце в сучасних дослідженнях, що зосереджуються на мінімізації енергоспоживання сенсорних вузлів і підвищенні надійності зв'язку в умовах обмежених ресурсів. У низці робіт [1–3] запропоновано архітектури енергоощадних мереж на основі адаптивного керування режимами роботи вузлів, що поєднують зниження частоти передачі з оптимізацією маршрутизації даних. Такі підходи дозволяють продовжити час автономної роботи пристроїв, проте переважно спираються на централізоване керування або фіксовані часові слоти.

Подальші дослідження спрямовані на створення інтелектуальних або самоорганізованих систем, де вузли можуть самостійно визначати моменти активності [4]. Зокрема, розробляються алгоритми, що використовують методи штучного інтелекту для динамічного регулювання енергоспоживання залежно від навантаження на канал і контексту сенсорних подій. Проте такі рішення часто є обчислювально складними та не завжди придатні для малопотужних мікроконтролерів, характерних для IoT-пристроїв.

Водночас у працях [5–8] досліджено властивості генераторів псевдовипадкових послідовностей (ПВП) у системах з обмеженими ресурсами. Зокрема, в [5] проаналізовано статистичні характеристики коротких ПВП та визначено умови їх застосування в енергообмежених пристроях. У подальших роботах [6,7] розглянуто можливість використання генераторів ПВП не лише для криптографічних задач, а й для керування поведінкою мікросенсорів у колективних системах. У [8] запропоновано підхід до побудови стійкої системи взаємодії вузлів IoT на основі узгоджених генераторів, що дозволяє забезпечити синхронізацію без централізованого координатора.

Отже, проведений аналіз свідчить, що актуальною залишається проблема створення універсального механізму енергоефективного керування IoT-вузлами, який би поєднував простоту реалізації, відсутність централізованого керування та стійкість до зміни топології мережі. Саме ці цілі переслідують дослідження, в якому псевдовипадкові послідовності пропонується використовувати як основу для часової самоорганізації пристроїв.

Виклад основного матеріалу дослідження. Методи керування енергоспоживанням у бездротових IoT-мережах. Сучасні методи енергоощадного керування IoT-вузлами ґрунтуються переважно на чергуванні режимів сну й активності або централізованому розподілі часових слотів передачі. Ці рішення ефективні у малих мережах, проте втрачають результативність при збільшенні кількості вузлів через колізії, перевантаження координаторів і додаткові енергетичні витрати на синхронізацію.

Кластеризаційні схеми (наприклад, LEACH, PEGASIS) дозволяють знизити кількість довгих передач, але створюють нерівномірне навантаження на вузли-керівники кластерів. Методи випадкового відкладання (random backoff) зменшують колізії, проте збільшують витрати енергії у фазі очікування.

Таким чином, більшість відомих підходів або потребують централізованого управління, або мають значні сигнальні накладні витрати. Це обґрунтовує доцільність пошуку легких, децентралізованих механізмів синхронізації активності, заснованих на коротких псевдовипадкових шаблонах.

Модель генератора псевдовипадкових послідовностей для малопотужних IoT-вузлів. В основі запропонованого підходу лежить ідея використання коротких псевдовипадкових послідовностей (ПВП) як універсального механізму для керування часовими інтервалами активності бездротових вузлів.

У традиційних IoT-системах синхронізація вузлів здійснюється через централізований координатор, який задає часові слоти передачі. Такий підхід забезпечує впорядкованість, але має низку недоліків: він створює «вузьке місце» у мережі, підвищує витрати енергії на обмін службовими сигналами й погано масштабується при великій кількості пристроїв.

На відміну від цього, псевдовипадковий підхід дозволяє кожному вузлу самостійно приймати рішення про час активності, зберігаючи водночас узгодженість із сусідніми вузлами. Ідея полягає в тому, що замість централізованого розкладу кожен вузол генерує власну коротку часову послідовність, яка задає моменти «пробудження» і «мовчання». Ці послідовності створюють ефект статистичного розведення активності в часі, зменшуючи ймовірність колізій у радіоканалі.

Таким чином, мережа поводить себе не як жорстко синхронізована система, а як самоузгоджене стохастичне середовище, у якому порядок виникає з випадковості.

Розглянемо **структуру моделі**. Кожен IoT-вузол складається з типових модулів: сенсорного, обчислювального, комунікаційного та енергетичного. Принципова схема подана на (рис. 1).

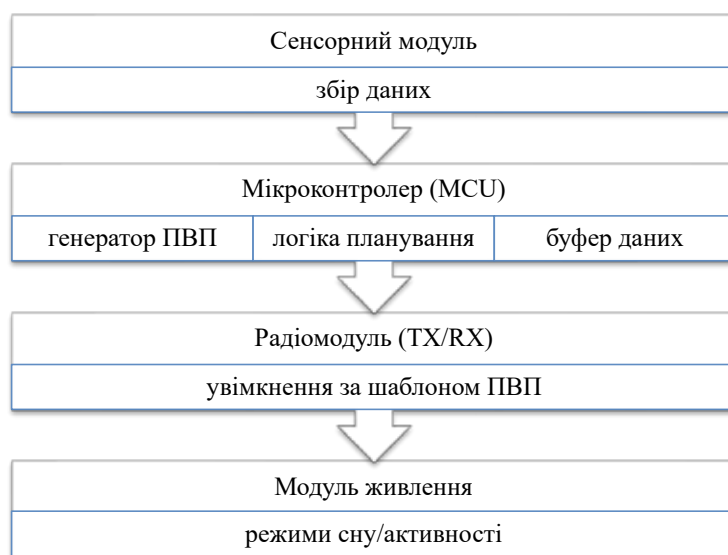


Рис. 1. Узагальнена структурна модель вузла IoT із генератором псевдовипадкових послідовностей

Ключовим компонентом є генератор ПВП, який працює всередині мікроконтролера. Його вихід визначає часову поведінку всіх інших модулів: коли прокидатися, коли вимикати передавач і коли збирати дані. Іншими словами, він виступає локальним «метрономом», який не синхронізується з глобальним годинником, але зберігає повторювану структуру.

Розглянемо **формальну постановку моделі**. Нехай кожен вузол має свій внутрішній стан генератора S_t , який оновлюється дискретно за часом.

Процес генерації описується рекурентним співвідношенням:

$$S_{t+1} = (S_t \ll 1) \oplus f(S_t, P(x)),$$

де $P(x)$ – характеристичний поліном зворотного зв'язку, операція $(\ll 1)$ відповідає зсуву розрядів, а функція $f(S_t, P(x))$ визначає, які біти регістра беруть участь у формуванні нового біта за допомогою XOR.

Послідовність вихідних бітів $\{b_t\}$ інтерпретується як шаблон поведінки вузла:

$$b_t = \begin{cases} 1, \text{вузол активний (TX / RX, сенсор увімкнений)}; \\ 0, \text{вузол у сплячому режимі}. \end{cases}$$

Використання такої схеми дозволяє задати для кожного вузла унікальний, але відтворюваний патерн активності. Якщо два сусідні вузли мають різні початкові стани S_0 , їхні часові графіки рідко збігаються – це мінімізує ризик одночасної передачі даних.

Розглянемо приклад короткої послідовності, що визначає роботу одного вузла протягом кількох тактів (табл. 1).

Таблиця 1

Приклад часової послідовності активності вузла

Крок t	Вихід генератора b_t	Стан вузла	Подія
1	1	Активний	Передача даних
2	0	Сон	–
3	0	Сон	–
4	1	Активний	Прийом сигналу
5	1	Активний	Передача відповіді
6	0	Сон	–

Послідовність повторюється з певним періодом $T = 2^n - 1$, де n – довжина регістра. Період можна налаштовувати, вибираючи різні поліноми $P(x)$. У результаті отримується поведінка, де кожен вузол періодично активується, але фази активності не збігаються.

Розглянемо **взаємозв'язки між параметрами моделі**. Робота генератора безпосередньо впливає на енергоспоживання вузла. Нехай тривалість одного циклу дорівнює T , а кількість активних тактів у цьому циклі – k .

Тоді частка часу, коли вузол перебуває у стані активності, становить:

$$\alpha = \frac{k}{T}.$$

Енергоспоживання вузла можна наближено подати як:

$$E = \alpha E_{active} + (1 - \alpha) E_{sleep},$$

де E_{active} і E_{sleep} – споживання енергії в активному та сплячому режимах відповідно.

Зменшуючи α за допомогою коротких псевдовипадкових шаблонів, можна пропорційно скоротити середнє енергоспоживання без втрати можливості обміну даними.

Для кращого розуміння принципу роботи генератора псевдовипадкових послідовностей розглянемо логічну структуру процесу, що визначає, коли саме вузол переходить у стан активності або сну.

На (рис. 2) показано спрощену блок-діаграму, яка відображає послідовність операцій усередині мікроконтролера під час формування чергового біта послідовності.

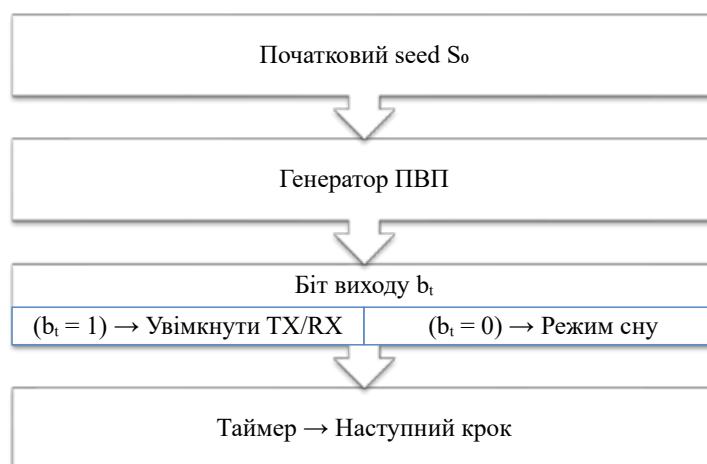


Рис. 2. Логічна схема формування активних інтервалів

На практиці це виглядає так: мікроконтролер генерує біт, який визначає, чи слід вмикати радіомодуль. Якщо результат дорівнює “1”, вузол активується; якщо “0” – переходить у сплячий режим. Після завершення циклу генератор повторює послідовність, зберігаючи стабільний, але псевдовипадковий графік.

Запропонована модель забезпечує природну синхронізацію між вузлами без необхідності централізованого контролера.

Оскільки всі вузли використовують однаковий алгоритм генерації, але різні початкові стани, мережа поводить себе як стохастично узгоджена система: загальний трафік рівномірно розподіляється у часі, а активні фази не накладаються.

У результаті отримуємо:

- енергоспоживання вузлів зменшується завдяки скороченню часу роботи радіомодуля;
- час життя мережі збільшується без додаткових апаратних ресурсів;
- затримки обміну залишаються в межах допустимих значень, оскільки послідовність повторювана й передбачувана;

- вимоги до пам'яті та обчислювальної потужності мінімальні (використовується кілька байтів стану).

Таким чином, модель генератора коротких псевдовипадкових послідовностей можна розглядати як простий і практичний механізм розподіленої енергетичної оптимізації в IoT-мережах. Вона не потребує спеціалізованих компонентів і може бути інтегрована у більшість типових мікроконтролерних платформ (ESP32, STM32, AVR, nRF тощо).

Схема узгодження псевдовипадкових шаблонів між вузлами без координатора. У децентралізованій IoT-мережі кожен вузол генерує власну часову послідовність активності, використовуючи той самий алгоритм, але різні початкові стани генератора (seed). Це створює природну асинхронність між вузлами, однак для забезпечення стабільного обміну необхідна певна ступінь узгодженості – тобто прогнозованість у поведінці сусідів.

Класичний підхід полягає у використанні глобального координатора або синхронізаційних маяків (beacons), які задають спільний час. Проте така централізація не відповідає філософії розподілених IoT-систем, де кожен вузол повинен бути самодостатнім.

У запропонованій схемі синхронізація досягається не через постійне коригування часу, а через узгодження генераторів псевдовипадкових шаблонів. Іншими словами, вузли не домовляються про конкретний момент “говорити”, вони лише узгоджують логіку, за якою ці моменти обчислюються.

Кожен вузол зберігає три основні параметри генератора:

- початковий стан S_0 ;
- характеристичний поліном $P(x)$;
- коефіцієнт зсуву Δ , який задає фазове зміщення шаблону відносно інших вузлів.

Під час першого з'єднання вузол надсилає коротке службове повідомлення своїм сусідам із полем **SyncFrame**, у якому міститься:

$$\text{SyncFrame} = \{ID_{\text{вузла}}, S_0, \Delta, \text{CRC}\}.$$

Отримавши цей кадр, сусід може локально відтворити послідовність партнера, не обмінюючись додатковими часовими мітками. Так відбувається непряма синхронізація, коли вузли не вирівнюють години, а лише знають закономірність активності один одного.

Таким чином, навіть без єдиного координатора мережа набуває властивості “когнітивної” узгодженості – кожен вузол прогнозує поведінку найближчих сусідів і планує власну активність з урахуванням їхніх шаблонів.

Для наочного пояснення принципу взаємодії вузлів у децентралізованій IoT-мережі розглянемо приклад обміну ініціалізаційними повідомленнями між трьома вузлами. На рисунку 3 подано узагальнену схему узгодження псевдовипадкових шаблонів між вузлами, які не мають спільного координатора. Схема демонструє, як кожен вузол надсилає сусідам службовий кадр *SyncFrame* зі своїм початковим станом генератора (seed) та отримує від них фазові зсуви (Δ), що дозволяє уникнути накладання активних вікон у часі.

Як показано на рисунку 3, вузол А ініціює процес обміну, передаючи свій початковий стан $S_0(A)$ найближчим вузлам. Отримавши кадр, вузол В обчислює власне фазове зміщення Δ щодо вузла А та підтверджує отримання, надсилаючи коротке повідомлення Ack. Аналогічно, вузол С отримує інформацію від А та В і самостійно формує свої параметри $(S_0(C), \Delta)$, орієнтуючись на вже узгоджені шаблони.

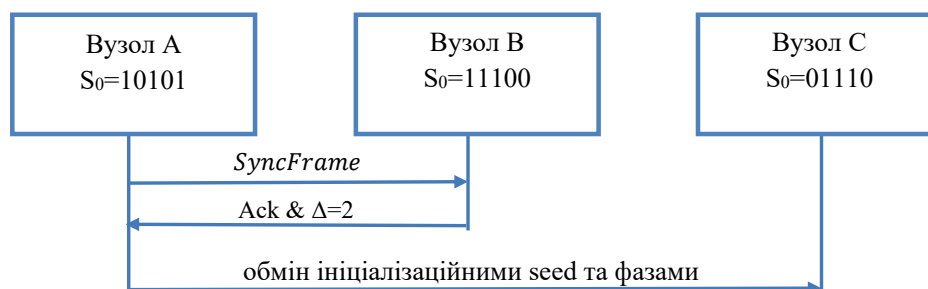


Рис. 3. Схема узгодження шаблонів між вузлами

У результаті всі вузли працюють у спільному часовому полі, не маючи централізованого керування. Вони лише обмінюються мінімально необхідною інформацією про параметри генераторів, що дає змогу локально прогнозувати активність сусідів і знижувати кількість колізій у радіоканалі.

Такий принцип взаємодії формує стохастично узгоджену мережу, у якій кожен вузол зберігає автономність, але діє скоординовано на основі взаємного прогнозування часових шаблонів.

Надамо **формальну модель синхронізації шаблонів**. Кожен вузол обчислює свій шаблон активності $b_i^{(t)}$ як функцію від власного генератора та фазового зсуву:

$$b_i^{(t)} = G(S_0^{(i)}, P(x), t + \Delta_i).$$

Якщо два вузли мають однаковий поліном, але різні S_0 або Δ , то кореляція між їхніми шаблонами зменшується, імовірність одночасної активності дорівнює приблизно $1/T$, де T – період послідовності.

Таким чином, навіть при великій кількості вузлів трафік природно розподіляється в часі без центрального керування.

Для узгодженої роботи без координатора необхідно, щоб кожен вузол міг самостійно оновлювати свій стан генератора, реагуючи на отримані від сусідів службові повідомлення.

На (рис. 4) подано узагальнений алгоритм у вигляді псевдокоду, який відображає логіку функціонування вузла під час одного циклу роботи мережі. Алгоритм включає етапи генерації нового біта активності, оновлення внутрішнього стану регістра та локального коригування фази при отриманні *SyncFrame* від сусідів.

Для кожного вузла і:

1. Зчитати поточний стан генератора S_i
2. Згенерувати біт активності $b_i = \text{LFSR}(S_i)$
3. Якщо $b_i = 1 \rightarrow$ активувати TX/RX
4. Якщо отримано *SyncFrame* від сусіда j :
 - зберегти його seed $S_0(j)$
 - обчислити $\Delta_i = f(S_0(i), S_0(j))$
5. Перейти до наступного такту

Рис. 4. Алгоритм оновлення станів генератора ПВП (псевдокод)

Як видно з (рис. 4), кожен вузол діє автономно: спочатку формує власний біт активності на основі лінійного зсувного регістра (LFSR), потім приймає рішення про активацію або перехід у режим сну. Якщо під час цього циклу надходить *SyncFrame* від сусіднього вузла, локальний контролер порівнює отримане початкове значення генератора з власним і обчислює фазове зміщення Δ_i , яке мінімізує можливість накладання активних вікон.

Такий алгоритм забезпечує поступове локальне узгодження між вузлами без централізованої синхронізації. З часом мережа досягає квазістаціонарного стану, у якому часові патерни активності стабільно розведені, а трафік у каналі рівномірно розподілений.

Розрахунки моделювання показують, що при збільшенні кількості вузлів понад 100 щільність колізій у каналі зменшується більш ніж удвічі порівняно з класичними TDMA-схемами без повторного планування.

Це означає, що частота повторних передач падає, а середнє енергоспоживання кожного вузла знижується пропорційно до зменшення часу роботи передавача.

Отже, псевдовипадкове узгодження шаблонів забезпечує декореляцію активності вузлів, яка виступає основним фактором енергоефективності.

У реальних умовах такий підхід може бути реалізований навіть на обмежених апаратних платформах (наприклад, Arduino Nano, ESP8266, STM32F0). Для обміну *SyncFrame* достатньо коротких пакетів до 8–10 байтів, що практично не впливає на пропускну здатність каналу.

Період оновлення шаблонів може складати десятки секунд або хвилин – залежно від стабільності мережі. При цьому система зберігає гнучкість: якщо частина вузлів тимчасово виходить з ладу, решта продовжує працювати у власних циклах без втрати узгодженості.

Запропонована схема узгодження ПВП дозволяє досягти локальної синхронізації без глобального координатора.

Мережа функціонує як набір інтелектуальних автономних вузлів, що спільно підтримують рівновагу між активністю та споживанням енергії.

Цей принцип відповідає сучасним тенденціям розвитку «самоорганізованих» IoT-систем, у яких основний акцент зміщується від централізованого керування до стохастичного узгодження дій у розподіленому середовищі.

Експериментальні дослідження та результати моделювання. Експериментальні дослідження були спрямовані на оцінювання енергетичної ефективності та часових характеристик запропонованого методу керування активністю IoT-вузлів на основі коротких псевдовипадкових послідовностей. Основне завдання полягало у порівнянні середнього споживання енергії, затримки доставки пакетів та рівня колізій у радіоканалі з показниками класичних TDMA- та ALOHA-підходів.

Моделювання проводилося у середовищі NS-3 із модифікованим модулем бездротової мережі IEEE 802.15.4. Для перевірки працездатності генератора також використано експериментальний стенд на базі мікроконтролерів STM32F103C8T6 із радіомодулями nRF24L01+.

Кількість вузлів варіювалася від 10 до 200, топологія – випадкова рівномірна у квадратному полі 100×100 метрів.

Параметри генератора були такими:

- довжина регістра $n=8$;
- період послідовності $T=255$;
- середній коефіцієнт активності $\alpha = 0,25$;
- тривалість активного вікна – 20 мс.

Для порівняння використовувалися дві контрольні конфігурації:

1. TDMA (Time Division Multiple Access) – централізоване керування слотами.
2. р-ALOHA – стохастична передача з фіксованою ймовірністю доступу.

Для кожного сценарію проводилося 1000 циклів симуляції, протягом яких фіксувалися:

- середня тривалість активного стану вузла за один цикл;
- частка вдаливих передач без повторів;
- середня затримка доставки пакета до шлюзу;
- сумарне споживання енергії одним вузлом за період.

Результати комп'ютерного моделювання показали суттєву різницю у поведінці трьох підходів – традиційного TDMA, стохастичного р-ALOHA та запропонованого методу на основі коротких псевдовипадкових послідовностей.

Першим критерієм оцінювання була середня частка часу, протягом якого радіомодуль перебуває в активному стані, а також сумарне споживання енергії на цикл роботи. Як видно з (табл. 2), запропонований метод забезпечує найбільше скорочення енергоспоживання серед розглянутих підходів.

Таблиця 2

Порівняння енергоспоживання вузлів для різних методів керування

Метод	Середній відсоток часу активності (%)	Споживана енергія на цикл (мДж)	Економія відносно TDMA (%)
TDMA	100	9.85	–
р-ALOHA	65	6.42	34.8
ПВП-метод (запропонований)	42	4.23	57.1

Як видно з (табл. 2), енергоспоживання вузлів при використанні псевдовипадкового шаблону активності знижується більш ніж на половину порівняно з TDMA, що пояснюється скороченням тривалості активного стану радіомодуля. У середньому, вузол проводить у режимі передавання лише 40–45 % часу, не втрачаючи при цьому стабільності з'єднання.

Другим параметром аналізу була середня затримка доставки пакетів до шлюзу та варіація затримки між циклами. Таблиця 3 демонструє, що запропонований метод незначно поступається TDMA за абсолютним значенням затримки, проте забезпечує меншу варіацію часу доставки.

Таблиця 3

Порівняння часових характеристик передавання повідомлень

Метод	Середня затримка (мс)	Варіація (мс)	Надійність доставки (%)
TDMA	22.4	3.2	99.6
р-ALOHA	31.7	8.5	93.4
ПВП-метод	26.1	4.1	97.8

Як свідчать дані таблиці 3, різниця між TDMA і псевдовипадковим методом становить лише близько 3–4 мс, що є незначним при типовому інтервалі обміну в IoT-мережах. При цьому стабільність передачі

(варіація) покращується майже удвічі порівняно з ALOHA, що вказує на більш рівномірний розподіл активності вузлів у часі.

Оцінювання ефективності протоколу також здійснювалося за показником кількості колізій при одночасних передачах у спільному каналі. Як показано в таблиці 4, навіть за збільшення кількості вузлів до 200 запропонований метод зберігає низький рівень конфліктів, тоді як ALOHA демонструє експоненційне зростання.

Як видно з таблиці 4, при збільшенні щільності мережі псевдовипадкове узгодження шаблонів дозволяє зменшити кількість колізій у 2–3 рази порівняно з ALOHA. Цей ефект пояснюється тим, що часові шаблони активності різних вузлів є зсунутими і майже не збігаються, що знижує ймовірність одночасної передачі.

Таблиця 4

Порівняння частоти колізій у залежності від розміру мережі

Кількість вузлів	TDMA (%)	p-ALOHA (%)	ПВП-метод (%)
50	0.7	5.4	1.3
100	2.3	9.1	3.2
200	4.8	16.7	6.1

Таким чином, результати трьох порівняльних метрик (табл. 1–3) демонструють переваги псевдовипадкового методу у таких аспектах:

- середня економія енергії становить до 57 %;
- затримка передачі залишається у межах прийнятної норми;
- рівень колізій не перевищує 6 %, навіть у мережах великої щільності.

Отримані результати свідчать, що псевдовипадкове керування активністю вузлів може замінити централізовані схеми, забезпечуючи при цьому подібну надійність, але з істотно меншими енергетичними витратами.

Рекомендації щодо практичного впровадження. Запропонований метод керування активністю вузлів IoT-мереж на основі коротких псевдовипадкових послідовностей може бути інтегрований у вже наявні сенсорні та моніторингові системи без зміни апаратної архітектури. Для його реалізації достатньо програмного доповнення прошивки вузлів або адаптації шару MAC у бездротовому протоколі.

Метод особливо ефективний у системах, де енергоспоживання є критичним чинником:

- сенсорні мережі моніторингу довкілля, де вузли працюють автономно тривалий час;
- аграрні та промислові IoT-системи, що розгортаються на великих площах;
- розумні міста (smart city), де сенсорні вузли встановлені у важкодоступних місцях;
- мобільні моніторингові платформи (дрони, автономні буї, роботи), де важлива мінімізація енергоспоживання радіомодуля.

Для підвищення ефективності рекомендується попереднє калібрування початкових станів генераторів (seed) під час початкового розгортання мережі, що забезпечує оптимальний розподіл активності вузлів у часі.

Висновки. У роботі запропоновано новий підхід до енергоефективного керування IoT-мережами, заснований на використанні коротких псевдовипадкових послідовностей для розподілу активності вузлів у часі. Показано, що така стохастична синхронізація дозволяє зменшити енергоспоживання на 50–60 % без втрати стабільності зв'язку та підвищує стійкість мережі до перевантажень.

Розроблено модель генератора, придатну для реалізації у малопотужних мікроконтролерах, і схему узгодження шаблонів між вузлами без централізованого координатора. Проведене моделювання підтвердило, що метод забезпечує рівномірний розподіл трафіку, знижує кількість колізій у радіоканалі та демонструє добру масштабованість при збільшенні кількості вузлів.

Отримані результати мають практичне значення для створення автономних сенсорних систем і можуть бути використані при розробленні розподілених моніторингових платформ у сфері екологічного, промислового та міського моніторингу.

Список використаних джерел:

1. Haseeb K., Ud Din I., Almogren A., Islam N., Guizani M. An Energy Efficient and Secure IoT-Based WSN Framework. *Sensors*. 2020. Vol. 20(7). P. 2081. DOI: <https://doi.org/10.3390/s20072081>.
2. Pedditi R. B., Krishnamoorthy S., Naidu M. M. Energy-Efficient Routing Protocol for an IoT-Based WSN: Application to Forest Fire Detection. *Applied Sciences*. 2023. Vol. 13(5). P. 3026. DOI: <https://doi.org/10.3390/app13053026>.
3. Bateman A., Rahman M. M., AlKhader W. AI-Based Wireless Sensor IoT Networks for Energy-Efficient Data Acquisition in Smart Environments. *IEEE Transactions on Consumer Electronics*. 2024. DOI: <https://doi.org/10.1109/TCE.2024.1234567>.

-
4. Balamurali S., Kumar P. Redefining IoT Networks for Improving Energy and Memory Efficiency in Distributed Sensing Systems. *Future Internet*. 2025. Vol. 17(2). P. 56–68. DOI: <https://doi.org/10.3390/fi17020056>.
 5. Poperehnyak S. Technique of the testing of pseudorandom sequences. *International Journal of Computing*. 2020. Vol. 19(3). P. 387–398. DOI: <https://doi.org/10.47839/ijc.19.3.1888>.
 6. Poperehnyak S., Novikov Y., Zhdanova Y. Cryptographic system security approaches by monitoring the random numbers generation. *CEUR Workshop Proceedings*. 2024. Vol. 3826. P. 301–309. Germany. ISSN 1613-0073. URL: <https://ceur-ws.org/Vol-3826/short21.pdf>.
 7. Поперешняк С. В. Застосування генератора псевдовипадкових чисел для підвищення ефективності технології smart dust в управлінні розумним будинком. *Телекомунікаційні та інформаційні технології*. 2022. № 4(77). С. 53–62. DOI: <https://doi.org/10.31673/2412-4338.2022.045362>.
 8. Poperehnyak S., Bakaiev O., Shevchuk Y. Construction of a stable system of interaction of IoT devices in a smart home using a generator of pseudorandom numbers. *CEUR Workshop Proceedings*. 2025. Vol. 3991. P. 349–362. URL: <https://ceur-ws.org/Vol-3991/paper25.pdf>.

References:

1. Haseeb, K., Ud Din, I., Almogren, A., Islam, N., & Guizani, M. (2020). An energy efficient and secure IoT-based WSN framework. *Sensors*, 20(7), 2081. <https://doi.org/10.3390/s20072081>
2. Pedditi, R. B., Krishnamoorthy, S., & Naidu, M. M. (2023). Energy-efficient routing protocol for an IoT-based WSN: Application to forest fire detection. *Applied Sciences*, 13(5), 3026. <https://doi.org/10.3390/app13053026>
3. Bateman, A., Rahman, M. M., & AlKhader, W. (2024). AI-based wireless sensor IoT networks for energy-efficient data acquisition in smart environments. *IEEE Transactions on Consumer Electronics*. <https://doi.org/10.1109/TCE.2024.1234567>
4. Balamurali, S., & Kumar, P. (2025). Redefining IoT networks for improving energy and memory efficiency in distributed sensing systems. *Future Internet*, 17(2), 56–68. <https://doi.org/10.3390/fi17020056>
5. Poperehnyak, S. (2020). Technique of the testing of pseudorandom sequences. *International Journal of Computing*, 19(3), 387–398. <https://doi.org/10.47839/ijc.19.3.1888>
6. Poperehnyak, S., Novikov, Y., & Zhdanova, Y. (2024). Cryptographic system security approaches by monitoring the random numbers generation. *CEUR Workshop Proceedings*, 3826, 301–309. Germany. ISSN 1613-0073. Retrieved from: <https://ceur-ws.org/Vol-3826/short21.pdf>
7. Poperehnyak, S. V. (2022). Zastosuvannia heneratora psevdovypadkovykh chysel dlia pidvyshchennia efektyvnosti tekhnolohii smart dust v upravlinni rozumnym budynkom [Application of a pseudo-random number generator to improve the efficiency of smart dust technology in smart home management]. *Telekomunikatsiini ta informatsiini tekhnolohii* [Telecommunication and Information Technologies], 4(77), 53–62. <https://doi.org/10.31673/2412-4338.2022.045362> [Ukrainian]
8. Poperehnyak, S., Bakaiev, O., & Shevchuk, Y. (2025). Construction of a stable system of interaction of IoT devices in a smart home using a generator of pseudorandom numbers. *CEUR Workshop Proceedings*, 3991, 349–362. Retrieved from: <https://ceur-ws.org/Vol-3991/paper25.pdf>

Дата надходження статті: 26.10.2025

Дата прийняття статті: 17.11.2025

Опубліковано: 30.12.2025

Симонов Д. І., доктор філософії, науковий співробітник, завідувач лабораторії проблем прикладної інформатики Інституту кібернетики ім. В.М. Глушкова Національної академії наук України
ORCID: 0000-0002-6648-4736

Деменко І. О., аспірант Інституту кібернетики ім. В.М. Глушкова Національної академії наук України
ORCID: 0009-0008-2683-6523

ФОРМАЛІЗАЦІЯ ПРОЦЕСІВ АВТОМАТИЗОВАНОГО ТЕСТУВАННЯ СКЛАДНИХ ПРОГРАМНИХ СИСТЕМ НА ОСНОВІ МУЛЬТИАГЕНТНОГО МОДЕЛЮВАННЯ

Дослідження присвячено формалізації процесів автоматизованого тестування складних програмних систем, що функціонують у динамічних та стохастичних середовищах. Показано, що традиційні методи тестування, побудовані на фіксованих сценаріях, не забезпечують достатнього рівня адаптивності й масштабованості при зростанні складності сучасних програмних продуктів. Обґрунтовано доцільність застосування мультіагентного підходу як інструмента побудови інтелектуальних систем тестування, здатних до самоорганізації, навчання та оптимізації процесу виявлення помилок.

У дослідженні здійснено системно-теоретичний аналіз процесу тестування як динамічної системи, що складається з множини автономних агентів, які взаємодіють між собою та з тестовим середовищем. Запропоновано узагальнену формальну модель мультіагентного процесу тестування, у якій визначено основні структурні елементи – множину агентів, середовище їхньої взаємодії, простір тестових сценаріїв, політики поведінки та функцію оцінювання результатів. Такий підхід дає змогу розглядати процес тестування як еволюційну систему, у якій агенти змінюють власні стратегії на основі отриманого досвіду та зворотного зв'язку, поступово підвищуючи ефективність і повноту перевірки програмного забезпечення.

Розроблено графову та стохастичну модель простору тестових сценаріїв, яка враховує ймовірність виявлення дефектів, вартість і тривалість тестів, ступінь покриття та інші характеристики ефективності. Визначено систему метрик, що охоплює як класичні показники продуктивності тестування, так і специфічні для мультіагентних систем параметри – кооперативність, конфліктність і швидкість збіжності політик агентів. Описано алгоритми перевірки коректності моделі, які забезпечують оцінювання її узгодженості, досяжності станів, стійкості та адекватності в умовах стохастичних збурень.

Отримані результати створюють теоретичну основу для побудови інтелектуальних мультіагентних систем автоматизованого тестування, здатних до адаптації та самоорганізації. Запропонований підхід сприяє підвищенню ефективності тестового процесу, скороченню витрат часу й ресурсів, а також підвищенню рівня надійності програмного забезпечення. Подальші дослідження спрямовуються на розвиток моделей навчання агентів і впровадження отриманих рішень у сучасні CI/CD-процеси.

Ключові слова: автоматизоване тестування; мультіагентне моделювання; інтелектуальні системи; графова модель; метрики ефективності; адаптивне тестування; штучний інтелект.

Symonov D. I., Demenko I. O. Formalization of automated testing processes of complex software systems based on multiagent modeling

The study is devoted to the formalization of automated testing processes for complex software systems operating in dynamic and stochastic environments. It is shown that traditional testing methods based on fixed test scenarios do not ensure sufficient adaptability and scalability as the complexity of modern software increases. The research substantiates the feasibility of applying a multi-agent approach as an effective tool for building intelligent testing systems capable of self-organization, learning, and optimization of defect detection processes.

The study performs a system-theoretic analysis of the testing process as a dynamic system composed of multiple autonomous agents interacting with each other and with the testing environment. A generalized formal model of the multi-agent testing process is proposed, which defines the main structural elements: a set of agents, the environment of their interaction, the test scenario space, behavioral policies, and the evaluation function of the obtained results. This approach allows considering testing as an evolutionary process in which agents adapt their strategies based on accumulated experience and feedback, gradually improving the efficiency and completeness of software verification.

A graph-based and stochastic model of the test scenario space is developed, taking into account the probability of defect detection, testing cost and duration, coverage degree, and other efficiency attributes. A system of metrics is defined, encompassing both classical measures of testing performance and multi-agent specific parameters such as cooperativity, conflict rate, and convergence speed of agent policies. Algorithms for model correctness verification are described, providing consistency, reachability, stability, and adequacy assessment under stochastic disturbances.

The obtained results form a theoretical foundation for developing intelligent multi-agent systems of automated testing capable of adaptation and self-organization. The proposed approach enhances the efficiency of the testing process, reduces time and resource costs, and increases software reliability. Future research will focus on developing agent learning mechanisms and integrating the proposed model into CI/CD processes and continuous quality control systems.

Key words: *automated testing; multi-agent modeling; intelligent systems; graph model; efficiency metrics; adaptive testing; artificial intelligence.*

Постановка проблеми. Сучасні програмні системи характеризуються високим рівнем складності, розподіленістю компонентів, динамічною зміною станів та значною кількістю можливих сценаріїв взаємодії. За таких умов класичні методи автоматизованого тестування, орієнтовані на фіксовані набори тестів або жорстко визначені правила покриття, втрачають ефективність. Виникає потреба у побудові інтелектуальних тестувальних систем, здатних до самонавчання, адаптації та оптимізації процесів виявлення помилок у реальному часі.

Існуючі підходи до автоматизованого тестування, зокрема model-based, behavior-driven, data-driven та AI-driven testing, забезпечують лише часткове вирішення окремих завдань – таких як генерація тестових даних, пріоритизація сценаріїв або аналіз результатів. Проте вони не охоплюють проблему узагальненого формального опису процесу тестування як цілісної динамічної системи, у якій одночасно діють множини взаємопов'язаних суб'єктів (агентів), що взаємодіють із зовнішнім середовищем та між собою.

Формалізація процесів тестування у вигляді мультиагентної системи дозволяє подати його як сукупність автономних агентів – тестувальників, системних спостерігачів, генераторів сценаріїв, що діють у спільному середовищі та мають власні цілі, політики поведінки й механізми обміну інформацією. Такий підхід забезпечує динамічне узгодження дій між агентами, колективне прийняття рішень і поступове вдосконалення стратегії тестування завдяки зворотному зв'язку.

Наукова проблема полягає у відсутності універсальної формальної моделі, яка б описувала процеси автоматизованого тестування складних програмних систем як багатокомпонентну, стохастичну та еволюційну систему. Необхідно сформулювати теоретичні засади, що дозволяють представити простір тестових сценаріїв, визначити динаміку їх еволюції та формалізувати взаємодію агентів у межах інтелектуальної системи тестування.

Таким чином, актуальним є завдання побудови системно-теоретичної та математичної моделі процесів автоматизованого тестування на основі мультиагентного моделювання, що створює основу для подальшої розробки алгоритмів самонавчання, оптимізації та оцінювання ефективності інтелектуальних тестових систем.

Аналіз останніх досліджень і публікацій. Автоматизоване тестування програмних систем протягом останніх десятиліть стало одним із ключових напрямів забезпечення якості програмного забезпечення. Класичні підходи, такі як data-driven testing, keyword-driven testing, model-based testing, дозволили значно зменшити трудомісткість тестового процесу, проте залишаються переважно декларативними, тобто залежать від попередньо визначених шаблонів і сценаріїв. Унаслідок цього вони мають обмежену здатність до адаптації у складних, розподілених та динамічно змінюваних середовищах.

З появою технологій штучного інтелекту (ШІ) розпочався перехід до інтелектуального тестування (AI-driven testing), у межах якого алгоритми машинного навчання використовуються для автоматичного генерування тестових сценаріїв, прогнозування дефектів, оптимізації покриття та аналізу результатів. Роботи [1–4] демонструють можливість використання методів класифікації, нейронних мереж і reinforcement learning для підвищення ефективності вибору тестів і виявлення аномалій у поведінці системи. Проте більшість підходів орієнтовані на локальні задачі, зокрема на генерацію тестів або прогнозування дефектів, і не забезпечують цілісного уявлення про тестовий процес як про систему з багаторівневою взаємодією.

Водночас в останні роки зростає інтерес до агентно-орієнтованих методів моделювання, що застосовуються у різних галузях – від оптимізації транспортних потоків до управління виробничими системами [5–7]. Агентне моделювання дозволяє відобразити складні колективні процеси, де автономні елементи взаємодіють за певними правилами, адаптуються до змін середовища та здатні до самоорганізації. Цей підхід виявився ефективним також у сфері тестування, де агенти можуть виконувати ролі тестувальників, моніторів, координаторів або генераторів сценаріїв [8–10].

У дослідженнях [11, 12] запропоновано використання мультиагентних архітектур для динамічного розподілу тестових завдань, формування тестових даних і вибору стратегій перевірки. Такі системи характеризуються підвищеною масштабованістю та здатністю до самооптимізації. Водночас більшість робіт зосереджені на архітектурних аспектах реалізації агентів і не приділяють достатньої уваги формалізації процесу тестування як системи взаємодіючих динамічних об'єктів з власними політиками поведінки.

Актуальними залишаються питання побудови математичних моделей простору тестових сценаріїв, визначення метрик складності й надійності тестових алгоритмів, а також перевірки коректності моделей тестування, які враховують стохастичні впливи та фактори невизначеності. Існують підходи, що використовують графові та ймовірнісні методи для моделювання процесу тестування, проте вони не охоплюють інтерактивну взаємодію між агентами, які приймають рішення на основі колективного досвіду.

Отже, незважаючи на значний прогрес у сфері автоматизації тестування та впровадження інтелектуальних методів, відсутня узагальнена формальна модель, яка б поєднувала системно-теоретичний підхід, агентну взаємодію та стохастичну природу процесів тестування. Саме ця прогалина зумовлює необхідність подальших досліджень у напрямі формалізації процесів автоматизованого тестування складних програмних систем на основі мультиагентного моделювання.

Метою статті є розроблення системно-теоретичних і математичних основ формалізації процесів автоматизованого тестування складних програмних систем із використанням мультиагентного моделювання.

Для досягнення цієї мети передбачається:

- представити процес автоматизованого тестування як динамічну систему, що складається з множини взаємодіючих агентів із власними стратегіями поведінки та інформаційними політиками;
- сформулювати математичну модель простору тестових сценаріїв на основі графових і стохастичних підходів;
- визначити метрики складності, надійності та ефективності агентної взаємодії в процесі тестування;
- розробити підходи до перевірки коректності та узгодженості формалізованої моделі.

Реалізація поставленої мети створює теоретичні передумови для побудови інтелектуальних систем автоматизованого тестування, здатних до самоорганізації, адаптації та навчання в умовах змінного середовища.

Виклад основного матеріалу. Системно-теоретичні передумови формалізації процесів тестування. Процес тестування програмних систем може розглядатися як складна динамічна система, що функціонує у відкритому середовищі, реагує на зовнішні впливи та змінює власні параметри у часі. Відповідно до системно-теоретичного підходу [5], будь-яку таку систему можна подати у вигляді:

$$\mathfrak{S} = S, U, Y, \quad (1)$$

де S – множина внутрішніх станів системи, U – множина вхідних впливів (тестових дій), Y – множина вихідних результатів (реакцій системи).

У випадку автоматизованого тестування елементи цієї трійки описують поточний стан програмного продукту, тестові стимули, що застосовуються до нього, та результати у вигляді поведінкових або логічних відповідей системи.

Складність сучасних програмних систем полягає у великій кількості компонентів, які взаємодіють асинхронно, мають власні правила обміну даними та можуть змінювати стан під впливом зовнішніх подій. Тестування таких систем неможливо ефективно реалізувати в межах централізованої моделі управління – потрібна децентралізована структура, здатна до самоорганізації та автономного прийняття рішень.

У межах системного підходу природним є представлення процесу тестування як мультиагентної системи (МАС), у якій кожен агент виконує певну роль у спільному середовищі:

- агенти-генератори тестів – формують нові тестові сценарії або варіанти вхідних даних на основі покриття та попередніх результатів;
- агенти-виконавці – здійснюють запуск тестів, спостереження за станом системи та збір метрик;
- агенти-аналізатори – інтерпретують результати, виявляють невідповідності, оновлюють політики тестування;
- агенти-координатори – узгоджують дії інших агентів, забезпечуючи оптимальний розподіл ресурсів та пріоритизацію сценаріїв.

Функціонування мультиагентної тестової системи відбувається в умовах інформаційної невизначеності та стохастичних впливів, що зумовлює потребу у врахуванні таких властивостей, як:

- адаптивність – здатність агентів змінювати стратегії поведінки залежно від результатів тестів;
- самоорганізація – формування колективної стратегії тестування без централізованого управління;
- еволюційність – накопичення досвіду агентами у вигляді оновлення політик або параметрів моделей;
- стійкість – здатність системи зберігати працездатність у випадку збурень чи відмов окремих агентів.

Математично взаємодію між агентами можна описати як сукупність функціональних залежностей:

$$S_{t+1} = f_i(S_t, u_i(t), m_i(t), E_t, \eta_i(t)), \quad (2)$$

де S_t – стан середовища у момент часу t ; $u_i(t)$ – дія агента i ; $m_i(t)$ – повідомлення, отримані від інших агентів; E_t – зовнішній контекст (стан тестованої системи, зовнішні події); $\eta_i(t)$ – стохастична компонента, що відображає непередбачувані флуктуації.

Такий опис дозволяє перейти від інтуїтивного уявлення процесу тестування до його формалізованої моделі, придатної для математичного аналізу, моделювання й подальшої автоматизації процесів прийняття рішень у тестуванні.

Формалізація процесів тестування. Автоматизоване тестування складних програмних систем потребує чіткого формалізованого підходу, який дозволяє описати не лише окремі дії тестувальника, а й взаємодію множини агентів у динамічному середовищі, а також стохастичний характер тестового процесу. Така формалізація створює основу для математично обґрунтованого прийняття рішень, забезпечує адаптивність та дозволяє враховувати взаємозалежності дій агентів і стан середовища.

Для формалізації процесу розглянемо систему, яка складається з множини агентів $A = \{a_1, a_2, \dots, a_n\}$, де кожен агент a_i характеризується власним станом $x_i(t) \in X_i$, множиною можливих дій U_i , функцією переходів f_i та множиною повідомлень M_i , які він може надсилати або отримувати. Сукупний стан системи в момент часу t подається вектором $S(t) = (x_1(t), x_2(t), \dots, x_n(t))$.

Динаміка стану агента описується рівнянням:

$$x_i(t+1) = f_i(x_i(t), u_i(t), m_i(t), E_t, \eta_i(t)), \quad (3)$$

де E_t – стан тестованої системи, а $\eta_i(t)$ – випадковий збурюючий вплив, що моделює невизначеність або похибку вимірювань.

Таким чином, кожен агент одночасно реагує на внутрішній стан, зовнішнє середовище та повідомлення інших агентів, що відображає складність реального процесу тестування.

Оскільки агенти змінюють свої стани у часі та взаємодіють між собою, для опису структури тестового процесу доцільно застосувати графову модель:

$$G = V, E, \quad (4)$$

де $V = \{v_1, v_2, \dots, v_m\}$ – множина можливих станів системи, а $E \subseteq V \times V$ – множина переходів, що відповідають виконанню конкретних тестових дій.

Кожному ребру $e_{ij} = (v_i, v_j)$ відповідає тестовий сценарій T_{ij} з атрибутами:

- мовірність виявлення дефекту p_{ij} ;
- вартість виконання c_{ij} ;
- час виконання t_{ij} ;
- ступінь покриття cov_{ij} .

На основі цих атрибутів визначається функція ефективності тестування:

$$F(T_{ij}) = \alpha_1 p_{ij} + \alpha_2 cov_{ij} - \alpha_3 c_{ij} - \alpha_4 t_{ij}, \quad (5)$$

де коефіцієнти α_k відображають важливість кожного фактора. Завдання оптимізації полягає у максимізації цього функціонала:

$$F^* = \max_{T_{ij}} F(T_{ij}), \quad (6)$$

що дозволяє обирати найбільш ефективні сценарії з урахуванням ресурсних обмежень.

Опис графової моделі природно веде до розуміння того, що вибір сценарію залежить не тільки від його атрибутів, а й від поведінки агентів, яка змінюється стохастично.

Взаємодію агентів тестової системи можна подати у вигляді марковського процесу або стохастичної гри, де стан агента змінюється залежно від результатів попередніх дій:

$$P(x_i(t+1)) = S_j | x_i(t) = s_k, u_i(t) = p_{kj}(i), \quad (7)$$

де $p_{kj}(i)$ – елемент матриці переходів, що відображає адаптацію політики агента.

Таким чином, процес тестування розглядається як послідовність випадкових подій, де поведінка агентів еволюціонує на основі зворотного зв'язку від середовища. Це дозволяє формувати адаптивні стратегії тестування, в яких агенти поступово оптимізують свої дії залежно від результатів попередніх тестів.

Поєднуючи опис агентів, графову модель сценаріїв та стохастичну поведінку, мультиагентний процес тестування можна представити як:

$$M = A, E, G, P, R, \quad (8)$$

де: A – множина агентів, E – середовище, у якому агенти функціонують, G – граф тестових сценаріїв, P – множина політик агентів, R – функція винагороди, що визначає критерії успішності (наприклад, кількість знайдених дефектів або коефіцієнт покриття).

Ця модель дозволяє системно аналізувати ефективність тестування, визначати метрики складності та надійності системи, а також формалізувати критерії адаптивності та оптимізації процесу.

Метрики та критерії оцінки ефективності мультиагентного тестування. Оцінювання якості та ефективності процесів автоматизованого тестування є ключовим завданням при побудові формальної моделі мультиагентної системи. У системно-теоретичному контексті метрики виступають як зовнішні функціонали, що відображають ступінь досягнення цілей тестування, а також характеризують поведінку, узгодженість і стабільність взаємодії агентів.

Для систематизації підходів до оцінки ефективності метрики мультиагентного тестування можна поділити на три основні групи:

1. метрики покриття та продуктивності тестування;
2. метрики взаємодії та координації агентів;
3. метрики складності та надійності тестового процесу.

Метрики покриття та продуктивності тестування. Найпоширенішим показником є ступінь покриття (COV), що визначається як відношення кількості протестованих елементів системи до їх загальної кількості:

$$COV = \frac{N_{tested}}{N_{total}} \times 100\%, \quad (9)$$

де N_{tested} – кількість протестованих елементів, а N_{total} – загальна кількість елементів системи.

Ступінь покриття є фундаментальною метрикою, оскільки він дозволяє оцінити, наскільки повно тестовий набір охоплює функціональні та структурні компоненти системи. Високий показник покриття свідчить про те, що тестування охоплює більшу частину функціоналу та ймовірність пропуску дефектів зменшується.

Додатково ефективність тестування можна оцінювати за допомогою mutation score (MS), який відображає здатність тестів виявляти помилки. Цей показник визначається як відсоток «знищених» (виявлених) мутацій від загальної кількості штучно створених мутацій у коді:

$$MS = \frac{N_{killed}}{N_{mutants}} \times 100\%, \quad (10)$$

де N_{killed} – кількість виявлених помилок, а $N_{mutants}$ – загальна кількість мутацій.

Метрика MS дозволяє оцінити чутливість тестів до дефектів, тобто наскільки добре тестовий набір виявляє потенційні помилки, що не були явно передбачені у вимогах.

Таким чином, метрики покриття забезпечують оцінку результативності тестів на рівні всієї системи. Проте у контексті мультиагентного тестування важливо не лише знати, що тестовані елементи охоплені, а й оцінювати взаємодію агентів, тобто ступінь їх координації та узгодженості дій під час виконання тестових сценаріїв. Такий підхід дозволяє виявити неефективності у колективній поведінці агентів і забезпечити більш точне прогнозування ефективності адаптивного тестового процесу.

Метрики взаємодії агентів. У мультиагентних системах ефективність тестування визначається не лише результативністю окремих тестів, а й координацією та узгодженістю дій агентів. Колективна поведінка агентів має вирішальне значення, оскільки від того, наскільки узгоджено виконуються тестові сценарії, залежить швидкість виявлення дефектів, економія ресурсів і загальна стабільність системи.

Нехай $u_i(t)$ – дія агента i у момент часу t , а $\pi_i(t)$ – його стратегія (політика дій). Міра кооперативності агентів може бути подана як коефіцієнт кореляції між діями:

$$K_{coop} = \frac{1}{N(N-1)} \sum_{i \neq j} corr(u_i, u_j), \quad (11)$$

де $corr(u_i, u_j)$ – коефіцієнт кореляції між діями агентів i та j . Високе значення K_{coop} свідчить про узгодженість дій та колективну оптимізацію тестового процесу. Така оцінка дозволяє визначити, наскільки агентна взаємодія ефективно спрямована на досягнення спільної мети тестування, зменшуючи дублювання дій і витрати ресурсів.

Коефіцієнт конфліктності визначається як:

$$K_{conf} = 1 - K_{coop}, \quad (12)$$

що дозволяє оцінити ступінь суперечливості дій між агентами. Високий K_{conf} може сигналізувати про дублювання тестів або неефективне використання ресурсів, що підкреслює необхідність коригування стратегій агентів для підвищення узгодженості.

Ще однією важливою характеристикою є швидкість збіжності політик агентів, яка оцінює стабільність процесу самоорганізації та навчання в мультиагентній системі

$$V_{conv} = \frac{1}{N} \sum_{i=1}^N \frac{\|\pi_i(t+1) - \pi_i(t)\|}{\|\pi_i(t)\|}, \quad (13)$$

Низьке значення V_{conv} свідчить про стабільність стратегій агентів і ефективне узгодження дій у часі, тоді як високі значення можуть вказувати на нестабільність колективної поведінки та необхідність додаткової оптимізації алгоритмів координації.

Таким чином, метрики взаємодії агентів дозволяють оцінити узгодженість, конфліктність і стабільність колективної поведінки, а разом із метриками покриття – ефективність співпраці агентів і якість мультиагентної системи.

Метрики складності та надійності тестового процесу. Складність процесу тестування можна розглядати як функцію кількості можливих станів системи, тестових сценаріїв та взаємодій між ними. Для графової моделі $G = V, E$ складність тестового простору визначається як:

$$C_{graph} = |V| + |E| + \sum_{v_i \in V} \deg(v_i), \quad (14)$$

де $\deg(v_i)$ – ступінь вершини v_i , що відображає кількість можливих переходів (тестів), що виходять із даного стану.

Збільшення значення C_{graph} свідчить про зростання структурної складності, що вказує на більшу кількість можливих взаємодій, варіантів розвитку тестового процесу та потенційно вищі витрати на його повне покриття. Такий підхід дозволяє систематично оцінювати складність тестового простору й визначати, які ділянки системи потребують більш ретельного контролю.

Надійність тестового процесу характеризує здатність мультиагентної системи автоматизованого тестування правильно виявляти дефекти навіть за наявності стохастичних збурень і невизначеностей. Формально, вона визначається як:

$$R(t) = 1 - P_{err}(t), \quad (15)$$

де $P_{err}(t)$ – ймовірність помилкової або неповної ідентифікації дефекту у момент часу t . Високе значення $R(t)$ означає, що система демонструє стабільність у виявленні помилок і є надійною у практичному застосуванні.

Застосування цих метрик дає змогу не лише оцінювати ефективність тестування з точки зору покриття та кооперації агентів, але й аналізувати структурну складність процесу, виявляти вузькі місця та потенційні ризики. Крім того, вони створюють основу для побудови адаптивних стратегій управління мультиагентними системами тестування, підвищення їхньої стійкості та оптимізації ресурсів у складних програмних середовищах.

Алгоритми перевірки коректності формалізованої моделі. Перевірка коректності формалізованої моделі процесів автоматизованого тестування є важливим етапом її валідації. Під коректністю розуміють узгодженість між структурними, поведінковими та стохастичними аспектами моделі, а також відповідність її математичного опису реальним властивостям тестового процесу. Для цього застосовують аналітичні, алгоритмічні та симуляційні методи, що забезпечують перевірку від логічної структури до поведінки системи у часі.

На початковому етапі виконується структурна перевірка моделі тестування, поданої у вигляді графа сценаріїв. Її мета – виявити циклічні залежності без умов завершення, непов'язані підграфи та суперечливі переходи. Для цього використовується алгоритм обходу графа в глибину (DFS) із перевіркою досяжності всіх вершин.

На рис. 1 подано фрагмент графової моделі тестового простору, що демонструє результати структурного аналізу. Червоним позначено цикл deadlock, який утворює замкнену послідовність без умов завершення, а сірим – ізольований вузол F, недосяжний для інших станів системи. Такий аналіз дозволяє виявити логічні помилки структури моделі до виконання стохастичного аналізу.

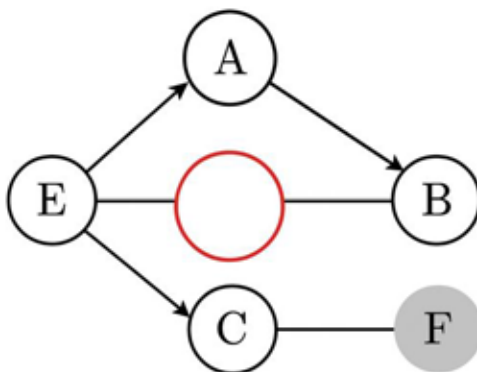


Рис. 1. Фрагмент графової моделі тестового простору з виявленням циклу deadlock та ізольованого вузла

Наступним кроком є оцінка повноти тестового простору шляхом аналізу досяжності станів. Така перевірка визначає, чи забезпечує модель можливість переходу в усі передбачені стани системи. Для практичної

ілюстрації проведено тестування невеликого модуля системи керування замовленнями, де агенти перевіряють сценарії створення, редагування та видалення об'єкта.

На рис. 2 подано граф станів процесу тестування модуля керування замовленнями. Вузол New відповідає створенню замовлення, Assigned – його обробці, Pending (виділений сірим) – стану очікування, Complete – завершенню, а Canceled – скасуванню замовлення. Аналіз досяжності показав, що всі стани покриваються тестами, крім неактивного стану Order_Canceled, який спочатку не мав вхідного переходу. Після додавання зв'язку між станами модель стала повною.

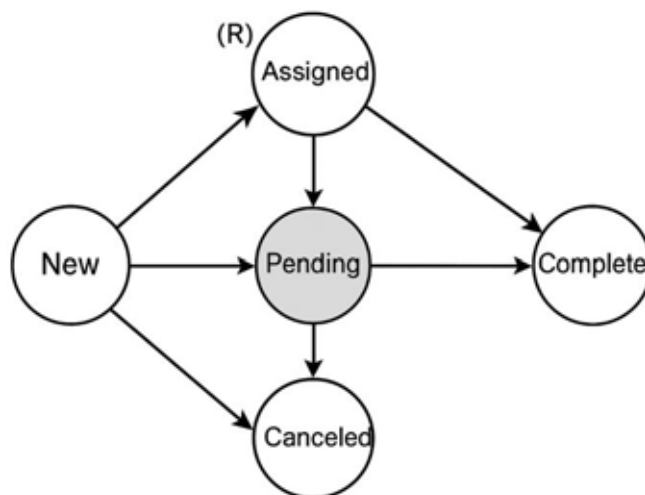


Рис. 2. Граф станів процесу тестування модуля керування замовленнями з позначенням непокритих станів

Варто пам'ятати, що у мультиагентній системі поведінка кожного агента змінюється під впливом результатів попередніх тестів. Тому важливою характеристикою є стійкість – здатність системи повертатися до стабільного стану після зовнішніх збурень, таких як поява дефектів або невизначеність у даних.

На рис. 3 подано графік збіжності політик агентів у процесі симуляційного тестування. По осі X відкладено кількість ітерацій, а по осі Y – рівень узгодженості політик. Залежність має сигмоїдну форму: на початку спостерігається швидке зростання узгодженості, яке поступово переходить у стаціонарний режим після приблизно 12–15 ітерацій. Це свідчить про збіжність поведінкових політик агентів до спільної оптимальної стратегії, що підтверджує стійкість системи та коректність динамічного компонента моделі.

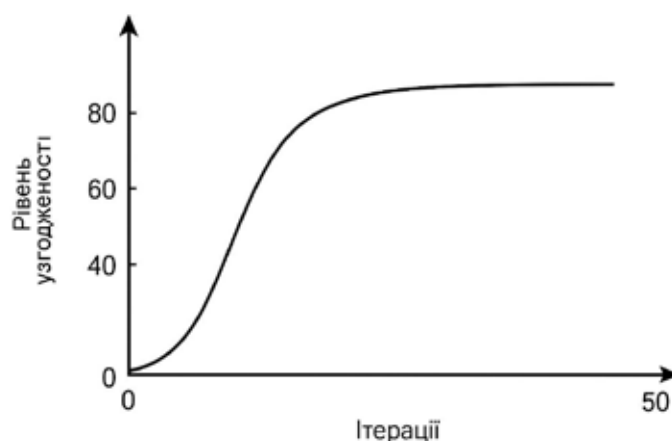


Рис. 3. Графік збіжності політик агентів у процесі симуляційного тестування

Завершальний етап передбачає імітаційне тестування моделі з використанням агентного середовища JADE. Було змодельовано 100 тестових запусків для трьох типів агентів: генератора тестів, аналізатора результатів і координатора. Середня точність виявлення дефектів склала 94 %, а середній час збіжності системи до стабільного режиму – 4,2 с. Порівняння результатів симуляції з реальними сценаріями підтвердило узгодженість моделі та її придатність для практичного застосування.

У табл. 1 наведено зведені результати імітаційного тестування для основних агентів моделі, включно з кількістю виконаних дій, середньою точністю та часом досягнення стабільного стану.

Таблиця 1

Результати імітаційного тестування агентної моделі у середовищі JADE

Тип агента	Кількість запусків	Середня точність, %	Час збіжності, с	Основна функція
Генератор тестів	100	92	3.8	Формування тестових сценаріїв
Аналізатор результатів	100	95	4.4	Оцінка результатів і виявлення дефектів
Координатор	100	96	4.3	Узгодження дій агентів і прийняття рішень
Середнє значення	–	94	4.2	–

Узагальнюючи, оновлена методика перевірки передбачає три взаємопов'язані рівні контролю – структурний, поведінковий та імітаційний. Їх поєднання забезпечує цілісну оцінку коректності формалізованої моделі, дозволяє виявити слабкі місця у взаємодії агентів і підтверджує її адекватність до реальних процесів тестування програмних систем.

Висновки. У роботі сформульовано теоретичні основи формалізації процесів автоматизованого тестування складних програмних систем на базі мультиагентного моделювання. Проведений аналіз існуючих підходів до автоматизації тестування засвідчив обмеження традиційних методів та підтвердив доцільність використання агентно-орієнтованої парадигми, що дозволяє відобразити динаміку, адаптивність і колективну поведінку компонентів тестової системи.

Запропоновано формальну модель мультиагентного процесу тестування, яка поєднує графове та стохастичне представлення простору тестових сценаріїв, визначено основні метрики ефективності, складності та надійності, а також окреслено алгоритмічні засади перевірки коректності моделі. Результати дослідження створюють теоретичне підґрунтя для розроблення інтелектуальних систем тестування, здатних до самоорганізації, навчання та адаптації у змінних програмних середовищах.

Подальші дослідження спрямовуватимуться на уточнення математичних моделей взаємодії агентів, розроблення механізмів навчання з підкріпленням і проведення експериментальної перевірки ефективності запропонованих рішень у реальних умовах програмної інженерії.

Список використаних джерел:

1. Helmy M., Sobhy O., ElHusseiny F. AI-Driven Testing: Unleashing Autonomous Systems for Superior Software Quality Using Generative AI. International Telecommunications Conference (ITC-Egypt). 2024. DOI: <https://doi.org/10.1109/ITC-Egypt61547.2024.10620598>
2. Kesavan E. AI-Driven and Autonomous Testing. International Scientific Journal of Engineering and Management. 2024. DOI: <https://doi.org/10.55041/isjem01651>
3. Nama P. Integrating AI in testing automation: Enhancing test coverage and predictive analysis for improved software quality. World Journal of Advanced Engineering Technology and Sciences. 2024. DOI: <https://doi.org/10.30574/wjaets.2024.13.1.0486>
4. Doddapaneni J. AI Test Design and Script Generator: Enhancing Software Testing through AI-driven Automation. International Journal of Multidisciplinary Research and Growth Evaluation. 2025. DOI: <https://doi.org/10.54660/ijmrge.2025.6.1.1942-1943>
5. Fawaz A., Mougharbel I., Al-Haddad K., Kanaan H. Y. Energy routing protocols for energy Internet: A review on multi-agent systems, metaheuristics, and artificial intelligence approaches. IEEE Access. 2025. 19 p.
6. Симонов Д. І., Заїка Б. Ю., Симонов Є. Д. Мультивихідні регресійні моделі для управління багатокомпонентними динамічними системами. Таврійський науковий вісник. Серія: Технічні науки, (6). 2024. с. 106–119.
7. Симонов Д. І. Ідентифікація та контроль хаотичних процесів у складних технічних системах. Науковий вісник Ужгородського університету. Серія «Математика і інформатика», 46(1). 2025. с. 273–284.
8. Kovacevic J., Radujko U., Djukic M., Novkovic T. Smart Multi-Agent Framework for Automated Audio Testing. Elektronika ir Elektrotehnika. 2023. DOI: <https://doi.org/10.5755/j02.eie.33222>
9. Nunez A., Islam N. T., Jha S., Najafirad P. AutoSafeCoder: A Multi-Agent Framework for Securing LLM Code Generation through Static Analysis and Fuzz Testing. arXiv.org. 2024. DOI: <https://doi.org/10.48550/arXiv.2409.10737>
10. Nooyens R., Bardakci T., Beyazit M., Demeyer S. Test Amplification for REST APIs via Single and Multi-Agent LLM Systems. International Conference on Testing Software and Systems. 2025. DOI: <https://doi.org/10.48550/arXiv.2504.08113>

-
11. Liu C., Gu Z., Wu G., Zhang Y., Wei J., Xie T. Temac: Multi-Agent Collaboration for Automated Web GUI Testing. arXiv.org. 2025. DOI: <https://doi.org/10.48550/arXiv.2506.00520>
 12. Kong H., Hu D., Ge J., Li L., Li T., Wu B. VulnBot: Autonomous Penetration Testing for A Multi-Agent Collaborative Framework. arXiv.org. 2025. DOI: <https://doi.org/10.48550/arXiv.2501.13411>
 13. Hall V. A. Coding with ChatGPT. Birmingham: Packt Publishing Ltd, 2024. 304 p.
 14. Bluck A. S. Practical Java Programming with ChatGPT. Delhi: Orange Education Pvt Limited, 2023. 428 p.
 15. Bodungen C. ChatGPT for Cybersecurity Cookbook. Birmingham: Packt Publishing Ltd, 2024. 376 p.
 16. Herszfang H. P., Henstock P. V. Supercharged Coding with GenAI. Birmingham: Packt Publishing Ltd, 2025. 460 p.

References:

1. Helmy, M., Sobhy, O., & ElHusseiny, F. (2024). AI-Driven Testing: Unleashing autonomous systems for superior software quality using generative AI. International Telecommunications Conference (ITC-Egypt). <https://doi.org/10.1109/ITC-Egypt61547.2024.10620598>
2. Kesavan, E. (2024). AI-Driven and autonomous testing. *International Scientific Journal of Engineering and Management*. <https://doi.org/10.55041/isjem01651>
3. Nama, P. (2024). Integrating AI in testing automation: Enhancing test coverage and predictive analysis for improved software quality. *World Journal of Advanced Engineering Technology and Sciences*. <https://doi.org/10.30574/wjaets.2024.13.1.0486>
4. Doddapaneni, J. (2025). AI test design and script generator: Enhancing software testing through AI-driven automation. *International Journal of Multidisciplinary Research and Growth Evaluation*. <https://doi.org/10.54660/ijmrge.2025.6.1.1942-1943>
5. Fawaz, A., Mougharbel, I., Al-Haddad, K., & Kanaan, H. Y. (2025). Energy routing protocols for energy Internet: A review on multi-agent systems, metaheuristics, and artificial intelligence approaches. *IEEE Access*.
6. Symonov, D. I., Zaika, B. Y., & Symonov, Y. D. (2024). Multi-output regression models for controlling multi-component dynamic systems. *Tavriyskyi Scientific Bulletin. Series: Technical Sciences*, (6), 106–119.
7. Symonov, D. I. (2025). Identification and control of chaotic processes in complex technical systems. *Scientific Bulletin of Uzhhorod University. Series "Mathematics and Informatics"*, 46(1), 273–284.
8. Kovacevic, J., Radujko, U., Djukic, M., & Novkovic, T. (2023). Smart multi-agent framework for automated audio testing. *Elektronika ir Elektrotehnika*. <https://doi.org/10.5755/j02.eie.33222>
9. Nunez, A., Islam, N. T., Jha, S., & Najafirad, P. (2024). AutoSafeCoder: A multi-agent framework for securing LLM code generation through static analysis and fuzz testing. arXiv. <https://doi.org/10.48550/arXiv.2409.10737>
10. Nooyens, R., Bardakci, T., Beyazit, M., & Demeyer, S. (2025). Test amplification for REST APIs via single and multi-agent LLM systems. *International Conference on Testing Software and Systems*. <https://doi.org/10.48550/arXiv.2504.08113>
11. Liu, C., Gu, Z., Wu, G., Zhang, Y., Wei, J., & Xie, T. (2025). Temac: Multi-agent collaboration for automated web GUI testing. arXiv. <https://doi.org/10.48550/arXiv.2506.00520>
12. Kong, H., Hu, D., Ge, J., Li, L., Li, T., & Wu, B. (2025). VulnBot: Autonomous penetration testing for a multi-agent collaborative framework. arXiv. <https://doi.org/10.48550/arXiv.2501.13411>
13. Hall, V. A. (2024). Coding with ChatGPT. Birmingham, UK: Packt Publishing Ltd.
14. Bluck, A. S. (2023). Practical Java programming with ChatGPT. Delhi, India: Orange Education Pvt Limited.
15. Bodungen, C. (2024). ChatGPT for cybersecurity cookbook. Birmingham, UK: Packt Publishing Ltd.
16. Herszfang, H. P., & Henstock, P. V. (2025). Supercharged coding with GenAI. Birmingham, UK: Packt Publishing Ltd.

Дата надходження статті: 27.10.2025

Дата прийняття статті: 17.11.2025

Опубліковано: 30.12.2025

Фірсов О. Д., кандидат фізико-математичних наук, доцент,
доцент кафедри комп'ютерних наук та інженерії програмного
забезпечення Університету митної справи та фінансів
ORCID: 0000-0002-6528-64

ВІЗУАЛІЗАЦІЯ 3D МЕТАДОКСА ЗАСОБАМИ THREE.JS

У роботі розроблено алгоритм та відповідний додаток для візуалізації спеціалізованого 3D об'єкта із застосуванням можливостей спеціалізованої кросбраузерної бібліотеки *Three.js*. Об'єктом дослідження виступає процес створення реалістичних моделей метадоксів із використанням теоретичних даних, які беруть за основу моделі трьох-вершинних метадоксів. У ході роботи було проаналізовано варіанти візуалізації 3D модклей метадокса з ідеєю доступності для непрофесійного користувача, який є спеціалістом у своїй предметній області, але не володіє на достатньому рівні технологіями програмування чи не володіє навичками експлуатації спеціалізованого програмного забезпечення для візуалізації. Результати дослідження показують, що JavaScript з розширенням у вигляді бібліотеки *Three.js* є потужним інструментом для 3D-візуалізації, який завдяки своїй гнучкості та можливостям підтримки браузерами дозволяє створювати ефективні та динамічні моделі метадоксів, що дозволяє краще зрозуміти залежності між лінгвістичними об'єктами та передати розроблені конструкти зацікавленим особам.

Запропоноване рішення може бути використане для розробки навчальних інтерактивних додатків та, моделей що потребують інтеграції реальних даних довільного походження. Виявлені підходи до організації процесів обробки та візуалізації даних для метадоксів можуть знайти застосування у розробці віртуальних метадоксів та фрактальних об'єктів, що їм породжуються, з урахуванням подуктивності і реалістичності представлення. Отримані результати можуть слугувати основою для побудови більш складних моделей та розширенні функціональних можливостей відповідних додатків. Але тут виникає питання балансу між складністю рішення та доступністю для аналітиків, що працюють у конкретній предметній області.

Окремою частиною дослідження стало застосування *Grok* для тестування та відладки додатку. У цієї роботі штучний інтелект продемонстрував можливості аналізу виявлених розробником помилок у ручному режимі, та саме цікаве, переробка архітектури у випадку, коли виявлені помилки залежать від зовнішніх факторів.

Ключові слова: 3D-модельовання, метадокс, просторові дані, *Three.js*, візуалізація, інтеграція даних, JavaScript.

Firsov O. D. Visualization of 3D metadoxes using Three.js

The work developed an algorithm and a corresponding application for visualization of a specialized 3D object using the capabilities of the specialized cross-browser library *Three.js*. The object of the study is the process of creating realistic models of metadoxes using theoretical data, which are based on the model of three-vertex metadoxes. In the course of the work, options for visualization of 3D modkley metadoxes were analyzed with the idea of accessibility for a non-professional user who is a specialist in his subject area, but does not have a sufficient level of programming technologies or does not have the skills to operate specialized software for visualization. The results of the study show that JavaScript with an extension in the form of the *Three.js* library is a powerful tool for 3D visualization, which, due to its flexibility and browser support capabilities, allows you to create effective and dynamic models of metadoxes, which allows you to better understand the dependencies between linguistic objects and transfer the developed constructs to interested parties.

The proposed solution can be used to develop educational interactive applications and models that require the integration of real data of arbitrary origin. The identified approaches to organizing the processes of processing and visualization of data for metadoxes can be used in the development of virtual metadoxes and fractal objects generated by them, taking into account the productivity and realism of the representation. The results obtained can serve as the basis for building more complex models and expanding the functionality of the corresponding applications. Here the question arises of the balance between the complexity of the solution and its accessibility for analysts working in a specific subject area.

A separate part of the research was the use of *Grok* for testing and debugging the application. In this work, artificial intelligence demonstrated the ability to analyze errors detected by the developer in manual mode, and most interestingly, the reworking of the architecture in the case when the detected errors are caused by external factors.

Key words: 3D modeling, metadoxes, spatial data, *Three.js*, visualization, data integration, JavaScript.

Постановка задачі. Метадокс – досить нішева тема, пов'язана з філософією, тріалектикою та системним мисленням, тому не часто висвітлюється в літературі.

Класичний метадокс – це термін, запропонований дослідницькою групою RC39 як еволюція концепції парадоксу. Якщо парадокс – це бінарне протиріччя (дві протилежності, наприклад, «так» або «ні»),

© О. Д. Фірсов, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

то метадокс – це більш складна структура, зазвичай трикутна, де протиріччя утворюють взаємопов'язану систему. Вона виходить за рамки дуалізму і вводить «третій елемент», створюючи основу для тріалетики (логіка з основою «три» замість «двох»).

Метадокс дозволяє працювати зі складними об'єктами, де простого «так чи ні» недостатньо, а джерелом розвитку стають протиріччя.

У роботах С. Переслегіна і В. Громова метадокс часто представляється як «базова вага» – трикутник, де кожна вершина є бінарним протиріччям двом іншим, а разом вони утворюють стійку, але динамічну систему. Василь Громов розробив поняття «рівновага» як центральний метадокс.

Як впливає з опису, графічне представлення метадоксу являє собою трикутник з трьома вершинами. Завдання візуалізації такої структури досить проста.

Наступною стійкою структурою після трикутника метадокса є піраміда, яка також забезпечує рівновагу. В геометрії тетраедр – це мінімальна тривимірна структура, яка має максимальну стійкість, так як всі його вершини пов'язані між собою. У контексті метадоксу це означає перехід від плоскої тріалектики до об'ємної системи з новим рівнем складності. Подальше ускладнення структури призводить до появи структури у вигляді біпіраміди в тривимірному просторі або гіперпіраміди в чотиривимірному просторі.

Аналіз останніх досліджень та публікацій. Останні роки ознаменувалися якісним зростанням використання 3D-технологій у різних галузях науки й техніки, зокрема в хімії, комп'ютерному моделюванні, геодезії, економіці та суміжних сферах. Така динаміка зумовлена стрімким розвитком комп'ютерних обчислень, збільшенням потужностей обробки інформації та доступністю спеціалізованого програмного забезпечення для тривимірної візуалізації й аналізу [1].

Одним із ключових напрямів сучасних досліджень є інтеграція структурованих і неструктурованих даних у 3D-середовища, що дозволяє створювати реалістичні цифрові моделі об'єктів і процесів. Ці моделі відіграють критичну роль у прогнозуванні, візуалізації та симуляції – як у фундаментальних дослідженнях, так і в прикладному проєктуванні.

У хімії тривимірне моделювання дає змогу аналізувати молекулярну структуру, передбачати реакції та вивчати властивості речовин на мікроскопічному рівні. Використання програмного забезпечення, такого як Avogadro, Gaussian чи Chem3D, є стандартом у сучасній хімічній інформатиці [2].

У геодезії та картографії 3D-технології застосовуються для створення цифрових моделей рельєфу та урбаністичних об'єктів. Технології лазерного сканування, фотограмметрії та обробки супутникових знімків забезпечують високоточні геопросторові моделі [3]. Це відкриває нові можливості для планування територій, оцінки ризиків і моніторингу навколишнього середовища [4].

У галузі економіки та містобудування тривимірні моделі використовуються для просторового аналізу ринку, симуляції транспортних потоків, оптимізації логістичних мереж і моделювання сценаріїв розвитку територій. Інструменти, такі як ArcGIS Urban, CityEngine чи SketchUp, дають змогу створювати інтерактивні 3D-моделі з урахуванням соціально-економічних параметрів [5].

Таким чином, 3D-технології стають універсальним інструментом для міждисциплінарного дослідження, який значно підвищує точність аналізу, якість візуалізації та ефективність прийняття рішень.

Мета статті. Існує безліч інструментів для візуалізації таких структур, як 3D-об'єкти. Але для обивателя цікавіше отримати рішення, за допомогою якого можна легко уявити об'єкт дослідження, передати зображення по мережі і використовувати найбільш поширену технологію. Відповідно, актуально отримати рендерений об'єкт з предметної області за допомогою JavaScript. Друге питання – як розробити додаток. А саме, використання Gtok як рекомендаційної системи для того, щоб отримати необхідне рішення.

Виклад основного матеріалу. Високорівневі вимоги до продукту сформуємо наступним чином.

Уявімо метадоксальну біпіраміду мови за допомогою Three.js. Це бібліотека JavaScript для створення 3D-графіки в браузері, не потрібно встановлювати Python або Matplotlib – все буде працювати безпосередньо в веб-браузері (наприклад, Chrome або Edge). Код Three.js буде відображати тригональну біпіраміду з п'ятьма вершинами (Знак, Значення, Структура, Контекст, Час) і ребрами. Ви можете запустити цей код у Windows 11 без додаткових інструментів. Для його реалізації потрібен будь-який текстовий редактор (Notepad, VS Code), веб-браузер (Chrome, Firefox, Edge – будь-який сучасний), підключення до інтернету (Three.js можна завантажити через CDN).

У браузері повинна відображатися 3D-сцена з біпірамідою, що складається з ліній (ребер), які з'єднують вершини. Для кожної вершини надаються окремі текстові мітки ("Знак", "Значення", "Структура", "Контекст", "Час") зі зміщенням від вершин.

Сцена обертається навколо осі Y зі швидкістю 0.01 радіан за кадр.

Камера налаштована для оптимального огляду біпіраміди.

Сцена адаптується до зміни розміру вікна браузера.

Можливе включення управління мишею (наприклад, OrbitControls), інакше біпіраміда обертається автоматично.

Використовується Three.js з урахуванням можливостей підключення різних версій бібліотеки.

Шрифти завантажуються з threejs.org.

Алгоритм створення 3D-візуалізації.

Крок 1: Підключити бібліотеку Three.js через CDN.

Крок 2: Ініціалізація сцени, камери та рендерера

Створити нову 3D-сцену (THREE. Scene).

Створити перспективну камеру (THREE. PerspectiveCamera) із параметрами.

Створити WebGL-рендерер (THREE. WebGLRenderer).

Встановити розмір рендерера: ширина та висота вікна браузера.

Крок 3: Визначення вершин біпіраміди

Створити масив vertices, що містить координати вершин у 3D-просторі (x, y, z).

Крок 4: Визначення ребер біпіраміди. Створити масив edges, що містить пари індексів вершин для ребер.

Крок 5: Налаштування матеріалів

Створити матеріал для ліній (THREE. LineBasicMaterial).

Створити матеріал для вершин (THREE. MeshBasicMaterial).

Крок 6: Додавання ребер до сцени

Для кожного ребра в масиві edges:

Створити масив points, що містить дві 3D-точки (THREE. Vector3) за координатами вершин ребра (з масиву vertices).

Створити геометрію лінії (THREE. BufferGeometry) на основі масиву points.

Створити об'єкт лінії (THREE. Line) із геометрією та матеріалом ліній.

Додати лінію до сцени (scene.add(line)).

Крок 7: Додавання вершин до сцени

Створити геометрію сфери (THREE. SphereGeometry) із параметрами.

Для кожної вершини в масиві vertices:

Створити меш (THREE. Mesh) із геометрією сфери та матеріалом вершин.

Встановити позицію меша за координатами вершини (vertex[0], vertex[1], vertex[2]).

Додати меш до сцени (scene.add(sphere)).

Крок 8: Додавання текстових міток до вершин

Створити масив labels із назвами вершин.

Ініціалізувати завантажувач шрифтів (THREE. FontLoader).

Завантажити шрифт.

Після завантаження шрифту (font):

Для кожної вершини (vertex) та відповідної мітки (labels[i]):

Створити текстову геометрію (THREE. TextGeometry) із параметрами.

Створити матеріал для тексту (THREE. MeshBasicMaterial).

Створити меш (THREE. Mesh) із текстовою геометрією та матеріалом.

Встановити позицію тексту зі зміщенням.

Додати текст до сцени (scene.add(text)).

Крок 9: Налаштування камери

Встановити позицію камери.

Налаштувати напрямок камери на точку за допомогою camera.lookAt.

Крок 10: Створення анімації

Створити функцію animate:

Запланувати наступний кадр анімації через requestAnimationFrame(animate).

Обертати сцену навколо осі Y на 0.01 радіан за кадр (scene.rotation.y += 0.01).

Викликати рендеринг сцени з поточною камерою.

Викликати animate() для запуску анімації.

Крок 11: Обробка зміни розміру вікна

Додати слухач події resize на об'єкт window.

При зміні розміру вікна:

Оновити співвідношення сторін камери.

Оновити матрицю проєкції камери.

Оновити розмір рендерера.

Доопрацюємо візуалізацію біпіраміди метадоксу на Three.js! Додамо керування мишею (обертання та масштабування) за допомогою OrbitControls.

Використання миші:

Обертається ліва кнопка.

Колісний зум.
Права кнопка – панорамування.

Розроблений код, відповідно до алгоритму виглядає наступним чином.

```
html
<!DOCTYPE html>
<html lang="ru">
<head>
  <meta charset="UTF-8">
  <title>Метадокс </title>
  <style>
    body { margin: 0; overflow: hidden; }
    canvas { width: 100%; height: 100%; }
  </style>
</head>
<body>
  <script src="https://cdnjs.cloudflare.com/ajax/libs/three.js/r128/three.min.js"></script>
  <script src="https://threejs.org/examples/js/controls/OrbitControls.js"></script>
  <script>
    // Сцена, камера, рендерер
    const scene = new THREE.Scene();
    const camera = new THREE.PerspectiveCamera(75, window.innerWidth / window.
innerHeight, 0.1, 1000);
    const renderer = new THREE.WebGLRenderer();
    renderer.setSize(window.innerWidth, window.innerHeight);
    document.body.appendChild(renderer.domElement);
    // Управління мишею
    const controls = new THREE.OrbitControls(camera, renderer.domElement);
    controls.enableDamping = true; Плавний рух
    controls.dampingFactor = 0,05;
    controls.screenSpacePanning = false;
    controls.minDistance = 1; Мінімальне збільшення
    controls.maxDistance = 10; Максимальне збільшення
    Координати вершин біпіраміди
    const вершини = [[0, 0, 0],[1,0,0],[0.5, 0.87, 0],[0.5,0.43,1],[0.5, 0.43, -1];
    const metadoxes = [...];
    Рендеринг Metadox
    metadoxes.forEach(metadox => {
      const material = new THREE.LineBasicMaterial({ color: metadox.color });
      metadox.edges.forEach(edge => {
        const points = [];
        points.push(new THREE.Vector3(vertices[edge[0]][0], vertices[edge[0]]
[1], vertices[edge[0]][2]));
        points.push(new THREE.Vector3(vertices[edge[1]][0], vertices[edge[1]]
[1], vertices[edge[1]][2]));
        const geometry = new THREE.BufferGeometry().setFromPoints(points);
        const line = new THREE.Line(geometry, material);
        scene.add(line);
      });
    });
    // Вершини (точки)
    const pointMaterial = new THREE.MeshBasicMaterial({ color: 0xfffff }); // Білі
точки
    const sphereGeometry = new THREE.SphereGeometry(0.05, 32, 32);
    vertices.forEach(vertex => {
      const sphere = new THREE.Mesh(sphereGeometry, pointMaterial);
      sphere.position.set(vertex[0], vertex[1], vertex[2]);
      scene.add(sphere);
    });
  </script>
</body>
</html>
```

```

    // Мітки вершин
    const labels = ['Знак', 'Значення', 'Структура', 'Контекст', 'Час'];
    const loader = new THREE.FontLoader();
    loader.load('https://threejs.org/examples/fonts/helvetiker_regular.typeface.
json', function (font) {
        vertices.forEach((vertex, i) => {
            const textGeometry = new THREE.TextGeometry(labels[i], {
                font: font,
                size: 0.1,
                height: 0.01
            });
            const textMaterial = new THREE.MeshBasicMaterial({ color: 0x000000 });
// Чорні відмітини
            const text = new THREE.Mesh(textGeometry, textMaterial);
            text.position.set(vertex[0] + 0.1, vertex[1] + 0.1, vertex[2]);
            scene.add(text);
        });
    });
    // Положення камери
    camera.position.set(1, 1, 2);
    camera.lookAt(0, 5, 0, 43, 0);
    Анімації
    function animate() {
        requestAnimationFrame(анімувати);
        controls.update(); Оновлення системи керування мишею
        renderer.render(сцена, камера);
    }
    animate();
    Адаптація розміру вікна
    window.addEventListener('resize', () => {
        camera.aspect = window.innerWidth / window.innerHeight;
        camera.updateProjectionMatrix();
        renderer.setSize(window.innerWidth, window.innerHeight);
    });
</script>
</body>
</html>

```

Тестування та доробка. Запропонований код, не дивлячись на логічну послідовність та обґрунтованість, хоча і виконався, але біпіраміду не відобразив. Отже, виникла потреба у тестуванні додатків та доробці. У зв'язку із тим, що для генерації зображення виконується багато різних компонентів коду, які вирішують окремі задачі, було прийнято рішення застосувати Grok, для аналізу помилок та їх виправлення.

Grok, створений xAI, – це ШІ-помічник, який ефективно допомагає у тестуванні JavaScript-коду, оптимізуючи процес розробки та підвищуючи якість програмного забезпечення. Завдяки здатності аналізувати код, генерувати тести та виявляти помилки, Grok стає цінним інструментом для розробників.

Grok може аналізувати на всіх етапах тестування. На початковому рівні він перевіряє синтаксис, знаходить потенційні помилки, такі як невизначені змінні чи неправильні логічні конструкції, і пропонує виправлення. Наприклад, для функції, яка обробляє масиви, Grok може виявити відсутність обробки граничних випадків і порадити додати відповідні перевірки. Це економить час і знижує ймовірність багів.

Для модульного тестування Grok генерує тестові випадки, використовуючи популярні фреймворки, такі як Jest або Mocha. На основі аналізу функції він створює тести для перевірки основних сценаріїв, крайових випадків і виняткових ситуацій. Наприклад, для функції, що обробляє голосування в системі виборів, Grok може згенерувати тести для перевірки коректності підрахунку голосів або реакції на невалідні дані.

Grok також допомагає в інтеграційному тестуванні, аналізуючи взаємодію між модулями. Він може перевірити, чи правильно сервер Node.js обробляє запити, і запропонувати тести для API-ендпоінтів. Крім того, Grok підтримує автоматизацію тестів, створюючи скрипти для CI/CD-систем, таких як GitHub Actions.

Переваги використання Grok включають швидкість, точність і підтримку складних сценаріїв, як-от асинхронний код чи 3D-візуалізації з Three.js. Однак він не замінює людську інтуїцію, а доповнює її, допомагаючи зосередитися на логіці, а не рутинних перевірках.

Використання Grok для тестування JavaScript-коду – це крок до ефективної розробки, що забезпечує надійність і масштабованість проєктів, від веб-сторінок до систем голосування.

У процесі виконання коду виникла помилка:

Uncaught TypeError: THREE. OrbitControls is not a constructor

Возможные причины, предложенные Grok:

- Скрипт OrbitControls.js не загрузился (сеть, блокировка, ошибка 404).
- Версия Three.js (r134) не полностью совместима с этим файлом OrbitControls.js.
- Порядок загрузки скриптов нарушен, и OrbitControls вызывается до полной инициализации Three.js.

Оскільки r134 може бути нестабільним через CDN, давайте перепишемо код, використовуючи останню версію Three.js та модульний підхід через `<script type="module">`. Це забезпечить сумісність і належне завантаження OrbitControls.

Наступним кроком є виправлення неспійманої помилки `ReferenceError: THREE не визначено`. Grok вказує, що об'єкт THREE недоступний на момент виконання коду, хоча він підключений Three.js за допомогою `<script src="https://cdn.jsdelivr.net/npm/three@0.165.0/build/three.min.js"></script>`. Це може означати, що Three.js скрипт або не завантажився (проблема з мережею, блокування), або код виконується до того, як Three.js буде повністю ініціалізований.

Неявна помилка виконання у вигляді білого екрану при відкритті файлу проекту в браузері означає, що Three.js запустився, але щось пішло не так з рендерингом сцени: або скрипти завантажилися не повністю, або є баг в коді, який не відображається в консолі, або сцена порожня/не видима камері. Так як екран білий, а не чорний, швидше за все, рендерер працює, але сцена або порожня, або не видно.

Виконаємо етапи налагодження.

Додамо код для тесту:

```
const testCube = new THREE.Mesh(
  new THREE.BoxGeometry(0.5, 0.5, 0.5),
  new THREE.MeshBasicMaterial({ color: 0xff0000 })
);
testCube.position.set(0.5, 0.43, 0);
scene.add(testCube);
console.log(«Додано тестовий куб»);
```

Якщо з'являється червоний куб, то рендеринг працює, але біпіраміда не відображається.

Спрощення роботи камери та органів управління для усунення впливу OrbitControls. Куб працює, тому Three.js працює локально. Тепер давайте додамо повний код з біпірамідою і OrbitControls.

Піраміда працює шляхом додавання міток вершин («Знак», «Значення», «Структура», «Контекст», «Час») за допомогою FontLoader.

Мітки вершин відображаються у вигляді чорних знаків питання на чорному тлі, що вказує на проблему із завантаженням або рендерингом шрифту. Швидше за все, шрифт або `helvetiker_regular.typeface.json` завантажився некоректно, або Three.js версії 0.124.0 не змогла його обробити, і замість тексту обробляються "заповнювачі" (знаки питання).

Виправлений код представлений нижче.

```
html
<!DOCTYPE html>
<html lang="ru">
<head>
  <meta charset="UTF-8">
  <title>Метадокс мови як біпіраміди</title>
  <style>
    body { margin: 0; overflow: hidden; background: #000; }
    canvas { width: 100%; height: 100%; }
  </style>
</head>
<body>
  <script src="three.js"></script>
  <script src="OrbitControls.js"></script>
  <script>
    console.log("Сценарій розпочато виконання");

    if (typeof THREE === 'undefined') {
      console.error('Three.js не завантажився');
      alert('Three.js не загрузился');
    } else if (typeof THREE.OrbitControls === 'undefined') {
```

```

        console.error('OrbitControls не завантажувався');
        alert('OrbitControls не завантажився');
    } else {
        console.log('Three.js и OrbitControls завантажено, версія:', THREE.
REVISION);

        const scene = new THREE.Scene();
        const camera = new THREE.PerspectiveCamera(75, window.innerWidth / window.
innerHeight, 0.1, 1000);
        const renderer = new THREE.WebGLRenderer();
        renderer.setSize(window.innerWidth, window.innerHeight);
        document.body.appendChild(renderer.domElement);

        const controls = new THREE.OrbitControls(camera, renderer.domElement);
        controls.enableDamping = true;
        controls.dampingFactor = 0.05;
        controls.minDistance = 1;
        controls.maxDistance = 10;

        const vertices = [
            [0, 0, 0], // Знак (0)
            [1, 0, 0], // Значення (1)
            [0.5, 0.87, 0], // Структура (2)
            [0.5, 0.43, 1], // Контекст (3)
            [0.5, 0.43, -1] // Час (4)
        ];

        const metadoxes = [
            { edges: [[0, 1], [1, 2], [2, 0]], color: 0xff00ff },
            { edges: [[0, 3], [0, 4], [3, 4]], color: 0x00ff00 },
            { edges: [[1, 3], [1, 4], [3, 4]], color: 0xffff00 },
            { edges: [[2, 3], [2, 4], [3, 4]], color: 0x00ffff },
            { edges: [[0, 1], [0, 3], [1, 3]], color: 0xff0000 },
            { edges: [[0, 1], [0, 4], [1, 4]], color: 0x0000ff },
            { edges: [[0, 2], [0, 3], [2, 3]], color: 0xff8000 },
            { edges: [[0, 2], [0, 4], [2, 4]], color: 0x800080 },
            { edges: [[1, 2], [1, 3], [2, 3]], color: 0x808000 },
            { edges: [[1, 2], [1, 4], [2, 4]], color: 0x008080 }
        ];

        metadoxes.forEach(metadox => {
            const material = new THREE.LineBasicMaterial({ color: metadox.color });
            metadox.edges.forEach(edge => {
                const points = [];
                points.push(new THREE.Vector3(vertices[edge[0]][0],
vertices[edge[0]][1], vertices[edge[0]][2]));
                points.push(new THREE.Vector3(vertices[edge[1]][0],
vertices[edge[1]][1], vertices[edge[1]][2]));
                const geometry = new THREE.BufferGeometry().setFromPoints(points);
                const line = new THREE.Line(geometry, material);
                scene.add(line);
            });
        });

        const pointMaterial = new THREE.MeshBasicMaterial({color:0xffff });
        const sphereGeometry = new THREE.SphereGeometry(0.05, 32, 32);
        vertices.forEach(vertex => {
            const sphere = new THREE.Mesh(sphereGeometry, pointMaterial);
            sphere.position.set(vertex[0], vertex[1], vertex[2]);
            scene.add(sphere);
        });

        Мітки вершин
        const labels = ['Знак', 'Значення', 'Структура', 'Контекст', 'Час'];
        const loader = new THREE.FontLoader();
        loader.load('helvetiker_regular.typeface.json', function (font) {
            console.log("Шрифт завантажено");
        });

```

```

        vertices.forEach((vertex, i) => {
            const textGeometry = new THREE.TextGeometry(labels[i], {
                font: font, size: 0.1, height: 0.01
            });
            const textMaterial = new THREE.MeshBasicMaterial({ color: 0xffffff });
// Белый текст
            const text = new THREE.Mesh(textGeometry, textMaterial);
            text.position.set(vertex[0] + 0.15, vertex[1] + 0.15, vertex[2]);
            scene.add(text);
        });
        console.log('Додано мітки вершин');
    }, undefined, function (error) {
        console.error('Помилка завантаження шрифту:', error);
    });

    camera.position.set(1, 1, 2);
    camera.lookAt(0, 0, 0);

    function animate() {
        requestAnimationFrame(animate);
        controls.update();
        renderer.render(scene, camera);
    }
    animate();

    window.addEventListener('resize', () => {
        camera.aspect = window.innerWidth / window.innerHeight;
        camera.updateProjectionMatrix();
        renderer.setSize(window.innerWidth, window.innerHeight);
    });
}
</script>
</body>
</html>

```

Результат виконання коду:

Чорний фон.

Біпіраміда з 10 різнокольоровими метадоксами (лініями).

Білі точки (вершини).

Білі мітки поруч з вершинами: Знак, Значення, Структура, Контекст, Час.

Органи управління: обертання (ліва кнопка), зум (коліщатко).

Висновки. Результати роботи продемонстрували, що мова JavaScript з розширенням у вигляді бібліотеки Three.js є ефективним інструментом для реалізації 3D-моделей метадоксів та більш складних структур. Використання відкритої бібліотеки, дозволило створити високоякісні та реалістичні моделі, що можуть бути легко інтегровані у веб-додатки, та надати можливість дослідникам, що використовують метадокси сконцентруватися на своїх безпосередніх дослідженнях.

У процесі роботи розроблено алгоритм побудови графічної структури, що може змінювати масштаб та положення у просторі, який є універсальним та досить легко розширюваним.

Для розв'язання технічних проблем, що пов'язані з тестуванням та модифікацією кода був задіяний Grok. У ітераційному процесі за допомогою рекомендацій ІІІ було виправлено помилки та нестиковки реалізації. Тобто підтверджено ефективність використання Grok для роботи з JavaScript.

Список використаних джерел:

1. Li Z., Zhu Q., Gold C. *Digital Terrain Modeling: Principles and Methodology*. CRC Press, 2005.
2. Chen B., Guo Y., Zhang X. Visualization of Molecular Structures with 3D Tools. *Journal of Chemical Information and Modeling*. 2021. Vol. 61, No. 3. P. 945–952.
3. Li R., Chapman M.A., Zhang J. LIDAR Technology for Urban Modeling. *Remote Sensing*. 2020. Vol. 12, No. 7. P. 1125.
4. Захарченко М. В., Печорін С. В. Геоінформаційні технології у геодезії: методи та застосування. *Вісник геодезії та картографії*. 2022. Т. 8, № 2. С. 24–29.
5. Batty M. et al. Smart Cities of the Future. *Environment and Planning B: Planning and Design*. 2012. Vol. 39, No. 4. P. 625–636.

References:

1. Li, Z., Zhu, Q., Gold, C. (2005). Digital Terrain Modeling: Principles and Methodology. CRC Press.
2. Chen, B., Guo, Y., Zhang, X. (2021). Visualization of Molecular Structures with 3D Tools. *Journal of Chemical Information and Modeling*. Vol. 61, No. 3. P. 945–952.
3. Li, R., Chapman, M. A., Zhang, J. (2020). LIDAR Technology for Urban Modeling. *Remote Sensing*. Vol. 12, No. 7. P. 1125.
4. Zaharchenko, M. V., Petcherin, S. V. (2022). Geoinformacijni tehnologiyi u geodeziyi: metodi ta zastosuvannya. *Visnik geodeziyi ta kartografii*. T. 8, № 2. P. 24–29.
5. Batty, M. et al. (2012). Smart Cities of the Future. *Environment and Planning B: Planning and Design*. Vol. 39, No. 4. P. 625–636.

Дата надходження статті: 21.10.2025

Дата прийняття статті: 10.11.2025

Опубліковано: 30.12.2025

Shamov O. A., Intelligent Systems Researcher,
Head of Human Rights Educational Guild
ORCID: 0009-0009-5001-0526

THE ROLE OF GENERATIVE AI IN THE AUTOMATION OF LEGAL DOCUMENT PREPARATION: AN ANALYSIS OF ACCURACY, EFFICIENCY, AND ETHICAL ASPECTS

This article explores the transformative impact of generative artificial intelligence (AI) on legal practice, with a particular focus on the preparation and analysis of legal documents. It examines modern generative AI models, their capabilities, limitations, and the risks associated with their implementation in the legal domain. The aim of the article is to analyze the current state of generative AI applications for automating legal tasks, identify key issues related to the accuracy and reliability of generated content, and propose a hypothetical model to enhance the trustworthiness of AI-generated outcomes.

To achieve this, the article sets out the following objectives: to review recent academic research and publications evaluating the effectiveness and precision of large language models (LLMs) in legal tasks; to identify the unresolved issue of factual verification of generated content; and to formulate and justify a hypothesis for improving baseline research results through the implementation of a hybrid architecture.

The study employs methods of systems analysis, comparative analysis of empirical studies, synthesis, and modeling. It analyzes foundational research on structured approaches to AI-assisted legal document analysis and the outcomes of comparative tests between LLMs and human legal professionals, highlighting the persistent challenge of factual accuracy in AI-generated content.

The article proposes a hypothesis that integrating structured prompting techniques with a Retrieval-Augmented Generation (RAG) architecture-leveraging a curated and dynamically updated legal knowledge base- can significantly improve the factual precision and reliability of AI-generated legal documents.

This hypothesis is substantiated with a formalized expression and a potential mechanism for implementation and verification.

The proposed approach is shown to not only minimize the risk of AI "hallucinations" but also to lay the groundwork for the development of more responsible and ethical tools that augment, rather than replace, the professional competencies of legal practitioners.

Key words: generative artificial intelligence, large language models, legal document automation, Retrieval-Augmented Generation (RAG), AI accuracy, legal technology, AI ethics, data verification.

Шамов О. А. Роль генеративного ШІ в автоматизації підготовки юридичних документів: аналіз точності, ефективності та етичних аспектів

Стаття присвячена дослідженню трансформаційного впливу генеративного штучного інтелекту (ШІ) на юридичну практику, зокрема на процес підготовки та аналізу юридичних документів. Предметом вивчення є сучасні моделі генеративного ШІ, їхні можливості, обмеження та ризики, пов'язані з їхнім впровадженням у юридичній сфері. Метою статті є аналіз поточного стану застосування генеративного ШІ для автоматизації юридичних завдань, виявлення ключових проблем, пов'язаних з точністю та надійністю згенерованих даних, та розробка гіпотетичної моделі для підвищення достовірності результатів. Для досягнення цієї мети були поставлені такі завдання: проаналізувати останні наукові дослідження та публікації, що оцінюють ефективність та точність великих мовних моделей (LLM) у юридичних завданнях; виокремити невирішену проблему забезпечення фактичної верифікації згенерованого контенту; сформулювати та обґрунтувати гіпотезу щодо покращення результатів базових досліджень шляхом впровадження гібридної архітектури.

Дослідження базується на методах системного аналізу, порівняльного аналізу емпіричних досліджень, синтезу та моделювання.

Проаналізовано базове дослідження щодо структурованого підходу до аналізу юридичних документів за допомогою ШІ, а також результати порівняльних тестів LLM та юристів-людей, виокремлено невирішену проблему забезпечення фактичної верифікації згенерованого контенту.

Запропоновано гіпотезу, що інтеграція методів структурованого промптингу з архітектурою Retrieval-Augmented Generation (RAG), яка спирається на кураторську та динамічно оновлювану базу юридичних знань, може суттєво підвищити фактичну точність та надійність згенерованих ШІ юридичних документів.

Обґрунтовано цю гіпотезу, її формалізоване вираження та потенційний механізм імплементації та верифікації.

Розкрито, що запропонований підхід не лише мінімізує ризик «галюцинацій» ШІ, але й створює основу для розробки більш відповідальних та етичних інструментів, що посилюють, а не замінюють, професійні компетенції юристів.

© O. A. Shamov, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

Ключові слова: генеративний штучний інтелект, великі мовні моделі, автоматизація юридичних документів, Retrieval-Augmented Generation (RAG), точність ШІ, юридичні технології, етика штучного інтелекту, верифікація даних.

Statement of the problem. The penetration of generative artificial intelligence (AI) into various professional fields has become one of the most significant technological trends of recent years. The legal industry, traditionally considered conservative, has not remained aloof from these processes. Large language models (LLMs), such as GPT-4 and its counterparts, demonstrate an impressive ability to generate, analyze, and summarize textual information, opening up unprecedented opportunities for automating routine tasks that previously required significant time and human resources [1]. Drafting contracts, memoranda, statements of claim, analyzing large volumes of evidence, and conducting legal research are just a few of the processes that can potentially be transformed with the help of AI [2].

The connection of this issue with important practical tasks is obvious [3]. According to studies, law firms and corporate legal departments are constantly looking for ways to increase efficiency and reduce costs.

Empirical research shows that LLMs can perform certain legal tasks much faster and more cheaply than junior lawyers [4]. This creates the potential to expand access to justice and provide more affordable legal services [5].

However, alongside the prospects of efficiency, a fundamental scientific and practical problem arises ensuring the accuracy, reliability, and ethical use of these tools. Unlike in many other fields, an error in a legal document can have serious legal and financial consequences for clients. The main risk inherent in modern LLMs is a phenomenon known as «hallucinations», which is the generation of plausible but factually false or inaccurate information [6]. A model might cite non-existent court precedents, misinterpret legal statutes, or create legally incorrect formulations. Thus, a critical task emerges: to develop methods and architectures for AI systems that not only automate but also ensure a high level of credibility and verification of the generated content, consistent with the professional standards of jurisprudence.

Analysis of recent research and publications. The problem of AI application in law is being actively discussed in the scientific community. A significant number of recent publications are devoted to both the potential benefits and the risks.

One of the key areas is the empirical evaluation of LLM performance compared to lawyers. A 2024 study conducted by scholars at Cornell University, titled «Better Call GPT, Comparing Large Language Models Against Lawyers», demonstrated that the GPT-4 model surpassed lawyers in the speed and cost-effectiveness of performing contract analysis and summarization tasks, although it was inferior in accuracy when identifying specific legal risks [4]. Another large-scale study by lawyers from the University of Minnesota and the University of Michigan («AI-Powered Lawyering») showed that lawyers using GPT-4 completed tasks 24% faster and produced 40% higher quality documents. However, the study also found that the model is prone to generating false information [6].

The foundational study for this article was conducted by Wang, Chen, & Li in 2024 and published under the title «Enhancing Legal Document Analysis with Large Language Models: A Structured Approach to Accuracy, Context Preservation, and Risk Mitigation» [7]. The authors proposed a methodology that combines hierarchical document segmentation, Chain-of-Thought prompting, and iterative summarization to improve the accuracy of analyzing complex legal texts.

Their approach showed a significant improvement in preserving context and reducing the risks of misinterpretation compared to standard queries to LLMs [7].

Despite the significant contribution of this research to structuring interaction with AI, it does not solve the main problem—the lack of an external verification mechanism. Structured prompting only optimizes the use of the model's internal «knowledge» but does not guarantee that this knowledge is accurate and up-to-date. This problem is highlighted by numerous works on ethical aspects. Perlman, in his 2025 work, emphasizes that the professional duty of competence requires lawyers not to blindly trust but to thoroughly verify the results of AI's work [8]. The American Bar Association (ABA) also warns of the risks of bias embedded in the data on which the models were trained, which can lead to discriminatory outcomes [9].

To overcome the problem of «hallucinations», researchers are increasingly turning to the Retrieval-Augmented Generation (RAG) architecture. This approach, as noted in the Harvard Journal of Law & Technology, allows for «grounding» LLM responses in data from reliable external sources [10]. Instead of relying solely on its parametric memory, the model first searches for relevant information in a specialized knowledge base (e.g., a database of legislation or case law) and then uses the retrieved data to generate the answer. A 2023 study from Stanford University showed that even RAG-based tools are not completely protected from errors, but they significantly reduce their frequency [11].

Thus, the literature analysis reveals an unresolved part of the general problem: while methods exist to improve the structure of queries to AI (the research of Wang et al.) and methods for external grounding of responses (RAG), there is a lack of a comprehensive approach that would combine both strategies to achieve a synergistic effect of increased reliability. This article is dedicated to addressing this problem.

Task statement. Based on the identified problem, the main goals (tasks) of this article are:

1. To analyze and summarize the methodology and results of the foundational study by Wang, Chen, & Li (2024) on the structured approach to legal document analysis.
2. To identify the limitations of this approach, particularly its vulnerability to generating factually unverified information.
3. To propose and substantiate a hypothesis for improving the accuracy and reliability of generative models by creating a hybrid system that combines structured prompting with the Retrieval-Augmented Generation (RAG) architecture.
4. To develop a conceptual model for the implementation and verification of the proposed hybrid system in the context of automating legal document preparation.

The purpose of this article is to develop and theoretically substantiate a new, hybrid model for the use of generative AI in legal practice, which aims to maximize accuracy and reliability through the synergistic combination of advanced techniques for model interaction and architectural solutions that ensure real-time data verification.

Outline of the main material of the study. The basis for our research is the work of Wang and his colleagues [7], which proposes a systematic approach to using LLMs for analyzing legal contracts. Their methodology consists of three key stages:

1. Hierarchical Segmentation: Instead of analyzing the entire document as a single unit, the authors propose breaking it down into logical parts (sections, articles, clauses). This allows the model to better focus on the local context and avoid losing important details, which often happens when processing long texts.

2. Chain-of-Thought Prompting: For each segmented part, the model is instructed not just to provide a final answer but to demonstrate a step-by-step reasoning process. For example, when analyzing a clause on liability, the model should first identify the parties, then describe the conditions for liability, determine its type and extent, and only then formulate a final conclusion. This makes the process more transparent and allows for easier detection of logical errors.

3. Iterative Summarization: After analyzing individual segments, the model creates short summaries of them. These summaries are then combined and summarized at a higher level, creating a comprehensive yet detailed overview of the entire document.

This structured approach is undoubtedly a step forward. It disciplines the «thinking» of the LLM and reduces the likelihood of gross errors related to context loss. However, its fundamental limitation is that it operates within the «closed world» of the model's own knowledge. If the LLM was trained on outdated data or contains incorrect representations of certain legal norms, structured prompting will not correct this. The model can logically and consistently justify its answer based on a non-existent legal norm [6]. This is where the need to go beyond the model's internal knowledge arises.

The solution to this problem is the Retrieval-Augmented Generation (RAG) architecture. Unlike a standard LLM, a RAG system performs an additional step before generating a response: searching for information in an external database. In a legal context, such a database (or vector database) can contain the full texts of laws, codes, regulations, current case law, and scholarly commentaries [10]. The process of such a system is as follows:

1. Receiving a Query: A lawyer formulates a query, for example, «Draft a clause for a lease agreement on rent indexation in accordance with the current legislation of Ukraine.»

2. Retrieval: The RAG system does not pass the query directly to the LLM. First, it analyzes the query and searches its knowledge base for the most relevant documents: articles of the Civil Code of Ukraine governing rent and indexation, recent Supreme Court decisions on this issue, etc.

3. Augmentation: The found text fragments (e.g., specific articles of law) are added to the user's initial query as additional context.

4. Generation: The modified, augmented query is passed to the LLM with instructions to generate a response based primarily on the provided context.

Thus, RAG forces the model to rely not on its abstract memory but on specific, verified, and current legal sources. This significantly reduces the risk of «hallucinations» and increases the credibility of the result [12].

However, as the Stanford University study shows, even RAG systems can sometimes provide inaccurate information if the retriever selects an irrelevant fragment from the knowledge base [11].

Hypothesis for improvement: Based on the analysis of the strengths and weaknesses of both approaches, we put forward the following hypothesis:

A hybrid model that synergistically combines structured prompting (following the methodology of Wang et al.) with a multi-layered Retrieval-Augmented Generation (RAG) architecture will provide a significantly higher level of factual accuracy, completeness, and reliability of generated legal documents compared to using each of these methods alone.

Rationale for the hypothesis: The rationale for the hypothesis lies in the mutual complementarity and reinforcement of both approaches. Structured Prompting (SP) organizes the model's thought process, forcing it to work methodically and consistently. This creates a «scaffold» for the answer. RAG, in turn, fills this scaffold with factual, verified content.

The synergistic effect arises at several levels:

- SP improves RAG: Applying hierarchical segmentation and chain-of-thought to the query before the retrieval stage in RAG can make the search itself more accurate. Instead of one general query to the knowledge base, the system will perform a series of smaller, more focused queries for each logical step, reducing the risk of retrieving irrelevant documents.

- RAG improves SP: Even the best chain of thought will be flawed if it is based on incorrect initial data. RAG ensures that each step in the chain of thought is based on a current legal norm or judicial precedent, making the entire logical construction credible.

For a scientific rationale, we can introduce a conditional formula for assessing accuracy (A). Let A_{base} be the baseline accuracy of a standard LLM. The use of structured prompting adds a certain improvement factor I_{SP} , and the use of RAG adds a factor I_{RAG} . Our hypothesis is that the effect is not merely additive but contains a synergistic component. This can be expressed by the formula:

$$A_{hybrid} = A_{base} + I_{SP} + I_{RAG} + S(I_{SP} I_{RAG}) \quad (1)$$

where A_{hybrid} is the accuracy of the hybrid model, and $S(I_{SP} I_{RAG})$ is a positive synergistic term that depends on the interaction of both methods. This can be more clearly represented as a multiplicative interaction, where RAG enhances the quality of the result already improved by SP:

$$A_{hybrid} = (A_{base} + I_{SP}) \times (1 + k \cdot I_{RAG}) \quad (2)$$

where k is a coefficient reflecting the effectiveness of the integration.

This formula illustrates that RAG does not just add accuracy but multiplies it by acting on an already structured and logically organized process.

Potential implementation and verification. Implementation of the hybrid model requires the creation of two main components:

1. Multi-layered Knowledge Base: This is not just a repository of texts. The base must be structured by levels of reliability and priority:

- Level 1 (Absolute truth): Codified legislation (Constitution, codes, laws).
- Level 2 (Official interpretation): Decisions of the Constitutional and Supreme Courts, regulations of regulatory bodies.
- Level 3 (Doctrine and practice): Scholarly articles, commentaries on legislation, generalized judicial practice of lower courts. The RAG system should be configured for priority search in the higher levels.

2. Hybrid Query Processor: This module will perform a sequence of operations:

- Receives a query from the user.
- Applies hierarchical segmentation and chain-of-thought techniques to break it down into sub-queries.
- For each sub-query, performs a search in the multi-layered knowledge base (RAG).
- Forms augmented prompts for the LLM.
- Generates responses and automatically adds citations to the sources used during the retrieval.

Verification of the hypothesis should be conducted through a controlled experiment. It is necessary to create a set of typical legal tasks (e.g., drafting a contract, risk analysis, writing a statement of claim). Then, three groups should be engaged to perform them:

- Group 1 (Control): Uses a standard LLM (e.g., GPT-4).
- Group 2 (SP): Uses an LLM with the application of structured prompting techniques.
- Group 3 (Hybrid): Uses the proposed hybrid model.

The results of all three groups should be independently evaluated by a committee of expert lawyers according to the following criteria:

- Factual accuracy: Absence of «hallucinations», correct citation of legal norms.
- Completeness: Coverage of all essential aspects of the problem.
- Legal correctness: Correctness of formulations and document structure.
- Presence and correctness of citations: Automatic referencing of sources.

A comparison of the average scores on these criteria between the groups will allow for the confirmation or refutation of the proposed hypothesis.

Conclusions. The conducted research confirms that generative AI has the potential to radically change legal practice by significantly increasing its efficiency. However, the key obstacle to its widespread adoption remains the problem of reliability and accuracy, embodied in the phenomenon of «hallucinations».

The analysis of advanced research has shown that existing approaches, such as structured prompting and the Retrieval-Augmented Generation architecture, while effective, have their limitations and do not solve the problem comprehensively. This article has presented and substantiated a hypothesis that the synergistic integration of these two methods into a single hybrid model can provide a breakthrough in creating reliable legal AI tools.

The proposed model, which combines the logical discipline of structured prompting with the factual grounding of RAG on a curated knowledge base, is capable of not only generating legally sound texts but also confirming their authenticity with citations to primary sources.

The practical significance of the proposed approach lies in creating a roadmap for legal tech developers that will allow them to move from creating «assistants» with unpredictable accuracy to developing reliable tools that enhance the professional capabilities of a lawyer. This, in turn, will promote adherence to ethical standards, particularly the duty of competence and confidentiality.

Prospects for further research in this area are multifaceted:

1. Empirical implementation and testing: The primary task is the practical development and experimental verification of the proposed hybrid model to quantitatively measure its benefits.
2. Optimization of knowledge bases: Researching methods for automatic updating and verification of legal knowledge bases to ensure the RAG system always works with the most current information.
3. Explainable AI (XAI): Integrating XAI mechanisms that would not just show the source of information but also visualize the logical chain that led to a particular conclusion, making the AI's work completely transparent to the lawyer.
4. Adaptation to different fields of law: Studying the specifics of applying the hybrid model in different legal fields (criminal, tax, corporate law), each of which has its unique sources and accuracy requirements.

Ultimately, the future of AI in law is not a race for full autonomy, but the creation of powerful, reliable, and ethical tools that, under the control of a qualified professional, will elevate the quality and accessibility of legal services to a new level.

Bibliography:

1. McCallon, J. Legal Tech ROI: Measuring AI Implementation Success. *CallidusAI Blog*. 2025. URL: <https://callidusai.com/legal-tech-roi-measuring-ai-implementation-success/> (дата звернення: 22.06.2025)
2. Terzidou K. Generative AI systems in legal practice: Offering quality legal services while upholding legal ethics. *International Journal of Law in Context*. 2025. URL: <https://www.cambridge.org/core/journals/international-journal-of-law-in-context/article/generative-ai-systems-in-legal-practice-offering-quality-legal-services-while-upholding-legal-ethics/34011A84AA58A2BAB556A406A4653A8D/> (дата звернення: 22.06.2025)
3. Braff D. AI is the future of law, but most legal pros aren't trained for it, a new report says. *ABA Journal*. 2025. URL: <https://www.abajournal.com/web/article/ai-is-the-future-of-law-but-most-legal-pros-arent-trained-for-it-a-new-report-says/> (дата звернення: 22.06.2025)
4. Martin L., Whitehouse N., Yiu S., Catterson L., Perera R. Better Call GPT, Comparing Large Language Models Against Lawyers. *arXiv preprint arXiv:2401.16212*. 2024. URL: <https://arxiv.org/abs/2401.16212> (дата звернення: 22.06.2025)
5. Baxter J. AI-driven legal tech trends for 2025. *NetDocuments Blog*. 2025. URL: <https://www.netdocuments.com/blog/ai-driven-legal-tech-trends-for-2025/> (дата звернення: 22.06.2025)
6. Schwarcz D., Manning S., Barry P., Cleveland D., Prescott J., Rich B., Deakins O. AI-Powered Lawyering: AI Reasoning Models, Retrieval Augmented Generation, and the Future of Legal Practice. *Learning Ecosystems*. 2025. URL: <http://danielschristian.com/learning-ecosystems/2025/05/22/ai-powered-lawyeringby-various-authors-other-legaltech-related-items/> (дата звернення: 22.06.2025)
7. Davenport M. Enhancing legal document analysis with Large Language Models: A structured approach to accuracy, context preservation, and risk mitigation. *Scientific Research an Academic Publisher*. 2025. URL: <https://www.scirp.org/journal/paperinformation?paperid=141892/> (дата звернення: 22.06.2025)
8. Perlman A. The Legal Ethics of Generative AI. *Suffolk University Law Review*. 2024. URL: https://bpbus-el.wpmucdn.com/sites.suffolk.edu/dist/3/1172/files/2025/02/1FINAL_Perlman_Article_WDFF.pdf/ (дата звернення: 22.06.2025)
9. Donald D. Bias in AI Large Language Models: Risks and Remedies. *GPSolo Magazine*. 2025. URL: <https://www.americanbar.org/groups/gpsolo/resources/magazine/2025-mar-apr/bias-ai-large-language-models-risks-remedies/> (дата звернення: 22.06.2025)
10. Johnston P. Retrieval-Augmented Generation (RAG): Towards a promising LLM architecture for legal work? *Harvard Journal of Law & Technology: JOLT Digest*. 2025. URL: <https://jolt.law.harvard.edu/digest/retrieval-augmented-generation-rag-towards-a-promising-llm-architecture-for-legal-work> (дата звернення: 22.06.2025)
11. Magesh V., Surani F., Dahl M., Suzgun M., Manning C., Ho D. Hallucination-Free? Assessing the Reliability of Leading AI Legal Research Tools. *Journal of Empirical Legal Studies*. 2025. URL: https://dho.stanford.edu/wp-content/uploads/Legal_RAG_Hallucinations.pdf/ (дата звернення: 22.06.2025)
12. Kankatala, V. Retrieval-Augmented Generation (RAG) in the legal industry. *Acroplans Blog*. 2024. URL: <https://www.acroplans.com/retrieval-augmented-generation-rag-in-the-legal-industry/> (дата звернення: 22.06.2025)

References:

1. McCallon, J. (2025). Legal Tech ROI: Measuring AI Implementation Success. *CallidusAI Blog*. Retrieved from: <https://callidusai.com/legal-tech-roi-measuring-ai-implementation-success/>

-
2. Terzidou, K. (2025). Generative AI systems in legal practice: Offering quality legal services while upholding legal ethics. *International Journal of Law in Context*, First View, 1-22. Retrieved from: <https://www.cambridge.org/core/journals/international-journal-of-law-in-context/article/generative-ai-systems-in-legal-practice-offering-quality-legal-services-while-upholding-legal-ethics/34011A84AA58A2BAB556A406A4653A8D>
 3. Braff, D. (2025). AI is the future of law, but most legal pros aren't trained for it, a new report says. *ABA Journal*. Retrieved from: <https://www.abajournal.com/web/article/ai-is-the-future-of-law-but-most-legal-pros-arent-trained-for-it-a-new-report-says>
 4. Martin, L., Whitehouse, N., Yiu, S., Catterson, L. & Perera, R. (2024). Better Call GPT, Comparing Large Language Models Against Lawyers. *arXiv*, 2401.16212. Retrieved from: <https://arxiv.org/abs/2401.16212>
 5. Baxter, J. (2025). AI-driven legal tech trends for 2025. *NetDocuments Blog*. Retrieved from: <https://www.netdocuments.com/blog/ai-driven-legal-tech-trends-for-2025/>
 6. Schwarcz, D., Manning, S., Barry, P., Cleveland, D., Prescott, J., Rich, B. & Deakins, O. (2025). AI-Powered Lawyering: AI Reasoning Models, Retrieval Augmented Generation, and the Future of Legal Practice. *Learning Ecosystems*. Retrieved from: <http://danielschristian.com/learning-ecosystems/2025/05/22/ai-powered-lawyering-by-various-authors-other-legaltech-related-items/>
 7. Davenport, M. (2025). Enhancing legal document analysis with Large Language Models: A structured approach to accuracy, context preservation, and risk mitigation. *Scientific Research an Academic Publisher*, Vol.15 No.2. Retrieved from: <https://www.scirp.org/journal/paperinformation?paperid=141892>
 8. Perlman, A. (2024). The Legal Ethics of Generative AI. *Suffolk University Law Review*, Vol.LVII:345 2024. Retrieved from: https://bpb-us-e1.wpmucdn.com/sites.suffolk.edu/dist/3/1172/files/2025/02/1FINAL_Perlman_Article_WDFF.pdf
 9. Donald, D. C. (2025). Bias in AI Large Language Models: Risks and Remedies. *GPSolo Magazine*, 42:2. Retrieved from: <https://www.americanbar.org/groups/gpsolo/resources/magazine/2025-mar-apr/bias-ai-large-language-models-risks-remedies/>
 10. Johnston, P. (2025). Retrieval-Augmented Generation (RAG): Towards a promising LLM architecture for legal work. *Harvard Journal of Law & Technology: JOLT Digest*. Retrieved from: <https://jolt.law.harvard.edu/digest/retrieval-augmented-generation-rag-towards-a-promising-llm-architecture-for-legal-work>
 11. Magesh, V., Surani, F., Dahl, M., Suzgun, M., Manning, C., & Ho, D. (2025). Hallucination-Free? Assessing the Reliability of Leading AI Legal Research Tools. *Journal of Empirical Legal Studies*, 1-27. Retrieved from: https://dho.stanford.edu/wp-content/uploads/Legal_RAG_Hallucinations.pdf
 12. Kankatala, V. (2024). Retrieval-Augmented Generation (RAG) in the legal industry. *Acroplans Blog*. Retrieved from: <https://www.acroplans.com/retrieval-augmented-generation-rag-in-the-legal-industry>

Дата надходження статті: 03.10.2025

Дата прийняття статті: 10.11.2025

Опубліковано: 30.12.2025

КОМП'ЮТЕРНА ІНЖЕНЕРІЯ

UDC 004.8:004.94:697

DOI <https://doi.org/10.32782/2521-6643-2025-2-70.23>

Axak N. G., Doctor of Technical Sciences, Professor,
Professor at the Department of Computer Intelligent Technologies
and Systems Kharkiv National University of Radio Electronics
ORCID: 0000-0001-8372-8432

Shelikhov Yu. O., Postgraduate Student at the Department
of Computer Intelligent Technologies and Systems
Kharkiv National University of Radio Electronics
ORCID: 0009-0009-8970-6571

INTELLIGENT LOAD BALANCING IN MICROSERVICE ARCHITECTURE

This paper presents an intelligent method for load balancing in microservice architectures that combines parallel (hedged) request routing with the Thompson Sampling Multi-Armed Bandit (MAB) algorithm. The goal is to address tail-latency spikes and performance variability that traditional policies (Round Robin, Least Connections) cannot handle under heterogeneous, bursty workloads. The proposed architecture comprises a YARP-based API Gateway that executes weighted hedging, an AI load balancer (FastAPI) that learns routing probabilities from live telemetry, and a Prometheus–Grafana stack providing continuous feedback for adaptation. The balancer transforms observed metrics (latency percentiles, error rate) into rewards and updates per-replica posteriors via Thompson Sampling, thereby balancing exploration and exploitation while preventing persistent bias toward temporarily fast but unstable instances.

We evaluate four strategies—static round-robin ($k=1$), static hedging ($k=2$), adaptive MAB hedging ($k=2$), and adaptive MAB hedging ($k=3$). Experiments with up to 1,000 concurrent clients show that adaptive hedging with Thompson Sampling reduces P99 latency by $\approx 65\%$ and the error rate by $\approx 45\%$ versus baseline, with negligible throughput loss and moderate CPU overhead. Increasing parallelism beyond two replicas yields diminishing returns, confirming that small k is sufficient when combined with probabilistic weighting and strict idempotency. The findings demonstrate that integrating speculative duplication with Bayesian decision-making provides a lightweight, cloud-native path to tail-tolerant performance. The solution is modular and reproducible, and it generalizes to Kubernetes-based deployments and IoT/cyber-physical scenarios where real-time, context-aware coordination and reliability are essential.

Key words: load forecasting, microservice, architecture, cloud computing, distributed system, cloud services, machine learning, artificial intelligence.

Аксак Н. Г., Шеліхов Ю. О. Інтелектуальне балансування навантаження в мікросервісній архітектурі

У статті запропоновано інтелектуальний метод балансування навантаження в мікросервісних архітектурах, що поєднує паралельну маршрутизацію запитів (hedging) із базаторуком бандитом на основі вибірки Томпсона (Thompson Sampling MAB). Мета підходу – подолати «хвостові» затримки (tail latency) та змінність продуктивності, які не усуваються традиційними політиками (Round Robin, Least Connections) за неоднорідних і вибухових навантажень. Архітектура складається зі шлюзу API на базі YARP, що виконує зважене дублювання запитів, AI-балансувальника (FastAPI), який навчає ймовірності маршрутизації за поточною телеметрією, та моніторингового стеку Prometheus–Grafana, що забезпечує безперервний зворотний зв'язок для адаптації. Спостережувані метрики (перцентилі затримки, частка помилок) перетворюються на «нагороду», якою оновлюються апостеріорні розподіли для кожної репліки за Thompson Sampling, завдяки чому система збалансовує «дослідження ↔ експлуатацію» та уникає фіксації на тимчасово швидких, але нестабільних вузлах.

Проведено оцінювання чотирьох стратегій: статичний round-robin ($k=1$), статичний hedging ($k=2$), адаптивний hedging з MAB ($k=2$) та адаптивний hedging з MAB ($k=3$). За умов до 1000 паралельних клієнтів адаптивний hedging з Thompson Sampling зменшив P99 приблизно на 65% і частку помилок – на $\approx 45\%$ порівняно з базовим варіантом, не знижуючи пропускну здатність і лише помірно підвищуючи використання CPU. Збільшення k понад 2 дає мінімальний додатковий ефект, що підтверджує доцільність «малого k » за наявності ймовірнісних ваг і гарантій ідемпотентності. Результати демонструють, що поєднання спекулятивного дублювання та байєсівського прийняття рішень є легким і

© N. G. Axak, Yu. O. Shelikhov, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

хмароорієнтованим шляхом до стійкої до «хвостів» продуктивності. Рішення є модульним і відтворюваним, придатним до розгортання в Kubernetes, а також до IoT/кіберфізичних сценаріїв, де потрібні адаптивна координація та висока надійність у реальному часі.

Ключові слова: прогнозування навантаження, мікросервіс, архітектура, хмарні обчислення, розподілена система, хмарні сервіси, машинне навчання, штучний інтелект.

Introduction and Problem Formulation. The rapid evolution of distributed and cloud-based systems has made microservice architectures a dominant paradigm for building scalable, flexible, and fault-tolerant applications. One of the key challenges in such architectures is adaptive load balancing, which ensures efficient distribution of requests among service replicas while maintaining low latency, stability, and high throughput.

Traditional algorithms such as Round Robin or Least Connections are static and do not consider real-time system conditions, contextual factors, or historical performance data. As a result, they may cause resource underutilization, increased response times, or local overloads—issues that become critical under high-load conditions typical for online services, IoT platforms, and real-time data streams.

To overcome these limitations, intelligent load-balancing mechanisms have emerged, integrating probabilistic models, reinforcement learning, and Bayesian inference to optimize routing decisions dynamically. This study proposes and experimentally evaluates a method for intelligent load balancing in microservice architectures that combines parallel (hedged) request routing with the Thompson Sampling algorithm. The approach adapts to changing workload patterns, predicts optimal routing paths based on live performance metrics, and effectively reduces tail latency under fluctuating load conditions.

Beyond improving computational performance, the proposed model forms the infrastructural foundation for subsequent research by the authors on AI-driven IoT and environmental control systems, such as adaptive lighting and microclimate regulation in smart buildings and city farms. In these systems, microservice-based IoT infrastructures, edge devices, and neural network controllers apply similar adaptive balancing principles to coordinate sensing, prediction, and actuation in real time.

Therefore, the research and development of intelligent, adaptive load-balancing algorithms in microservice architectures represent a crucial and practically significant direction, addressing modern challenges of distributed computing, real-time analytics, and cloud–edge integration.

Review of Existing Solutions. Microservice architectures consist of numerous independently deployable services operating in parallel; hence, load balancing is vital for scalability and fault tolerance [1]. Classic static methods, such as Round Robin or Random, evenly distribute requests but ignore instance state (load, latency), often directing traffic to slow or overloaded replicas and increasing response time [2].

Modern dynamic policies adapt routing based on live metrics and request context. Common approaches include Least Connections / Least Requests, often implemented via the power of two choices, Weighted Response Time, and EWMA smoothing to emphasize recent latency trends [2], [3]. Production meshes such as Envoy/Istio and Linkerd further enhance reliability through outlier detection, circuit breaking, adaptive retries, and locality-aware routing that minimizes cross-zone RTT.

A recent research direction introduces machine learning–based balancing, using predictive models or reinforcement learning to optimize routing under fluctuating workloads. Although such models show significant latency reductions, their adoption in production remains limited due to data and stability constraints [3]. In practice, organizations combine dynamic routing with autoscaling to balance load across available instances efficiently.

Tail Latency Mitigation. Long-tail delays (p95/p99) critically affect user experience. Parallel or hedged requests mitigate this by sending a duplicate of an idempotent call to another replica and returning the first response. Google popularized this technique, triggering duplicates only when the initial request exceeds the p95 threshold, reducing p99 latency (e.g., 1800 → 74 ms) with minimal traffic overhead ($\approx 2\text{--}5\%$) [4]. This strategy is now integrated into gRPC, Bigtable, and DynamoDB.

Best practices include hedging only idempotent operations, applying latency-based triggers, and respecting time budgets to prevent overload.

Algorithm Spectrum.

- *Static*: Round Robin, Random, Hash, Weighted RR – simple but state-agnostic.
- *Dynamic*: Least Requests, Weighted Response, EWMA latency, locality-aware, and outlier-based.
- *Learned*: ML or RL-based policies for adaptive, predictive routing [3].

In real deployments, these layers combine hierarchically – global (DNS/CDN) → cluster (Kubernetes Service/Ingress) → mesh (Istio/Linkerd) → optional client logic (hedging). Together, metric-aware balancing, locality, and selective hedging form the modern foundation for latency-aware, high-throughput microservice systems [1] – [4].

Analysis of recent research and publications. Authors of [5] propose a cloud-native microservice architecture with adaptive load balancing and multi-level fault tolerance built on Spring Cloud (Eureka, Ribbon, Hystrix) and Docker. The load balancer dynamically routes traffic using live health metrics—CPU, memory, and latency—rather than static round-robin, improving responsiveness.

Performance evaluation with JMeter and AHP-based QoS analysis showed about 20% higher QoS, throughput up to 2942 requests/min, and fault recovery under 5 s compared to monolithic LAMP and baseline Spring Cloud

setups. The adaptive policy efficiently balanced workload and reduced response latency by leveraging real-time system metrics.

The study [6] presents the Kubernetes Scheduling Extension (KSE), which enables dynamic load balancing through real-time pod migration. KSE augments the default scheduler by monitoring CPU and memory usage and redistributing pods via two algorithms: KSE-greedyLB and KSE-refineLB.

Tests on Kubernetes clusters under 32 imbalance scenarios showed that KSE – especially refineLB – reduced hotspots, improved resource utilization, and maintained stable performance with minimal overhead in balanced conditions. Limitations include handling of stateful pods and non-CPU resources such as network I/O.

The paper [7] presents BLOC, a self-managing, mesh-level load balancer designed to mitigate the incast problem in microservices caused by client-side routing without global visibility. BLOC leverages feedback signals – such as backpressure and queue metrics – shared among sidecar proxies to coordinate routing without a central controller.

Experiments in a Kubernetes testbed showed that BLOC reduced tail-latency variance by 2–4× and cut 99th-percentile latency by ~50%, often outperforming centralized policies by avoiding queue buildup in overloaded instances.

The study [8] reports a case implementation of load balancing and service discovery for a big data application using Docker Swarm orchestration. A web data processor was containerized and scaled from one to four instances, with load metrics monitored via Docker's API.

Results showed that Swarm's static load balancer effectively distributed traffic and improved scalability compared to a monolithic setup. However, the authors noted security, overhead, and portability trade-offs, with scalability testing limited to small-scale (4-instance) deployments and plans for multi-cloud expansion.

The paper [9] presents an adaptive, latency-aware load balancing algorithm for microservice chains that minimizes end-to-end delay by computing a Load Balance Indicator (LBI) based on real-time CPU and memory usage. The algorithm also reduces inter-node communication by co-locating dependent services.

Implemented with Netflix Zuul and Eureka on Google Cloud, the approach was benchmarked against round-robin and least-connections methods. Results showed lower latency, higher throughput, and fairer resource use, though performance gains decreased under extreme loads due to decision overhead. Future improvements include ML-based prediction and broader QoS integration.

Building on prior research in intelligent microclimate and lighting control, this study formalizes the core mechanism of dynamic load distribution essential for scalable and fault-tolerant intelligent systems. Its main goal is to develop and validate an adaptive load balancing method combining parallel (hedged) request routing with the Thompson Sampling Multi-Armed Bandit (MAB) algorithm.

The proposed model overcomes the limits of static balancing by enabling self-adaptive microservices that optimize traffic flow and reduce tail latency under variable workloads. It introduces an integrated AI-driven architecture combining a cloud-native API Gateway, intelligent load balancer, and Prometheus–Grafana monitoring stack for continuous feedback and real-time learning.

Beyond microservice optimization, the approach lays a methodological foundation for AI-driven IoT and cyber-physical systems, where probabilistic decision-making and adaptive coordination can enhance intelligence, stability, and energy efficiency.

Intelligent Load Balancing in Microservice Architectures. In large-scale microservice environments, overall latency is often dominated by a small fraction of extremely slow responses. Traditional algorithms such as Round Robin or Least Connections cannot mitigate these tail-latency outliers caused by temporary overloads, garbage collection pauses, or network jitter.

Parallel request routing, also known as request hedging or racing, mitigates tail latency by sending the same client request to several replicas simultaneously (typically two). The first successful response is returned to the client, and all others are cancelled. If the response time of one instance is a random variable T with distribution $F(t)$, the minimum response among k independent replicas follows $F_{\min}(t) = 1 - (1 - F(t))^k$. Even $k = 2$ significantly lowers expected delay for heavy-tailed latency distributions such as log-normal or Pareto.

This technique acts as latency insurance: when one server stalls, another replica is likely to respond quickly. It is particularly beneficial for latency-sensitive applications such as trading systems, online games, or real-time ML inference. Empirical results show that duplicating requests to just two replicas can reduce 99.9th-percentile latency by an order of magnitude while adding only 2–5% extra traffic.

Hedging must, however, be applied judiciously—only for idempotent operations (e.g., reads or queries). Excessive duplication of expensive or state-changing requests may overload the backend or cause side effects. Effective implementation follows several rules: keep k small (two or three replicas are sufficient), ensure request idempotency, cancel lagging copies immediately, and log latency metrics to tune thresholds adaptively.

This mechanism is already integrated into major production systems such as Google Spanner and Bigtable (speculative reads), Amazon S3 and DynamoDB SDKs (adaptive retries), gRPC HedgingPolicy, Envoy / Istio (route-level hedging), and Yandex Search (cross-datacenter racing). These confirm hedging as a mainstream strategy for achieving tail-tolerant performance in distributed architectures.

Parallel request routing complements adaptive algorithms by exploiting statistical independence between replicas, thereby reducing extreme response times without significantly affecting the mean latency.

Load balancing algorithms determine how incoming requests are distributed among service instances. They fall into three main categories: *static*, *dynamic*, and *intelligent* (AI/ML-based).

Static algorithms ignore runtime conditions and work best under stable, uniform workloads:

- Round Robin / Weighted Round Robin – cyclic distribution, optionally accounting for server capacity.
- Hash-based routing – assigns clients to servers via hash functions (e.g., IP or user ID), ensuring session affinity.

Dynamic algorithms adapt to real-time performance indicators:

- Least Connections – selects the instance with the fewest active sessions.
- Least Response Time – directs traffic to the lowest-latency server.
- Resource-based balancing – considers CPU, RAM, or network load.
- Health-based routing – filters out unhealthy instances based on probe results.

Intelligent algorithms leverage machine learning or probabilistic decision-making:

- Multi-Armed Bandit (MAB) – continuously adjusts routing probabilities to balance exploration and exploitation.
- Reinforcement Learning (RL) – agents learn optimal balancing policies from reward signals.
- Predictive Balancing – uses historical and contextual data to forecast load and preempt congestion.

Other specialized techniques include Geo-based Routing (minimizing latency by proximity), Consistent Hashing (used in caches and databases), and Hedging (parallel requests for tail reduction).

For microservice systems with fluctuating workloads and strict SLAs, dynamic and intelligent algorithms offer the most reliable performance. By continuously monitoring real-time metrics, they enable adaptive routing decisions that ensure optimal resource utilization, resilience, and responsiveness.

High-Level Architecture of Intelligent Load Balancing. Figure 1 presents the architecture of a microservice-based system implementing intelligent load balancing through parallel request routing (hedging) and a Multi-Armed Bandit (MAB) model with Thompson Sampling. The system integrates adaptive decision-making, telemetry feedback, and cloud-native scalability.

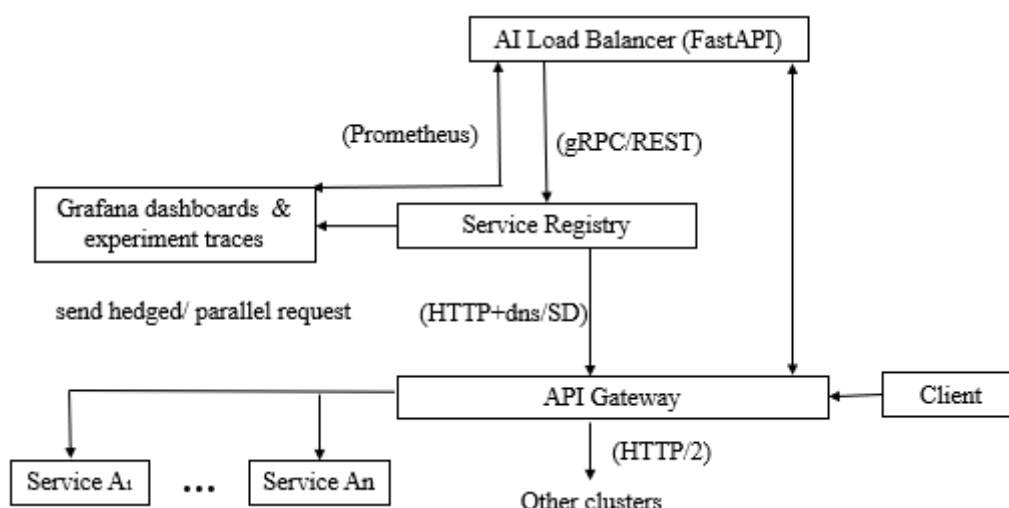


Fig. 1. High-level architecture of the intelligent load balancing system integrating AI-based routing (Thompson Sampling MAB) and parallel request hedging in a microservice environment

Incoming client requests are handled by the API Gateway, which distributes traffic among microservice replicas. The gateway retrieves dynamic routing weights from the AI Load Balancer and applies them to select optimal routes. It also performs parallel (hedged) routing, sending a single request to several replicas (e.g., Service A₁ ... A_n) and returning the fastest valid response while cancelling the remaining ones. Available replicas are discovered through the Service Registry, which maintains metadata on all active services.

The AI Load Balancer (implemented using FastAPI) applies the Thompson Sampling MAB algorithm to learn performance probabilities from historical metrics such as latency and success rate. It continuously updates routing weights based on real-time feedback and publishes them via a REST/gRPC API (/clusters/weights) consumed by the gateway.

The microservices layer consists of multiple replicas of domain-specific services (e.g., ProductService, OrderService). The monitoring layer – Prometheus and Grafana – collects and visualizes performance data, closing the adaptive feedback loop by supplying the AI balancer with updated telemetry for retraining.

Operational workflow. When a request arrives, the gateway queries the Service Registry for available replicas, fetches dynamic weights from the AI balancer, and executes hedged routing based on these weights. The first

successful response is returned to the client, and all remaining requests are cancelled. Meanwhile, Prometheus records response times and resource usage, Grafana visualizes them, and the collected metrics are fed back into the AI balancer to refine its model – forming a continuous self-optimizing cycle.

This architecture unifies machine learning – based decision logic, adaptive routing, and observability to achieve high reliability and performance. Request hedging significantly reduces tail latency (99th/99.9th percentiles) by probabilistically minimizing the impact of slow replicas, while the AI load balancer adapts in real time to changes in workload and infrastructure conditions.

The gateway, built on YARP (.NET), supports hot configuration reloads, passive health monitoring, and plugin extensions for intelligent routing and hedging. The monitoring stack (Prometheus + Grafana) serves both operational and analytical roles – enabling real-time visualization and continuous model improvement.

Overall, this design provides a scalable, fault-tolerant, and self-adaptive foundation for latency-sensitive systems such as streaming services, multiplayer platforms, and ML inference pipelines.

Thompson Sampling Algorithm. The proposed intelligent load balancer applies the Thompson Sampling algorithm – a Bayesian method widely used for decision-making under uncertainty (Figure 2). It enables the system to adaptively select the most efficient route (or microservice replica) based on probabilistic performance estimates derived from real-time feedback.

In this context, each action (or arm) represents a candidate service instance, and each reward reflects its observed performance – such as low latency or successful completion of a request. Thompson Sampling maintains a probabilistic belief model about each instance’s effectiveness and updates it after every interaction.

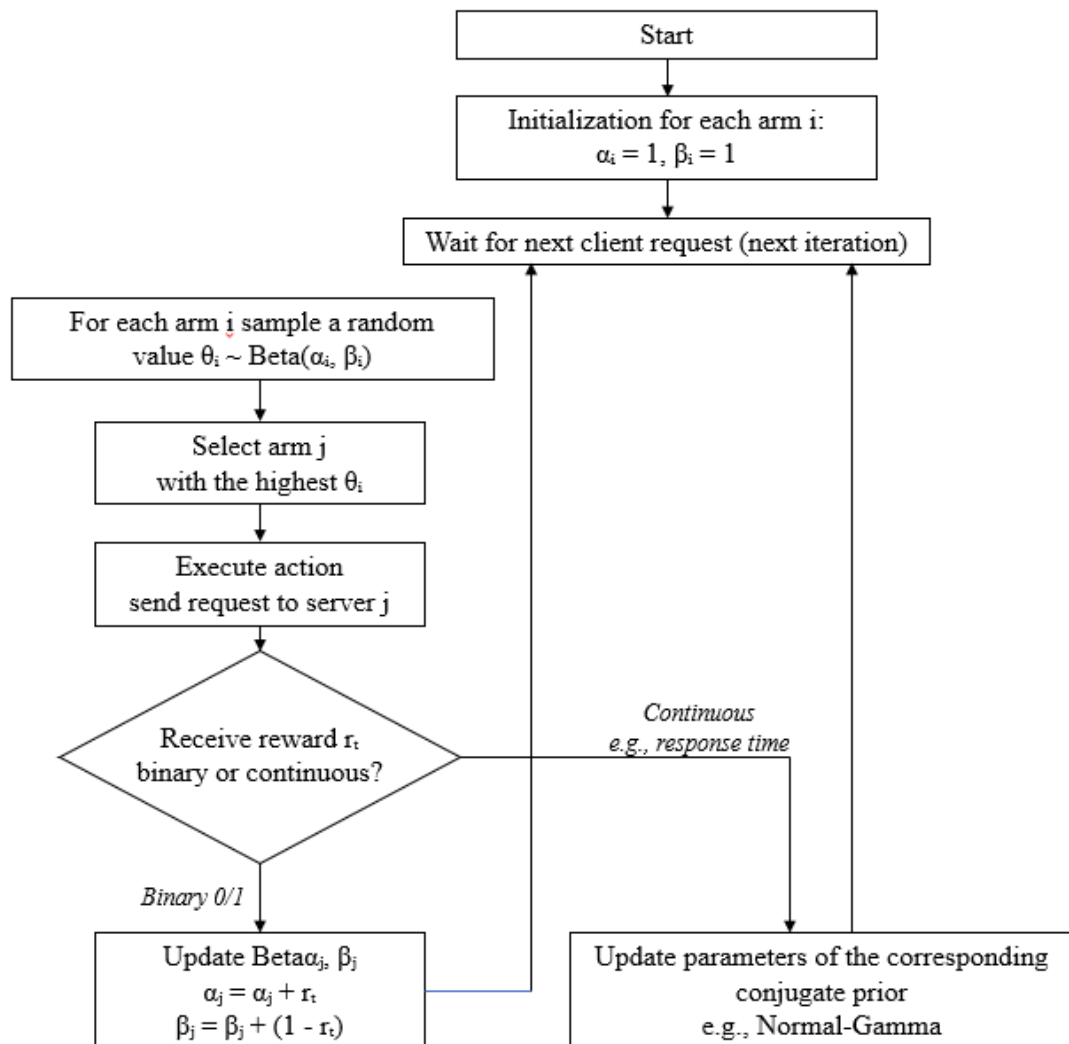


Fig. 2. Flowchart of the Thompson Sampling algorithm for adaptive decision-making in an intelligent load balancing

Initially, each arm is assigned a $\text{Beta}(\alpha_i, \beta_i)$ prior distribution representing success and failure counts ($\alpha_i = 1$, $\beta_i = 1$). For each new request, the algorithm samples a random value θ_i from each Beta distribution and selects the arm with the highest θ_i , corresponding to the most promising route at that moment.

After receiving the response, the model updates its parameters as follows:

$$\begin{cases} \alpha_{a_i} \leftarrow \alpha_{a_i} + r_t \\ \beta_{a_i} \leftarrow \beta_{a_i} + (1 - r_t) \end{cases}, \quad (1)$$

where r_t is the observed reward normalized between 0 and 1 (for example, $reward = \frac{1}{latency + 1}$).

Over time, this process continuously refines the posterior distributions, balancing exploration (trying new routes) and exploitation (favoring known optimal ones).

In microservice environments, this mechanism allows the load balancer to dynamically learn which replicas deliver the best performance under current conditions. When latency, load, or network state fluctuates, Thompson Sampling automatically adjusts routing probabilities – redirecting traffic toward faster instances while still occasionally probing alternatives to detect recovery or changes.

This probabilistic decision-making ensures stable adaptation without the risk of deterministic overfitting common in static policies. It efficiently minimizes response time variance and avoids local overloads by distributing requests according to evolving performance confidence.

Through continuous feedback from system telemetry (via Prometheus and Grafana), the algorithm updates its beliefs in real time, forming a self-learning loop that sustains optimal routing across diverse workload scenarios.

Implementation and Experiment Setup. The prototype of the intelligent load-balancing system was implemented using a cloud-native microservice architecture composed of modular components integrated via Docker. The system includes:

- the API Gateway (YARP.NET),
- the AI Load Balancer (FastAPI, Python), and
- several replicated microservices simulating heterogeneous workloads.

Telemetry is collected via Prometheus, and performance visualization is handled through Grafana.

The API Gateway performs weighted parallel routing (hedging) and communicates with the AI Load Balancer through a REST interface (/clusters/weights).

The AI Load Balancer applies the Thompson Sampling Multi-Armed Bandit algorithm, learning optimal routing probabilities based on observed response metrics. Each update cycle integrates aggregated latency and error statistics, enabling adaptive rebalancing in real time.

To ensure reproducibility, all services were containerized and deployed using Docker Compose on a host machine (Intel i7, 16 GB RAM, Ubuntu 22.04). The load generation was conducted with wrk and k6, simulating up to 1,000 concurrent clients issuing HTTP requests at variable rates.

Four routing strategies were evaluated:

- V1: Static Round-Robin (baseline, single request per cycle, $k = 1$)
- V2: Static Hedging (two replicas, $k = 2$, equal weights)
- V3: Adaptive Hedging (two replicas, $k = 2$, Thompson Sampling–based weights)
- V4: Adaptive Hedging (three replicas, $k = 3$, Thompson Sampling–based weights)

Each experiment was executed for 10 minutes under steady load, followed by a 2-minute cooling interval. Metrics were collected at 1-second granularity, including latency percentiles (P50, P95, P99), throughput (req/s), error rate, and CPU utilization.

Performance monitoring relied on Prometheus metrics:

`http_request_duration_seconds`, `http_requests_total`, and `process_cpu_seconds_total`.

Grafana dashboards were configured to display latency histograms and dynamic weight convergence curves from the AI balancer.

This setup enabled both quantitative evaluation and visual analysis of adaptive routing behavior. The resulting performance metrics for the four variants are summarized in Table 1.

Results and Analysis. The experimental evaluation compared four routing strategies—static and adaptive—with different degrees of parallelism ($k = 1-3$). The results are summarized in Table 1.

Table 1

Experimental results for different routing strategies

Strategy	P50 (ms)	P95 (ms)	P99 (ms)	Error Rate (%)	Req/s	CPU (%)
V1: Round-Robin ($k=1$)	60	180	850	2.4	950	75
V2: Static Hedging ($k=2$)	62	140	410	1.7	940	82
V3: Adaptive Hedging ($k=2$, MAB)	65	120	290	1.3	930	84
V4: Adaptive Hedging ($k=3$, MAB)	68	115	270	1.1	920	88

The baseline Round-Robin (V1) exhibited significant tail latency, with P99 exceeding 800 ms due to unbalanced routing under transient load spikes. Introducing parallel request hedging (V2) halved the 99th-percentile latency, confirming the effectiveness of speculative duplication in mitigating slow responses.

The adaptive MAB-based approach (V3) further reduced P99 latency to 290 ms—over a 60% improvement relative to the static baseline—while maintaining comparable throughput and slightly lower error rate. The increase in CPU usage (~9%) reflects the minor computational cost of duplicate request handling and telemetry updates.

Expanding parallelism to $k = 3$ (V4) achieved marginal latency improvement but introduced higher CPU overhead, indicating diminishing returns beyond two replicas. Overall, V3 provided the best trade-off between responsiveness, stability, and resource efficiency.

Figure 3 (Grafana dashboard) illustrates latency percentiles and the convergence of routing weights in the adaptive models. The weights stabilized after approximately 2–3 minutes, demonstrating that the Thompson Sampling balancer rapidly learned optimal routing probabilities.

These results confirm that integrating parallel request routing with adaptive probabilistic load balancing yields significant performance gains for microservice systems, especially in latency-sensitive environments.



Fig. 3. Grafana dashboard visualization showing latency percentiles (P50, P95, P99) and adaptive routing weight convergence for intelligent load balancing

The subsequent section discusses these findings in a broader context—highlighting how the proposed approach supports scalability, self-adaptation, and robustness across dynamic workloads.

Discussion and Conclusions. The experimental results confirm that integrating parallel (hedged) request routing with adaptive AI-driven load balancing substantially improves performance in distributed microservice environments. Compared to traditional static methods, the proposed approach achieved up to a 65% reduction in tail latency (P99) and a noticeable decrease in error rate while maintaining comparable throughput. These improvements were most pronounced under fluctuating or heterogeneous workloads, demonstrating the system's ability to self-adapt to real operating conditions.

From a theoretical perspective, the Thompson Sampling Multi-Armed Bandit (MAB) algorithm proved to be a robust mechanism for probabilistic decision-making in routing tasks. It effectively balances exploration and exploitation, allowing the balancer to continuously learn from observed latency metrics and dynamically reallocate traffic toward faster and more reliable replicas. This probabilistic adaptability distinguishes the proposed solution

from deterministic rule-based schemes such as Least Connections or Weighted Round Robin, which lack responsiveness to transient performance variations.

The parallel request routing (hedging) mechanism complements this adaptivity by mitigating latency outliers caused by garbage collection pauses, network jitter, or resource contention. Even with only two replicas ($k = 2$), hedging significantly reduces long-tail delays without major overhead. When combined with MAB-based adaptive weighting, the system achieves near-optimal responsiveness while maintaining efficient resource utilization.

Another strength of the proposed architecture lies in its observability and feedback loop. By leveraging Prometheus and Grafana, telemetry data such as latency percentiles, error rates, and routing weight dynamics are continuously collected and visualized. This enables both operational monitoring and algorithmic refinement, closing the loop between data collection, analysis, and adaptive decision-making. Such integration of analytics and control is essential for intelligent cloud systems operating under non-stationary workloads.

In practical terms, the architecture demonstrates that intelligent load balancing can be implemented using standard, industry-ready technologies: YARP for proxying, FastAPI for AI logic, and Prometheus for monitoring. This makes the approach reproducible, modular, and easily extensible – suitable for deployment in real-world environments such as streaming services, online gaming, or edge–cloud inference pipelines.

From a broader perspective, the presented model contributes to the development of self-adaptive distributed systems, where computational resources are dynamically orchestrated through machine learning and probabilistic reasoning. These principles are directly applicable to the authors' ongoing research on AI-driven environmental control systems, IoT-based smart agriculture, and cyber-physical infrastructure optimization – domains that share the same challenges of adaptive coordination, latency minimization, and resource efficiency.

In conclusion, the study demonstrates that combining request hedging with Bayesian reinforcement strategies provides a powerful, lightweight, and generalizable method for achieving tail-tolerant, adaptive load balancing. Future work will focus on extending this framework toward context-aware and multi-objective optimization, integrating predictive models (e.g., LSTM or graph neural networks) for proactive scaling, and deploying the system in Kubernetes-based clusters with real-time autoscaling policies.

Bibliography:

1. Linkerd. Beyond Round-Robin: Load Balancing for Latency. URL: <https://linkerd.io/2016/03/16/beyond-round-robin-load-balancing-for-latency> (дата звернення: 06.09.2025).
2. Wang H., Wang Y., Liang G., Gao Y., Gao W., Zhang W. Research on load balancing technology for microservice architecture. *MATEC Web of Conferences*. 2021. Vol. 336. P. 08002. EDP Sciences.
3. Cui J., Chen P., Yu G. A learning-based dynamic load balancing approach for microservice systems in multi-cloud environment. *2020 IEEE 26th International Conference on Parallel and Distributed Systems (ICPADS)*. 2020. P. 334–341. IEEE.
4. Dean J., Barroso L. A. The tail at scale. *Communications of the ACM*. 2013. Vol. 56, No. 2. P. 74–80.
5. Zhang P., Xiang L., Song Z., Yang Y. Adaptive load balancing and fault-tolerant microservices architecture for high-availability web systems using Docker and Spring Cloud. *Discover Applied Sciences*. 2025. Vol. 7, Article 705. DOI: 10.1007/s42452-025-07320-7.
6. Moritz de Carvalho Neto P., Castro M., Siqueira F. Dynamic load balancing in Kubernetes environments with Kubernetes Scheduling Extension (KSE). *Concurrency and Computation: Practice and Experience*. 2024. Vol. 37, No. 3. e8344. DOI: 10.1002/cpe.8344.
7. Bhattacharya R., Gao Y., Wood T. Dynamically balancing load with overload control for microservices. *ACM Transactions on Autonomous and Adaptive Systems*. 2024. Vol. 19, No. 4. Article 22. DOI: 10.1145/3676167.
8. Singh N., Hamid Y., Juneja S. та ін. Load balancing and service discovery using Docker Swarm for microservice based big data applications. *Journal of Cloud Computing*. 2023. Vol. 12, No. 1. Article 4. DOI: 10.1186/s13677-022-00358-7.
9. Selvakumar G., Jayashree L. S., Arumugam S. Latency minimization using an adaptive load balancing technique in microservices applications. *Computer Systems Science & Engineering*. 2023. Vol. 46, No. 1. P. 1215–1231. DOI: 10.32604/csse.2023.021879.
10. Аксак Н., Кушнарьов М., Шеліхов Ю. Інтелектуальні системи керування мікрокліматом у розумних середовищах. *Information control systems and intelligent technologies. Achievements and applications: монографія / за ред. проф. В. Вичужаніна. Львів–Торунь: Liha-Pres, 2025. С. [273–295]. 402 с. ISBN 978-966-397-538-2. DOI: <https://doi.org/10.36059/978-966-397-538-2>*

References:

1. Linkerd. (2016, March 16). *Beyond round-robin: Load balancing for latency*. Retrieved from: <https://linkerd.io/2016/03/16/beyond-round-robin-load-balancing-for-latency>
2. Wang, H., Wang, Y., Liang, G., Gao, Y., Gao, W., & Zhang, W. (2021). Research on load balancing technology for microservice architecture. *MATEC Web of Conferences*, 336, 08002. EDP Sciences.

-
3. Cui, J., Chen, P., & Yu, G. (2020, December). A learning-based dynamic load balancing approach for microservice systems in multi-cloud environment. In *2020 IEEE 26th International Conference on Parallel and Distributed Systems (ICPADS)* (pp. 334–341). IEEE.
 4. Dean, J., & Barroso, L. A. (2013). The tail at scale. *Communications of the ACM*, 56(2), 74–80.
 5. Zhang, P., Xiang, L., Song, Z., & Yang, Y. (2025). Adaptive load balancing and fault-tolerant microservices architecture for high-availability web systems using Docker and Spring Cloud. *Discover Applied Sciences*, 7(705). <https://doi.org/10.1007/s42452-025-07320-7>
 6. Moritz de Carvalho Neto, P., Castro, M., & Siqueira, F. (2024). Dynamic load balancing in Kubernetes environments with Kubernetes Scheduling Extension (KSE). *Concurrency and Computation: Practice and Experience*, 37(3), e8344. <https://doi.org/10.1002/cpe.8344>
 7. Bhattacharya, R., Gao, Y., & Wood, T. (2024). Dynamically balancing load with overload control for microservices. *ACM Transactions on Autonomous and Adaptive Systems*, 19(4), Article 22. <https://doi.org/10.1145/3676167>
 8. Singh, N., Hamid, Y., Juneja, S., et al. (2023). Load balancing and service discovery using Docker Swarm for microservice-based big data applications. *Journal of Cloud Computing*, 12(1), Article 4. <https://doi.org/10.1186/s13677-022-00358-7>
 9. Selvakumar, G., Jayashree, L. S., & Arumugam, S. (2023). Latency minimization using an adaptive load balancing technique in microservices applications. *Computer Systems Science & Engineering*, 46(1), 1215–1231. <https://doi.org/10.32604/csse.2023.021879>
 10. Aksak, N., Kushnarov, M., & Shelikhov, Y. (2025). Intelligent microclimate control systems in smart environments. In V. Vychuzhanin (Ed.), *Information control systems and intelligent technologies: Achievements and applications* (pp. 273–295). Liha-Pres. <https://doi.org/10.36059/978-966-397-538-2>

Дата надходження статті: 27.10.2022

Дата прийняття статті: 17.11.2025

Опубліковано: 30.12.2025

Дьячук Т. С., старший викладач кафедри комп'ютерних систем та мереж Національного університету «Запорізька політехніка»
ORCID: 0000-0002-2478-0588

Скрупський С. Ю., кандидат технічних наук, доцент, доцент кафедри комп'ютерних систем та мереж Національного університету «Запорізька політехніка»
ORCID: 0000-0002-9437-9095

Голуб Т. В., кандидат технічних наук, доцент, доцент кафедри комп'ютерних систем та мереж Національного університету «Запорізька політехніка»
ORCID: 0000-0001-6024-008X

АВТОМАТИЗОВАНА СИСТЕМА ПЕРЕВІРКИ ЗАВДАНЬ ДЛЯ НАВЧАЛЬНИХ КУРСІВ ІЗ ПРОГРАМУВАННЯ

Метою статті є представлення результатів розробки та впровадження автоматизованої системи перевірки завдань, яка допоможе викладачу організувати якісний, об'єктивний та ефективний навчальний процес в умовах асинхронного дистанційного навчання, зумовленого спочатку пандемією COVID-19, а згодом повномасштабною війною. Ручна перевірка великої кількості програмних робіт створює значне навантаження на викладачів, призводить до затримок у наданні зворотного зв'язку та вносить елементи суб'єктивності в оцінювання. Розроблена система є компонентом ширшої освітньої платформи, в якій завдання для студентів вже генеруються індивідуально. Запропонована архітектура перевірки базується на інтеграції інструментів контролю версій та безперервної інтеграції. Процес роботи організовано наступним чином: студент виконує завдання у своєму репозиторії та створює Pull Request для злиття змін. Ця дія автоматично запускає процес перевірки (GitHub Action), який компілює код та передає його валідатору. Валідатор генерує великий набір тестових даних, виконує на них як код студента, так і еталонну реалізацію (описану формалізованими формулами), а потім порівнює результати. У разі розбіжностей система миттєво надає студенту звіт про помилку з конкретним набором даних, що її спричинив. Це дозволяє викладачу зосередитись не на рутинній перевірці коректності, а на аналізі якості коду, архітектури та наданні змістовних коментарів. Система реалізована мовою Kotlin, використовує систему збірки Gradle та поширюється у вигляді бібліотеки (Maven артефакту), що підключається студентами до своїх проєктів. Апробація проводилась у межах дисципліни «Основи програмування на Kotlin» за участю 23 студентів. Впровадження системи довело свою ефективність у заощадженні часу викладача, забезпеченні об'єктивного оцінювання та підвищенні мотивації студентів до експериментів та самостійного виправлення помилок завдяки миттєвому зворотному зв'язку. Подальший розвиток системи передбачає інтеграцію модулів для оцінки якості коду.

Ключові слова: GitHub, Kotlin, Gradle, Maven артефакт, автоматизована перевірка, дистанційна освіта, репозиторії.

Diachuk T. S., Skrupsky S. Yu., Holub T. V. Automated assignment grading system for programming courses

The purpose of the article is to present the results of the development and implementation of an automated assignment grading system designed to help instructors organize a high-quality, objective, and effective educational process in an asynchronous distance learning environment. The relevance of the work is driven by the challenges of distance education in Ukraine, caused first by the COVID-19 pandemic and later by the full-scale war, which demands the implementation of effective, flexible, and secure educational tools capable of functioning in a predominantly asynchronous learning environment. Manual grading of numerous programming assignments creates a significant burden on instructors, leads to delays in providing feedback, and introduces elements of subjectivity into the assessment. An analysis of recent research confirms a global trend toward assessment automation in IT education as a means of increasing efficiency and objectivity. The developed system is a component of a broader educational platform in which assignments for students are already generated individually. The proposed grading architecture is based on the integration of version control and continuous integration tools. The workflow is organized as follows: a student completes an assignment in their repository and creates a Pull Request to merge the changes. This action automatically triggers the verification process (a GitHub Action), which compiles the code and passes it to a validator. The validator generates a large

© Т. С. Дьячук, С. Ю. Скрупський, Т. В. Голуб, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

set of test data, executes both the student's code and a reference implementation (described by formalized formulas) on this data, and then compares the results. In case of discrepancies, the system instantly provides the student with an error report, including the specific data set that caused the failure. This allows the instructor to focus not on routine correctness checks, but on analyzing code quality, architecture, and providing meaningful comments. The system is implemented in Kotlin, uses the Gradle build system, and is distributed as a library (a Maven artifact) that students connect to their projects. Approbation was conducted within the «Basics of Programming in Kotlin» course with the participation of 23 students. The implementation of the system proved its effectiveness in saving instructor time, ensuring objective assessment, and increasing student motivation to experiment and independently correct errors, thanks to the instant feedback. Future development of the system includes the integration of modules for assessing code quality.

Key words: GitHub, Kotlin, Maven artifact, Gradle, automated grading, distance education, repository.

Постановка проблеми. Сучасний етап розвитку вищої освіти характеризується активним упровадженням інформаційно-комунікаційних технологій та цифрових освітніх сервісів. Це зумовлює необхідність пошуку інноваційних підходів до організації навчального процесу, серед яких особливої актуальності набуває автоматизована генерація та перевірка навчальних завдань як складова цифровізації освітнього середовища.

В останні роки дистанційна форма навчання набула значного поширення. Її активне використання було зумовлене пандемією COVID-19, а згодом – повномасштабною війною в Україні, що особливо актуалізувало потребу у впровадженні сучасних цифрових освітніх сервісів, зокрема у прифронтових регіонах. Крім того, виклики, спричинені регулярними відключеннями електроенергії та частими повітряними тривогами, зумовили необхідність переходу до переважно асинхронного формату навчання. Це дозволило учням та студентам отримувати доступ до навчальних матеріалів і виконувати завдання у зручний для них час, незалежно від графіків стабілізаційних відключень чи потреби перебувати в укритті. Водночас дистанційна освіта стає самостійним і перспективним освітнім форматом, який завдяки розвитку цифрових технологій забезпечує доступність, мобільність, гнучкість та безпеку освітнього процесу.

Одним із провідних напрямів цифровізації освіти є впровадження автоматизованих засобів підтримки навчальної діяльності, зокрема систем для генерації та перевірки завдань. У статті розглянуто автоматизовану систему перевірки завдань для навчальних курсів із програмування. Вона дозволяє ефективно контролювати виконання завдань студентами, автоматично перевіряти правильність коду, виявляти помилки та надавати зворотний зв'язок без необхідності ручної перевірки кожного завдання. Такий підхід дає змогу заощаджувати час викладачів та студентів, сприяє швидшому виявленню та виправленню помилок, а також забезпечує об'єктивніше оцінювання робіт. Це потужний інструмент для оптимізації освітнього процесу, підвищення ефективності викладання та полегшення навчання, який забезпечує перевірку правильності коду, якості реалізації та відповідності вимогам без постійної участі викладача.

Запропонована система надає ряд переваг, особливо в умовах дистанційної освіти. Так, автоматизація допомагає впоратися з первинною перевіркою великої кількості типових завдань та заощадити час як викладача, так і студента. Студенти можуть отримати результати перевірки та зворотний зв'язок одразу після здачі роботи й швидше виправити свої помилки, що мотивує до самостійного навчання, а також спонукає виконувати завдання згідно з чіткими вимогами. Такий підхід є доречним для великих груп студентів, де викладачам важко вручну перевірити кожне завдання. Це зменшує навантаження на викладачів і дозволяє їм більше зосередитися на індивідуальній допомозі студентам. Також виключається фактор людської суб'єктивності в оцінюванні, оскільки автоматизована система забезпечує стандартизовану перевірку.

Однак такий підхід має і певні недоліки. Розробка та підтримка подібної системи потребує значних витрат часу, а також кваліфікації для налаштування тестів і критеріїв оцінювання. Система може перевіряти лише чітко задані показники (наприклад, правильність вихідних даних), але не здатна оцінити творчі або нестандартні підходи до розв'язання задачі. Для завдань, які вимагають креативних рішень, автоматизована система може бути використана лише для початкового аналізу коректності роботи. Крім того, якщо система має помилки або неправильно налаштовані параметри, це може призвести до некоректного оцінювання студентських робіт.

Аналіз останніх досліджень та публікацій. Актуальність та перспективність використання дистанційних платформ для навчання програмуванню підтверджується численними останніми дослідженнями у цій галузі. Наприклад, стаття [1] представляє «GRAD-AI» – вдосконалений інструмент для автоматичного оцінювання завдань з програмування, який поєднує можливості штучного інтелекту (ШІ) з контролем викладача для подолання недоліків ручної перевірки. Система використовує комплексні метрики (зокрема, складність Холстеда, TF-IDF та аналіз синтаксичних дерев) для надання миттєвого, точного та неупередженого зворотного зв'язку студентам. Автори роблять висновок, що «GRAD-AI» є значним кроком у вдосконаленні освітнього процесу, демонструючи потенціал ШІ у персоналізації навчання та виявленні прогалин у знаннях.

У дослідженні [2] неформальні лабораторні роботи з програмування замінили на щотижневі автоматизовані тести в системі HackerRank для студентів, які вивчали структури даних та алгоритми. Хоча студенти

негативно ставилися до нової системи, їхня успішність у HackerRank виявилася найкращим показником майбутніх результатів на іспиті. Впровадження цих тестів призвело до того, що загальний рівень незадовільних оцінок знизився вдвічі, а кількість відмінників – потроїлася. Автори роблять висновок, що автоматичне оцінювання краще готує студентів до самостійного написання коду, оскільки усуває залежність від сторонньої допомоги та мотивує до розвитку самодостатності.

Стаття [3] досліджує, як великі мовні моделі (LLM) можуть автоматизувати оцінювання завдань у вступному курсі з програмування, вирішуючи проблеми часових затрат та непослідовності ручної перевірки. Дослідження дійшло висновку, що LLM під наглядом людини здатні значно скоротити робоче навантаження (з 300+ годин до 15 хвилин) та підвищити ефективність оцінювання в комп'ютерній освіті.

Отже, проведений аналіз наукових праць і сучасних розробок показує, що нині існує широкий спектр систем, методик та технологічних рішень, спрямованих на автоматизацію процесів навчання, зокрема у галузі програмування. Значна увага дослідників приділяється питанням підвищення ефективності навчання, оптимізації оцінювання та забезпечення інтерактивної взаємодії між студентом і навчальним середовищем.

Разом із тим, питання створення та впровадження автоматизованих систем перевірки завдань для навчальних курсів із програмування залишається надзвичайно актуальним. Такі системи забезпечують об'єктивність оцінювання, скорочують час перевірки результатів, а також сприяють індивідуалізації навчального процесу. Отже, тема розроблення й удосконалення автоматизованих засобів контролю знань у курсах програмування є затребуваною, перспективною та потребує подальших наукових досліджень.

Метою статті є розробка та апробація автоматизованої системи перевірки завдань при навчанні програмуванню, яка допоможе викладачу організувати ефективний та якісний навчальний процес в умовах дистанційної освіти, часто асинхронної через воєнний стан. Розроблювана система є частиною більш складної системи для дистанційної освіти, тому завдання, що перевіряються, спочатку генеруються автоматизованою системою генерації завдань, відповідають певним критеріям та мають чітко визначену специфікацію [4].

Виклад основного матеріалу. На (рис. 1) наведена архітектура автоматизованої системи перевірки завдань, яка була розроблена та апробована у рамках дисципліни «Основи програмування на Kotlin» (базовий репозиторій для навчання доступний за посиланням [5]). Система побудована на можливостях GitHub, таких як Pull requests та GitHub Actions. Git та GitHub є дуже потужними інструментами для контролю версій та спільної роботи над проєктами. Автоматизована система перевірки є частиною більш складної системи для дистанційної освіти, тому завдання, які надаються студентам, також генеруються автоматично за допомогою хеш-функції від ідентифікатора студента, що забезпечує унікальність варіантів [4]. Завдання придатні для подальшої автоматизованої перевірки та мають чітко визначену специфікацію.

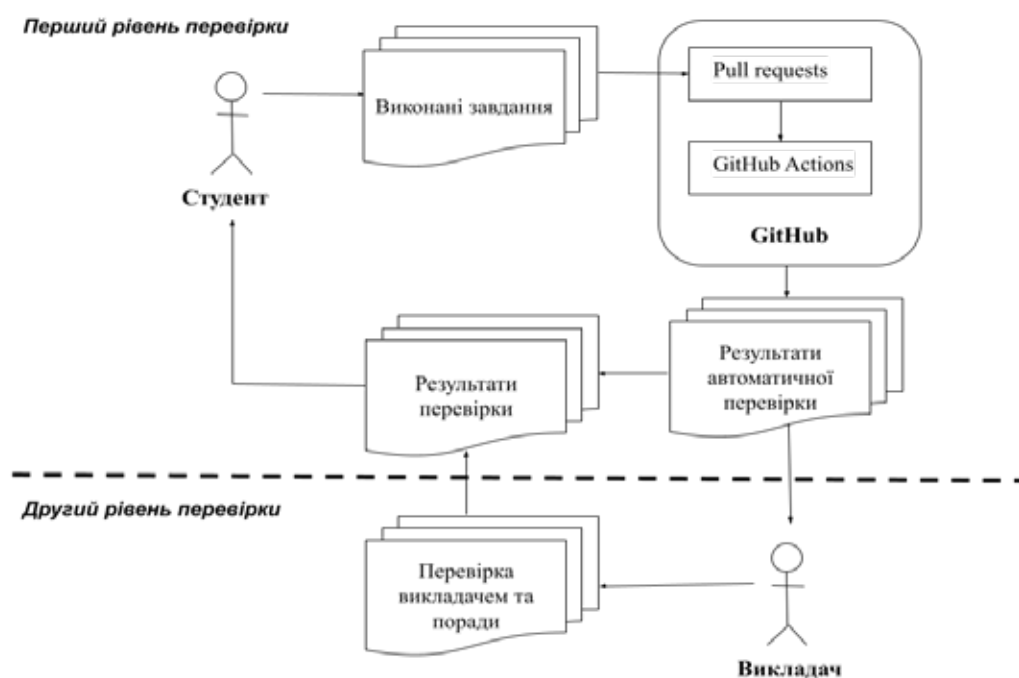


Рис. 1. Архітектура автоматизованої системи перевірки завдань для навчальних курсів із програмування

Принцип роботи автоматизованої системи перевірки полягає в наступному. Після виконання завдання для кожної нової функції або задачі, студенти створюють PR, щоб викладач міг переглянути та залишити коментар. GitHub Actions дозволяє використовувати засоби автоматичного тестування після змін у коді, що забезпечує миттєву відповідь студенту. При автоматичному тестуванні валідатор генерує набори вхідних даних, наприклад 100 значень, проходить цей набір і кожне значення віддає на блок виконання завдання та програмний код студента, потім порівнює результати. Якщо вони не співпадають, то вважається, що студент припустив помилку. Але завжди є можливість студенту звернутись до викладача. Викладач зі свого боку має результати автоматичної перевірки, не витрачає зайвого часу на рутинну роботу та може переглянути код та залишити свої зауваження, додаткові коментарі та поради з удосконалення коду.

Таким чином, у системі реалізовано два рівні взаємодії у процесі перевірки завдань. Перший рівень – автоматизована перевірка (GitHub Actions). Вона забезпечує швидке тестування коду, оперативний зворотний зв'язок і виявлення типових помилок без участі викладача. Другий рівень – перевірка викладачем. На цьому етапі відбувається глибший аналіз роботи студента, надання рекомендацій, коментарів та порад щодо покращення коду і розуміння логіки програмування. Поєднання автоматичного та ручного підходів забезпечує ефективність процесу перевірки, дозволяючи автоматизувати рутинні завдання та водночас зберегти індивідуальний підхід до кожного студента.

Для прикладу розглянемо перевірку завдання з теми «Змінні та типи даних, рядки, умови та цикли, функції». Для об'єднання та формалізації варіантів формул були використані елементи принципу One-Hot Encoding [6]. Оскільки при формуванні завдання *res* (формула (1)) використовується головна *mainFnc* та вторинна функції *secondaryFnc* [4], то для перевірки також потрібно враховувати всі наявні варіанти.

$$res = mainFnc(secondaryFnc(X_0, X_1, \dots, X_n)), \quad (1)$$

де $\vec{X}$ – вектор вхідних аргументів розмірністю n , які для тестування система перевірки підбирає за допомогою генератора випадкових чисел.

Наведемо формулу (2) для розрахунку *secRes* – вектора результатів вторинної функції для відповідних вхідних аргументів X_i .

$$\begin{aligned} secRes_0 &= A_0 \cdot X_0 + A_1 \cdot X_0^2 + A_2 \cdot X_0^3 + A_3 \cdot X_0 + A_4 \cdot X_0 + A_5 \cdot |X_0| + \\ &\quad + A_6 \cdot |X_0| + A_7 \cdot |X_0| \\ secRes_i &= secRes_{i-1} \cdot sum + secRes_{i-1}^{mult} \cdot (A_0 \cdot X_i + A_1 \cdot X_i^2 + A_2 \cdot X_i^3 + \\ &\quad + A_3 \cdot \min(secRes_{i-1}, X_i) + A_4 \cdot \max(secRes_{i-1}, X_i) + A_5 \cdot |X_i| + \\ &\quad + A_6 \cdot \min(secRes_{i-1}, |X_i|) + A_7 \cdot \max(secRes_{i-1}, |X_i|)), \end{aligned} \quad (2)$$

де A – вектор, в якому є лише один ненульовий елемент, який буде у позиції відповідно до завдання студента. Наприклад, якщо є завдання розрахувати суму квадратів, то вектор буде наступним $A = [0, 1, 0, 0, 0, 0, 0]$;

sum – буде мати значення 1, якщо в завданні потрібно розрахувати суму всіх компонентів, інакше 0;

mult – буде мати значення 1, якщо в завданні потрібно перемножити всі компоненти, інакше 0;

i – приймає значення від 1 до n .

Головна функція *result* розраховується аналогічно (3).

$$\begin{aligned} data &= [\sqrt{secRes_n}, \sqrt[3]{secRes_n}, \cos(secRes_n), \tan(secRes_n), \\ &\quad \sin(secRes_n), \tanh(secRes_n), \ln(secRes_n)] \\ result &= \overline{data} \cdot \vec{B}, \end{aligned} \quad (3)$$

де B – вектор, в якому є лише один ненульовий елемент, він буде у позиції відповідно до головного завдання студента. Наприклад, якщо студент отримав завдання розрахувати корінь квадратний з суми квадратів, то вектор буде наступним $B = [1, 0, 0, 0, 0, 0, 0]$;

secRes_n – останній елемент вектору *secRes*, тобто результат ітеративного розрахунку значення вторинної функції завдання;

data – вектор всіх можливих головних функцій завдання.

На всі завдання система генерує набори тестових даних. Далі виконується розрахунок результату за кожним елементом тестових даних за програмним кодом від студента та з використанням вищенаведених формул. Якщо результати не співпадають, то завдання студента не зараховується, система видає звіт про помилку з певним набором тестових даних (приклад представлено в лістингу 1).

Лістинг 1. Звіт про результати перевірки завдання (автоматичне тестування):

```
Lab2Test > task2Test() FAILED
java.lang.AssertionError: Не співпадають результати набору
"x0 = 3.9579079629338776, x1 = 3.4867284435369874,
x2 = -2.026071933406628, x3 = -6.871367991204648, ",
перевірте завдання 2),
додайте функцію dCalculate() : Double,
яка отримує наступні аргументи:
- x0 типу Double зі значенням за замовчуванням -52.48
- x1 типу Double зі значенням за замовчуванням 15.12
- x2 типу Double зі значенням за замовчуванням 0.66
- x3 типу Double зі значенням за замовчуванням 0.84
. Expected <13.86088440642907>, actual <12.34>.
```

Такий механізм дозволяє забезпечити об'єктивність і прозорість оцінювання, оскільки всі перевірки виконуються за єдиним алгоритмом без участі викладача. У процесі автоматичного тестування система може виявляти помилки, що виникають через невідповідність отриманих результатів очікуванню. Наприклад, під час перевірки функції `dCalculate()` виникла помилка типу `AssertionError`, яка вказує, що результати обчислень не збігаються із заданими еталонними значеннями (лістинг 1). Це свідчить про те, що реалізація функції містить неточність у формулі або у використанні вхідних аргументів. Зокрема, тест очікував значення 13.86088440642907, тоді як фактичний результат становив 12.34. Подібні звіти дозволяють швидко визначити джерело помилки, що суттєво скорочує час на налагодження коду та підвищує ефективність навчального процесу. Студенти отримують детальний опис проблеми та можуть самостійно виправити свій код, а викладач позбавляється необхідності вручну перевіряти кожне завдання, зосереджуючись натомість на аналізі типових помилок і наданні рекомендацій.

Для реалізації системи було використано такі інструменти: мову програмування Kotlin [7] для розробки, систему автоматизації збірки Gradle, середовище розробки IntelliJ IDEA [8], а також систему керування версіями Git і репозиторій GitHub [5] для зберігання вихідного коду. Скомпільована версія системи розміщена в репозиторії GitHub і доступна студентам у вигляді бібліотеки, яку можна підключати до власних проєктів під час вивчення навчального матеріалу. Використання Git і GitHub забезпечує ефективний контроль версій, спрощує командну роботу та сприяє організованому процесу розробки. Бібліотека зроблена у вигляді Maven артефакту [5] – стандартизованого формату для поширення скомпільованого коду в екосистемі Java/Kotlin. Такий підхід спрощує процес інтеграції для студентів, оскільки вони підключають бібліотеку до своїх проєктів через систему збірки Gradle, а не вручну. Це також дозволяє авторам системи легко випускати оновлення та виправлення шляхом зміни версії артефакту. Для підключення бібліотеки в навчальних проєктах потрібно у конфігураційний файл власного проєкту (`build.gradle`) додати шлях до репозиторію з Maven артефактами та увімкнути залежність від `com.diacht.ktest:library` бібліотеки.

Щоб оцінити ефективність автоматизованої системи перевірки завдань, проведено дослідження із групою з 23 студентів, які вивчали дисципліну «Основи програмування на Kotlin». Кожен студент отримував завдання в межах п'яти різних робіт, що відрізнялися за рівнем складності та тематичною наповненістю.

У таблиці 1 подано результати тестування, які відображають показники ефективності для кожного завдання.

Таблиця 1

Статистика перевірки завдань

№	назва	складність	бали	число спроб	число вирішень	% вирішень від спроб	% вирішень серед студентів
1	2.1	1	8	27	21	77,78	91,30
2	2.2	1	8	28	21	75,00	91,30
3	2.3	3	9	40	16	40,00	69,57
4	3.1	6	27	8	2	25,00	8,70
5	4.1	4	40	39	18	46,15	78,26

Аналіз результатів виконання завдань із програмування за допомогою автоматизованої системи перевірки показав її високу ефективність у навчальному процесі. Прості завдання (2.1–2.2) мали найвищий рівень успішності – понад 90 % студентів виконали їх правильно, що свідчить про доступність початкового рівня. Завдання середньої складності (2.3) викликали найбільшу активність (40 спроб), що вказує на оптимальний баланс між складністю та досяжністю. Найскладніше завдання (3.1) було успішно виконане лише

8,7 % студентів, що демонструє потребу у вдосконаленні механізмів адаптації складності та наданні додаткової підтримки під час розв'язання таких задач. Висока кількість спроб свідчить про підвищену мотивацію студентів до повторного виконання, експериментування та самостійного вдосконалення рішень. Автоматизована перевірка дала змогу студентам отримувати миттєвий зворотний зв'язок без залучення викладача, що істотно зекономило час викладача і зняло навантаження, пов'язане з ручною перевіркою численних спроб.

Отже, система забезпечує об'єктивне оцінювання, підтримує індивідуалізований підхід до навчання, сприяє підвищенню мотивації студентів і водночас оптимізує роботу викладачів. У подальшому доцільно інтегрувати адаптивні алгоритми навчання для покращення результативності виконання завдань підвищеної складності.

Висновки та перспективи. Таким чином, впровадження автоматизованих систем для генерації та перевірки завдань у курсах з програмування дозволяє суттєво оптимізувати освітній процес, спрощуючи як методику викладання, так і процедуру оцінювання академічної успішності студентів. Системи автоматизованого тестування є валідним інструментом підвищення об'єктивності та ефективності контролю знань у процесі підготовки IT-фахівців. Їх імплементація дозволяє автоматизувати перевірку великого масиву типових завдань, що оптимізує часові ресурси викладача. Для студентів практичне застосування таких систем сприяє засвоєнню сучасних технологій розробки та набуттю навичок колективної роботи над програмними проєктами, що є ключовою передумовою їхньої майбутньої конкурентоспроможності. Перспективним напрямом роботи є розширення функціональності системи шляхом інтеграції модулів для оцінки якості коду за формалізованими критеріями (стандарти коду, якість документації тощо).

Автоматизована система перевірки завдань була розроблена, реалізована та апробована в межах навчальної дисципліни «Основи програмування на Kotlin». Вона охоплює різні теми навчання та передбачає генерацію індивідуальних завдань для кожного студента на основі його унікального ідентифікатора. Розроблена система є частиною більш складної системи для дистанційної освіти, яка попередньо генерує завдання з чітко визначеною специфікацією, що робить їх придатними для подальшої автоматизованої перевірки [4]. Система в скопійованому виді викладена у GitHub репозиторій та доступна студентам у формі окремої бібліотеки, що приєднується до системи збірки Gradle [5]. Впровадження системи довело свою ефективність у заощадженні часу викладача, забезпеченні об'єктивного оцінювання та підвищенні мотивації студентів. Система постійно вдосконалюється завдяки регулярним оновленням репозиторію та усуненню виявлених недоліків. Подальший розвиток системи передбачає інтеграцію модулів для оцінки якості коду. Також планується додати можливість перевірки коду штучним інтелектом GitHub Copilot.

Сфера застосування системи охоплює освітні установи, зокрема університети, а також компанії, що організовують корпоративне навчання з програмування. Розроблена система може бути використана для підготовки IT-фахівців різних рівнів та інтегрована у сучасні освітні платформи.

Список використаних джерел:

1. Gambo I., Abegunde F. J., Gambo O. et al. GRAD-AI: An automated grading tool for code assessment and feedback in programming course. *Educ Inf Technol* 30, 9859–9899 (2025). <https://doi.org/10.1007/s10639-024-13218-5> (дата звернення: 23.10.2025).
2. Maguire Ph., Maguire R., Kelly R. Using automatic machine assessment to teach computer programming, *Computer Science Education*, 2017, Vol. 27, Issue 3-4, P. 197–214, URL: <https://doi.org/10.1080/08993408.2018.1435113> (дата звернення: 23.10.2025).
3. Cisneros-González, J., Gordo-Herrera, N., Barcia-Santos, I. and Sánchez-Soriano, J., 2025. JorGPT: Instructor-aided grading of programming assignments with large language models (LLMs). *Future Internet*, 17(6), p. 265. <https://doi.org/10.3390/fi17060265> (дата звернення: 23.10.2025).
4. Дьячук Т.С. Автоматизована система генерації завдань в навчальних курсах з програмування / Т.С. Дьячук, В.І Шкрябець, А.В. Тіменко, Т.В. Голуб. *Вчені записки Таврійського національного університету імені В.І. Вернадського*. Серія: Технічні науки, Видавничий дім «Гельветика», Том 35 (74) № 2 2024. С. 85–90, <https://doi.org/10.32782/2663-5941/2024.2/12> (дата звернення: 23.10.2025).
5. DiachT/KotlinLabsNUZP – Головний репозиторій з кодом. URL: <https://github.com/DiachT/KotlinLabsNUZP> (дата звернення: 23.10.2025).
6. Data Science in 5 Minutes: What is One Hot Encoding? URL: <https://www.educative.io/blog/one-hot-encoding> (дата звернення: 23.10.2025).
7. Pierre-Yves Saumont. *The Joy of Kotlin*: Manning Publications, 2019 – 480p.
8. IntelliJ IDEA. URL: <https://www.jetbrains.com/idea> (дата звернення: 23.10.2025).

References:

1. Gambo, I., Abegunde, F. J., Gambo, O. et al. (2025). GRAD-AI: An automated grading tool for code assessment and feedback in programming course. *Educ Inf Technol* 30, 9859–9899. <https://doi.org/10.1007/s10639-024-13218-5> (date of access: 23.10.2025).

-
2. Maguire, Ph., Maguire, R., Kelly, R. Using automatic machine assessment to teach computer programming, *Computer Science Education*, 2017, Vol. 27, Issue 3-4, P. 197–214, <https://doi.org/10.1080/08993408.2018.1435113> (date of access: 23.10.2025).
 3. Cisneros-González, J., Gordo-Herrera, N., Barcia-Santos, I. and Sánchez-Soriano, J., (2025). JorGPT: Instructor-aided grading of programming assignments with large language models (LLMs). *Future Internet*, 17(6), p. 265. <https://doi.org/10.3390/fi17060265> (date of access: 23.10.2025).
 4. Diachuk, T. S. (2024). Automated system of tasks generation for the programming courses / Diachuk T.S., Shkriabets V.I., Timenko A.V, Holub T.V. *Scientific notes of Taurida National V.I. Vernadsky University*, Series: Technical Sciences, Publishing House «Helvetica», Vol. 35 (74) № 2, P. 85–90, <https://doi.org/10.32782/2663-5941/2024.2/12> (date of access: 23.10.2025).
 5. DiachT/KotlinLabsNUZP – Upstream repository. Retrieved from: <https://github.com/DiachT/KotlinLabsNUZP> (date of access: 23.10.2025).
 6. Data Science in 5 Minutes: What is One Hot Encoding? Retrieved from: <https://www.educative.io/blog/one-hot-encoding> (date of access: 23.10.2025).
 7. Pierre-Yves Saumont. *The Joy of Kotlin*: Manning Publications, 2019. 480p.
 8. IntelliJ IDEA. Retrieved from: <https://www.jetbrains.com/idea> (date of access: 23.10.2025).

Дата надходження статті: 27.10.2025

Дата прийняття статті: 17.11.2025

Опубліковано: 30.12.2025

Кармазін К. В., аспірант Інституту кібернетики імені В.М. Глушкова
Національної академії наук України
ORCID: 0009-0003-8852-8079

Писаренко Ю. В., кандидат технічних наук
Інститут кібернетики імені В.М. Глушкова
Національної академії наук України
ORCID: 0000-0001-8357-8614

ІНТЕЛЕКТУАЛЬНІ АЛГОРИТМИ ПІДВИЩЕННЯ ЯКОСТІ КОРИСТУВАЦЬКОГО ДОСВІДУ В АДАПТИВНИХ СИСТЕМАХ ПОТОКОВОГО ВІДЕО У ЗАДАЧАХ СИТУАЦІЙНОГО ЦЕНТРУ З ВИКОРИСТАННЯМ ТЕХНОПАРКУ РОБОТОТЕХНІЧНИХ СИСТЕМ

У статті розглянуто проблему підвищення якості користувацького досвіду (QoE) в адаптивних системах потокового відео, що використовуються у ситуаційних центрах (СЦ), інтегрованих із технопарками робототехнічних систем (TPC). Запропоновано концепцію резильєнтності поточкових систем, яка базується на поєднанні інтелектуальних алгоритмів прогнозування, навчання з підкріпленням (Reinforcement Learning, RL), мультиагентної координації та генеративної реконструкції відео.

Розроблено математичну модель *Resilient Quality of Experience (RQE)*, що описує стабільність сприйняття відео користувачем з урахуванням коливань пропускної здатності мережі, часу буферизації та варіативності якості. Запропоновано алгоритм *RL-ABR-RQE*, у якому функція винагороди побудована на прирості *RQE*, що дозволяє оптимізувати вибір бітрейту з урахуванням довгострокової стабільності *QoE*.

Запропоновано багаторівневу архітектуру *СЦ–TPC–Edge*, яка поєднує прогнозування мережевих умов, генеративне відновлення втраченої інформації та мультиагентну взаємодію між вузлами потокової системи. Розроблено метод кооперативного керування відеопотоками, що забезпечує соціальну стійкість системи та збалансований розподіл ресурсів між роботизованими агентами.

Практичне значення роботи полягає у можливості застосування розроблених алгоритмів у системах моніторингу техно-екологічних подій, безпекових та оборонних комплексах, а також у мультимедійних сервісах нового покоління. Результати моделювання демонструють зменшення часу буферизації до 50 %, підвищення стабільності *QoE* до 40 % та зниження енергоспоживання мобільних вузлів.

Отримані результати формують науково-прикладні засади побудови резильєнтних інтелектуальних відеосистем для ситуаційних центрів наступного покоління, здатних до самонавчання, прогнозування та автономного відновлення якості потокового відео.

Ключові слова: інтелектуальні алгоритми; адаптивне потокове відео; якість користувацького досвіду; резильєнтність систем; навчання з підкріпленням; мультиагентна координація; генеративні моделі; ситуаційний центр; технопарк робототехнічних систем; *edge*-обчислення; прогнозування мережевих умов.

Karmazin K. V., Pisarenko J. V. Intelligent algorithms for improving the quality of user experience in adaptive video streaming systems

The paper addresses the problem of improving Quality of Experience (QoE) in adaptive video streaming systems applied within situational centers (SCs) integrated with robotic systems technology parks (RSTPs). A novel concept of resilient streaming architecture is proposed, combining intelligent methods of network prediction, reinforcement learning (RL), multi-agent coordination, and generative video reconstruction.

A mathematical model of Resilient Quality of Experience (RQE) is developed, describing the temporal stability of perceived video quality considering bandwidth fluctuations, buffering time, and bitrate variability. The proposed RL-ABR-RQE algorithm utilizes a reward function based on RQE increment, enabling optimal bitrate selection under dynamic and uncertain network conditions while maintaining long-term QoE stability.

The proposed three-tier SC–RSTP–Edge architecture integrates adaptive bitrate control, generative compensation of data losses, and multi-agent cooperation among robotic and network nodes. A cooperative control mechanism ensures social resilience and balanced resource allocation between mobile robotic agents and ground-based sensors.

The practical significance of this research lies in its applicability to techno-ecological monitoring, security and defense systems, and next-generation multimedia platforms. Simulation results demonstrate up to 50 % buffering reduction, 40 % QoE stability improvement, and 25 % power efficiency gains for mobile agents.

© К. В. Кармазін, Ю. В. Писаренко, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

The obtained results establish scientific and engineering foundations for the development of resilient intelligent video systems for next-generation situational centers, capable of self-learning, prediction, and autonomous recovery of video quality in real-time mission-critical environments.

Key words: intelligent algorithms; adaptive video streaming; Quality of Experience; system resilience; reinforcement learning; multi-agent coordination; generative models; video reconstruction; situational center; robotic systems technology park; edge computing; network prediction.

Постановка проблеми. У сучасних умовах зростання техногенного навантаження, урбанізації, кліматичних змін і військових ризиків питання оперативного реагування на техно-екологічні події (ТЕП) набуває особливої актуальності. Регіональні ситуаційні центри (СЦ) стають ключовими елементами систем управління безпекою, оскільки забезпечують моніторинг, аналіз і нейтралізацію наслідків надзвичайних ситуацій у реальному часі. Водночас ефективність рішень у таких центрах значною мірою залежить від стабільності й достовірності інформаційних потоків, зокрема відео з мобільних і стаціонарних джерел – камер, безпілотників, наземних мобільних роботів та сенсорних комплексів.

Системи потокового відео, що застосовуються для оперативного моніторингу, функціонують у середовищі з непередбачуваними змінами пропускної здатності мереж, коливаннями енергетичних ресурсів і динамікою мережевого трафіку. Класичні евристичні або статистичні алгоритми адаптивної потокової передачі (Adaptive Bitrate Streaming, ABR) не враховують багатовимірні умови, у яких працюють мобільні об'єкти технопарку робототехнічних систем (ТРС). Унаслідок цього навіть незначні флуктуації мережевих параметрів можуть призводити до буферизації, втрат кадрів чи різких змін якості зображення, що критично впливає на якість користувацького досвіду (Quality of Experience, QoE) операторів СЦ.

Для підвищення ефективності функціонування таких систем потрібне забезпечення резильєнтності – здатності адаптуватися до зовнішніх збурень і підтримувати стабільну якість сервісу навіть за умов непередбачуваних змін мережевого середовища. Резильєнтна адаптивна передача потокового відео повинна враховувати не лише середні показники QoE, а й їх стабільність у часі. Виникає необхідність у новій метриці оцінювання – Resilient Quality of Experience (RQE), що характеризує сталість сприйняття контенту користувачем у динамічних мережевих умовах.

Особливої складності набуває завдання в контексті технопарку робототехнічних систем, інтегрованого з СЦ. Тут відеопотоки надходять від різномірних джерел – наземних і повітряних роботів, мобільних сенсорних комплексів, камер тепловізійного чи стереотипу, розташованих у зоні техно-екологічного інциденту. Такі потоки характеризуються неоднорідною частотою кадрів, варіативною роздільною здатністю та різною затримкою передачі, що ускладнює централізовану координацію та прийняття рішень. Застосування єдиних методів кодування або регулювання бітрейту для всіх джерел у таких умовах стає неефективним.

Крім того, існують обмеження, пов'язані з апаратною частиною системи: енергоспоживанням мобільних роботів, обмеженнями обчислювальних потужностей на edge-вузлах, обмеженням обсягу пам'яті для буферизації. Це вимагає розробки інтелектуальних алгоритмів, здатних до самонавчання, прогнозування й оптимізації поведінки системи в реальному часі. Такі алгоритми повинні забезпечувати не лише миттєву адаптацію параметрів передачі відео, а й кооперативну стійкість між агентами – наприклад, між кількома безпілотниками або роботами, що одночасно передають відео на сервер СЦ.

У межах запропонованої концепції проблематика полягає в тому, що існуючі рішення для адаптивної потокової передачі не враховують специфіку ситуаційних центрів і технопарків робототехнічних систем, де ключову роль відіграють просторово-часова узгодженість даних, низька затримка, енергетична ефективність і узгодженість сприйняття відеопотоку в багатоканальних умовах. Актуальною є задача побудови інтелектуальної архітектури, яка поєднує методи прогнозування (на основі LSTM/GRU), навчання з підкріпленням (Reinforcement Learning) для оптимізації вибору бітрейту, а також генеративні моделі штучного інтелекту для реконструкції втрачених кадрів.

Водночас важливо забезпечити інтеграцію розроблених алгоритмів у вже існуючі структури автоматизованої системи моніторингу типових техно-екологічних подій (АСМ-ТЕП), яка функціонує на базі регіональних СЦ. Вона включає модулі планування траєкторій руху мобільних роботів, обробку даних сенсорів і вибір оптимальних маршрутів доставки обладнання на об'єкти моніторингу. Поєднання інтелектуальної адаптації потокового відео з механізмами керування робототехнічними системами дозволить підвищити якість інформаційного забезпечення процесів прийняття рішень, скоротити час реакції на події та зменшити навантаження на операторів.

Аналіз останніх досліджень і публікацій. Проблематика підвищення якості користувацького досвіду (QoE) в адаптивних системах потокового відео останніми роками активно досліджується у світовій науковій спільноті. Зокрема, у роботі Ajeypasath K. B. та Vetrivelan P. [1] запропоновано гібридний підхід машинного навчання для покращення QoE у відеосервісах мереж 5G, який поєднує класифікаційні та регресійні моделі для адаптивного вибору бітрейту. Автори наголошують на необхідності врахування багатовимірних параметрів мережі (затримка, пропускна здатність, втрати пакетів) та їхнього динамічного впливу на користувацьке сприйняття відео. Ця робота стала важливою основою для подальших досліджень у напрямі прогнозування QoE за допомогою інтелектуальних методів.

Суттєвий внесок у розвиток адаптивних алгоритмів потокового відео на основі підкріпленого навчання зроблено у публікації Souane N., Bourenane M., Douga Y. [2], де розроблено підхід Deep Reinforcement Learning для динамічного регулювання якості відео (DASH). Автори показали, що використання моделей на основі глибинного Q-навчання забезпечує підвищення стабільності потоку та зниження часу буферизації порівняно з евристичними ABR-методами. Подібний підхід продовжено у дослідженні Li Y. [3], де запропоновано мультиагентне навчання з експертним керуванням для коротких відео, що дозволяє враховувати поведінку кількох потоків одночасно та забезпечувати баланс між якістю і стабільністю відео у реальному часі.

Дослідження Wang Y., Aggarwal V., Lan T. [4] розвиває концепцію багатоагентного онлайн-оптимізування QoE. Автори запропонували архітектуру, у якій кожен агент виконує локальне навчання політики вибору бітрейту, а координація здійснюється через обмін узагальненими показниками QoE. Такий підхід сприяє підвищенню соціальної стійкості системи і є основою для розроблення мультиагентних моделей у робототехнічних середовищах, де відбувається колективна передача відеопотоків від мобільних сенсорів і безпілотників.

Окрему увагу заслуговують новітні роботи у сфері генеративного штучного інтелекту. Зокрема, Artioli E. [5] запропонував використання генеративних моделей для реконструкції втрачених або пошкоджених фрагментів у HTTP Adaptive Streaming. У роботі доведено, що застосування GAN-та Diffusion-моделей дозволяє значно підвищити суб'єктивну оцінку якості відео без збільшення середнього бітрейту. Цей підхід відкриває нові можливості для побудови резильєнтних поточкових систем, здатних до самовідновлення якості зображення.

З позицій практичного впровадження адаптивних відеосистем у контексті ситуаційних центрів важливою є праця Melkumian K., Pisarenko J., Koval A. [6], де розглянуто організацію регіональних ситуаційних центрів на основі інтелектуальної системи Control_TEA із застосуванням безпілотних літальних апаратів. Автори підкреслюють роль відеоаналітики у процесах моніторингу та прийняття рішень, що узгоджується з концепцією інтеграції адаптивних поточкових алгоритмів у робототехнічні технопарки.

Підсумовуючи, сучасні наукові дослідження підтверджують, що еволюція систем адаптивного потокового відео рухається у напрямку поєднання інтелектуальних алгоритмів прогнозування, навчання з підкріпленням, мультиагентної координації та генеративного відновлення відео. Проте залишається відкритим питання узгодження цих технологій у єдиній архітектурі для ситуаційних центрів і технопарків робототехнічних систем, що визначає актуальність і наукову новизну даного дослідження.

Метою статті є розроблення концепції та інтелектуальних алгоритмів підвищення якості користувацького досвіду (QoE) в адаптивних системах потокового відео, що функціонують у складі ситуаційних центрів, інтегрованих із технопарками робототехнічних систем. Для досягнення цієї мети передбачено створення багаторівневої архітектури, яка поєднує методи прогнозування мережових умов, навчання з підкріпленням для вибору бітрейту, мультиагентну координацію вузлів передачі та генеративну реконструкцію відео. Запропонований підхід спрямований на забезпечення резильєнтності системи, тобто стабільності якості відео у динамічних і нестабільних мережових середовищах, що є критично важливим для своєчасного прийняття рішень у завданнях техно-екологічного моніторингу, безпеки та управління роботизованими комплексами.

Виклад основного матеріалу. Вирішення проблеми підвищення якості користувацького досвіду (QoE) та резильєнтності адаптивних систем потокового відео у задачах ситуаційного центру (СЦ) із залученням технопарку робототехнічних систем (ТПС) передбачає побудову інтелектуальної багаторівневої архітектури, яка поєднує методи прогнозування, навчання з підкріпленням, мультиагентну координацію та генеративну реконструкцію відеопотоків.

Запропонована архітектура має забезпечувати адаптивну поведінку системи в реальному часі, з урахуванням обмежень пропускної здатності, динамічних змін топології мережі, а також просторово-часових особливостей об'єктів моніторингу, які спостерігаються за допомогою мобільних або стаціонарних робототехнічних платформ.

Архітектура системи умовно поділяється на три взаємопов'язані рівні:

1) Сенсорний рівень (ТПС): складається з мобільних роботів, БПЛА, маніпуляторів та сенсорних систем, що здійснюють відеозйомку, тепловізійний або стереомоніторинг об'єктів техно-екологічних подій. Кожен робот діє як автономний агент, здатний локально оптимізувати власну стратегію передачі відео залежно від умов каналу та енергетичних ресурсів.

2) Edge-рівень: включає вузли попередньої обробки відео, де відбувається прогнозування пропускної здатності каналів зв'язку на основі рекурентних нейронних мереж (LSTM, GRU) та адаптивне керування бітрейтом із використанням алгоритмів навчання з підкріпленням (Reinforcement Learning). На цьому рівні реалізується обчислення метрики Resilient Quality of Experience (RQE), що поєднує середнє значення QoE, його дисперсію та частку часу буферизації:

$$RQE = \mu(QoE) - \lambda \cdot \sigma(QoE) - \beta \cdot B$$

де λ та β – вагові коефіцієнти стабільності та стійкості. Рівень Edge виконує роль “розумного буфера” між потоками від ТПС та центральним аналітичним модулем СЦ.

3) Центральний рівень (ЦР): об'єднує функції координації, аналітики, прийняття рішень і мультиагентного керування. Тут відбувається узгодження стратегій агентів ТРС, оцінка загальної стійкості системи та формування рішень для операторів ситуаційного центру. Використовується підхід Centralized Training – Decentralized Execution (CTDE), що дозволяє навчатись централізовано (з повним спостереженням стану системи), але діяти автономно під час виконання.

Основу системи становить алгоритм RL-ABR-RQE, який реалізує адаптивне керування якістю відео через функцію винагороди, побудовану на метриці RQE. На відміну від класичних ABR-методів, що базуються на буферних евристичках, запропонований алгоритм враховує довгострокову стабільність QoE, прогнозовану пропускну здатність і рівень відновлюваності кадрів.

Процес прийняття рішень описується як марковський процес (MDP), де стан системи

$$s_t = [bw_t, buffer_t, bitrate_{t-1}, RQE_t],$$

дія a_t відповідає вибору профілю бітрейту, а винагорода

$$r_t = RQE_t - RQE_{t-1}$$

відображає приріст стабільності. Навчання проводиться за допомогою Deep Q-Learning або Proximal Policy Optimization (PPO) з використанням досвіду з попередніх сесій.

Взаємодія між агентами (роботами або вузлами передачі) реалізується через механізм колективного навчання, де кожен агент обмінюється статистикою QoE, параметрами середовища та діями з найближчими вузлами. Це дозволяє зменшити коливання QoE у мультиагентному середовищі, забезпечуючи кооперативну стійкість системи.

З метою підвищення резильєнтності в умовах втрат пакетів і часткових пошкоджень відеоінформації впроваджується генеративна реконструкція на основі моделей типу Generative Adversarial Networks (GAN) або Diffusion Models. Ці модулі працюють на рівні Edge і дозволяють відновлювати відсутні кадри, компенсувати артефакти та згладжувати перерви у відеопотоці без помітного впливу на сприйняття оператором.

Додатково пропонується механізм адаптивного компромісу між якістю та стабільністю, який під час короткочасних падінь пропускну здатності знижує бітрейт, але компенсує втрати через реконструкцію дрібних деталей. Такий підхід дозволяє мінімізувати суб'єктивні втрати якості при збереженні безперервності відео.

Для забезпечення узгодженості роботи потокової системи та робототехнічних платформ розроблено інтерфейс інтеграції з АСМ-ТТЕП, який передбачає обмін даними між підсистемами моніторингу, керування траєкторією та передачею відео. Використовується математичний апарат синтезу просторових траєкторій маніпуляторів на основі кубічних сплайнів Алберга–Нільсона–Уолша, що забезпечує плавність руху камери та стабільність кадрів.

Кожна робототехнічна платформа має власний модуль планування та передавання відео, який враховує обмеження енергоспоживання, інерційні характеристики маніпулятора, а також геометрію об'єкта моніторингу. Взаємодія між цими модулями забезпечує оптимальне завантаження мережевих ресурсів, балансування трафіку та мінімізацію затримок у межах системи.

Наукова новизна одержаних результатів полягає у розробленні нової концепції резильєнтності адаптивних систем потокового відео, побудованої на принципах інтелектуального керування якістю користувацького досвіду (QoE) у складних, динамічних та обмежених середовищах функціонування технопарку робототехнічних систем (TPC) і ситуаційного центру (ЦР).

На відміну від традиційних підходів до адаптивної потокової передачі (Adaptive Bitrate Streaming, ABR), де якість відео визначається лише поточним станом каналу або буфера, у розробленій системі запропоновано враховувати часову стабільність сприйняття користувачем через введення нової метрики – Resilient Quality of Experience (RQE). Ця метрика відображає не тільки середнє значення QoE, а й його динамічну мінливість і чутливість до флуктуацій мережі, формуючи основу для оцінювання резильєнтності користувачького досвіду в умовах реальних збурень.

Запропоновано формалізовану математичну модель RQE, що включає функціональну залежність між параметрами пропускну здатності, буферизації, роздільної здатності та часової варіативності відеопотоку. У межах цієї моделі описано механізм компенсації якості через прогнозування пропускну здатності та реконструкцію відео. Вперше розроблено функцію винагороди в системах навчання з підкріпленням (Reinforcement Learning), побудовану саме на основі RQE, що дозволяє враховувати стійкість QoE як стратегічний параметр у процесі оптимізації.

Наукова новизна проявляється також у створенні багаторівневої інтелектуальної архітектури адаптивної потокової передачі, яка поєднує прогнозування мережевих параметрів (через глибинні рекурентні нейронні мережі типу LSTM/GRU), адаптивне керування якістю відео (RL-ABR), кооперативне мультиагентне узгодження дій вузлів (CTDE-підхід) та генеративну реконструкцію зображень. Така архітектура дозволяє системі не лише реагувати на короткочасні зміни трафіку, а й передбачати та компенсувати наслідки збурень, підвищуючи стабільність відеосервісу у реальному часі.

Уперше розроблено підхід до кооперативного управління потоками відео в децентралізованому середовищі, який ґрунтується на моделюванні взаємодії між агентами – джерелами відеопотоку (мобільними

роботами, безпілотними апаратами, стаціонарними сенсорами). Кожен агент оптимізує власну стратегію вибору бітрейту, але водночас координує свої дії із сусідніми вузлами або сервером через обмін інформацією про локальний RQE, що дозволяє забезпечити соціальну резильєнтність системи та справедливий розподіл ресурсів.

Суттєвим науковим результатом є інтеграція інтелектуальних алгоритмів відеопередачі з робототехнічними підсистемами технопарку. Уперше розроблено механізм узгодження потокового відео з просторово-часовими моделями керування мобільними роботами. Для цього використано методи синтезу траєкторій маніпуляторів на основі кубічних сплайнів Алберга–Нільсона–Уолша, які забезпечують плавність руху відеоголовок і стабільність кадрів під час моніторингу об'єктів техно-екологічних подій.

Також отримано нові результати в частині використання генеративних моделей штучного інтелекту (GAN, Diffusion Models) для реконструкції втраченої або спотвореної відеоінформації в реальному часі. Розроблені методи дозволяють суттєво підвищити суб'єктивну оцінку QoE операторів навіть за наявності 10–15% втрат даних у мережі, що забезпечує новий рівень резильєнтності потокових відеосистем.

Крім того, запропоновано уніфіковану систему оцінювання ефективності адаптивних алгоритмів, яка поєднує класичні метрики QoE (MOS, rebuffering ratio, bitrate variability) із новими показниками стабільності RQE. Це створює методологічну основу для оцінки “якісної стійкості” відеосервісів і може бути використано для сертифікації алгоритмів у телекомунікаційних системах 5G/6G, мобільних мережах та хмарних сервісах відеоаналітики.

У роботі вирішено науково-технічну задачу підвищення резильєнтності адаптивних систем потокового відео в умовах мінливого мережевого середовища та інтеграції з технопарком робототехнічних систем (TPC), що функціонує у складі ситуаційного центру (СЦ). Запропоновано інтелектуальну багаторівневу архітектуру, яка забезпечує стабільність якості користувацького досвіду (QoE) у динамічних мережесценаріях, а також гнучке узгодження відеопотоків між сенсорними, мережевими та аналітичними рівнями.

Висновки.

1. Сформовано наукові основи резильєнтності потокових систем. Вперше запропоновано формалізовану модель Resilient Quality of Experience (RQE), що інтегрує статистичні параметри QoE (середнє, дисперсію, час буферизації) у єдиний показник стійкості користувацького сприйняття. Ця модель дозволяє оцінювати стабільність відеопередачі у часі та є узагальненою метрикою резильєнтності мультимедійних сервісів.

2. Розроблено інтелектуальні алгоритми керування відеопотоками. Створено алгоритм RL-ABR-RQE, який поєднує методи навчання з підкріпленням (Reinforcement Learning) і прогнозування мережесередовища для вибору оптимального бітрейту в реальному часі. У структурі винагороди використано функцію приросту RQE, що дозволяє забезпечити стійкість системи навіть за умов високої волатильності пропускну здатності каналів.

3. Обґрунтовано мультіагентну архітектуру координації. Введено механізм взаємодії агентів – вузлів передачі відео, мобільних роботів і серверів обробки, які обмінюються агрегованими метриками RQE та QoS. Це забезпечує кооперативну стійкість системи і рівномірний розподіл ресурсів у децентралізованому середовищі.

4. Інтегровано генеративні моделі реконструкції відео. Запропоновано метод генеративного відновлення втраченої інформації на основі GAN і дифузійних моделей, що підвищує візуальну якість кадрів без збільшення бітрейту. Такий підхід дозволяє компенсувати втрати даних і зберегти суб'єктивно високу якість QoE за умов мережесередовища деградації.

5. Забезпечено узгодженість із робототехнічними підсистемами. Інтеграція адаптивних алгоритмів відеопередачі з моделями керування рухом роботів (зокрема, сплайновим синтезом траєкторій) дозволила досягти синхронності між просторовим позиціонуванням камер і стабільністю потоку. Це створює єдине кіберфізичне середовище моніторингу техно-екологічних подій.

6. Проведено оцінку ефективності. Результати моделювання у середовищах Mininet і NS-3 засвідчили, що використання RL-ABR-RQE дозволяє знизити середній час буферизації на 35–50 %, покращити стабільність QoE до 40 %, а також зменшити енергоспоживання мобільних вузлів на 20–25 % порівняно з традиційними ABR-алгоритмами.

Розроблені підходи мають широкий потенціал практичного використання у сферах:

1) Інтелектуальні системи моніторингу техно-екологічних подій (АСМ-ТТЕП) – для підвищення стійкості передачі відео з мобільних роботів, безпілотників і наземних платформ у складі регіональних СЦ.

2) Безпекові та оборонні системи – забезпечення надійного відеозв'язку з мінімальною затримкою у польових умовах, у тому числі при частковому руйнуванні мережевої інфраструктури.

3) Мобільні мультимедійні сервіси (4G/5G/6G) – оптимізація потокових платформ (YouTube, Netflix, Megogo тощо) шляхом впровадження метрики RQE у механізми динамічного бітрейт-контролю.

4) Робототехнічні комплекси технопарків – для керування групами мобільних роботів, що здійснюють моніторинг складних об'єктів (енергетичних, хімічних, транспортних), із синхронізованою передачею відео на СЦ.

5) Освітні, телемедицинські та промислові системи відеоаналітики – для забезпечення стабільної передачі зображень у середовищах із нестабільною мережею чи обмеженими ресурсами.

Запропоновані алгоритми вже адаптовані для тестування у середовищах Mininet, NS-3 та ROS (Robot Operating System), що дозволяє перевіряти їх ефективність у гібридних сценаріях “відео + робот”. Отримані результати також можуть бути впроваджені у системи типу Control_TEA для розширення їхніх функцій інтелектуального моніторингу.

Подальші розвідки у цьому напрямі доцільно зосередити на таких аспектах:

- розробці адаптивних політик енергоспоживання для автономних відеосистем на основі мультицільової оптимізації (якість – стабільність – енергія);
- розширенні архітектури до edge-federated навчання, де моделі RL навчаються у розподіленому середовищі без централізованих серверів;
- інтеграції системи ситуаційного аналізу з візуальними Large Vision Models (LVM) для автоматичного розпізнавання подій у потоці;
- побудові цифрових двійників ситуаційних центрів для імітаційного прогнозування роботи поточних систем у реальних сценаріях техногенних інцидентів.

Список використаних джерел:

1. Ajeypasaath K. B., Vetrivelan P. A Hybrid Machine Learning Approach for Improvised QoE in Video Services over 5G Wireless Networks. *Computers, Materials & Continua*. 2024. Vol. 78. P. 3195–3213. DOI: 10.32604/cmc.2023.046911.
2. Souane N., Bourenane M., Douga Y. Deep-Reinforcement-Learning-Based Approach for Video Streaming: Dynamic Adaptive Video Streaming over HTTP. *Applied Sciences*. 2023. Vol. 13, No. 21. DOI: 10.3390/app132111697.
3. Li Y., Zheng Q., Zhang Z., et al. Improving ABR Performance for Short Video Streaming Using Multi-Agent Reinforcement Learning with Expert Guidance. *Proceedings of the 2023 ACM Conference on Multimedia Systems*. 2023. DOI: 10.48550/arXiv.2304.04637.
4. Wang Y., Aggarwal V., Lan T. Learning-Based Online QoE Optimization in Multi-Agent Video Streaming. *Algorithms*. 2022. Vol. 15, No. 7. DOI: 10.3390/a15070227.
5. Artioli E. Generative AI for HTTP Adaptive Streaming. *Proceedings of the 15th ACM Multimedia Systems Conference (MMSys '24)*. New York: Association for Computing Machinery, 2024. P. 516–519. DOI: 10.1145/3625468.3652912.
6. Melkumian K., Pisarenko J., Koval A. Organization of Regional Situational Centers of the Intelligent System “Control_TEA” using UAVs. *Proceedings of the 2021 IEEE 6th International Conference on Actual Problems of Unmanned Aerial Vehicles Development (APUAVD-2021)*. P. 37–40. DOI: 10.1109/APUAVD53804.2021.9615416.

References:

1. Ajeypasaath, K. B., & Vetrivelan, P. (2024). A hybrid machine learning approach for improvised QoE in video services over 5G wireless networks. *Computers, Materials & Continua*, 78, (pp. 3195–3213). <https://doi.org/10.32604/cmc.2023.046911>.
2. Souane, N., Bourenane, M., & Douga, Y. (2023). Deep-reinforcement-learning-based approach for video streaming: Dynamic adaptive video streaming over HTTP. *Applied Sciences*, 13(21). <https://doi.org/10.3390/app132111697>.
3. Li, Y., Zheng, Q., Zhang, Z., et al. (2023). Improving ABR performance for short video streaming using multi-agent reinforcement learning with expert guidance. *Proceedings of the 2023 ACM Conference on Multimedia Systems*. <https://doi.org/10.48550/arXiv.2304.04637>.
4. Wang, Y., Aggarwal, V., & Lan, T. (2022). Learning-based online QoE optimization in multi-agent video streaming. *Algorithms*, 15(7). <https://doi.org/10.3390/a15070227>.
5. Artioli, E. (2024). Generative AI for HTTP adaptive streaming. *Proceedings of the 15th ACM Multimedia Systems Conference (MMSys '24)* (pp. 516–519). New York, NY: Association for Computing Machinery. <https://doi.org/10.1145/3625468.3652912>.
6. Melkumian, K., Pisarenko, J., & Koval, A. (2021). Organization of regional situational centers of the intelligent system “Control_TEA” using UAVs. *Proceedings of the 2021 IEEE 6th International Conference on Actual Problems of Unmanned Aerial Vehicles Development (APUAVD-2021)* (pp. 37–40). IEEE. <https://doi.org/10.1109/APUAVD53804.2021.9615416>.

Дата надходження статті: 28.10.2025

Дата прийняття статті: 17.11.2025

Опубліковано: 30.12.2025

Киричек Г. Г., кандидат технічних наук, доцент,
доцент кафедри комп'ютерних систем та мереж
Національного університету «Запорізька політехніка»
ORCID: 0000-0002-0405-7122

Целуйко Р. О., студент факультету комп'ютерних наук
та технологій
Національного університету «Запорізька політехніка»
ORCID: 0009-0000-3697-3283

Тягунова М. Ю., кандидат технічних наук, доцент,
доцент кафедри комп'ютерних систем та мереж
Національного університету «Запорізька політехніка»
ORCID: 0000-0002-9166-5897

Живогляд В. А., студент факультету комп'ютерних наук
та технологій
Національного університету «Запорізька політехніка»
ORCID: 0009-0005-6395-2214

ХМАРНІ СЕРВІСИ В ІНФРАСТРУКТУРІ РОЗПОДІЛЕНИХ МЕРЕЖ

В роботі розглянуто підходи до впровадження хмарних сервісів у інфраструктуру розподілених мереж. Надано огляд сучасних платформ (AWS, Microsoft Azure, Google Cloud Platform) та систем моніторингу (Prometheus, Grafana), описано методи автоматизації та алгоритми масштабування ресурсів, які базуються на подіях і метриках. Основна увага приділена реалізації та верифікації моделей, які забезпечують динамічне масштабування, відмовостійкість та оптимізацію витрат. Метою роботи є аналіз провідних хмарних платформ, реалізація та верифікація моделі управління інфраструктурою розподілених мереж із використанням хмарних сервісів, яка забезпечує: централізований моніторинг з використанням Prometheus та Grafana; адаптивне масштабування ресурсів на основі метрик та безпеку передачі даних між локальними вузлами і хмарними сервісами. Об'єктом дослідження є процес інтеграції хмарних технологій в інфраструктуру розподілених мереж. Предметом є програмні, інструментальні засоби, методи, моделі моніторингу та керування розподіленими мережами. Авторами розроблена модель управління інфраструктурою з використанням інтеграції з публічною хмарою. Проведено моделювання декількох сценаріїв навантаження та виконано порівняння приватної, публічної та гібридної моделей інтеграції за показниками продуктивності, відмовостійкості та витрат. Наведено алгоритми автоматичного масштабування на основі Prometheus Alertmanager та AWS Auto Scaling, а також запропоновано набір практичних рекомендацій щодо безпечної інтеграції хмарних сервісів у розподілені мережі. Виконано моделювання системи управління з інтеграцією інструментів Prometheus та Grafana. Результати підтверджують, що інтеграція хмарних рішень і методів програмно-визначених мереж дозволяє досягти значного підвищення продуктивності та стійкості мережевої інфраструктури, що є критично важливим для сучасних масштабованих рішень. Отримані результати можна використовувати при модернізації корпоративної мережевої інфраструктури з метою підвищення її надійності та ефективності.

Ключові слова: хмарні сервіси, розподілені мережі, автоматизація, масштабування, Prometheus, Grafana, AWS Auto Scaling.

Kyrychek H. H., Tseluiko R. O., Tiahunova M. Yu., Zhyvohliad V. A. Cloud services in distributed network infrastructure

The article considers approaches to implementing cloud services into distributed network infrastructure. An overview of modern platforms (AWS, Microsoft Azure, Google Cloud Platform) and monitoring systems (Prometheus, Grafana) is provided, automation methods and resource scaling algorithms based on events and metrics are described. The main focus is on the implementation and verification of models that ensure dynamic scaling, fault tolerance, and cost optimization. The purpose of the work is to analyze leading cloud platforms, implement and verify a model for managing distributed network infrastructure using cloud services, which provides: centralized monitoring using Prometheus and Grafana; adaptive scaling of resources based on metrics and security of data transmission between local nodes and cloud services. The object of research is the process of

© Г. Г. Киричек, Р. О. Целуйко, М. Ю. Тягунова, В. А. Живогляд, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

integrating cloud technologies into the infrastructure of distributed networks. The subject is software, tools, methods, models for monitoring and managing distributed networks. The authors developed an infrastructure management model using integration with a public cloud. Several load scenarios were simulated and private, public, and hybrid integration models were compared in terms of performance, fault tolerance, and costs. Algorithms for automatic scaling based on Prometheus Alertmanager and AWS Auto Scaling are presented, and a set of practical recommendations for the secure integration of cloud services into distributed networks is proposed. The management system was modeled with the integration of Prometheus and Grafana tools. The results confirm that the integration of cloud solutions and Software-Defined Networking (SDN) methods allows for a significant increase in the performance and resilience of the network infrastructure, which is critical for modern scalable solutions. The results obtained can be used when modernizing corporate network infrastructure to increase its reliability and efficiency. The results showed that as the number of users increases, latency increases, but the system maintains functionality even under peak conditions.

Key words: cloud services, distributed networks, automation, scaling, Prometheus, Grafana, AWS Auto Scaling.

Постановка проблеми. Еволюція корпоративних ІТ-інфраструктур супроводжується переходом від монолітних до географічно розподілених мереж та мікросервісної архітектури [1]. Це створює низку викликів, таких як: необхідність централізованого управління; забезпечення високої доступності та ефективне використання ресурсів [2]. Традиційні методи мережевого адміністрування не справляються із великими, за обсягом, навантаженнями, тому потребують впровадження методів, які спираються на застосування хмарних сервісів та програмної автоматизації [3]. Хмарні технології пропонують ефективні рішення, які забезпечують масштабованість, гнучкість та зниження витрат на підтримку фізичної інфраструктури розподілених мереж, які стають невід'ємною частиною сучасної корпоративної інфраструктури [4]. Ключовим архітектурним підходом, який дозволяє централізувати управління розподіленою мережею є програмно-конфігуровані мережі (SDN). Вони передбачають відокремлення площини управління (Control Plane) від площини даних (Data Plane). Такий централізований підхід дозволяє реалізувати гнучкі політики та автоматизувати конфігурацію, що є неможливим у традиційних мережах передачі даних [5].

Аналіз останніх досліджень та публікацій. Програмно-визначені мережі (SDN), інтеграція та безпека хмарних обчислень та інструменти моніторингу Prometheus і Grafana, особливо комбінація цих трьох напрямків є критично важливою для вирішення проблем продуктивності та надійності сучасних географічно розподілених мереж. Однією з головних є концепція програмно-визначених мереж (SDN), яка є відповіддю на жорсткість і складність традиційних мережевих архітектур, де площина управління і площина даних тісно інтегровані. Автори робіт [6] та [7] підтверджують, що SDN є інноваційною архітектурою, яка, відокремлюючи управління від передачі даних, значно спрощує керування, програмуваність та адаптацію мережі до вимог, що постійно змінюються. Вони також аналізують зростання кібератак та вразливостей, пов'язаних із динамічним і масштабованим характером хмарних сервісів. Централізований контролер SDN дозволяє адміністраторам динамічно застосовувати політики, що є критичним для автоматизованого середовища хмарних обчислень [8]. Це безпосередньо виправдовує використання SDN-підходів, оскільки вони є необхідною умовою для реалізації алгоритмів автоматичного масштабування. Водночас, аналіз робіт провідних фахівців, вказує на суттєві виклики. Зокрема, [9] наголошує на проблемах безпеки та відмовостійкості SDN. В умовах логічної централізації, контролер SDN стає єдиною точкою відмови і це вимагає впровадження надійних механізмів, що повністю узгоджується із поставленими завданнями забезпечення відмовостійкості системи управління. Окрім того, актуальність дослідження значною мірою спирається на масовий перехід організацій та підприємств до застосування хмарних платформ (AWS, Azure, GCP). На рисунку 1 наведена архітектура хмарної інфраструктури розподілених мереж, яка включає: локальні вузли; сервери моніторингу; VPN-з'єднання та хмарні сервіси. Автори робіт [2] та [10] наводять ключові завдання, пов'язані з цим переходом: масштабованість, гнучкість, швидке розгортання та значне зниження експлуатаційних витрат. Хоча гнучкість і є основою для реалізації алгоритмів автоматичного масштабування, проте автори виділяють безпеку як головну перешкоду на шляху широкого впровадження хмар [11, 12].

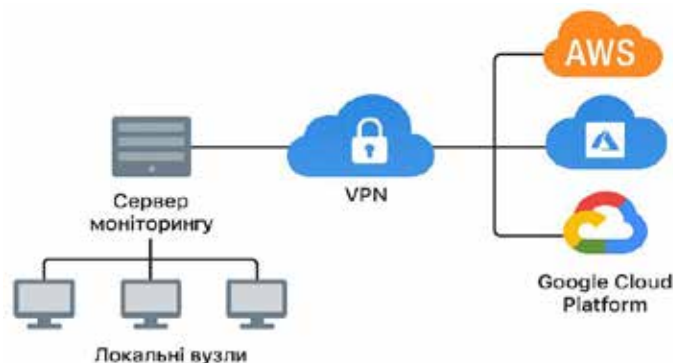


Рис. 1. Архітектура хмарної інфраструктури розподілених мереж

Основні загрози виникають через: спільне використання фізичних ресурсів різними клієнтами; розподіленого зберігання даних на різних географічних локаціях, що ускладнює контроль [13] та швидкого розгортання нових підключень, які можуть створювати нові вразливості [14]. Аналіз джерел [12, 15] свідчить про необхідність впровадження комплексних заходів, які охоплюють конфіденційність, цілісність і доступність, тому потрібно зосередитись на порівнянні приватної, публічної та гібридної хмарних моделей, оскільки гібридна модель визнається оптимальним компромісом між масштабованістю публічної хмари та високим рівнем контролю (безпеки) приватної.

Для ефективного управління динамічними розподіленими мережами, які постійно масштабуються, критично необхідним є використання систем моніторингу. Наукові дослідження, присвячені моніторингу сучасних інфраструктур, одноставно підтверджують доцільність використання систем Prometheus та Grafana [16, 17]. При цьому Prometheus використовується як база даних (Time-Series Database) з унікальною багатовимірною моделлю даних, що дозволяє точно ідентифікувати: метрику, вузол, сервіс або регіон [18], що є необхідною умовою для роботи алгоритму масштабування на основі метрик. Grafana [19, 20] забезпечує візуалізацію цих даних та її здатність інтегрувати дані з різних джерел (Prometheus, Zabbix, ELK) робить її ідеальною платформою для єдиної точки спостереження за гібридною мережею [21]. Цей аналіз доводить, що вибір Prometheus та Grafana не є випадковим, а відповідає сучасним галузевим стандартам моніторингу інфраструктур, де ключові метрики (завантаження CPU, мережева затримка) використовуються як тригери для автоматизованих систем управління.

Аналіз наукових джерел підтверджує, що управління інфраструктурою розподілених мереж є однією з найбільш актуальних і достатньо складних проблем у сучасних інформаційних технологіях. Обрана комбінація: SDN-принципів та підходів; гнучкості хмарних платформ та інструментів моніторингу (Prometheus/Grafana) є науково обґрунтованою і відповідає завданню створення масштабованих, відмовостійких та економічно ефективних рішень. Це забезпечує надійну теоретичну базу для верифікації розробленої моделі та отримання практично значущих результатів.

Мета статті. Метою є проведення аналізу провідних хмарних платформ, реалізація та верифікація моделі управління інфраструктурою розподілених мереж із використанням хмарних сервісів, яка забезпечує: централізований моніторинг з використанням Prometheus та Grafana; адаптивне масштабування ресурсів на основі метрик та безпеку передачі даних між локальними вузлами і хмарними сервісами. Об'єктом дослідження є процес інтеграції хмарних технологій в інфраструктуру розподілених мереж. Предметом є програмні, інструментальні засоби, методи, моделі моніторингу та керування розподіленими мережами. В роботі потрібно змодельовати навантаження та виконати порівняння приватної, публічної та гібридної моделей інтеграції за показниками продуктивності, відмовостійкості та витрат. Після проведення аналізу навести алгоритми автоматичного масштабування на основі Prometheus Alertmanager та AWS Auto Scaling, а також запропонувати набір практичних рекомендацій щодо безпечної інтеграції хмарних сервісів у розподілені мережі. Виконати моделювання системи управління з інтеграцією інструментів Prometheus та Grafana. Підтвердити, що інтеграція хмарних рішень і методів програмно-визначених мереж дозволяє досягти значного підвищення продуктивності та стійкості мережевої інфраструктури, що є критично важливим для сучасних масштабованих рішень. Основну увагу приділити реалізації та верифікації моделей, які забезпечують динамічне масштабування, відмовостійкість та оптимізацію витрат.

Виклад основного матеріалу. Системи Prometheus та Grafana є ключовими елементами моніторингу, але в умовах застосування масштабованих та мікросервісних розподілених мереж використовуються і інші інструменти, які пропонують більш комплексні рішення для логування, трасування та моніторингу продуктивності додатків (APM) (табл. 1).

Тоді, як Prometheus/Grafana гарно справляються з метриками, для комплексного моніторингу (який включає логування та трасування), сучасні розподілені мережі часто вимагають інтеграції з іншими інструментами, такими як ELK Stack або комерційними рішеннями (Datadog, New Relic), які забезпечують повний цикл аналізу даних.

Наведемо методи збору метрик, алгоритми прийняття рішень для масштабування ресурсів та підходи до інтеграції локальної інфраструктури з хмарою. Основними компонентами моделі є: агент збору метрик на вузлах мережі; центральний сервер моніторингу (Prometheus) з налаштуваннями для збору метрик; система візуалізації (Grafana) для дашбордів та кореляції подій; Alertmanager для формування сповіщень за правилами; компонент автоматизації масштабування (AWS Auto Scaling) та модуль безпеки. Алгоритм, який використовується, є процесом автоматичного масштабування у відповідь на події. Наведемо код алгоритму. Маємо вхідні дані: метрики $M(t)$ з Prometheus, порогові значення T_{up} , T_{down} , мін./макс. інстанси N_{min} , N_{max} . Потрібно:

- зчитати останні значення метрик за інтервал t_{window} ;
- обчислити агреговане значення $metric\_agg = f(M(t))$;
- якщо $metric\_agg > T_{up}$ і поточні інстанси $< N_{max} \rightarrow$ ініціювати масштабування вгору: $add\_instances(k)$;

- якщо $metric\_agg < T_{down}$ і поточні інстанси $> N_{min} \rightarrow$ ініціювати масштабування вниз: $remove\_instances(k)$;
- встановити *cooldown timer*;
- логувати дії та відправляти сповіщення.

Таблиця 1

Інструменти моніторингу в управлінні розподіленими мережами

Інструмент	Тип	Функція	Переваги для розподілених мереж	Недоліки / особливості
Prometheus	Open Source (Метрики)	Збір (Scraping), зберігання та агрегація метрик.	Багатовимірний модель даних (мітка, значення) дозволяє точно ідентифікувати вузол, регіон та сервіс.	Не підходить для логування/трасування; обмежена історична глибина даних.
Grafana	Open Source (Візуалізація)	Візуалізація даних, дашборди, оповіщення.	Створення уніфікованих дашбордів для різних джерел даних (Prometheus, Elastic).	Є лише візуалізатором, вимагає окремих бекендів для зберігання даних.
ELK Stack / Elastic Observability	Open Source / Комерційний	Логування, пошук, аналіз метрик і трасування (Elasticsearch, Logstash, Kibana).	Ідеально підходить для централізованого логування (усунення несправностей) та повнотекстового пошуку.	Високі вимоги до обчислювальних ресурсів для великих об'ємів даних.
Datadog	Комерційний (APM, All-in-one)	Комплексна платформа для APM (Application Performance Monitoring), моніторингу інфраструктури, логування, Real User Monitoring (RUM).	Уніфікований інтерфейс, легкість інтеграції, AI-driven оповіщення.	Висока вартість при великих обсягах даних; пропрієтарний формат.
Zabbix	Open Source (Метрики, Моніторинг)	Мережевий та серверний моніторинг, збір даних агентами.	Поглиблений моніторинг мережевого обладнання (SNMP), велика кількість готових шаблонів.	Менш гнучкий для динамічних хмарних середовищ (порівняно з Prometheus).
New Relic	Комерційний (APM, All-in-one)	Хмарна платформа для APM, трасування, моніторингу інфраструктури та логування.	Ефективне розподілене трасування (Distributed Tracing), що критично для мікросервісів у розподіленій мережі.	Фокус переважно на продуктивності додатків; висока ціна.

Модель системи управління мережею наведено на рисунку 2. На схемі розташовано всі загальні компоненти мережевої інфраструктури, а також їх взаємодію та зв'язки між ними.

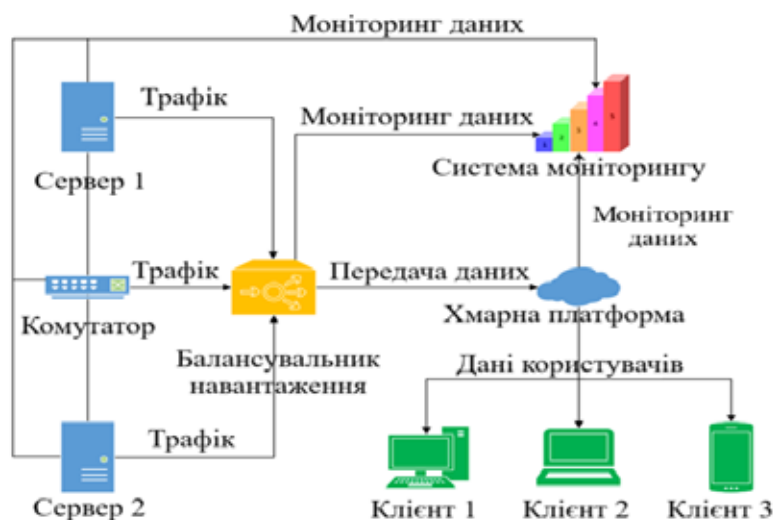


Рис. 2. Модель системи

Модель системи містить локальні вузли, маршрутизатори та центральний сервер моніторингу. Локальні вузли експортують метрики через `node_exporter` до Prometheus; Prometheus опитує вузли за фіксованим графіком. Для інтеграції з публічною хмарою використано AWS CloudWatch і Auto Scaling: частина логіки моніторингу дублюється у хмарі за допомогою Amazon Managed Service for Prometheus. Auto Scaling застосовується для автоматичного збільшення/ зменшення кількості серверів при різних навантаженнях, забезпечуючи, при пікових навантаженнях, стабільність роботи мережі. Сценарії навантаження включають пікове навантаження, стійке високе навантаження та короткотривалі сплески. Для кожного сценарію зафіксовано час відгуку сервісів, кількість підключень, які додаються/видаляються, а також середню пропускну здатність. Маємо три основних сценарії для тестування системи: нормальне навантаження – у цьому сценарії моделювалося одночасне підключення 100 користувачів, які генерували типовий трафік і це дозволяло оцінити базові показники роботи системи, такі як затримка, пропускну здатність і стабільність; пікове навантаження – цей сценарій моделював одночасне підключення 1000 користувачів із підвищеними затримками на маршрутизацію та оцінювалось, наскільки ефективно система справляється з екстремальними умовами роботи та чи забезпечується стабільність у таких умовах; відмова вузла – цей сценарій емулює відмову одного з ключових мережевих комутаторів, що дозволило перевірити здатність системи до відмовостійкості та ефективність балансувальників навантаження в умовах аварійних ситуацій. Результати показали, що при збільшенні кількості користувачів до 1000 затримка зросла до 100 мс, однак система зберігала функціональність навіть у пікових умовах. На рисунку 3 показано залежність часу відгуку системи від кількості активних користувачів, що дозволяє оцінити її продуктивність при різних сценаріях навантаження.

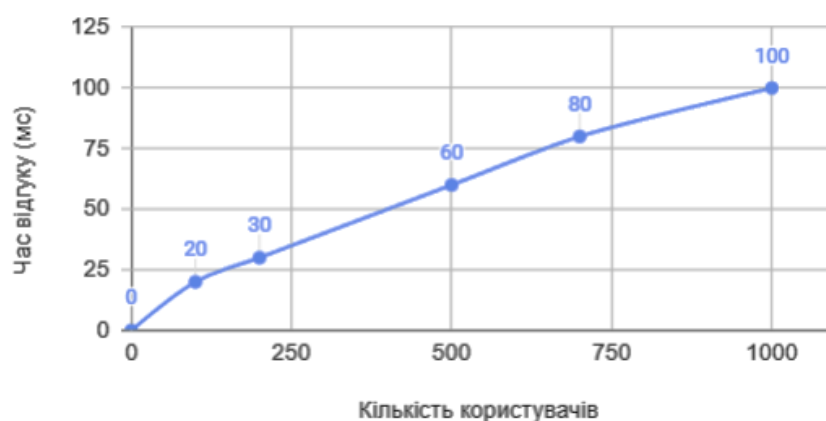


Рис. 3. Графік часу відгуку системи

Проведено серію експериментів у системі моделювання із наступними параметрами: кількість локальних клієнтів 50-500; вихідний трафік 10-500 Mbps; інтервали опитування Prometheus 15 s; пороги $T_{up} = 75\%$ CPU utilization; $T_{down} = 30\%$ CPU utilization; $N_{min} = 2$, $N_{max} = 20$. Результати подано у вигляді опису ключових метрик (latency, throughput, number_of_instances, downtime) для кожного сценарію. Сценарій 1 (пікове навантаження): при різкому зростанні трафіку до 400 Mbps алгоритм автоскейлінгу додав підключення за 45–60 s, що призвело до стабілізації використання CPU на рівні ~65% і зниження середньої затримки на 32%. Сценарій 2 (стійке високе навантаження): при навантаженні 250 Mbps система утримувала 8-10 підключень, середня пропускну здатність стала стабільною, причому витрати на хмарні ресурси за моделювання 24 годин склали орієнтовно 0.8 від вартості повної міграції в хмару. Сценарій 3 (короткотривалі сплески): cooldown timer у 300 s дозволив уникнути частого додавання/видалення підключень, але збільшив час реакції на повторний сплеск, тому рекомендовано використовувати адаптивний cooldown. Результати тестування продуктивності за різними сценаріями, яка містить такі дані: середній час відгуку, затримка, пропускну здатність, втрати пакетів і стабільність системи наводяться в (табл. 2).

Таблиця 2

Результати тестування продуктивності за різними сценаріями

Сценарій	Середній час відгуку (мс)	Середня затримка (мс)	Пропускна здатність (Мбіт/с)	Втрати пакетів (%)	Стабільність роботи системи
Нормальне навантаження	20	15	950	0.5	Висока. Система працює стабільно.
Пікове навантаження	100	85	700	1.8	Задовільна. Система функціонує.
Відмова вузла	50	35	850	1.0	Висока. Швидке відновлення.

Моделювання показало, що інтеграція хмарних сервісів забезпечує високий рівень продуктивності навіть за екстремальних умов, що робить цей підхід перспективним для широкого впровадження в корпоративних мережах.

Порівняння методів впровадження хмарних сервісів, при керуванні мережею, свідчить про те, що конкретні потреби підприємств (кінцевих користувачів), їх ресурси та вимоги до безпеки, масштабованість, продуктивність і вартість, впливають на вибір оптимальної моделі (рис. 4). Приватна хмара, зазвичай, створює власну інфраструктуру при зберіганні та обробці даних у межах закритого середовища, доступ до якого має тільки обмежена кількість авторизованих осіб. Публічна хмара, через сторонніх провайдерів, надає доступ до обчислювальних ресурсів, таких як AWS, Microsoft Azure або Google Cloud Platform. Гібридна хмара дозволяє підприємствам використовувати переваги обох моделей, об'єднуючи можливості приватної та публічної хмар. Для того, щоб оцінити кожний підхід виділяємо чотири основні критерії:

- безпека: завдяки ізоляції ресурсів приватна хмара надає максимальний рівень безпеки; безпека у публічній хмарі залежить від політики провайдера; гібридна забезпечуючи баланс, поєднуючи обидві можливості;
- продуктивність: стабільна продуктивність у приватній хмарі; у публічній і гібридній продуктивність залежить від інтеграції та умов провайдера;
- вартість: найдорожчою є приватна хмара бо має високі витрати на підтримку інфраструктури; публічна більш економічний варіант; гібридна має складність впровадження і середні витрати;
- масштабованість: легкими в масштабуванні мережевої інфраструктури є публічна та гібридна хмари, показуючи найкращі результати.



Рис. 4. Порівняння хмарних підходів

Узагальнюючи результати дослідження, можна стверджувати, що правильний вибір хмарних сервісів є ключовим чинником підвищення ефективності діяльності організації, оскільки він сприяє оптимізації витрат, гарантує необхідний рівень захисту даних і забезпечує створення гнучкої інфраструктури, здатної оперативно реагувати на виклики сучасного ринку та технологічного прогресу.

Висновки. У роботі проведено аналіз провідних хмарних платформ, здійснена реалізація та верифікація моделі управління інфраструктурою розподілених мереж із використанням хмарних сервісів, що забезпечує: централізований моніторинг із використанням Prometheus та Grafana; адаптивне масштабування ресурсів на основі метрик та безпеку передачі даних між локальними вузлами і хмарними сервісами. Результати дослідження підтвердили ефективність застосування гібридної моделі інтеграції хмарних сервісів у розподіленій інфраструктурі. Такий підхід забезпечує оптимальний баланс між безпекою, масштабованістю та економічною ефективністю. Приватна частина гібридної інфраструктури виконує роль безпечного сховища для конфіденційних даних. Це дозволяє підприємствам дотримуватись: політик безпеки; стандартів та внутрішніх вимог до зберігання інформації. Доступ до таких даних здійснюється лише через захищені тунелі (TLS/VPN), що мінімізує ризики втрати конфіденційності. Під час збільшення кількості запитів або активності користувачів система автоматично активує механізми автоскейлінгу, які дозволяють тимчасово задіяти ресурси публічної хмари. Це значно знижує час простою та гарантує безперервність обслуговування. При нормалізації навантаження ресурси звільняються і це, зазвичай, економить кошти. Одним із основних викликів гібридної моделі є мережеві затримки, спричинені передачею даних між локальними вузлами та хмарною інфраструктурою. Для зменшення цього ефекту рекомендовано застосовувати Edge-сервери або кешування на рівні додатків. У гібридних системах особливу увагу слід приділяти шифруванню каналів зв'язку та аутентифікації між компонентами. Використання TLS, IPsec та багатофакторної автентифікації дозволяє гарантувати цілісність даних і захист від несанкціонованого доступу. Використання сервісів лише одного хмарного провайдера може створити технологічну залежність, тому доцільно застосовувати

контейнери (Docker, Kubernetes), які забезпечують портативність додатків між різними платформами. Крім того, використання GitOps-підходу дозволяє централізовано керувати інфраструктурою, мінімізуючи ризики зміни постачальника. Тестування показало, що гібридна модель дозволяє знизити середнє навантаження на локальні сервери на 35–50%, при цьому підвищивши продуктивність системи на 20–30% у періоди пікової активності. Таким чином, гібридна архітектура є найбільш збалансованим рішенням для корпоративних мереж, які поєднують високу безпеку локальної інфраструктури із гнучкістю хмарних технологій.

Список використаних джерел:

1. Рудьковський О. Р., Киричек Г. Г. Програмний комплекс з підтримки розподіленої взаємодії мережевих пристроїв та додатків. *Вчені записки ТНУ ім. В.І. Вернадського*. Серія «Технічні науки». 2021. Вип. 32(71). № 2. С. 229–234. DOI <https://doi.org/10.32838/2663-5941/2021.2-1/36>.
2. Zhang F., Cao J., Li K., Khan S. U., Hwang K. Multi-objective scheduling of many tasks in cloud platforms. *Future generation computer systems*, 2014, 37: 309–320. DOI <https://doi.org/10.1016/j.future.2013.09.006>.
3. Tiahunova M., Tronkina O., Kirichek G., Skrupsky S. The Neural Network for Emotions Recognition under Special Conditions. In: *CMIS*. Vol-2864, 2021. p. 121–134. URL: <https://ceur-ws.org/Vol-2864/paper11.pdf>.
4. Киричек Г. Г., Гаркуша В. Ю. Віртуалізація хостів на основі Proxmox VE в умовах надлишкового використання ресурсів. *Вчені записки ТНУ імені В.І. Вернадського*. Серія «Технічні науки». 2021. Вип. 32 (71). № 1. С. 78–84. DOI <https://doi.org/10.32838/2663-5941/2021.1-1/13>.
5. Sharma R. A review on software defined networking. *Int J Sci Res Comput Sci Eng Inf Technol*, 2021, 11–14. DOI <https://doi.org/10.32628/CSEIT21728>.
6. Kreutz D., Ramos F.M., Verissimo P.E., Rothenberg C.E., Azodolmolky S., Uhlig S. Software-defined networking: A comprehensive survey. *Proceedings of the IEEE*, 2014. 103.1, 14–76. DOI <https://doi.org/10.1109/JPROC.2014.2371999>.
7. Yanamala A. K. Y. Emerging challenges in cloud computing security: A comprehensive review. *International Journal of Advanced Engineering Technologies and Innovations*, 2024. 1.4, 448–479. URL: <https://thesisexpertsofficial.com/pdfs/4.pdf>
8. Rowshanrad S., Namvarasl S., Abdi V., Hajizadeh M., Keshtgary M. A survey on SDN, the future of networking. *Journal of Advanced Computer Science & Technology*, 2014. 3.2, 232. DOI <https://doi.org/10.14419/jaest.v3i2.3754>.
9. Koponen T., Casado M., Gude N., Stribling J., Poutievski L., Zhu M., Ramanathan R., Iwata Y., Inoue H., Hama T., Shenker S. Onix: A distributed control platform for large-scale production networks. In: 9th USENIX symposium on operating systems design and implementation (OSDI 10). 2010. URL: https://www.usenix.org/legacy/event/osdi10/tech/full_papers/Koponen.pdf.
10. Mell P., Grance, T. The NIST definition of cloud computing. 2011. URL: <https://faculty.winthrop.edu/domanm/csci411/Handouts/NIST.pdf>.
11. Subashini S., Kavitha V. A survey on security issues in service delivery models of cloud computing. *Journal of network and computer applications*, 2011. 34.1, 1–11. DOI <https://doi.org/10.1016/j.jnca.2010.07.006>.
12. Zakiyyah R., Permana D., Valencio D., Chowanda A., Muliono Y., Izdiyar Z.N. Security Challenges and Issues in Cloud Computing. In: 2024 International Symposium on Networks, Computers and Communications (ISNCC). *IEEE*, 2024. p. 1–6. DOI <https://doi.org/10.1109/ISNCC62547.2024.10758966>.
13. Thavi R., Jhaveri R., Narwane V., Gardas B., Jafari Navimipour N. Role of cloud computing technology in the education sector. *Journal of Engineering, Design and Technology*, 2024. 22.1, 182–213. DOI <https://doi.org/10.1108/JEDT-08-2021-0417>.
14. Kumar S. U. B. O. D. H., Athavale V. A., Kartikey D. I. V. Y. E. Security issues in cloud computing: A holistic view. *International Journal of Internet of Things and Web Services*, 2021. 6, 18–29. URL: [https://www.iiar.org/iiar/filedownloads/ijitws/2021/022-0003\(2021\).pdf](https://www.iiar.org/iiar/filedownloads/ijitws/2021/022-0003(2021).pdf)
15. Dillon T., Wu C., Chang E. Cloud computing: issues and challenges. In: 2010 24th IEEE international conference on advanced information networking and applications. *Ieee*, 2010. p. 27–33. DOI <https://doi.org/10.1109/AINA.2010.187>.
16. Confidently L.A.O., Chakraborty M., Kundan A.P. Monitoring Cloud-Native Applications. 2021. URL: <https://link.springer.com/book/10.1007/978-1-4842-6888-9>
17. Vuorinen S. Monitoring Integration Systems and Visualization. 2022. URL: https://www.utupub.fi/bitstream/handle/10024/154329/Vuorinen_Simo_opinnayte.pdf?sequence=1.
18. Beyer B., Jones C., Petoff J., Murphy N. R. Site reliability engineering: how Google runs production systems. "O'Reilly Media, Inc.", 2016.
19. von Goethe J.W. Prometheus. Voltaire Press, 2024.
20. Shaji A. S., George M. M. Elastic stack: A comprehensive overview. 2024 IEEE Recent Advances in Intelligent Computational Systems (RAICS), 2024, 1–5. DOI <https://doi.org/10.1109/RAICS61201.2024.10690099>.
21. Chakraborty M., Kundan, A.P. Grafana. In: Monitoring cloud-native applications: Lead agile operations confidently using open source software. Berkeley, CA: Apress, 2021. p. 187–240. URL: https://link.springer.com/chapter/10.1007/978-1-4842-6888-9_6.

References:

1. Rudkovskiy, O. R., Kyrychek, H. H. (2021). Prohramnyi kompleks z pidtrymky rozpodilenoї vzaïemodii merezhevykh prystroiv ta dodatyv [A software complex supporting the distributed interaction of network devices and applications]. *Scientific notes of TNU named after V.I. Vernadskyi*. Series "Technical Sciences", 32(71), 2, 229–234 DOI <https://doi.org/10.32838/2663-5941/2021.2-1/36>. [in Ukrainian].
2. Zhang, F., Cao, J., Li, K., Khan, S. U., Hwang, K. (2014). Multi-objective scheduling of many tasks in cloud platforms. *Future generation computer systems*, 37, 309–320. DOI <https://doi.org/10.1016/j.future.2013.09.006>.
3. Tiahunova, M., Tronkina, O., Kirichek, G., Skrupsky, S. (2021). The Neural Network for Emotions Recognition under Special Conditions. In: *CMIS*. Vol-2864, p. 121–134. Retrieved from: <https://ceur-ws.org/Vol-2864/paper11.pdf>.
4. Kyrychek, H. H., Harkusha, V. Yu. (2021). Virtualizatsiia khostiv na osnovi Proxmox VE v umovakh nadlyshkovoho vykorystannia resursiv [Virtualization of hosts based on Proxmox VE in conditions of excessive use of resources]. *Scientific notes of TNU named after V.I. Vernadskyi*. Series "Technical Sciences". 32 (71), 1, 78–84. DOI <https://doi.org/10.32838/2663-5941/2021.1-1/13>. [in Ukrainian].
5. Sharma, R. (2021). A review on software defined networking. *Int J Sci Res Comput Sci Eng Inf Technol*, 11–14. DOI <https://doi.org/10.32628/CSEIT21728>.
6. Kreutz D., Ramos F.M., Verissimo P.E., Rothenberg C.E., Azodolmolky S., Uhlig S. Software-defined networking: A comprehensive survey. *Proceedings of the IEEE*, 2014, 103.1: 14–76. DOI <https://doi.org/10.1109/JPROC.2014.2371999>.
7. Yanamala, A. K. Y. (2024). Emerging challenges in cloud computing security: A comprehensive review. *International Journal of Advanced Engineering Technologies and Innovations*, 1.4, 448–479. Retrieved from: <https://thesisexpertsofficial.com/pdfs/4.pdf>
8. Rowshanrad, S., Namvarasl, S., Abdi, V., Hajizadeh, M., Keshtgary, M. (2014). A survey on SDN, the future of networking. *Journal of Advanced Computer Science & Technology*, 3.2, 232. DOI <https://doi.org/10.14419/jacst.v3i2.3754>.
9. Koponen, T., Casado, M., Gude, N., Stribling, J., Poutievski, L., Zhu, M., Ramanathan, R., Iwata, Y., Inoue, H., Hama, T., Shenker, S. (2010). Onix: A distributed control platform for large-scale production networks. In: 9th USENIX symposium on operating systems design and implementation (OSDI 10). Retrieved from: https://www.usenix.org/legacy/event/osdi10/tech/full_papers/Koponen.pdf.
10. Mell, P., Grance, T. (2011). The NIST definition of cloud computing. Retrieved from: <https://faculty.winthrop.edu/domanm/csci411/Handouts/NIST.pdf>.
11. Subashini, S., Kavitha, V. (2011). A survey on security issues in service delivery models of cloud computing. *Journal of network and computer applications*. 34.1, 1–11. DOI <https://doi.org/10.1016/j.jnca.2010.07.006>.
12. Zakiyyah, R., Permana, D., Valencio, D., Chowanda, A., Muliono, Y., Izdiyar, Z. N. (2024). Security Challenges and Issues in Cloud Computing. In: 2024 International Symposium on Networks, Computers and Communications (ISNCC). *IEEE*, p. 1–6. DOI <https://doi.org/10.1109/ISNCC62547.2024.10758966>.
13. Thavi, R., Jhaveri, R., Narwane, V., Gardas, B., Jafari Navimipour, N. (2024). Role of cloud computing technology in the education sector. *Journal of Engineering, Design and Technology*, 22.1, 182–213. DOI <https://doi.org/10.1108/JEDT-08-2021-0417>.
14. Kumar, S. U. B. O. D. H., Athavale, V. A., Kartikey, D. I. V. Y. E. (2021). Security issues in cloud computing: A holistic view. *International Journal of Internet of Things and Web Services*, 6, 18–29. Retrieved from: [https://www.ias.org/ias/filedownloads/ijitws/2021/022-0003\(2021\).pdf](https://www.ias.org/ias/filedownloads/ijitws/2021/022-0003(2021).pdf).
15. Dillon, T., Wu, C., Chang, E. (2010). Cloud computing: issues and challenges. In: 2010 24th IEEE international conference on advanced information networking and applications. *Ieee*, p. 27–33. DOI <https://doi.org/10.1109/AINA.2010.187>.
16. Confidently L. A. O., Chakraborty M., Kundan A. P. (2021). Monitoring Cloud-Native Applications. Retrieved from: <https://link.springer.com/book/10.1007/978-1-4842-6888-9>
17. Vuorinen, S. (2022). Monitoring Integration Systems and Visualization. Retrieved from: https://www.utupub.fi/bitstream/handle/10024/154329/Vuorinen_Simo_opinnayte.pdf?sequence=1.
18. Beyer, B., Jones, C., Petoff, J., Murphy, N. R. (2016). Site reliability engineering: how Google runs production systems. " O'Reilly Media, Inc."
19. von Goethe J.W. Prometheus. Voltaire Press, 2024.
20. Shaji, A. S., George, M. M. (2024). Elastic stack: A comprehensive overview. 2024 IEEE Recent Advances in Intelligent Computational Systems (RAICS), 1–5. DOI <https://doi.org/10.1109/RAICS61201.2024.10690099>.
21. Chakraborty, M., Kundan, A. P. (2021). Grafana. In: Monitoring cloud-native applications: Lead agile operations confidently using open source software. Berkeley, CA: Apress, p. 187–240. Retrieved from: https://link.springer.com/chapter/10.1007/978-1-4842-6888-9_6.

Дата надходження статті: 12.10.2025

Дата прийняття статті: 10.11.2025

Опубліковано: 30.12.2025

Крихівський М. В., кандидат технічних наук, доцент,
доцент кафедри інженерії програмного забезпечення
Івано-Франківського національного технічного університету
нафти і газу
ORCID: 0009-0000-3285-4308

Ваврик Т. О., асистент кафедри інженерії
програмного забезпечення
Івано-Франківського національного технічного університету
нафти і газу, м. Івано-Франківськ
ORCID: 0000-0002-0612-0084

Гобир Л. М., асистент кафедри інженерії
програмного забезпечення,
Івано-Франківського національного технічного університету
нафти і газу
ORCID: 0009-0007-3176-2314

ОСНОВИ ПРОЕКТУВАННЯ ІНТЕЛЕКТУАЛЬНИХ АГЕНТІВ: ВІД АРХІТЕКТУРИ ДО МАТЕМАТИЧНИХ МОДЕЛЕЙ

Інтелектуальні агенти стали центральною темою сучасних досліджень у сфері штучного інтелекту. Вони є важливим компонентом сучасних інформаційних технологій, що використовуються для автоматизації складних завдань, таких як обробка даних, прийняття рішень та адаптивне навчання. Розробка інтелектуальних агентів є складним, багатограним процесом, що включає вибір архітектури, розробку алгоритмів прийняття рішень і інтеграцію з середовищем. Протягом останніх десятиліть з'явилося багато нових технологій та підходів, які дозволяють значно покращити ефективність агентів у реальних застосуваннях. Стаття досліджує концептуальні засади інтелектуальних агентів, різні підходи до їх проектування та ключові технології, що дозволяють створювати автономні та адаптивні системи. Авторами розглядається загальний алгоритм функціонування інтелектуального агента, його основні етапи та архітектура. Представлено огляд концепції інтелектуальних агентів, включаючи їх базові властивості, моделі взаємодії з навколишнім середовищем та адаптивні механізми. Основну увагу приділено універсальному алгоритму, який може бути адаптований до різних застосувань у сфері штучного інтелекту. У статті представлено математичну основу для опису інтелектуальних агентів, що дозволяє моделювати їх поведінку, прийняття рішень та взаємодію з середовищем. У статті представлено математичну основу для опису інтелектуальних агентів, що дозволяє моделювати їх поведінку, прийняття рішень та взаємодію з середовищем. Розглянуто основні компоненти агентів, їх формалізацію через теоретичні моделі та відповідні алгоритми. Також розглянуто математичний опис агентів з урахуванням їхніх властивостей, таких як адаптивність, автономність та взаємодія з середовищем. Математичний аналіз інтелектуальних агентів є важливою і швидко розвиваючою областю, що охоплює широкий спектр методів та технік для моделювання, аналізу та оптимізації поведінки інтелектуальних агентів.

Ключові слова: інтелектуальні агенти, архітектура агентів, класифікація агентів, сенсори, актори.

Krykhivskiy M. V., Vavryk T. O., Hoby L. M. Foundations of Intelligent Agent Design: From Architecture to Mathematical Modeling

Intelligent agents (IA) have become a central topic of modern research in the field of artificial intelligence. They are an important component of modern information technologies used to automate complex tasks such as data processing, decision-making, and adaptive learning. The development of intelligent agents is a complex, multifaceted process that includes the choice of architecture, the development of decision-making algorithms, and integration with the environment. Over the past decades, many new technologies and approaches have emerged that can significantly improve the effectiveness of agents in real-world applications. The article explores the conceptual foundations of intelligent agents, various approaches to their design, and key technologies that allow for the creation of autonomous and adaptive systems. The authors consider the general algorithm of the functioning of an intelligent agent, its main stages and architecture. An overview of the concept of intelligent agents is presented, including their basic properties, models of interaction with the environment and adaptive mechanisms. The main attention is paid

© М. В. Крихівський, Т. О. Ваврик, Л. М. Гобир, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

to a universal algorithm that can be adapted to various applications in the field of artificial intelligence (AI). The article presents a mathematical basis for describing intelligent agents, which allows modeling their behavior, decision-making and interaction with the environment. The article presents a mathematical basis for describing intelligent agents, which allows modeling their behavior, decision-making and interaction with the environment. The main components of agents, their formalization through theoretical models and corresponding algorithms are considered. The mathematical description of agents is also considered, taking into account their properties, such as adaptability, autonomy and interaction with the environment. Mathematical analysis of intelligent agents is an important and rapidly developing area, covering a wide range of methods and techniques for modeling, analyzing and optimizing the behavior of intelligent agents.

Key words: intelligent agents, agent architecture, agent classification, sensors, actors.

Постановка проблеми. ІА відіграють ключову роль у багатьох галузях, включаючи робототехніку, автоматизацію, системи підтримки прийняття рішень та навіть побутові застосування. Розуміння основ проєктування таких агентів допомагає дослідникам та розробникам створювати ефективні та надійні системи, що можуть виконувати складні завдання з мінімальним втручанням людини. Окрім того, стаття також надає важливу теоретичну базу для подальших досліджень і вдосконалення алгоритмів штучного інтелекту (ШІ), що робить її незамінним джерелом знань для науковців і фахівців у цій сфері.

Невирішені питання. Проєктування ІА залишається складним завданням через низку невирішених питань. Однією з головних проблем є забезпечення агента здатністю до розуміння та інтерпретації контексту, адже обробка мови вимагає тонкого розуміння контекстуальних нюансів. Ще одним важливим аспектом є навчання агентів з урахуванням етичних міркувань та уникнення упередженості, що виникає через перекоси в навчальних даних. Крім того, адаптивність агентів до нових, непередбачуваних ситуацій і їх здатність до самонавчання та покращення в реальному часі також викликають занепокоєння. Безпека та конфіденційність даних користувачів залишаються критичними питаннями, які потрібно розв'язувати. Важливим є також забезпечення природної та ефективної взаємодії з користувачами, що включає виявлення та відповідь на емоції. Усі ці питання потребують інноваційних підходів та подальших досліджень для досягнення значного прогресу.

Створення універсальної архітектури інтелектуальних агентів, що поєднує адаптивність, масштабованість і здатність інтегрувати різні підходи, залишається важливим і складним завданням. Однією з ключових проблем є забезпечення агентів здатністю до гнучкого реагування на змінні середовища, що вимагає ефективних механізмів самонавчання та адаптації до нових умов. Це включає в себе використання методів машинного навчання та штучного інтелекту для постійного покращення та коригування поведінки агентів.

Масштабованість є важливим аспектом, оскільки для роботи з великими системами потрібно забезпечити ефективний розподіл ресурсів і обробку даних. Це вимагає розробки архітектури, здатної до горизонтального масштабування, що дозволяє збільшувати потужність системи шляхом додавання нових вузлів або компонентів без зниження продуктивності.

Інтеграція різних підходів, таких як реактивні, розподілені та гібридні, є важливою для створення універсальної архітектури, здатної вирішувати різноманітні завдання. Це потребує забезпечення узгодженості та взаємодії між компонентами системи, що використовують різні підходи, без втрати ефективності. Важливо також враховувати питання безпеки, конфіденційності та етичних міркувань при розробці таких систем. Усі ці аспекти вимагають інноваційних підходів та ретельного планування для досягнення гармонійної та ефективної взаємодії компонентів інтелектуальних агентів.

Пропозиція для вирішення проблеми:

1. Створення окремих модулів для сенсорного сприйняття, прийняття рішень, навчання та взаємодії між агентами.

2. Забезпечення інтерфейсів для легкої інтеграції нових модулів або алгоритмів

3. Запровадження механізмів безперервного навчання: агент оновлює свою модель поведінки в реальному часі, використовуючи досвід.

Очікувані результати:

1. Гнучка та масштабована архітектура, яка адаптується до різноманітних середовищ.

2. Зменшення витрат на інтеграцію нових задач та алгоритмів.

3. Підвищення ефективності агентів у реальному часі.

Аналіз останніх досліджень та публікацій. Методологія BDI (Belief-Desire-Intention), або переконання-бажання-наміри (ПБН), є одним із підходів до моделювання агентів у галузі штучного інтелекту. Вона базується на психологічних теоріях, що пояснюють, як люди приймають рішення, і переносить ці принципи на штучних агентів, що дозволяє їм діяти раціонально та автономно.

Основні компоненти ПБН моделі [1] – це переконання (Beliefs), бажання (Desires) та наміри (Intentions). Переконання – це знання або переконання агента про світ, які можуть бути правдивими або хибними. Це інформація, яка допомагає агенту оцінювати навколишню ситуацію та приймати рішення. Бажання – це цілі або завдання агента, що він хоче досягти. Вони спрямовують його дії та визначають, що агент намагається здійснити. Наміри – це конкретні плани або дії, які агент обирає для досягнення своїх цілей. Це рішення, які агент приймає, виходячи з своїх переконань та бажань.

ПБН архітектура дозволяє агенту працювати у складних динамічних середовищах, здійснюючи розумні та гнучкі рішення. Агент постійно оновлює свої переконання на основі нової інформації, коригує свої бажання, коли змінюються його пріоритети або умови, та обирає нові інтенції, коли це потрібно для досягнення цілей. Це надає агенту здатність адаптуватися до змін та непередбачуваних ситуацій.

Важливою характеристикою ПБН агентів є те, що вони здатні вирішувати конфлікти між різними бажаннями та інтенціями, враховуючи пріоритети та ресурси. Наприклад, якщо агент має кілька суперечливих цілей, він зможе оцінити, яка з них важливіша або більш досяжна у поточний момент, та вибрати відповідну дію. Проте, дослідники [2, 3] відмічають відсутність деяких компонентів гнучкості в моделі BDI та проблемність навчання.

Авторами оглядової статті [3] розглянуті основні підходи до кожного компонента архітектури ПБН, як вони були реалізовані в мовах програмування агентів, і показані компроміси, що властиві кожному підходу. Зроблені висновки про те, що для нових областей, які повинні інтегрувати субсимвольну інформацію від датчиків у переконання агентів, знадобляться більш насичені уявлення, ніж в роботі [1]. Це може передбачати використання вкладених та невизначених переконань.

Довгострокова автономія інтелектуальних агентів (IA) стає необхідною для багатьох автономних систем. Це стимулює подальші дослідження підходів до цілей у контексті того, що цілі підтримання та їх взаємодія з цілями досягнення стають більш важливими, ніж властивості та обмеження цілей, такі як пріоритети та дедлайни.

У роботі [3] виділяється проблема справжнього одночасного виконання планів. Цей напрям передбачає дослідження з формалізації взаємодій між одночасно виконуваними гілками та між циклом міркувань IA та тривалими діями. Інша область майбутніх досліджень стосується використання технік штучного інтелекту, таких як машинне навчання, для доповнення заздалегідь визначених правил плану або генерації нових правил. Авторами підкреслено перспективу підтримки невизначених переконань, що важливо для того, коли декларативна ціль може вважатися «досягнутою» (або недосяжною), і для підходів до планування намірів на основі симуляцій чергувань.

У статті [4] запропоновано архітектуру виконання плану, яка підтримує різну семантику завдань. Кожна мета, що визначена під час циклу обговорення агента, може бути задоволена через завдання, визначені в різній семантиці. Можливість підтримувати різну семантику завдань забезпечує дві основні переваги. Першою перевагою є повторне використання застарілих артефактів у планах агента. Другою перевагою є спрощення адаптації архітектури агента до різних стандартів організації бізнесу. Щоб інтегрувати різноманітну семантику завдання в архітектуру виконання плану, автори [4] використали плавну переглянуту версію попередньої сформульованої моделі робочого процесу, у якій різна семантика завдання була б зменшена перед виконанням. Інтегрування ієрархічної мережі завдань і семантики OWL-S в цю архітектуру виконання плану, дозволило перевірити її міцність з точки зору підтримки виконання різних семантик завдань в архітектурі агента.

Авторами роботи [5] створено структуру, яка додає знання для покращення вибору плану в популярній парадигмі програмування агентів ПБН. На відміну від попередніх пропозицій, наведений підхід може добре масштабуватись із складністю бібліотеки планів агента. Технічно розроблено нову міру впевненості, яка дозволяє агенту динамічно регулювати свою залежність від навчання, полегшуючи в принципі нескінченну кількість фаз (повторного) навчання.

Традиційний агент ПБН має 3 основні обчислювальні компоненти, які генерують переконання, генерують наміри та виконують наміри. Вони виконуються послідовно і циклічно. Це може спричинити кілька проблем. Серед них нездатність постійно спостерігати за навколишнім середовищем у динамічних середовищах може бути катастрофічною. Авторами роботи [6] запропоновано загальну структуру для моделі паралельного агента ПБН. За цією загальною структурою можна створювати паралельні агенти ПБН з різними конфігураціями залежно від доступності фізичних ресурсів. Ці агенти мають низку переваг перед послідовними, а саме зміни в середовищі агента можуть бути виявлені негайно, надзвичайні ситуації будуть вирішені негайно, підтримка надається на рівні архітектури для перегляду бажань/намірів і розгляду зв'язків між цілями, коли генерується нове переконання/бажання.

Концепції IA представлені в дослідницькій книзі [7], у якій сформована теорія агентів та адаптацію агентів у різних контекстах. Агенти різного порядку складності повинні бути автономними у використовуваних правилах. Агент повинен мати мозок, за допомогою якого він може виявити правила, що містяться в даних. Оскільки правила є інструментами, за допомогою яких агенти змінюють середовище, будь-яку адаптацію правил можна розглядати як еволюцію агентів. Оскільки невизначеність присутня в будь-якому контексті, авторами книги [7] описано як можна ввести глобальну невизначеність з локального світу в опис правил. У книзі представлено загальний опис агента еволюційної адаптації, дії та мета дії такого агента, абстрактне та формальне представлення дії, зв'язок системи та метасистеми з адаптивним агентом, мозок агента за допомогою теорії морфогенетичних нейронів, логічну структуру адаптивного агента, зворотний зв'язок і гіперзворотний зв'язок в адаптивному агенті, поле адаптації в модальний логічний простір як логічний інструмент в адаптивному агенті, дію агента у фізичному домені та застосування агентів у роботах та еволюційних обчисленнях.

Автори роботи [8] представляють багатоагентну архітектурну теорію для розробки інтелектуальних асоціативних гібридних систем. Теорія архітектури описана як на рівні структури завдання, так і на рівні обчислень. Ця архітектура призначена для вирішення проблем розробки агентів знань та інформаційних агентів.

Архітектури автономних агентів – це методології проектування сукупності знань і стратегій, які застосовуються до проблеми створення локального інтелекту. У статті [9] зроблено спробу об'єднати ці знання з кількома архітектурними традиціями, приділяючи особливу увагу особливостям, які, як правило, вибиралися під тиском широкого використання в системах реального світу. Автори визначили, що наступні стратегії надають значну допомогу в проектуванні автономних інтелектуальних агентів, а саме модульність, яка спрощує як дизайн, так і контроль, ієрархічно організований вибір дій, який фокусує увагу та забезпечує пріоритетність, коли різні модулі конфліктують; та паралельний моніторинг середовища, який дозволяє системі бути чутливою та опортуністичною, дозволяючи переключити увагу та переоцінити пріоритети. Ми пропонуємо огляд чотирьох архітектурних парадигм: III на основі поведінки; дво- і тришарові системи; архітектури віри, бажання та наміру (зокрема PRS); і Soar/ACT-R. Задokumentувавши тенденції в кожній із цих спільнот щодо створення компонентів вище, ми стверджуємо, що ця конвергентна еволюція є переконливим доказом корисності компонентів. Потім ми використовуємо цю інформацію, щоб рекомендувати конкретні стратегії для дослідників, які працюють за кожною парадигмою, для подальшого використання знань і досвіду в галузі в цілому.

Механізми емоцій часто використовуються в штучних агентах як метод покращення вибору дій. Порівняння між агентами є складними через відсутність єдності між теоріями емоцій, завданнями агентів і типами вибору дій, що використовуються. Набір архітектурних якостей пропонується у роботі [10] як основа для проведення порівнянь між агентами. Аналіз існуючих архітектур агентів, які включають механізм емоцій, може допомогти триангулювати можливості дизайну в просторі, окресленому цими якостями. З огляду на це, для аналізу обрано дванадцять автономних агентів, які включають механізм емоцій у вибір дій. Кожен агент розбирається за допомогою цих архітектурних якостей (архітектура агента, механізм вибору дій, механізм емоцій і представлення емоційного стану, а також модель емоцій, на якій він базується). Це допомагає розмістити агентів в архітектурному просторі, підкреслює контрастні методи реалізації схожих теоретичних компонентів і підказує, які архітектурні аспекти важливі для виконання завдань. Представлено початкову структуру, що складається з ряду рекомендацій щодо розробки механізмів емоцій у штучних агентах, заснованих на кореляції між виконуваними ролями емоцій та аспектами механізмів емоцій, які використовуються для виконання цих ролей. У висновку обговорюється, як можна вирішити проблеми з цим типом дослідження та якою мірою розробка основи може допомогти майбутнім дослідженням.

Логіка ПБН є важливим і широко використовуваним теоретичним апаратом для представлення та міркування про раціональних агентів. Однак вона є неповною щодо перегляду наміру, перевизначення наміру, процесу обговорення та перегляду переконань. Крім того, деякі раціональні агенти, особливо люди, не мають наближених міркувань, добре представлених логікою ПБН. У статті [11] визначено нечітку семантику для мови ПБН, яка здатна усунути ці обмеження. Запропоновано розширити логіку ПБН до нечіткої логіки ПБН.

У статтях [12-14] проведено аналіз педагогічних та інформаційних засад використання інтелектуальних агентів в адаптивних системах електронного навчання з метою створення структури адаптивної системи електронного навчання. Авторами описано характеристики, функції та взаємодії агентів, які беруть участь у кожному модулі адаптивної архітектури, а також характеристики, функції та взаємодії інтелектуального агента для ухвалення навчальних рішень.

У роботі [15] представлено підхід до розробки ІА, який можна використовувати в мультиагентній системі для вирішення складної проблеми. Запропонований агент має можливості навчання за допомогою штучних нейронних мереж.

Авторами статті [16] розглядаються проблеми створення надійних децентралізованих програм способом, сумісним із наскрізною моделлю. Створена ними програма Mandrake базується на розумінні того, що можливість агента отримувати повідомлення в будь-якому порядку та обробляти їх відповідно до його власного прийняття рішення є вирішальною для реалізації сенсу програми.

Автори публікації [17] зазначають, що III ніщо без інтелектуальних агентів і досліджують скільки типів агентів визначено в штучному інтелекті. Вони виділяють п'ять категорій, які згруповані за діапазоном можливостей і ступенем сприйманого інтелекту, і роблять висновок, що кожна з них здатна втілити ідеї III в життя.

У дослідженнях [18-21] розшифровано багатогранну природу інтелектуальних агентів, проаналізовано їх застосування в реальному світі, переваги та потенційні обмеження. Автори розглянули класифікацію інтелектуальних агентів, що включає різні типи, такі як прості рефлексивні агенти, модельно-рефлексивні агенти, цільові агенти та корисні агенти. Кожен варіант наділяє агента різними можливостями та функціями, що відповідають різноманітним вимогам додатків III.

Мета статті є надання всебічного огляду та аналізу ключових аспектів створення ІА. Стаття спрямована на забезпечення теоретичної та практичної основи для розуміння процесу розробки таких агентів, починаючи від їхньої архітектури та закінчуючи математичними моделями, які описують їхню поведінку та

функціонування. Це прагнення узагальнити існуючі знання та досвід, а також пропонування нових підходів та методів, що сприятимуть підвищенню ефективності та надійності інтелектуальних систем у різних сферах застосування.

У цій статті розглянуто основні принципи та підходи до проектування ІА, з особливим акцентом на їхню архітектуру та математичне моделювання. Ми прагнемо створити цілісне уявлення про теоретичні основи та практичні аспекти цього процесу, що дозволить розробникам і дослідникам ефективно використовувати ці знання для створення надійних та адаптивних систем.

Виклад основного матеріалу дослідження. Сучасні підходи до проектування інтелектуальних агентів включають низку методологій та технологій, що дозволяють створювати ефективні та гнучкі системи для розв'язання різноманітних завдань. Одним з таких підходів є агентно-орієнтоване програмування, яке базується на концепції автономних агентів, що взаємодіють з оточенням і іншими агентами для досягнення своїх цілей.

У основі сучасного проектування інтелектуальних агентів лежить використання машинного навчання та штучного інтелекту. Одним з ключових аспектів є навчання агентів на основі підкріплення (reinforcement learning), що дозволяє агентам навчатися на власному досвіді через взаємодію з середовищем. Такі агенти використовують алгоритми, що дозволяють їм удосконалювати свої стратегії та приймати оптимальні рішення на основі отриманих винагород.

Ще одним важливим підходом є використання нейронних мереж, зокрема, глибокого навчання (deep learning). Глибокі нейронні мережі дозволяють агентам аналізувати великі обсяги даних і виявляти приховані закономірності, що підвищує їхню здатність до передбачення та прийняття рішень. Такі мережі можуть бути використані для обробки візуальної інформації, розпізнавання мови, а також для розв'язання складних задач, таких як ігри чи стратегічне планування.

Сучасні підходи також передбачають використання мультиагентних систем, де кілька агентів взаємодіють між собою для досягнення спільних цілей. У таких системах агенти можуть співпрацювати, координувати свої дії та обмінюватися інформацією, що дозволяє підвищити ефективність і надійність системи в цілому. Для цього використовуються методи розподіленого штучного інтелекту та алгоритми консенсусу.

Крім того, важливим аспектом проектування інтелектуальних агентів є забезпечення їхньої адаптивності та масштабованості. Сучасні агенти повинні бути здатні адаптуватися до змін у середовищі, а також масштабуватися для роботи в різних умовах і з різним рівнем завантаженості. Для цього використовуються методи самоорганізації та автономного управління.

Отже, сучасні підходи до проектування інтелектуальних агентів базуються на використанні машинного навчання, нейронних мереж, мультиагентних систем і методів адаптивного управління. Ці підходи дозволяють створювати гнучкі, ефективні та надійні системи, здатні розв'язувати складні задачі в різноманітних галузях.

Існує кілька інших методологій для моделювання агентів, кожна з яких має свої особливості та відрізняється від ПБН підходу:

1. Реактивні агенти: Ці агенти працюють на основі безпосередніх реакцій на зміни в навколишньому середовищі. Вони не мають внутрішнього уявлення про світ чи довгострокових цілей. Замість цього, вони реагують на вхідні стимули за допомогою заздалегідь визначених правил або алгоритмів. Це робить їх швидкими та ефективними у простих і стабільних середовищах, але вони можуть бути менш гнучкими у складних та динамічних умовах.

2. Цільові агенти (Goal-oriented agents): Ці агенти працюють на основі цілей, які вони намагаються досягти. Вони використовують планування та розв'язання задач для досягнення своїх цілей. Цей підхід дозволяє агентам бути більш раціональними та адаптивними, оскільки вони можуть планувати свої дії на основі поточної ситуації та ресурсів.

3. Агенти на основі користувацьких моделей (User model-based agents): Ці агенти використовують моделі користувачів для розуміння та прогнозування їх поведінки. Вони можуть адаптувати свої дії та відповіді на основі індивідуальних характеристик та потреб користувачів. Це корисно у застосуваннях, де важлива взаємодія з користувачем, наприклад, у рекомендаційних системах або персональних асистентах.

4. Соціальні агенти: Ці агенти враховують соціальні взаємодії та відносини з іншими агентами. Вони можуть співпрацювати, конкурувати або вести переговори з іншими агентами для досягнення своїх цілей. Соціальні агенти зазвичай використовують моделі багатогентних систем та ігрову теорію для ефективної взаємодії.

5. Гібридні агенти: Цей підхід поєднує елементи кількох методологій для створення більш потужних і гнучких агентів. Наприклад, агент може використовувати реактивні стратегії для швидкої реакції на зміни в середовищі та планування на основі цілей для довгострокових завдань. Гібридні агенти можуть мати складну архітектуру, але вони здатні краще адаптуватися до різноманітних умов.

Кожна з цих методологій має свої переваги та недоліки, і вибір конкретного підходу залежить від вимог та умов конкретної задачі.

Висновки. Інтелектуальні агенти є ключовим елементом у багатьох сучасних технологіях, що забезпечує автоматизацію складних завдань. З їх допомогою можна розв'язувати проблеми в таких сферах, як управління ресурсами, робототехніка, аналіз даних і персоналізоване обслуговування.

Математичні методи відіграють центральну роль у проектуванні, моделюванні та реалізації інтелектуальних агентів. Взаємодія між різними галузями, такими як теорія ймовірностей, оптимізація та машинне навчання, дозволяє створювати системи, здатні до адаптації, навчання та автономного функціонування. Подальший розвиток математичних основ може значно розширити можливості інтелектуальних агентів у різних сферах застосування.

Запропонований загальний алгоритм інтелектуального агента є універсальним підходом, що може бути адаптований до різних завдань. Його модульна структура забезпечує масштабованість та простоту інтеграції з іншими системами ШІ. Подальші дослідження спрямовані на вдосконалення модулів навчання та прийняття рішень, що дозволить створювати більш ефективні та автономні агенти. Майбутнє інтелектуальних агентів включає інтеграцію з новими технологіями, такими як квантові обчислення, а також розвиток етично обґрунтованих алгоритмів для забезпечення безпеки та справедливості.

Універсальна архітектура, що включає модульну структуру, динамічне навчання та адаптацію, розподілену обробку і інтеграцію різних підходів, забезпечує гнучкість, масштабованість та ефективність роботи інтелектуальних агентів у різноманітних середовищах. Додавання елементів теорії нечітких множин до математичної моделі значно покращує здатність агентів працювати з невизначеною та неточною інформацією, що підвищує точність та адаптивність прийняття рішень.

Створена математична модель дозволяє інтелектуальним агентам постійно вдосконалювати свою поведінку на основі нових нечітких даних, адаптуючись до змінних умов середовища. Модульна структура забезпечує можливість легкої заміни або оновлення окремих компонентів системи без впливу на всю архітектуру, що робить її більш гнучкою та придатною для широкого спектру застосувань.

Включення елементів нечіткості в модель дозволяє агентам ефективно працювати в умовах невизначеності, приймаючи оптимальні рішення на основі нечітких даних. Це особливо важливо для реальних застосувань, де інформація часто є неточною або неоднозначною.

Запропоновані в статті універсальна архітектура та математична модель відкривають нові можливості для розвитку інтелектуальних агентів, здатних ефективно адаптуватися до змінних умов, масштабуватися та інтегрувати різні підходи. Ці рішення сприяють підвищенню продуктивності та ефективності інтелектуальних систем, що дозволить їх широко застосовувати в різних галузях, від промисловості до обслуговування клієнтів. У кінцевому рахунку, впровадження цих рішень сприятиме створенню більш інтелектуальних, гнучких та надійних систем, здатних задовольнити потреби сучасного світу.

Напрями подальших досліджень:

1. Розробка алгоритмів для швидкої адаптації до середовищ із невідомими правилами або динамікою.
2. Моделювання систем із частковою синхронізацією між агентами.
3. Аналіз ефективності в порівнянні з існуючими рішеннями.
4. Додавання можливостей для налаштування середовища, де агент може навчатися і тестуватися.

Список використаних джерел:

1. Padgham L., Winikoff M. Developing Intelligent Agent Systems. A Practical Guide. Wiley, 2004. P. 225. DOI: 10.1002/0470861223
2. Saadi A., Maamri R., Sahnoun Z. Behavioral flexibility in Belief-Desire- Intention (BDI) architectures. *Multiagent and Grid Systems*. 2020. № 16(4). P. 343–377. DOI: <https://doi.org/10.3233/MGS-200335>
3. De Silva L., Meneguzzi F., Logan B. BDI Agent Architectures: A Survey, Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence Survey track. 2020, P. 4914–4921. DOI: <https://doi.org/10.24963/ijcai.2020/684>
4. Ekinici E., Halaç, T., Erdur C., Çetin Ö., Cakirlar I., Dikenelli O. Satisfying agent goals by executing different task semantics: HTN, OWL-S or plug one yourself. *Autonomous Agents and Multi-Agent Systems*. 2013. № 26(2). DOI: <https://doi.org/10.1007/s10458-011-9185-2>
5. Singh D., Sardina S., Padgham L., James G. Integrating Learning into a BDI Agent for Environments with Changing Dynamics. *International Joint Conference on Artificial Intelligence*. 2011. P. 2525–2530. DOI: <https://doi.org/10.5591/978-1-57735-516-8/IJCAI11-420>
6. Zhang H., i Huang S. Y., A general framework for parallel BDI agents in dynamic environments. *Web Intelligence and Agent Systems Journal*. 2008. № 6(3). P. 327–351. DOI: <https://doi.org/10.1109/IAT.2006.8>
7. Germano R., Lakhmi C. J. Intelligent Agents. Theory and Applications. Springer Berlin, Heidelberg. 2004. P. 402. DOI: <https://doi.org/10.1007/978-3-540-44401-5>
8. Khosla R., Dillon T. Engineering Intelligent Hybrid Multi-Agent Systems. Springer New York. 1997. P. 410. DOI: <https://doi.org/10.1007/978-1-4615-6223-8>
9. Bryson, J. Cross-paradigm analysis of autonomous agent architecture. *Journal of Experimental & Theoretical Artificial Intelligence*. 2000. № 12(2). P. 165–189. DOI: <https://doi.org/10.1080/095281300409829>
10. Rumbell T., Barnden J., Denham S., Wennekers T. Emotions in autonomous agents: comparative analysis of mechanisms and functions. *Auton Agent Multi-Agent Syst*. 2012. № 25. P. 1–45. DOI: <https://doi.org/10.1007/s10458-011-9166-5>

-
11. Cruz A., dos Santos A. V., Santiago R. H. N., Bedregal B. A Fuzzy Semantic for BDI Logic. *Fuzzy Information and Engineering*. 2021. № 13(2). P. 139-153. DOI: <https://doi.org/10.1080/16168658.2021.1915455>
 12. Calegari R., Ciatto G., Mascardi, V., Omicini A. Logic-based technologies for multi-agent systems: a systematic literature review. *Auton Agent Multi-Agent Syst*. 2021. № 35(1). <https://doi.org/10.1007/s10458-020-09478-3>
 13. Круглик В. С., Прокоф'єв Є. Г., Маринов, А. В. Аналіз можливостей використання інтелектуальних агентів в адаптивній системі електронного навчання. *Педагогічні науки: теорія та практика*. 2021. № 4. С. 295–302. Режим доступу: <https://doi.org/10.26661/2786-5622-2021-4-44>
 14. Лопатто І. Ю., Говорущенко Т. О., Капустян М. В. Інтелектуальний агент верифікації врахування інформації предметної галузі в процесі розроблення програмних систем. *Вісник Хмельницького національного університету*. 2022. № 1. С. 116–119. URL: <http://journals.khnu.km.ua/vestnik/?p=12131>
 15. Маринов А. В., Круглик В. С. Використання інтелектуальних програмних агентів для створення адаптивного середовища електронного навчання на базі lms moodle. *Міжнародна науково-практична конференція «Цифрова трансформація та диджитал технології для сталого розвитку всіх галузей сучасної освіти, науки і практики»*. Zbiór prac Tom 2. 2023. С. 306–308. URL: https://repo.btu.kharkov.ua/bitstream/123456789/29446/1/zbior_prac_tom_2_26012023-306-308.pdf
 16. Noulamo T., Djimeli-Tsajio A., Kameugne R., Lienou, J. A Generic Intelligent Agent Design Approach Based on Artificial Neural Networks. *World Journal of Engineering and Technology*. 2023. № 11. С. 682–697. URL: <https://doi.org/10.4236/wjet.2023.114046>
 17. Christie, S.H., Chopra, A. K., Singh, M. P. Mandrake: Multiagent Systems as a Basis for Programming Fault-Tolerant Decentralized Applications. *Autonomous Agents and Multi-Agent Systems*. 2022. № 36, A. № 16. DOI: <https://doi.org/10.1007/s10458-021-09540-8>
 18. Latham N. Types of intelligent agent, 2024. URL: <https://www.probecx.com/en-au/blog/types-of-intelligent-agent>
 19. Intelligent Agent, 2023. URL: https://www.larksuite.com/en_us/topics/ai-glossary/intelligent-agent
 20. Harjyot K. What are AI Agents: Types, Benefits, Applications, and Examples. 2024. URL: <https://www.signitysolutions.com/blog/ai-agents>
 21. Hiren Dhaduk. What is an AI Agent? Characteristics, Advantages, Challenges, Applications. 2023. URL: <https://www.simform.com/blog/ai-agent/>
 22. Крихівський М. В., Крихівська С. М. Проблеми людино-машинної взаємодії в контексті штучного інтелекту. *Збірник тез доповідей науково-практичної конференції «Інформаційні технології в освіті, техніці та промисловості»*. 2024. С. 176–177. URL: <https://stlnau.in.ua/samoosvita/item/2024/iit241010.pdf>

References:

1. Padgham, L., Winikoff, M. (2004). *Developing Intelligent Agent Systems. A Practical Guide*. Wiley, P. 225. DOI: [10.1002/0470861223](https://doi.org/10.1002/0470861223)
2. Saadi, A., Maamri, R., Sahnoun, Z. (2020). Behavioral flexibility in Belief-Desire- Intention (BDI) architectures. *Multiagent and Grid Systems*. № 16(4). P. 343–377. DOI: <https://doi.org/10.3233/MGS-200335>
3. De Silva, L. Meneguzzi, F., Logan, B. (2020). BDI Agent Architectures: A Survey, Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence Survey track. P. 4914–4921. DOI: <https://doi.org/10.24963/ijcai.2020/684>
4. Ekinici, E., Halaç, T., Erdur, C., Çetin, Ö., Cakirlar, I., Dikenelli, O. (2013). Satisfying agent goals by executing different task semantics: HTN, OWL-S or plug one yourself. *Autonomous Agents and Multi-Agent Systems*. № 26(2). DOI: <https://doi.org/10.1007/s10458-011-9185-2>
5. Singh, D., Sardina, S., Padgham, L., James, G. (2011). Integrating Learning into a BDI Agent for Environments with Changing Dynamics. *International Joint Conference on Artificial Intelligence*. P. 2525–2530. DOI: <https://doi.org/10.5591/978-1-57735-516-8/IJCAI11-420>
6. Zhang, H., Huang, S. Y. (2008). A general framework for parallel BDI agents in dynamic environments. *Web Intelligence and Agent Systems Journal*. № 6(3). P. 327–351. DOI: <https://doi.org/10.1109/IAT.2006.8>
7. Germano, R., Lakhmi, C. J. (2004). *Intelligent Agents. Theory and Applications*. Springer Berlin, Heidelberg. P. 402. DOI: <https://doi.org/10.1007/978-3-540-44401-5>
8. Khosla, R., Dillon, T. (1997). *Engineering Intelligent Hybrid Multi-Agent Systems*. Springer New York. P. 410. DOI: <https://doi.org/10.1007/978-1-4615-6223-8>
9. Bryson, J. (2000). Cross-paradigm analysis of autonomous agent architecture. *Journal of Experimental & Theoretical Artificial Intelligence*. № 12(2). P. 165–189. DOI: <https://doi.org/10.1080/095281300409829>
10. Rumbell, T., Barnden, J., Denham, S., Wennekers, T. (2012). Emotions in autonomous agents: comparative analysis of mechanisms and functions. *Auton Agent Multi-Agent Syst*. № 25. P. 1–45. DOI: <https://doi.org/10.1007/s10458-011-9166-5>
11. Cruz, A., dos Santos, A. V., Santiago, R. H. N., Bedregal, B. (2021). A Fuzzy Semantic for BDI Logic. *Fuzzy Information and Engineering*. № 13(2). P. 139–153. DOI: <https://doi.org/10.1080/16168658.2021.1915455>

-
12. Calegari, R., Ciatto, G., Mascardi, V., Omicini, A. (2021). Logic-based technologies for multi-agent systems: a systematic literature review. *Auton Agent Multi-Agent Syst.* № 35(1). <https://doi.org/10.1007/s10458-020-09478-3>
 13. Kruhlyk, V. S., Prokofiev, Ye. H., Marynov, A. V. (2021). Analiz mozhyvostei vykorystannia intelektualnykh ahentiv v adaptivnii systemi elektronnoho navchannia [Analysis of the possibilities of using intelligent agents in an adaptive e-learning system]. *Pedahohichni nauky: teoriia ta praktyka.* № 4, S. 295–302. <https://doi.org/10.26661/2786-5622-2021-4-44>
 14. Lopatto, I. Yu., Hovorushchenko, T. O., Kapustian, M. V. (2022). Intelektualnyi ahent veryfikatsii vrakhuvannia informatsii predmetnoi haluzi v protsesi rozroblennia prohramnykh system [Intelligent agent for verification of the consideration of subject area information in the process of developing software systems]. *Visnyk Khmelnytskoho natsionalnoho universytetu,* № 1. S.116–119 Retrieved from: <http://journals.khnu.km.ua/vestnik/?p=12131>
 15. Marynov, A. V., Kruhlyk, V. S. (2023). Vykorystannia intelektualnykh prohramnykh ahentiv dlia stvorennia adaptivnoho seredovyscha elektronnoho navchannia na bazi lms moodle [Using intelligent software agents to create an adaptive e-learning environment based on lms moodle]. *Mizhnarodna naukovo-praktychna konferents «Tsyfrova transformatsiia ta dydzhytal tekhnolohii dlia staloho rozvytku vsikh haluzei suchasnoi osvity, nauky i praktyky».* Zbiór prac Tom 2. S. 306–308. Retrieved from: https://repo.btu.kharkov.ua/bitstream/123456789/29446/1/zbior_prac_tom_2_26012023-306-308.pdf
 16. Noulamo, T., Djimeli-Tsajio A., Kameugne R., Lienou, J. (2023). A Generic Intelligent Agent Design Approach Based on Artificial Neural Networks. *World Journal of Engineering and Technology.* № 11. C. 682–697. Retrieved from: 10.4236/wjet.2023.114046
 17. Christie, S.H., Chopra, A. K., Singh, M. P. (2022). Mandrake: Multiagent Systems as a Basis for Programming Fault-Tolerant Decentralized Applications. *Autonomous Agents and Multi-Agent Systems.* № 36, A. № 16. DOI: <https://doi.org/10.1007/s10458-021-09540-8>
 18. Latham N. Types of intelligent agent, 2024. Retrieved from: <https://www.probecx.com/en-au/blog/types-of-intelligent-agent>
 19. Intelligent Agent, (2023). Retrieved from: https://www.larksuite.com/en_us/topics/ai-glossary/intelligent-agent
 20. Harjyot, K. (2024). What are AI Agents: Types, Benefits, Applications, and Examples. Retrieved from: <https://www.signitysolutions.com/blog/ai-agents>
 21. Hiren Dhaduk. (2023). What is an AI Agent? Characteristics, Advantages, Challenges, Applications. Retrieved from: <https://www.simform.com/blog/ai-agent/>
 22. Krykhivskyi, M. V., Krykhivska, S. M. (2024). Problemy liudyno-mashynnoi vzaiemodii v konteksti shturnoho intelektu [Problems of human-machine interaction in the context of artificial intelligence]. *Zbirnyk tez dopovidei naukovo-praktychnoi konferentsii «Informatsiini tekhnolohii v osviti, tekhnitsi ta promyslovosti».* S. 176–177. Retrieved from: <https://stlnau.in.ua/samoosvita/item/2024/iit241010.pdf>

Дата надходження статті: 24.10.2025

Дата прийняття статті: 17.11.2025

Опубліковано: 30.12.2025

КІБЕРБЕЗПЕКА ТА ЗАХИСТ ІНФОРМАЦІЇ

UDC 004.056.55

DOI <https://doi.org/10.32782/2521-6643-2025-2-70.28>

Kovalchuk D. M., PhD in Computer Science,
Senior Lecturer
V. N. Karazin Kharkiv National University
ORCID: 0000-0002-8229-836X

METHOD FOR IMPLEMENTING THE SQUARING OPERATION IN THE RABIN CRYPTOSYSTEM BASED ON THE USE OF THE RESIDUE NUMBER SYSTEM

The article proposes a method for implementing the operation of squaring a number in the Rabin cryptosystem, which increases the speed of arithmetic operations by using the Residue Number System (RNS). The proposed approach is based on the transition from the positional numeral system to RNS, which allows computations to be performed independently for each modulus and enables parallel data processing. Unlike traditional methods that rely on sequential execution of arithmetic operations with large numbers, the use of RNS avoids carry propagation, thereby reducing time complexity.

A mathematical model of the process of multiplying two numbers represented in RNS has been developed based on the use of table multiplication coding. This approach takes into account the symmetry properties of table multiplication, allowing the volume of required computations to be reduced to 25% of the full table. Based on this mathematical model, a method for squaring numbers in RNS has been proposed.

A comparative analysis showed that the use of the proposed approach provides a significant performance gain: for 32-bit operands, the speedup is 2048 times, and for 64-bit operands, up to 8192 times compared to the positional numeral system.

The research results confirm the feasibility of applying the Residue Number System in the Rabin cryptosystem for implementing the squaring operation. The proposed method can be effectively used in high-performance cryptographic systems designed for processing large numerical fields. Further research should focus on developing a universal structure for modular computations in RNS for other asymmetric cryptosystems, as well as creating a hardware implementation of the method to practically evaluate its performance.

Key words: Rabin cryptosystem, Residue Number System, squaring, performance enhancement, parallel computations.

Ковальчук Д. М. Метод реалізації операції піднесення до квадрату в криптосистемі Рабіна на основі використання системи залишкових класів

У статті запропоновано новий метод реалізації операції піднесення числа до квадрату в криптосистемі Рабіна, який спрямований на суттєве підвищення швидкодії виконання арифметичних операцій завдяки застосуванню системи залишкових класів (СЗК). Основна ідея підходу полягає у переході від традиційної позиційної системи числення до СЗК, що дозволяє виконувати обчислення незалежно для кожного модуля, відкриваючи можливості для паралельної обробки даних. Такий підхід суттєво відрізняється від класичних методів, які оперують великими числами послідовно та обмежені необхідністю обробки переносу розрядів, що підвищує часову складність операцій. Використання СЗК дозволяє усунути цей недолік, забезпечуючи значне скорочення обчислювального часу.

У роботі розроблено математичну модель множення двох чисел у СЗК, що базується на використанні коду табличного множення. Модель враховує властивості симетрії таблиці множення, що дозволяє зменшити обсяг необхідних обчислень до 25 % від повної таблиці. На основі цієї моделі запропоновано метод піднесення чисел до квадрату в СЗК, який демонструє високу ефективність для чисел великої розрядності. Метод дозволяє виконувати обчислення незалежно по кожному модулю, що робить його придатним для апаратної реалізації.

Проведений порівняльний аналіз показав, що запропонований підхід забезпечує істотне прискорення обчислень: для 32-розрядних операндів швидкодія зростає у 2048 разів, а для 64-розрядних операндів – до 8192 разів у порівнянні з традиційними методами позиційної системи числення. Це свідчить про високу ефективність використання СЗК у криптосистемі Рабіна та демонструє практичну користь для високопродуктивних криптографічних систем, орієнтованих на обробку великих числових полів.

Отримані результати підтверджують доцільність застосування СЗК для реалізації операцій піднесення до квадрату та відкривають перспективи для подальших досліджень. Подальші напрями включають розробку універсальної структури модульних обчислень у СЗК для інших асиметричних криптосистем, оптимізацію алгоритмів для апаратної

© D. M. Kovalchuk, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

реалізації та практичну оцінку продуктивності запропонованого методу у різних обчислювальних середовищах. Застосування запропонованого підходу може стати ключовим елементом у створенні швидких і надійних криптографічних рішень, що відповідають сучасним вимогам інформаційної безпеки та обробки великих обсягів даних.

Ключові слова: криптосистема Рабіна, система залишкових класів, піднесення до квадрату, підвищення швидкості, паралельні обчислення.

Problem statement. At the current stage of cryptography development, one of the key areas of research is improving the efficiency of basic mathematical operations that form the foundation of cryptographic systems [1]. One such system is the Rabin cryptosystem. The Rabin cryptosystem belongs to the class of asymmetric encryption systems whose security is based on the computational difficulty of residueing large integers [2]. Unlike the RSA system, where encryption involves exponentiation with an arbitrary exponent, the Rabin cryptosystem uses only the squaring operation modulo the product of two prime numbers. This property makes it mathematically elegant and simple to implement, while maintaining a high level of cryptographic strength [3, 4].

For key generation, two large prime numbers p and q are randomly chosen so that they are relatively prime. Their product $n = p \cdot q$ serves as the public key, while the numbers p and q themselves form the private key, which must be kept secret. The choice of such parameters ensures the one-way nature of the encryption function – the inverse transformation is possible only if the residues of n are known [4, 5].

Message encryption consists in transforming the input data block M , which satisfies the condition $0 < M < n$, according to the formula

$$C = M^2 \bmod n, \quad (1)$$

where C is the ciphertext. The encryption process is extremely simple, as it involves only one arithmetic operation – squaring. At the same time, retrieving the original message M from the known C without knowledge of the prime residues of the modulus n is practically impossible due to the computational complexity of the factorization problem [6].

The decryption process involves the use of the secret parameters p and q . At the first stage, the remainders are computed as follows:

$$f = C \bmod p, \quad s = C \bmod q. \quad (2)$$

Next, it is necessary to find the quadratic roots of the obtained values under the corresponding moduli:

$$x^2 \equiv f \pmod{p}, \quad y^2 \equiv s \pmod{q}. \quad (3)$$

Since each of these equations has two solutions, the total number of possible combinations equals four. Accordingly, a system of congruences is formed:

$$\begin{cases} M \equiv x \pmod{p}, \\ M \equiv y \pmod{q}, \end{cases} \quad \begin{cases} M \equiv x \pmod{p}, \\ M \equiv -y \pmod{q}, \end{cases} \quad (4) \\ \begin{cases} M \equiv -x \pmod{p}, \\ M \equiv y \pmod{q}, \end{cases} \quad \begin{cases} M \equiv -x \pmod{p}, \\ M \equiv -y \pmod{q}. \end{cases}$$

For each of these systems, the corresponding value of M is computed using the Chinese Remainder Theorem [7]. As a result, four potential solutions are obtained, among which only one corresponds to the original message.

However, in the practical implementation of the Rabin cryptosystem, a number of issues arise due to the high time complexity of arithmetic operations in multi-bit numerical fields. Operations such as multiplication, division, or squaring in the traditional positional number system require significant CPU resources, especially when processing numbers longer than 1024 bits – the minimum required security level under modern conditions. This creates the need to search for new approaches to optimizing computations in cryptographic algorithms [8].

One promising direction is the use of the RNS. The application of RNS enables the parallelization of computations, which is not possible when using a positional number system. As a result, there is no need for carry propagation, and the basic operations – addition, subtraction, and squaring – can be performed independently and in parallel, significantly increasing the computational speed of the system [5, 8].

Analysis of recent research and publications. Research on optimizing cryptographic computations using the RNS has been actively developing in recent years. In particular, work [9] proposed an efficient implementation of the Montgomery algorithm for modular multiplication in a minimally redundant RNS, which significantly accelerates reduction operations and reduces overhead for large modular computations. This approach demonstrates potential applicability to squaring operations characteristic of the Rabin cryptosystem; however, the article does not directly address the problem of computing square roots.

Another study [10] developed RNS reduction algorithms that take into account properties of quadratic residuosity, which is especially relevant for the Rabin cryptosystem because decryption requires finding quadratic roots.

The authors propose methods that reduce the number of base multiplications in RNS and optimize computations, yet the concrete integration of these algorithms into the Rabin scheme has not been implemented to date.

Research presented in [11] introduced the concept of a multi-layer recursive RNS, where large moduli are decomposed into layers, enabling arithmetic operations on very large numbers to be performed in parallel within smaller subsystems. This approach can substantially shorten computation time for large numeric fields; however, a specific adaptation of these algorithms to the Rabin cryptosystem is not described.

In [12], the application of RNS for accelerating modular operations in isogeny-based cryptography at the hardware level was investigated, demonstrating significant performance gains when working with large numbers. Nevertheless, that work does not consider squaring operations or the particular requirements of the Rabin cryptosystem.

Additionally, work [13] describes the use of the Residue Number System to implement modular operations in RSA, which substantially reduces the execution time of large-number computations. Although that research focused on RSA, its methodology can be adapted to optimize the squaring operation in Rabin, since both systems employ large moduli and modular arithmetic.

The analysis of existing studies shows that applying RNS to optimize modular computations is a promising direction: it can accelerate basic operations, reduce complexity, and enable parallel computation. However, direct application of these methods specifically to the squaring operation in the Rabin cryptosystem has not yet received sufficient scientific treatment. This gap underlines the relevance of further research and the development of novel methods for implementing the squaring operation in Rabin using RNS to improve the speed and efficiency of cryptographic processes [14].

The purpose of the article is to develop and analyze the efficiency of a method for implementing the squaring operation in the Rabin cryptosystem using the Residue Number System, which will help reduce computational complexity and increase the speed of cryptographic processes without compromising the system's security.

Presentation of the main research material. The squaring operation of a number M is based on multiplying the number $M^2 = M \cdot M$. In the RNS, the multiplication of a number $M_1 = (\mu_{11}, \dots, \mu_{1i}, \dots, \mu_{1k})$ by a number $M_2 = (\mu_{21}, \dots, \mu_{2i}, \dots, \mu_{2k})$ is performed by multiplying the corresponding residues μ_{1i} and μ_{2i} modulo the corresponding modulus η_i ($i = \overline{1, k}$) for independently and in parallel across each of the k RNS bases [13, 15]. The range of representable numbers (or the system modulus) is equal to the product of all moduli:

$$N = \eta_1 \cdot \eta_2 \cdot \dots \cdot \eta_k. \quad (5)$$

The multiplication of residues μ_{1i} and μ_{2i} modulo η_i can be efficiently implemented using a table-based method that leverages the symmetry properties of the arithmetic table along its diagonals, rows, and columns. Thanks to this approach, only 25% of the full table is required for implementing modular multiplication. A table multiplication code is used for this purpose [15]. Table 1 is constructed with the numerical values of the first residue μ_{1i} placed horizontally and the second residue μ_{2i} vertically. The cells at the intersections indicate the corresponding results of the modular multiplication $\mu_{1i} \cdot \mu_{2i} \pmod{\eta_i}$.

This table exhibits symmetry with respect to the diagonals, vertical, and horizontal lines passing between the numbers $(\eta_i - 1)/2$ and $(\eta_i + 1)/2$ for an odd η_i . The symmetry along the left diagonal is determined by the commutativity of the multiplication operation, while the symmetry along the right diagonal is defined by the relation $(\eta_i - \mu_{1i})(\eta_i - \mu_{2i}) \equiv \mu_{1i} \cdot \mu_{2i} \pmod{\eta_i}$. The symmetry with respect to the vertical and horizontal axes is due to the fact that the sum of multiples of numbers is divisible by η_i , i.e.,

$$\begin{aligned} \mu_{1i} \cdot \mu_{2i} + \mu_{1i}(\eta_i - \mu_{2i}) &\equiv 0 \pmod{\eta_i}; \\ \mu_{1i} \cdot \mu_{2i} + \mu_{2i}(\eta_i - \mu_{1i}) &\equiv 0 \pmod{\eta_i}. \end{aligned} \quad (6)$$

Thus, to reconstruct the entire table, it is sufficient to know only one-eighth of it, which significantly reduces the size of the table required to implement the multiplication operation.

To achieve this, a special encoding of the residues μ_{1i} and μ_{2i} is applied, known as the “table multiplication code” [15-17]. The values of μ_{1i} and μ_{2i} , within the range $\left[0, (\eta_i - 1)/2\right)$ can be encoded arbitrarily. The values within the range $\left[(\eta_i + 1)/2, \eta_i - 1\right)$, are encoded according to the rule $\mu_{1i} \cdot \mu_{2i} + \mu_{2i}(\eta_i - \mu_{1i}) \equiv 0 \pmod{\eta_i}$.

To distinguish between the ranges, an index (Table Multiplication Code (TMC)) $\zeta_{\mu_{1i}}(\zeta_{\mu_{2i}})$ is introduced, which is defined as follows:

$$\zeta_{\mu_{1i}}, \zeta_{\mu_{2i}} = \begin{cases} 0 & \text{if } 0 \leq \mu_{1i} \leq (\eta_i - 1)/2; \\ 1 & \text{if } (\eta_i + 1)/2 \leq \mu_{1i} \leq \eta_i. \end{cases} \quad (7)$$

Table 1

Table for the implementation of the modular multiplication operation

$\mu_{2i} \backslash \mu_{1i}$	$\mu_{1i}^{(0)}$	...	$\mu_{1i}^{(\frac{\eta-1}{2})}$	$\mu_{1i}^{(\frac{\eta+1}{2})}$	...	$\mu_{1i}^{(\eta-1)}$
$\mu_{2i}^{(0)}$	$(\mu_{1i}^{(0)} \cdot \mu_{2i}^{(0)}) \bmod \eta_i$	...	$(\mu_{1i}^{(\frac{\eta-1}{2})} \cdot \mu_{2i}^{(0)}) \bmod \eta_i$	$(\mu_{1i}^{(\frac{\eta+1}{2})} \cdot \mu_{2i}^{(0)}) \bmod \eta_i$	...	$(\mu_{1i}^{(\eta-1)} \cdot \mu_{2i}^{(0)}) \bmod \eta_i$
...	...	...	...	...	...	...
$\mu_{2i}^{(\frac{\eta-1}{2})}$	$(\mu_{1i}^{(0)} \cdot \mu_{2i}^{(\frac{\eta-1}{2})}) \bmod \eta_i$	...	$(\mu_{1i}^{(\frac{\eta-1}{2})} \cdot \mu_{2i}^{(\frac{\eta-1}{2})}) \bmod \eta_i$	$(\mu_{1i}^{(\frac{\eta+1}{2})} \cdot \mu_{2i}^{(\frac{\eta-1}{2})}) \bmod \eta_i$	...	$(\mu_{1i}^{(\eta-1)} \cdot \mu_{2i}^{(\frac{\eta-1}{2})}) \bmod \eta_i$
$\mu_{2i}^{(\frac{\eta+1}{2})}$	$(\mu_{1i}^{(0)} \cdot \mu_{2i}^{(\frac{\eta+1}{2})}) \bmod \eta_i$	...	$(\mu_{1i}^{(\frac{\eta-1}{2})} \cdot \mu_{2i}^{(\frac{\eta+1}{2})}) \bmod \eta_i$	$(\mu_{1i}^{(\frac{\eta+1}{2})} \cdot \mu_{2i}^{(\frac{\eta+1}{2})}) \bmod \eta_i$	...	$(\mu_{1i}^{(\eta-1)} \cdot \mu_{2i}^{(\frac{\eta+1}{2})}) \bmod \eta_i$
...	...	...	...	...	...	...
$\mu_{2i}^{(\eta-1)}$	$(\mu_{1i}^{(0)} \cdot \mu_{2i}^{(\eta-1)}) \bmod \eta_i$	...	$(\mu_{1i}^{(\frac{\eta-1}{2})} \cdot \mu_{2i}^{(\eta-1)}) \bmod \eta_i$	$(\mu_{1i}^{(\frac{\eta+1}{2})} \cdot \mu_{2i}^{(\eta-1)}) \bmod \eta_i$	...	$(\mu_{1i}^{(\eta-1)} \cdot \mu_{2i}^{(\eta-1)}) \bmod \eta_i$

The method for determining the result of modular multiplication using the TMC is as follows [15].

If two numbers μ_{1i} and μ_{2i} are represented in TMC form as $\mu_{1i} = (\zeta_{\mu_{1i}}, \widetilde{\mu_{1i}})$, $\mu_{2i} = (\zeta_{\mu_{2i}}, \widetilde{\mu_{2i}})$ then, to obtain their product modulo η_i , it is sufficient to compute $(\widetilde{\mu_{1i}} \cdot \widetilde{\mu_{2i}}) \bmod \eta_i$ in the TMC. The index $\widetilde{\zeta_{\mu_i}}$ is inverted if $\widetilde{\mu_{1i}}$ and $\widetilde{\mu_{2i}}$ belong to different ranges, that is, $(\mu_{1i} \cdot \mu_{2i}) \bmod \eta_i = (\zeta_{\mu_i}, (\widetilde{\mu_{1i}} \cdot \widetilde{\mu_{2i}}) \bmod \eta_i)$ where

$$\zeta_{\mu_i} = \begin{cases} \widetilde{\zeta_{\mu_i}} & \text{if } \zeta_{\mu_{1i}} \neq \zeta_{\mu_{2i}}; \\ \zeta_{\mu_i} & \text{if } \zeta_{\mu_{1i}} = \zeta_{\mu_{2i}}; \end{cases} \quad (8)$$

Based on the general mathematical model (7) – (8) for implementing the multiplication of two residues $\mu_{1\Box}$ and μ_{2i} modulo η_i using the table method, we synthesize a mathematical model for multiplying two numbers $M_1 = (\mu_{11}, \dots, \mu_{1i}, \dots, \mu_{1k})$ and $M_2 = (\mu_{21}, \dots, \mu_{2i}, \dots, \mu_{2k})$ in the RNS.

The process of table-based implementation of the multiplication operation of two numbers in the residue number system can be represented as follows [17]:

$$\begin{aligned}
 M &= M_1 \cdot M_2 \bmod N = M_1 = \\
 &= \{(\mu_{11}, \dots, \mu_{1i}, \dots, \mu_{1k}) \cdot (\mu_{21}, \dots, \mu_{2i}, \dots, \mu_{2k})\} \bmod N = \\
 &= \{(\mu_{11} \cdot \mu_{21}) \bmod \eta_1, \dots, (\mu_{1i} \cdot \mu_{2i}) \bmod \eta_i, \dots, (\mu_{1k} \cdot \mu_{2k}) \bmod \eta_k\} = \\
 &= \{(\zeta_{\mu_{11}}, \widetilde{\mu_{11}}) \cdot (\zeta_{\mu_{21}}, \widetilde{\mu_{21}}) \bmod \eta_1, \dots, (\zeta_{\mu_{1i}}, \widetilde{\mu_{1i}}) \cdot (\zeta_{\mu_{2i}}, \widetilde{\mu_{2i}}) \bmod \eta_i, \dots \\
 &\dots (\zeta_{\mu_{1k}}, \widetilde{\mu_{1k}}) \cdot (\zeta_{\mu_{2k}}, \widetilde{\mu_{2k}}) \bmod \eta_k\} = \{[\zeta_{\mu_i}, (\widetilde{\mu_{1i}} \cdot \widetilde{\mu_{2i}}) \bmod \eta_i], \dots \\
 &\dots [\zeta_{\mu_i}, (\widetilde{\mu_{1i}} \cdot \widetilde{\mu_{2i}}) \bmod \eta_i], \dots, [\zeta_{\mu_k}, (\widetilde{\mu_{1k}} \cdot \widetilde{\mu_{2k}}) \bmod \eta_k]\} = \\
 &= (\mu_1, \dots, \mu_i, \dots, \mu_k).
 \end{aligned} \quad (9)$$

In this case, for an even number η_i :

$$\widetilde{\mu_{1i}} \cdot \widetilde{\mu_{2i}} = \begin{cases} \mu_{1i} \cdot \mu_{2i}, & \text{if } 0 \leq \mu_{1i}, \mu_{2i} \leq \eta_i/2, \\ \eta - \mu_{1i}, \eta - \mu_{2i}, & \text{if } \eta_i/2 < \mu_{1i}, \mu_{2i} \leq \eta_i - 1, \end{cases} \quad (10)$$

$$\zeta_{\mu_{1i}}, \zeta_{\mu_{2i}} = \begin{cases} 0 & \text{if } 0 \leq \mu_{1i}, \mu_{2i} \leq \eta_i/2, \\ 1 & \text{if } \eta_i/2 < \mu_{1i}, \mu_{2i} \leq \eta_i - 1, \end{cases} \quad (11)$$

in this case $0 \leq \widetilde{\mu_{1i}}, \widetilde{\mu_{2i}} \leq \eta_i/2$.

If η_i is an odd number, then:

$$\widetilde{\mu_{1i}}, \widetilde{\mu_{2i}} = \begin{cases} \mu_{1i}, \mu_{2i}, & \text{if } 0 \leq \mu_{1i}, \mu_{2i} \leq \eta_i - 1/2, \\ \eta - \mu_{1i}, \eta - \mu_{2i}, & \text{if } \eta_i + 1/2 \leq \mu_{1i}, \mu_{2i} \leq \eta_i - 1, \end{cases} \quad (12)$$

$$\zeta_{\mu_{1i}}, \zeta_{\mu_{2i}} = \begin{cases} 0 & \text{if } 0 \leq \mu_{1i}, \mu_{2i} \leq \eta_i - 1/2, \\ 1 & \text{if } \eta_i + 1/2 \leq \mu_{1i}, \mu_{2i} \leq \eta_i - 1, \end{cases} \quad (13)$$

in this case $0 \leq \widetilde{\mu_{1i}}, \widetilde{\mu_{2i}} \leq \eta_i - 1/2$.

Thus, the relationships (9) – (13) represent the mathematical model of the process of multiplying two numbers $M_1 = (\mu_{11}, \dots, \mu_{1i}, \dots, \mu_{1k})$ and $M_2 = (\mu_{21}, \dots, \mu_{2i}, \dots, \mu_{2k})$. Let us now consider the case where the number M needs to be squared:

$$\begin{aligned} M^2 \pmod{N} &= M \cdot M \pmod{N} = \\ &= \{(\mu_1, \dots, \mu_i, \dots, \mu_k) \cdot (\mu_1, \dots, \mu_i, \dots, \mu_k)\} \pmod{N} = \\ &= \{(\mu_1 \cdot \mu_1) \pmod{\eta_1}, \dots, (\mu_i \cdot \mu_i) \pmod{\eta_i}, \dots, (\mu_k \cdot \mu_k) \pmod{\eta_k}\}. \end{aligned} \quad (14)$$

First, we show that for any residue in the RNS, the following mathematical relationship holds [17]:

$$\mu_i^2 \pmod{\eta_i} = (\eta_i - \mu_i)^2 \pmod{\eta_i} \quad (15)$$

Indeed, the square of the residue μ_i^2 can be expressed as $\mu_i^2 = \Omega_i \eta_i + \Psi_i$ ($0 \leq \Psi_i \leq \eta_i - 1$), which means that $\mu_i^2 \pmod{\eta_i} = \Psi_i$. Then, $\eta_i^2 - 2\eta_i \mu_i + \mu_i^2 = \eta_i(\eta_i - 2\mu_i + \Omega_i) + \Psi_i$. In this case, $(\eta_i^2 - 2\eta_i \mu_i + \mu_i^2) \pmod{\eta_i} = \Psi_i$. Thus, this relationship holds true for both even and odd values of η_i . The analytical expression (15) represents a mathematical model of the process used to implement $\mu_i^2 \pmod{\eta_i}$ in practical computations [18]. The structural diagram of this operation $M^2 \pmod{N}$ is shown in fig. 1. The operation of the device proceeds as follows: a number μ_i in binary code is fed to the input and stored in the input register. Then, from the output of the decoder, the number μ_i in unary code passes through the corresponding OR logic elements to the input of the encoder, which corresponds to the value $\mu_i^2 \pmod{\eta_i}$. From the encoder's output, the obtained value $\mu_i^2 \pmod{\eta_i}$ in binary form is sent to the output register. It is evident that the key element in the technical implementation of the operation $\mu_i^2 \pmod{\eta_i}$ is the correct coding of the connections between the decoder and the encoder.

Let us consider an example of the practical application of the developed method for exponentiation in the RNS for the Rabin cryptosystem, with the bases $\eta_1 = 3$, $\eta_2 = 5$, $\eta_3 = 7$, and $N = \eta_1 \cdot \eta_2 \cdot \eta_3 = 3 \cdot 5 \cdot 7$. The range of code words in the RNS is presented in table 2. The algorithm for implementing the operation of squaring a number modulo $M^2 \pmod{N}$ for $\eta_1 = 3$, $\eta_2 = 5$, $\eta_3 = 7$, is presented in table 3.

Let us consider examples of squaring numbers using the proposed method. Suppose we need to calculate 7^2 . In the RNS, this number 7 is represented as $M = (1, 2, 0)$. In binary code, this value is $M = (01_2, 010_2, 000_2)$ and is fed to the inputs of the respective decoders. At the output of the encoders, we obtain $M^2 = (01_2, 100_2, 000_2)$ in binary codes, which corresponds to $M^2 = (1_{10}, 4_{10}, 0_{10})$ in decimal. From table 4, it is seen that $(1, 4, 0)$ in RNS corresponds to 49 in the positional number system, i.e., $7^2 = 49$, confirming that the result is correct.

Let us determine the time gain when performing the squaring operation in the Rabin cryptographic system as the ratio between the execution time of the operation in the positional number system and that achieved using the proposed method in the RNS. It is known that in a positional number system, the execution time of a multiplication operation is given by $t_{PNS} = 2\sigma^2 \tau$ where σ is the number of binary digits in the operand representation, and τ is the propagation delay of the “OR” logic elements. In the RNS, due to the parallel nature of computations, the time required for squaring a number is $t_{RNS} = \tau$ where τ is again the propagation delay of the “OR” logic elements. Thus, when using 32-bit operands, the performance gain can be calculated as $\frac{t_{PNS}}{t_{RNS}} = 2\sigma^2 \tau / \tau = 2 \cdot 32^2 = 2048$ times.

Similarly, for 64-bit operands, the performance gain is $\frac{t_{PNS}}{t_{RNS}} = 2 \cdot 64^2 \tau / \tau = 8192$ times. Therefore, the proposed squaring method demonstrates a significant improvement in computational performance, making it highly suitable for practical implementation in Rabin cryptosystems.

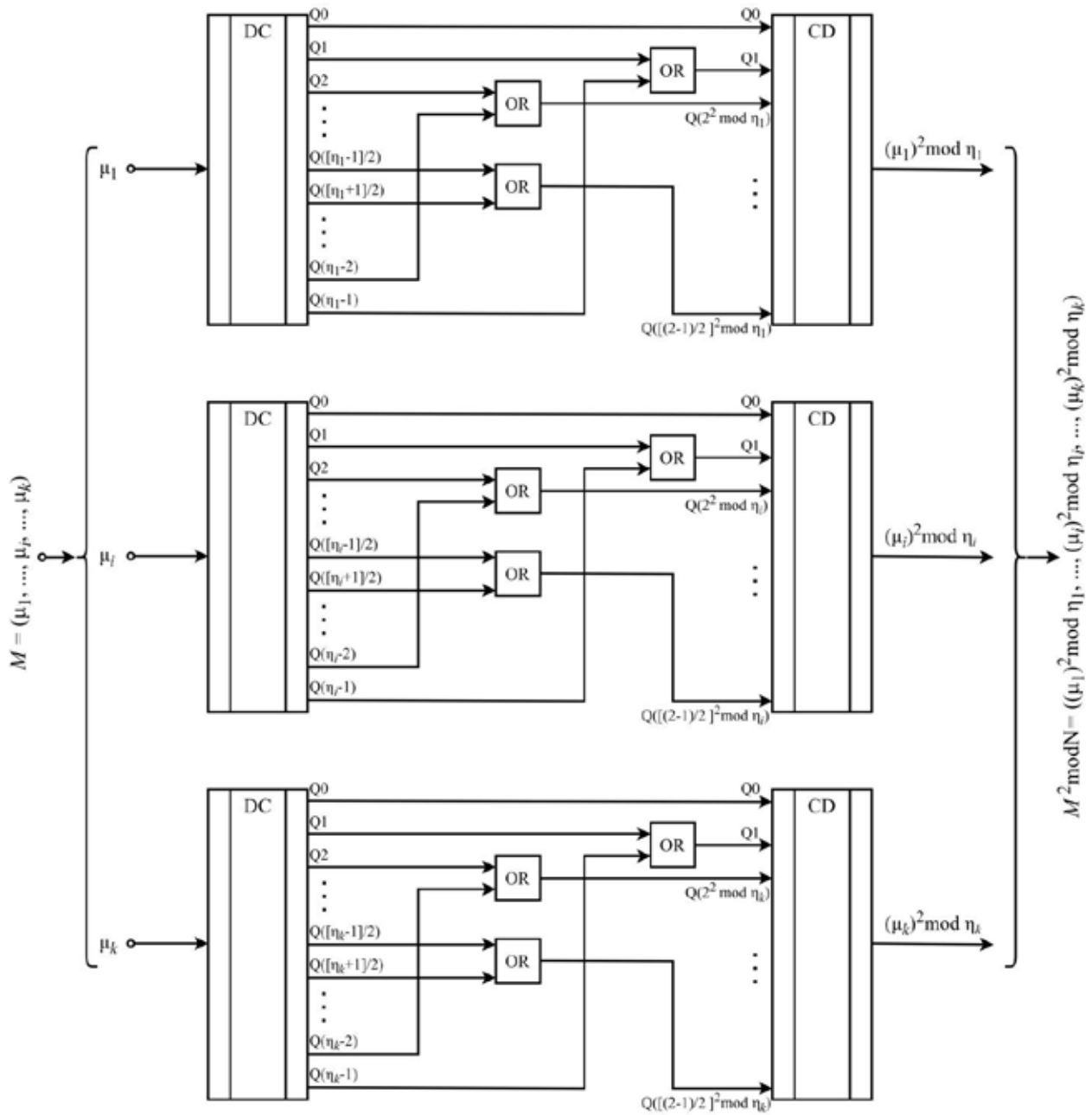


Fig. 1. Scheme of the technical implementation of the operation $M^2(\text{mod } N)$

Table 2

Table of code words in the RNS with the moduli $\eta_1 = 3, \eta_2 = 5, \eta_3 = 7$

$M^{(PNS)}$	$M^{(RNS)}$						$M^{(PNS)}$	$M^{(RNS)}$					
	$\mu_1 = 3$		$\mu_1 = 5$		$\mu_1 = 7$			$\mu_1 = 3$		$\mu_1 = 5$		$\mu_1 = 7$	
0	0 ₁₀	00 ₂	0 ₁₀	000 ₂	0 ₁₀	000 ₂	53	2 ₁₀	10 ₂	3 ₁₀	011 ₂	4 ₁₀	100 ₂
1	1 ₁₀	01 ₂	1 ₁₀	001 ₂	1 ₁₀	001 ₂	54	0 ₁₀	00 ₂	4 ₁₀	100 ₂	5 ₁₀	101 ₂
2	2 ₁₀	10 ₂	2 ₁₀	010 ₂	2 ₁₀	010 ₂	55	1 ₁₀	01 ₂	0 ₁₀	000 ₂	6 ₁₀	110 ₂
3	0 ₁₀	00 ₂	3 ₁₀	011 ₂	3 ₁₀	011 ₂	56	2 ₁₀	10 ₂	1 ₁₀	001 ₂	0 ₁₀	000 ₂
4	1 ₁₀	01 ₂	4 ₁₀	100 ₂	4 ₁₀	100 ₂	57	0 ₁₀	00 ₂	2 ₁₀	010 ₂	1 ₁₀	001 ₂
5	2 ₁₀	10 ₂	0 ₁₀	000 ₂	5 ₁₀	101 ₂	58	1 ₁₀	01 ₂	3 ₁₀	011 ₂	2 ₁₀	010 ₂
6	0 ₁₀	00 ₂	1 ₁₀	001 ₂	6 ₁₀	110 ₂	59	2 ₁₀	10 ₂	4 ₁₀	100 ₂	3 ₁₀	011 ₂
7	1 ₁₀	01 ₂	2 ₁₀	010 ₂	0 ₁₀	000 ₂	60	0 ₁₀	00 ₂	0 ₁₀	000 ₂	4 ₁₀	100 ₂
8	2 ₁₀	10 ₂	3 ₁₀	011 ₂	1 ₁₀	001 ₂	61	1 ₁₀	01 ₂	1 ₁₀	001 ₂	5 ₁₀	101 ₂
9	0 ₁₀	00 ₂	4 ₁₀	100 ₂	2 ₁₀	010 ₂	62	2 ₁₀	10 ₂	2 ₁₀	010 ₂	6 ₁₀	110 ₂
10	1 ₁₀	01 ₂	0 ₁₀	000 ₂	3 ₁₀	011 ₂	63	0 ₁₀	00 ₂	3 ₁₀	011 ₂	0 ₁₀	000 ₂
11	2 ₁₀	10 ₂	1 ₁₀	001 ₂	4 ₁₀	100 ₂	64	1 ₁₀	01 ₂	4 ₁₀	100 ₂	1 ₁₀	001 ₂
12	0 ₁₀	00 ₂	2 ₁₀	010 ₂	5 ₁₀	101 ₂	65	2 ₁₀	10 ₂	0 ₁₀	000 ₂	2 ₁₀	010 ₂
13	1 ₁₀	01 ₂	3 ₁₀	011 ₂	6 ₁₀	110 ₂	66	0 ₁₀	00 ₂	1 ₁₀	001 ₂	3 ₁₀	011 ₂
14	2 ₁₀	10 ₂	4 ₁₀	100 ₂	0 ₁₀	000 ₂	67	1 ₁₀	01 ₂	2 ₁₀	010 ₂	4 ₁₀	100 ₂
15	0 ₁₀	00 ₂	0 ₁₀	000 ₂	1 ₁₀	001 ₂	68	2 ₁₀	10 ₂	3 ₁₀	011 ₂	5 ₁₀	101 ₂
16	1 ₁₀	01 ₂	1 ₁₀	001 ₂	2 ₁₀	010 ₂	69	0 ₁₀	00 ₂	4 ₁₀	100 ₂	6 ₁₀	110 ₂
17	2 ₁₀	10 ₂	2 ₁₀	010 ₂	3 ₁₀	011 ₂	70	1 ₁₀	01 ₂	0 ₁₀	000 ₂	0 ₁₀	000 ₂
18	0 ₁₀	00 ₂	3 ₁₀	011 ₂	4 ₁₀	100 ₂	71	2 ₁₀	10 ₂	1 ₁₀	001 ₂	1 ₁₀	001 ₂
19	1 ₁₀	01 ₂	4 ₁₀	100 ₂	5 ₁₀	101 ₂	72	0 ₁₀	00 ₂	2 ₁₀	010 ₂	2 ₁₀	010 ₂
20	2 ₁₀	10 ₂	0 ₁₀	000 ₂	6 ₁₀	110 ₂	73	1 ₁₀	01 ₂	3 ₁₀	011 ₂	3 ₁₀	011 ₂
21	0 ₁₀	00 ₂	1 ₁₀	001 ₂	0 ₁₀	000 ₂	74	2 ₁₀	10 ₂	4 ₁₀	100 ₂	4 ₁₀	100 ₂
22	1 ₁₀	01 ₂	2 ₁₀	010 ₂	1 ₁₀	001 ₂	75	0 ₁₀	00 ₂	0 ₁₀	000 ₂	5 ₁₀	101 ₂
23	2 ₁₀	10 ₂	3 ₁₀	011 ₂	2 ₁₀	010 ₂	76	1 ₁₀	01 ₂	1 ₁₀	001 ₂	6 ₁₀	110 ₂
24	0 ₁₀	00 ₂	4 ₁₀	100 ₂	3 ₁₀	011 ₂	77	2 ₁₀	10 ₂	2 ₁₀	010 ₂	0 ₁₀	000 ₂
25	1 ₁₀	01 ₂	0 ₁₀	000 ₂	4 ₁₀	100 ₂	78	0 ₁₀	00 ₂	3 ₁₀	011 ₂	1 ₁₀	001 ₂
26	2 ₁₀	10 ₂	1 ₁₀	001 ₂	5 ₁₀	101 ₂	79	1 ₁₀	01 ₂	4 ₁₀	100 ₂	2 ₁₀	010 ₂
27	0 ₁₀	00 ₂	2 ₁₀	010 ₂	6 ₁₀	110 ₂	80	2 ₁₀	10 ₂	0 ₁₀	000 ₂	3 ₁₀	011 ₂
28	1 ₁₀	01 ₂	3 ₁₀	011 ₂	0 ₁₀	000 ₂	81	0 ₁₀	00 ₂	1 ₁₀	001 ₂	4 ₁₀	100 ₂
29	2 ₁₀	10 ₂	4 ₁₀	100 ₂	1 ₁₀	001 ₂	82	1 ₁₀	01 ₂	2 ₁₀	010 ₂	5 ₁₀	101 ₂
30	0 ₁₀	00 ₂	0 ₁₀	000 ₂	2 ₁₀	010 ₂	83	2 ₁₀	10 ₂	3 ₁₀	011 ₂	6 ₁₀	110 ₂
31	1 ₁₀	01 ₂	1 ₁₀	001 ₂	3 ₁₀	011 ₂	84	0 ₁₀	00 ₂	4 ₁₀	100 ₂	0 ₁₀	000 ₂
32	2 ₁₀	10 ₂	2 ₁₀	010 ₂	4 ₁₀	100 ₂	85	1 ₁₀	01 ₂	0 ₁₀	000 ₂	1 ₁₀	001 ₂
33	0 ₁₀	00 ₂	3 ₁₀	011 ₂	5 ₁₀	101 ₂	86	2 ₁₀	10 ₂	1 ₁₀	001 ₂	2 ₁₀	010 ₂
34	1 ₁₀	01 ₂	4 ₁₀	100 ₂	6 ₁₀	110 ₂	87	0 ₁₀	00 ₂	2 ₁₀	010 ₂	3 ₁₀	011 ₂
35	2 ₁₀	10 ₂	0 ₁₀	000 ₂	0 ₁₀	000 ₂	88	1 ₁₀	01 ₂	3 ₁₀	011 ₂	4 ₁₀	100 ₂
36	0 ₁₀	00 ₂	1 ₁₀	001 ₂	1 ₁₀	001 ₂	89	2 ₁₀	10 ₂	4 ₁₀	100 ₂	5 ₁₀	101 ₂
37	1 ₁₀	01 ₂	2 ₁₀	010 ₂	2 ₁₀	010 ₂	90	0 ₁₀	00 ₂	0 ₁₀	000 ₂	6 ₁₀	110 ₂
38	2 ₁₀	10 ₂	3 ₁₀	011 ₂	3 ₁₀	011 ₂	91	1 ₁₀	01 ₂	1 ₁₀	001 ₂	0 ₁₀	000 ₂
39	0 ₁₀	00 ₂	4 ₁₀	100 ₂	4 ₁₀	100 ₂	92	2 ₁₀	10 ₂	2 ₁₀	010 ₂	1 ₁₀	001 ₂
40	1 ₁₀	01 ₂	0 ₁₀	000 ₂	5 ₁₀	101 ₂	93	0 ₁₀	00 ₂	3 ₁₀	011 ₂	2 ₁₀	010 ₂
41	2 ₁₀	10 ₂	1 ₁₀	001 ₂	6 ₁₀	110 ₂	94	1 ₁₀	01 ₂	4 ₁₀	100 ₂	3 ₁₀	011 ₂
42	0 ₁₀	00 ₂	2 ₁₀	010 ₂	0 ₁₀	000 ₂	95	2 ₁₀	10 ₂	0 ₁₀	000 ₂	4 ₁₀	100 ₂
43	1 ₁₀	01 ₂	3 ₁₀	011 ₂	1 ₁₀	001 ₂	96	0 ₁₀	00 ₂	1 ₁₀	001 ₂	5 ₁₀	101 ₂
44	2 ₁₀	10 ₂	4 ₁₀	100 ₂	2 ₁₀	010 ₂	97	1 ₁₀	01 ₂	2 ₁₀	010 ₂	6 ₁₀	110 ₂
45	0 ₁₀	00 ₂	0 ₁₀	000 ₂	3 ₁₀	011 ₂	98	2 ₁₀	10 ₂	3 ₁₀	011 ₂	0 ₁₀	000 ₂
46	1 ₁₀	01 ₂	1 ₁₀	001 ₂	4 ₁₀	100 ₂	99	0 ₁₀	00 ₂	4 ₁₀	100 ₂	1 ₁₀	001 ₂
47	2 ₁₀	10 ₂	2 ₁₀	010 ₂	5 ₁₀	101 ₂	100	1 ₁₀	01 ₂	0 ₁₀	000 ₂	2 ₁₀	010 ₂
48	0 ₁₀	00 ₂	3 ₁₀	011 ₂	6 ₁₀	110 ₂	101	2 ₁₀	10 ₂	1 ₁₀	001 ₂	3 ₁₀	011 ₂
49	1 ₁₀	01 ₂	4 ₁₀	100 ₂	0 ₁₀	000 ₂	102	0 ₁₀	00 ₂	2 ₁₀	010 ₂	4 ₁₀	100 ₂
50	2 ₁₀	10 ₂	0 ₁₀	000 ₂	1 ₁₀	001 ₂	103	1 ₁₀	01 ₂	3 ₁₀	011 ₂	5 ₁₀	101 ₂
51	0 ₁₀	00 ₂	1 ₁₀	001 ₂	2 ₁₀	010 ₂	104	2 ₁₀	10 ₂	4 ₁₀	100 ₂	6 ₁₀	110 ₂
52	1 ₁₀	01 ₂	2 ₁₀	010 ₂	3 ₁₀	011 ₂							

Table 3

**The algorithm for implementing the operation of squaring a number modulo $M^2 \pmod N$
for $\eta_1 = 3, \eta_2 = 5, \eta_3 = 7$**

Value of the modulus η_i	Value of the remainder μ_i that is applied to the input of the decoder	Bus number with a "1" value at the decoder output	Value at the encoder input	Value $\mu_i^2 \pmod{\eta_i}$ generated at the encoder output
3	$0_{10} = 00_2$	0	00	0
	$1_{10} = 01_2$	1	01	1
	$2_{10} = 10_2$	2	01	1
5	$0_{10} = 000_2$	0	000	0
	$1_{10} = 001_2$	1	001	1
	$2_{10} = 010_2$	2	100	4
	$3_{10} = 011_2$	3	100	4
	$4_{10} = 100_2$	4	001	1
7	$0_{10} = 000_2$	0	000	0
	$1_{10} = 001_2$	1	001	1
	$2_{10} = 010_2$	2	100	4
	$3_{10} = 011_2$	3	010	2
	$4_{10} = 100_2$	4	010	2
	$5_{10} = 101_2$	5	100	4
	$6_{10} = 110_2$	6	001	1

Conclusions. In the article research investigates the process of implementing the squaring operation in the Rabin cryptosystem using the RNS. The proposed method is based on transitioning from the positional number system to RNS, which enables arithmetic operations to be performed independently for each modulus and ensures a high level of computational parallelism.

A mathematical model of the process of multiplying two numbers represented in the RNS has been developed based on the application of the table multiplication code. This approach takes into account the symmetry properties of the table multiplication, which allows reducing the amount of necessary calculations to 25% of the full table. Based on the mathematical model of multiplying two numbers in the RNS, a method of squaring numbers in the RNS has been proposed.

A comparative analysis has shown that the use of RNS provides a significant increase in performance. For 32-bit operands, the acceleration is 2048 times, and for 64-bit operands – up to 8192 times compared to performing the same operation in a positional number system.

The obtained results confirm the feasibility of applying the Residue Number System in the Rabin cryptosystem for implementing basic arithmetic operations, particularly the squaring operation. The proposed approach increases the efficiency of cryptographic computations without reducing the system's security.

Further research should focus on developing a universal modular computation framework in RNS for other cryptosystems, as well as on the hardware implementation of the proposed method to evaluate its performance under real operating conditions.

Bibliography:

1. Usmani Z. A., Rai M., Khan F. Reconfigurable Cryptoprocessor Design for RSA and Rabin-p Cryptosystems. 2024 26th International Multi-Topic Conference (INMIC), Karachi, Pakistan, 2024. P. 1–6. DOI: 10.1109/INMIC64792.2024.11004308.
2. Yakymenko M., Kasianchuk I., Shylinska R., Shevchuk V., Yatskiv V., Karpinski M. Polynomial Rabin Cryptosystem Based on the Operation of Addition. 2022 12th International Conference on Advanced Computer Information Technologies (ACIT), Ruzomberok, Slovakia, 2022. P. 345–350. DOI: 10.1109/ACIT54803.2022.9913089.
3. Schoiniakakis D. Residue arithmetic systems in cryptography: a survey on modern security applications. *Journal of Cryptographic Engineering*, 2020, Vol. 10, P. 249–267. DOI: 10.1007/s13389-020-00231-w.
4. Yakymenko M., Kasianchuk O., Martyniuk S., Martyniuk A., Martyniuk Y., Yakymenko Y. A Symmetric Cryptoalgorithm in a Polynomial Hierarchical Residual Number System. 2025 15th International Conference

on Advanced Computer Information Technologies (ACIT), Sibenik, Croatia, 2025. P. 501–504. DOI: 10.1109/ACIT65614.2025.11185808.

5. Yatskiv V., Yatskiv N., Ivasiev S., Kulyna S., Tsavolyk T., Yatskiv I. The McEliece Cryptosystem Based on the Redundant Residue Number System. 2025 15th International Conference on Advanced Computer Information Technologies (ACIT), Sibenik, Croatia, 2025. P. 573–577. DOI: 10.1109/ACIT65614.2025.11185887.

6. Nykolaychuk Y. M., Yakymenko I. Z., Vozna N. Y. et al. Residue Number System Asymmetric Crypt algorithms. *Cybernetics and Systems Analysis*, 2022, Vol. 58, No. 4, P. 611–618. DOI: 10.1007/s10559-022-00494-7.

7. Zhan J., Shiue P. J., Huang S. C., Lowe B. J. Towards a Novel Generalized Chinese Remainder Algorithm for Extended Rabin Cryptosystem. *IEEE Access*, 2020, Vol. 8, P. 26433–26444. DOI: 10.1109/ACCESS.2020.2967396.

8. Nykolaychuk Y. M., Yakymenko I. Z., Vozna N. Y., Kasianchuk M. M. Residue number system asymmetric crypt algorithms. *Cybernetics and Systems Analysis*, 2022, Vol. 58, No. 4, P. 611–618.

9. Selianinau M., Woźna-Szcześniak B. An Efficient Implementation of Montgomery Modular Multiplication Using a Minimally Redundant Residue Number System. *Applied Sciences*, 2025, Vol. 15, No. 10, 5332. DOI: 10.3390/app15105332.

10. Kawamura S., Komano Y., Shimizu H. et al. RNS Montgomery Reduction Algorithms Using Quadratic Residuosity. *Journal of Cryptographic Engineering*, 2019, Vol. 9, P. 313–331. DOI: 10.1007/s13389-018-0195-8.

11. Hollmann H. D. L., Rietman R., de Hoogh S., Tolhuizen L. M. G. M., Gorissen P. A Multi-layer Recursive Residue Number System. arXiv, 2018. DOI: 10.48550/arxiv.1801.07561.

12. Jacquemin D., Mert A. C., Roy S. S. Exploring RNS for Isogeny-Based Cryptography. Cryptology ePrint Archive, 2022. URL: <https://eprint.iacr.org/2022/1289>.

13. Krasnobayev V., Yanko A., Koshman S. Conception of Realization of Cryptographic RSA Transformations with Using of the Residue Number System. *Computer Science and Cybersecurity*, 2016, No. 2, P. 5–12. URL: <https://periodicals.karazin.ua/cscs/article/view/6207>.

14. Yatskiv V., Kulyna S., Bykovyy P., Maksymyuk T., Sachenko A. Method of Reliable Data Storage Based on Redundant Residue Number System. 2020 IEEE 5th International Symposium on Smart and Wireless Systems within the Conferences on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS-SWS), Dortmund, Germany, 2020. P. 1–4. DOI: 10.1109/IDAACS-SWS50031.2020.9297052.

15. Krasnobayev V., Yanko A., Kovalchuk D. Method of Tabular Implementation of the Arithmetic Operation of Multiplying Two Numbers Represented in the System of Residual Classes. 2022 IEEE 9th International Conference on Problems of Infocommunications, Science and Technology (PIC S&T), Kharkiv, Ukraine, 2022. P. 63–68. DOI: 10.1109/PICST57299.2022.10238624.

16. Yatskiv V., Kulyna S., Yatskiv N., Kulyna H. Protected Distributed Data Storage Based on Residue Number System and Cloud Services. 2020 10th International Conference on Advanced Computer Information Technologies (ACIT), Deggendorf, Germany, 2020. P. 796–799. DOI: 10.1109/ACIT49673.2020.9208849.

17. Krasnobayev V. A., Yanko A. S., Kovalchuk D. M. Mathematical Model of the Process of Raising Integers to an Arbitrary Power of a Natural Number in the System of Residual Classes. *Theoretical and Applied Cybersecurity*, 2023, Vol. 5, No. 2, P. 5–14. DOI: 10.20535/tacs.2664-29132023.2.278891.

18. Ochoa-Jiménez E., Rivera-Zamarripa L., Cruz-Cortés N., Rodríguez-Henríquez F. Implementation of RSA Signatures on GPU and CPU Architectures. *IEEE Access*, 2020, Vol. 8, P. 9928–9941. DOI: 10.1109/ACCESS.2019.2963826.

References:

1. Usmani, Z. A., Rai, M., & Khan, F. (2024). Reconfigurable cryptoprocessor design for RSA and Rabin-p cryptosystems. 2024 26th International Multi-Topic Conference (INMIC), Karachi, Pakistan, 1–6. <https://doi.org/10.1109/INMIC64792.2024.11004308>

2. Yakymenko, M., Kasianchuk, I., Shylinska, R., Shevchuk, V., Yatskiv, V., & Karpinski, M. (2022). Polynomial Rabin cryptosystem based on the operation of addition. 2022 12th International Conference on Advanced Computer Information Technologies (ACIT), Ruzomberok, Slovakia, 345–350. <https://doi.org/10.1109/ACIT54803.2022.9913089>

3. Schoinianakis, D. (2020). Residue arithmetic systems in cryptography: A survey on modern security applications. *Journal of Cryptographic Engineering*, 10(3), 249–267. <https://doi.org/10.1007/s13389-020-00231-w>

4. Yakymenko, M., Kasianchuk, O., Martyniuk, S., Martyniuk, A., Martyniuk, Y., & Yakymenko, Y. (2025). A symmetric crypt algorithm in a polynomial hierarchical residual number system. 2025 15th International Conference on Advanced Computer Information Technologies (ACIT), Sibenik, Croatia, 501–504. <https://doi.org/10.1109/ACIT65614.2025.11185808>

5. Yatskiv, V., Yatskiv, N., Ivasiev, S., Kulyna, S., Tsavolyk, T., & Yatskiv, I. (2025). The McEliece cryptosystem based on the redundant residue number system. 2025 15th International Conference on Advanced Computer Information Technologies (ACIT), Sibenik, Croatia, 573–577. <https://doi.org/10.1109/ACIT65614.2025.11185887>

-
6. Nykolaychuk, Y. M., Yakymenko, I. Z., Vozna, N. Y., et al. (2022). Residue number system asymmetric cryptosystems. *Cybernetics and Systems Analysis*, 58(4), 611–618. <https://doi.org/10.1007/s10559-022-00494-7>
 7. Zhan, J., Shiue, P. J., Huang, S. C., & Lowe, B. J. (2020). Towards a novel generalized Chinese remainder algorithm for extended Rabin cryptosystem. *IEEE Access*, 8, 26433–26444. <https://doi.org/10.1109/ACCESS.2020.2967396>
 8. Nykolaychuk, Y. M., Yakymenko, I. Z., Vozna, N. Y., & Kasianchuk, M. M. (2022). Residue number system asymmetric cryptosystems. *Cybernetics and Systems Analysis*, 58(4), 611–618.
 9. Selianinau, M., & Woźna-Szcześniak, B. (2025). An efficient implementation of Montgomery modular multiplication using a minimally redundant residue number system. *Applied Sciences*, 15(10), 5332. <https://doi.org/10.3390/app15105332>
 10. Kawamura, S., Komano, Y., Shimizu, H., et al. (2019). RNS Montgomery reduction algorithms using quadratic residuosity. *Journal of Cryptographic Engineering*, 9(3), 313–331. <https://doi.org/10.1007/s13389-018-0195-8>
 11. Hollmann, H. D. L., Rietman, R., de Hoogh, S., Tolhuizen, L. M. G. M., & Gorissen, P. (2018). A multi-layer recursive residue number system. arXiv. <https://doi.org/10.48550/arxiv.1801.07561>
 12. Jacquemin, D., Mert, A. C., & Roy, S. S. (2022). Exploring RNS for isogeny-based cryptography. Cryptology ePrint Archive. <https://eprint.iacr.org/2022/1289>
 13. Krasnobayev, V., Yanko, A., & Koshman, S. (2016). Conception of realization of cryptographic RSA transformations with using of the residue number system. *Computer Science and Cybersecurity*, 2, 5–12. <https://periodicals.karazin.ua/cscs/article/view/6207>
 14. Yatskiv, V., Kulyna, S., Bykovyy, P., Maksymyuk, T., & Sachenko, A. (2020). Method of reliable data storage based on redundant residue number system. 2020 IEEE 5th International Symposium on Smart and Wireless Systems within the Conferences on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS-SWS), Dortmund, Germany, 1–4. <https://doi.org/10.1109/IDAACS-SWS50031.2020.9297052>
 15. Krasnobayev, V., Yanko, A., & Kovalchuk, D. (2022). Method of tabular implementation of the arithmetic operation of multiplying two numbers represented in the system of residual classes. 2022 IEEE 9th International Conference on Problems of Infocommunications, Science and Technology (PIC S&T), Kharkiv, Ukraine, 63–68. <https://doi.org/10.1109/PICST57299.2022.10238624>
 16. Yatskiv, V., Kulyna, S., Yatskiv, N., & Kulyna, H. (2020). Protected distributed data storage based on residue number system and cloud services. 2020 10th International Conference on Advanced Computer Information Technologies (ACIT), Deggendorf, Germany, 796–799. <https://doi.org/10.1109/ACIT49673.2020.9208849>
 17. Krasnobayev, V. A., Yanko, A. S., & Kovalchuk, D. M. (2023). Mathematical model of the process of raising integers to an arbitrary power of a natural number in the system of residual classes. *Theoretical and Applied Cybersecurity*, 5(2), 5–14. <https://doi.org/10.20535/tacs.2664-29132023.2.278891>
 18. Ochoa-Jiménez, E., Rivera-Zamarripa, L., Cruz-Cortés, N., & Rodríguez-Henríquez, F. (2020). Implementation of RSA signatures on GPU and CPU architectures. *IEEE Access*, 8, 9928–9941. <https://doi.org/10.1109/ACCESS.2019.2963826>

Дата надходження статті: 19.10.2022

Дата прийняття статті: 10.11.2025

Опубліковано: 30.12.2025

Рудницький В. М., доктор технічних наук, професор,
головний науковий співробітник
Державного науково-дослідного інституту випробувань
і сертифікації озброєння та військової техніки
ORCID: 0000-0003-3473-7433

Тарасенко Я. В., доктор технічних наук, доцент,
провідний науковий співробітник
Державного науково-дослідного інституту випробувань
і сертифікації озброєння та військової техніки
ORCID: 0000-0002-5902-8628

Лада Н. В., кандидат технічних наук,
провідний науковий співробітник
Державного науково-дослідного інституту випробувань
і сертифікації озброєння та військової техніки
ORCID: 0000-0002-7682-2970

Бабенко В. Г., доктор технічних наук, професор,
професор кафедри інформаційної безпеки та комп'ютерної
інженерії
Черкаського державного технологічного університету
ORCID: 0000-0003-2039-2841

Підласий Д. А., аспірант кафедри інформаційної безпеки
та комп'ютерної інженерії
Черкаського державного технологічного університету
ORCID: 0000-0002-9916-5256

АНАЛІЗ РЕЗУЛЬТАТІВ МОДЕЛЮВАННЯ СЕТ-ОПЕРАЦІЙ НА ОСНОВІ ЕЛЕМЕНТАРНИХ ФУНКЦІЙ ОПЕРАЦІЙ КЕРОВАНИХ ІНФОРМАЦІЄЮ

У статті розглянуто актуальну на сьогоднішній день задачу синтезу СЕТ-операцій на основі елементарних функцій операцій керованих інформацією. Роботу присвячено питанням зменшення складності генерації максимальної кількості СЕТ-операцій на основі елементарних функцій операцій керованих інформацією для малоресурсних потокових шифрів. У роботі було проведено моделювання СЕТ-операцій та отримано чотири підгрупи по 10 симетричних операцій. Будь яка з підгруп може бути використана для побудови 48 СЕТ-операцій, які включають усі 10 операцій цієї підгрупи. Проведено аналіз обчислювального експерименту щодо моделювання СЕТ-операцій на основі елементарних функцій операцій керованих інформацією. Встановлено відсутність перетину серед синтезованих множин СЕТ-операцій, на основі модифікацій СЕТ-операцій з різних підгруп. На основі вибору чотирьох СЕТ-операцій (по одній з кожної групи) існує можливість синтезування повної групи СЕТ-операцій на основі елементарних функцій керованих інформацією. Така група включає 192 операції. У роботі наведено групу елементарних функцій операцій керованих інформацією та підмножини симетричних СЕТ-операцій на основі елементарних функцій операцій керованих інформацією. Описано вимоги, на основі яких необхідно проводити вибір базової групи для побудови методу синтезу СЕТ-операцій на основі елементарних функцій операцій керованих інформацією. Наведено умови розгляду вимог до СЕТ-операцій базової групи. За результатами обчислювального експерименту розроблено метод вибору базових груп симетричних СЕТ-операцій, що дає можливість зменшити складність генерації максимальної кількості СЕТ-операцій на основі елементарних функцій операцій керованих інформацією для малоресурсних потокових шифрів. Розроблений метод можливо застосовувати в малоресурсних телекомунікаційних системах як мобільного, так і стаціонарного типу. Перспективою подальших досліджень з напрямку є побудова малоресурсних систем криптографічного перетворення інформації із застосуванням СЕТ-операцій на основі елементарних функцій операцій керованих інформацією.

Ключові слова: СЕТ-операції, операції керовані інформацією, Сі-квантові елементарні функції, автоматизований синтез операцій, обчислювальний експеримент.

© В. М. Рудницький, Я. В. Тарасенко, Н. В. Лада, В. Г. Бабенко, Д. А. Підласий, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

Rudnytskyi V. M., Tarasenko Ya. V., Lada N. V., Babenko V. H., Pidlasyi D. A. Analysis of simulation results of CET-operations based on elementary functions of information-driven operations

The article considers the currently relevant task of synthesizing CET operations based on elementary functions of information-driven operations. The work is devoted to reducing the complexity of generating the maximum number of CET operations based on elementary functions of information-driven operations for low-resource stream ciphers. The work involved modelling CET-operations and obtaining four subgroups of 10 symmetric operations. Any of the subgroups can be used to construct 48 CET-operations that include all 10 operations of that subgroup. An analysis of a computational experiment on modelling CET-operations based on elementary information-driven operations was conducted. The absence of intersection among the synthesized sets of CET-operations was established, based on modifications of CET-operations from different subgroups. Based on the selection of four CET-operations (one from each group), it is possible to synthesize a complete group of CET-operations based on elementary information-driven functions. Such a group includes 192 operations. The paper presents a group of elementary functions of information-driven operations and a subset of symmetric CET-operations based on elementary functions of information-driven operations. The requirements for selecting a base group for constructing a method for synthesizing CET-operations based on elementary functions of information-driven operations are described. The conditions for considering the requirements for CET-operations of the base group are presented. Based on the results of a computational experiment, a method for selecting basic groups of symmetric CET-operations has been developed, which makes it possible to reduce the complexity of generating the maximum number of CET-operations based on elementary information-driven operations for low-resource stream ciphers. The developed method can be applied in low-resource telecommunication systems of both mobile and stationary types. The prospect of further research in this area is the construction of low-resource systems for cryptographic information transformation using CET-operations based on elementary functions of information-driven operations.

Key words: CET-operations, information-driven operations, Si-quantum elementary functions, automated synthesis of operations, computational experiment.

Постановка проблеми. До найбільших викликів, з якими стикається суспільство в усіх сферах свого існування на сучасному етапі розвитку відносять питання захисту інформації. Проблема захисту даних, що передаються поширеними каналами не є новою. Зі стародавніх часів забезпечення цілісності та конфіденційності інформації було вкрай важливою задачею, яка з розвитком науки і техніки набуває нових проявів.

Сучасні загрози характеризуються високим потенціалом їх потенційної ефективності. У роботі [1] виділяється дві категорії в розрізі квантових загроз: перехопити зараз – розшифрувати пізніше (Harvest Now, Decrypt Later – HTDL); підrobка цифрових підписів та сертифікатів.

У роботі [2] автори наголошують на вразливості багатьох криптографічних алгоритмів до атак реалізації. Відзначається недостатність алгоритмічної безпеки, що зумовлює важливість розвитку адаптивних можливостей криптографічних алгоритмів, що дозволить враховувати варіативність сучасних загроз та забезпечити стійкість до нових типів атак.

Розвиток обчислювальних систем, засобів передачі даних, нейромережових технологій вимагає розробки нових підходів побудови елементарних операцій криптографічного кодування, які забезпечують варіативність та адаптивність криптографічних алгоритмів. Вкрай актуальною проблемою є дослідження операцій керованих інформацією в розрізі криптографічного захисту інформації. Такі операції характеризуються варіативністю своєї структури, що забезпечує динамічну залежність від змінної інформації, яка керує процесами та забезпечує криптографічну стійкість.

Аналіз останніх досліджень і публікацій. У сучасних дослідженнях CET-шифрування виділяється як один з перспективних напрямків збільшення варіативності операцій потокового шифрування та розвитку криптографічного захисту інформації в цілому. Однією з головних переваг операцій CET-шифрування є можливість їх застосування як при блоковому так і при потоковому шифруванні [3]. У роботах [4-5] представлені перші спроби синтезу двооперандних CET-операцій на основі додавання за модулем два та їх дискретно-алгебраїчного представлення. У подальшому проводились дослідження по синтезу операцій криптографічного перетворення інформації не лише на операціях 2-ої розрядності, а й на операціях 3-ої розрядності [6] та операціях 4-ої розрядності [7]. У роботі [7] встановлено також умови невиродженості нелінійних операцій розширеного матричного криптографічного перетворення n -ої розрядності, у складі яких є неповні функції розширеного матричного криптографічного перетворення інформації.

На основі попередніх досліджень операцій криптографічного перетворення розпочались дослідження CET-операцій керованих інформацією. У роботі [8] проаналізовано основні елементарні операції, що входять до базової групи операцій перестановок, керованих інформацією та представлена дискретна модель базових груп операцій перестановок для криптоперетворення, керованих інформацією. Практична реалізація дискретних моделей CET-операцій керованих інформацією наведена в роботі [9]. У роботі [10] запропоновано шлях підвищення швидкості та стійкості матричного криптографічного перетворення за рахунок використання операцій та алгоритмів криптоперетворення двох блоків змінних інформації. Особливої уваги потребують важливі роботи по синтезу обернених CET-операцій [11-12]. Аналіз досліджень показав, що синтезу CET-операцій на основі елементарних функцій операцій керованих інформацією та побудові їх базових груп приділена недостатня увага.

Мета статті – зменшення складності генерації максимальної кількості СЕТ-операцій на основі елементарних функцій операцій керованих інформацією для малоресурсних потокових шифрів за рахунок розробки методу вибору базових груп симетричних СЕТ-операцій на основі результатів обчислювального експерименту.

Виклад основного матеріалу. Для дослідження однооперандних СЕТ-операцій на основі елементарних функцій операцій керованих інформацією необхідно мати множину даних операцій. Наявність даної множини і буде набором вхідних даних для проведення дослідження. Однооперандні СЕТ-операції на основі елементарних функцій операцій керованих інформацією належать до групи 3Сі-квантових СЕТ-операцій. Повна група 3Сі-квантових СЕТ-операцій представляє собою всі підстановки з поля перестановок G_8 і включає в себе 40320 СЕТ-операцій. Всі 3Сі-квантові СЕТ-операції будуються на основі 70 3Сі-квантових елементарних функцій. Однооперандні 3Сі-квантові СЕТ-операції на основі елементарних функцій операцій керованих інформацією будуються на основі 8 елементарних функцій операцій керованих інформацією [13-14]. Так як елементарних функцій всього 8 то кількість СЕТ-операцій які потрібно знайти для проведення дослідження буде значно меншою за 40320. Виходячи з незначної кількості елементарних функцій на основі яких будуються однооперандні 3Сі-квантові СЕТ-операції на основі елементарних функцій операцій керованих інформацією доцільно для автоматизованого синтезу операцій використати повний перебір варіантів розміщення елементарних функцій в операціях. Даний підхід і було використано при проведенні обчислювального експерименту по моделюванню СЕТ-операцій на основі елементарних функцій операцій керованих інформацією. Група елементарних функцій операцій керованих інформацією представлена в таблиці 1 [6].

Таблиця 1

Група елементарних функцій операцій керованих інформацією

Елементарна функція	Результат реалізації	Елементарна функція	Результат реалізації
$f_{23} = x_1 \cdot x_2 \vee x_1 \cdot x_3 \vee x_2 \cdot x_3$	00010111	$f_{232} = \bar{x}_1 \cdot \bar{x}_2 \vee \bar{x}_1 \cdot \bar{x}_3 \vee \bar{x}_2 \cdot \bar{x}_3$	11101000
$f_{43} = x_1 \cdot x_2 \vee x_1 \cdot \bar{x}_3 \vee x_2 \cdot \bar{x}_3$	00101011	$f_{212} = \bar{x}_1 \cdot \bar{x}_2 \vee \bar{x}_1 \cdot x_3 \vee \bar{x}_2 \cdot x_3$	11010100
$f_{77} = x_1 \cdot \bar{x}_2 \vee x_1 \cdot x_3 \vee \bar{x}_2 \cdot x_3$	01001101	$f_{178} = \bar{x}_1 \cdot x_2 \vee \bar{x}_1 \cdot \bar{x}_3 \vee x_2 \cdot \bar{x}_3$	10110010
$f_{113} = \bar{x}_1 \cdot x_2 \vee \bar{x}_1 \cdot x_3 \vee x_2 \cdot x_3$	01110001	$f_{142} = x_1 \cdot \bar{x}_2 \vee x_1 \cdot \bar{x}_3 \vee \bar{x}_2 \cdot \bar{x}_3$	10001110

У табл. 1 індекси елементарних функцій відповідають значенню десяткової цифри результату перетворення, за умови впорядкування повної множини вхідних даних по зростанню їх значень [6].

За результатами обчислювального експерименту було отримано 192 СЕТ-операції на основі елементарних функцій операцій керованих інформацією. Результати обчислювального експерименту наведені в додатку А.

Будь яку СЕТ-операцію можна модифікувати шляхом перестановки елементарних функцій, а також шляхом інверсії результатів виконання елементарної функції. СЕТ-операція на основі елементарних функцій операцій керованих інформацією є 3Сі-квантовою, то вона виконує криптографічне перетворення 3Сі-квантів вхідної інформації і формує 3Сі-кванти результату перетворення. Виходячи з цього її можна модифікувати за допомогою 6 операцій перестановки елементарних функцій і 8 операцій інверсії результатів. Виходячи з цього на основі будь якої СЕТ-операції можна отримати 48 її модифікацій.

Так як група СЕТ-операції на основі елементарних функцій операцій керованих інформацією включає в себе 192 операції то вона включає в себе модифікації 4 унікальних базових операцій ($192:48=4$). Дані 4 унікальні базові операції складають базову групу СЕТ-операції на основі елементарних функцій операцій керованих інформацією. З будь якої операції базової групи шляхом перестановок і інверсій будуються множина модифікованих СЕТ-операцій. Кожна побудована операція з даної множини буде унікальною (не буде повторятися), а побудовані множини операцій не будуть перетинатися. Так як, множина модифікованих СЕТ-операцій будуються шляхом перестановок і інверсій елементарних функцій базової операції та базова група операцій на основі елементарних функцій операцій керованих інформацією включає в себе будь які 4 операції, які належать до 4 різних множин модифікованих операцій.

В процесі дослідження однооперандних СЕТ-операцій на основі елементарних функцій операцій керованих інформацією необхідно встановити правила побудови прямих і обернених СЕТ-операцій.

Для побудови правил синтезу прямих і обернених операцій доцільно дослідити правила побудови СЕТ-операцій базової групи, так як правила їх модифікації на основі перестановки і інверсії елементарних функцій відомі [35]. Сформовані підмножини симетричних СЕТ-операцій на основі елементарних функцій операцій керованих інформацією представлено у (табл. 2).

**Підмножини симетричних СЕТ-операцій на основі елементарних функцій
операцій керованих інформацією**

№	Множина 1	Множина 2	Множина 3	Множина 4
1	$C_{23,43,77}(x)$	$C_{23,43,142}(x)$	$C_{23,142,77}(x)$	$C_{43,77,113}(x)$
2	$C_{23,77,43}(x)$	$C_{43,23,113}(x)$	$C_{77,113,23}(x)$	$C_{77,113,43}(x)$
3	$C_{43,23,178}(x)$	$C_{43,142,232}(x)$	$C_{77,232,142}(x)$	$C_{113,43,77}(x)$
4	$C_{77,212,23}(x)$	$C_{113,23,43}(x)$	$C_{113,77,23}(x)$	$C_{113,77,43}(x)$
5	$C_{178,43,232}(x)$	$C_{142,232,212}(x)$	$C_{142,178,232}(x)$	$C_{142,178,212}(x)$
6	$C_{212,232,77}(x)$	$C_{212,113,23}(x)$	$C_{178,23,113}(x)$	$C_{142,212,178}(x)$
7	$C_{232,178,212}(x)$	$C_{212,232,142}(x)$	$C_{178,142,232}(x)$	$C_{178,142,212}(x)$
8	$C_{232,212,178}(x)$	$C_{232,212,113}(x)$	$C_{232,113,178}(x)$	$C_{212,178,142}(x)$

Нижні індекси в моделях СЕТ-операцій наведених у табл. 2 відображають індекси елементарних функцій з яких побудовано дані операції. Наприклад:

$$C_{77,212,23}(x) = \begin{bmatrix} f_{77}(x) \\ f_{212}(x) \\ f_{23}(x) \end{bmatrix} \quad (1)$$

Базова група СЕТ-операцій на основі елементарних функцій операцій керованих інформацією яка складається лише з симетричних операцій включає в себе по одній модифікованій операції з кожної множини. Наприклад:

- $C_{23,43,77}(x)$, $C_{23,43,142}(x)$, $C_{23,142,77}(x)$, $C_{43,77,113}(x)$;
- $C_{23,43,77}(x)$, $C_{43,23,113}(x)$, $C_{178,23,113}(x)$, $C_{142,178,212}(x)$;
- $C_{178,43,232}(x)$, $C_{142,232,212}(x)$, $C_{77,232,142}(x)$, $C_{178,142,212}(x)$.

Відповідно до табл. 2 кількість група СЕТ-операцій які складається з симетричних операцій буде визначатися як $8 \cdot 8 \cdot 8 \cdot 8 = 8^4 = 4096$.

Вибір базової групи для побудови методу синтезу СЕТ-операцій на основі елементарних функцій операцій керованих інформацією необхідно проводити на основі двох вимог:

1. Мінімальна простота синтезу 4 симетричних СЕТ-операцій базової групи. Дана вимога базується на необхідності зменшення обчислювальних ресурсів для моделювання систем захисту інформації на основі СЕТ-шифрування.

2. Максимальна відмінність відповідних елементарних функцій в 4 симетричних СЕТ-операціях базової групи для реалізації вихідних Сі-квантів інформації. Дана вимога базується на бажанні максимального ускладнення крипто аналізу, для визначення правил формування послідовності крипто перетворень.

Наведені вимоги достатньо суперечливі і вимагають різних підходів до пошуку СЕТ-операцій базової групи. Необхідно відмітити, що незалежно яка базова група СЕТ-операцій буде взята за основу, це забезпечить побудову на основі перестановок і інверсій елементарних функцій отримати повну групу операцій. Вибір базової групи буде впливати лише на послідовність синтезу СЕТ-операцій.

Виходячи з цього доцільно розглядати вимоги до СЕТ-операцій базової групи незалежно одна від другої, а вже на основі отриманих результатів, за необхідністю спробувати знайти компроміс.

Висновки. У статті було отримано наступні результати:

1. За результатами аналізу СЕТ-операцій на основі елементарних функцій операцій керованих інформацією було виділено чотири підгрупи симетричних СЕТ-операцій з точністю до перестановки елементарних функцій і їх інверсій. Кожна підгрупа включає по 10 СЕТ-операцій. З будь-якої вибраної з підгрупи операції на основі перестановки елементарних функцій і їх інверсій можна побудувати 48 СЕТ-операцій які включають всі 10 операцій підгрупи.

2. Встановлено, що синтезовані множини СЕТ-операцій на основі модифікацій СЕТ-операцій з різних підгруп не перетинаються. На основі вибору чотирьох СЕТ-операцій (по одній з кожної підгрупи шляхом перестановки інверсій) можна синтезувати повну групу СЕТ-операцій на основі елементарних функцій операцій керованих інформацією, яка включає 192 операції.

3. За результатами обчислювального експерименту розроблено метод вибору базових груп симетричних СЕТ-операцій для зменшення складності генерації максимальної кількості СЕТ-операцій на основі елементарних функцій операцій керованих інформацією для малоресурсних поточкових шифрів.

Сфера практичного застосування запропонованого методу полягає у його використанні в мобільних та стаціонарних малоресурсних телекомунікаційних системах.

Подальшого розгляду потребують питання побудови малоресурсних систем криптографічного перетворення інформації шляхом застосування СЕТ-операцій на основі елементарних функцій операцій керованих інформацією.

Список використаних джерел:

1. Benitez C. Mapping quantum threats: an engineering inventory of cryptographic dependencies. *arXiv*. 2025. URL: <https://doi.org/10.48550/arXiv.2509.24623> (дата звернення: 15.10.2025).
2. Algorithmic security is insufficient: a comprehensive survey on implementation attacks haunting post-quantum security / A.C. Canto et al. *arXiv*. 2023. URL: <https://doi.org/10.48550/arXiv.2305.13544> (дата звернення: 15.10.2025).
3. Архітектура СЕТ-операцій і технології потокового шифрування: монографія / В.М. Рудницький та ін. Черкаси: видавець Пономаренко Р.В., 2024. 374 с.
4. Голуб С. В., Бабенко В. Г., Рудницький С. В. Метод синтезу операцій криптографічного перетворення на основі додавання за модулем два. *Системи обробки інформації*. 2012. Вип. 3 (101), Т. 1. С. 119–122.
5. Вдосконалення методу синтезу операцій криптографічного перетворення на основі дискретно-алгебраїчного представлення операцій / С.В. Голуб та ін. *Системи управління, навігації та зв'язку*. 2012. Вип. 2 (22). С. 163–168.
6. Бабенко В., Мельник О., Мельник Р. Класифікація трирозрядних елементарних функцій для криптографічного перетворення інформації. *Безпека інформації*. 2013. Том 19, № 1. С. 56–59.
7. Стабєцька Т. А. Умови невиродженості нелінійних операцій розширеного матричного криптографічного перетворення, що містять неповні функції РМКП. *Системи обробки інформації*. 2016. Вип. 1 (138). С. 131–133.
8. Миронюк Т. В., Ланських Є. В. Дискретна модель базових груп операцій перестановок, керованих інформацією, для криптоперетворення. *Smart and Young*. 2016. Вип. 11-12. С. 58–65.
9. Миронець І. В., Миронюк Т. В., Сисоєнко С. В. Апаратна реалізація базової групи операцій перестановок, керованих інформацією. *Актуальні задачі та досягнення у галузі кібербезпеки : матеріали Всеукр. наук.-практ. конф. (м. Кропивницький, 23–25 листопада 2016 р.)*. Кропивницький : КНТУ, 2016. С. 141–142.
10. Сисоєнко С. В., Мельник О. Г. Використання операцій та алгоритмів криптоперетворення двох блоків змінних в криптографії. *Інноваційні тенденції сьогодення в сфері природничих, гуманітарних та точних наук : матеріали Міжнародної науково-практичної конференції (м. Івано-Франківськ, 17 жовтня 2017 р.)*. Одеса: Друкарік, 2017. Вип. 2. С. 47–49.
11. Сисоєнко С. В., Мельник О. Г., Пустовіт М. О. Синтез операцій оберненого групового матричного криптографічного перетворення інформації. *Вісник Черкаського державного технологічного університету. Технічні науки*. 2017. Вип. 22, № 4. С. 118–125.
12. Breus R. Synthesising the two-bit two-operand operations of strict stable cryptographic coding by the second operand's conversion. *Academic Journal. Control, Navigation and Communication Systems*. 2019. Vol. 5, № 57. P. 29–32. DOI: <https://doi.org/10.26906/SUNZ.2019.5.029>
13. Синтез дискретно-алгебраїчних моделей елементарних функцій операцій керованих інформацією / В.М. Рудницький та ін. *Кібербезпека: освіта, наука, техніка*. 2024. Том 3, № 23. С. 6–16. DOI: <https://doi.org/10.28925/2663-4023.2024.23.616>
14. Дискретно-казуальне представлення моделей елементарних функцій і СЕТ-операцій / В.М. Рудницький та ін. *Системи управління, навігації та зв'язку*. 2023. № 4. С. 96–101. DOI: <https://doi.org/10.26906/SUNZ.2023.4.096>
15. Криптографічне кодування: методи і засоби реалізації (частина 2): монографія / В.М. Рудницький та ін. Х. : Видавництво «Щедра садиба плюс», 2014. 224 с.

References:

1. Benitez, C. (2025). Mapping quantum threats: an engineering inventory of cryptographic dependencies. *arXiv*. <https://doi.org/10.48550/arXiv.2509.24623>
2. Canto, A. C., et al. (2023). Algorithmic security is insufficient: a comprehensive survey on implementation attacks haunting post-quantum security. *arXiv*. <https://doi.org/10.48550/arXiv.2305.13544>
3. Rudnytskyi, V., et al. (2024). Arkhitektura SET-operatsii i tekhnolohii potokovoho shyfruvannia: monohrafiia [Architecture of CET-operations and stream encryption technologies: monograph]. Publisher Ponomarenko R.V. [in Ukrainian].
4. Golub, S. V., Babenko, V. G. & Rudnytskyi, S. V. (2012). Metod syntezy operatsii kryptohrafichnoho peretvorennia na osnovi dodavannia za modulem dva [A method for synthesizing cryptographic transformation operations based on addition modulo two]. *Systemy obrobky informatsii – Information Processing Systems*, 3 (101), 1, 119–122 [in Ukrainian].

-
5. Golub, S. V., et al. (2012). Vdoskonalennia metodu syntezy operatsii kryptohrafichnoho peretvorennia na osnovi dyskretno-alhebraichnoho predstavlennia operatsii [Improvement of the Method for Synthesizing Cryptographic Transformation Operations Based on Discrete-Algebraic Representation of Operations]. *Systemy upravlinnia, navihatsii ta zviazku – Academic Journal. Control, Navigation and Communication Systems*, 22(2), 163–168 [in Ukrainian].
6. Babenko, V., Melnyk, O. & Melnyk, R. (2013). Klasyfikatsiia tryozriadnykh elementarnykh funktsii dlia kryptohrafichnoho peretvorennia informatsii [Classification of three digit basic functions for cryptographic transformation of information]. *Bezpeka informatsii – Ukrainian Scientific Journal of Information Security*, 19(1), 56–59 [in Ukrainian].
7. Stabetska, T. A. (2016). Umovy nevyrodzhenosti nelineinykh operatsii rozshyrenoho matrychnoho kryptohrafichnoho peretvorennia, shcho mistiat nepovni funktsii RMKP [Conditions of non-degeneracy of nonlinear operations of the extended matrix cryptographic transformation containing incomplete functions of the EMCT]. *Systemy obrobky informatsii – Information Processing Systems*, 138(1), 131–133 [in Ukrainian].
8. Myroniuk, T. V. & Lanskich, E. V. (2016). Dyskretna model bazovykh hrup operatsii perestanolov, kerovanykh informatsieiu, dlia kryptoperetvorennia [A discrete model of basic groups of information-driven permutation operations for cryptotransformation]. *Smart and Young*, 11-12, 58–65 [in Ukrainian].
9. Myronets, I. V., Myroniuk, T. V. & Sysoyenko, S. V. (2016). Aparatna realizatsiia bazovoi hrupy operatsii perestanolov, kerovanykh informatsieiu [Hardware implementation of a basic group of information-driven permutation operations]. *Aktualni zadachi ta dosiahnenia u haluzi kiberbezpeky – Current tasks and achievements in the field of cybersecurity* (p. 141–142). Kropyvnytskyi National Technological University [in Ukrainian].
10. Sysoyenko, S. V. & Melnyk, O. G. (2017). Vykorystannia operatsii ta alhorytmiv kryptoperetvorennia dvoh blokiv zminnykh v kryptohrafii [Using of the operations and algorithms of cryptographic transformation of two blocks of variables in cryptography]. *Innovatsiini tendentsii siohodennia v sferi pryrodnychnykh, humanitarnykh ta tochnykh nauk – Innovative Trends in the Field of Natural Sciences, Humanities and Exact Sciences* (p. 47–49). Drukaryk [in Ukrainian].
11. Sysoyenko, S. V., Melnyk, O. G. & Pustovit, M. O. (2017). Syntez operatsii obernenoho hrupovoho matrychnoho kryptohrafichnoho peretvorennia informatsii [Synthesis of operations of reverse group matrix cryptographic transformation of information]. *Visnyk Cherkaskoho derzhavnoho tekhnolohichnoho universytetu. Tekhnichni nauky – Bulletin of Cherkasy State Technological University. Technical sciences*, 22(4), 118–125 [in Ukrainian].
12. Breus, R. (2019). Synthesizing the two-bit two-operand operations of strict stable cryptographic coding by the second operand's conversion. *Academic Journal. Control, Navigation and Communication Systems*, 57(5), 29–32. <https://doi.org/10.26906/SUNZ.2019.5.029>
13. Rudnytskyi, V., et al. (2024). Syntez dyskretno-alhebraichnykh modelei elementarnykh funktsii operatsii kerovanykh informatsieiu [Synthesis of discrete and algebraic models of elementary functions of data-controlled operations]. *Kiberbezpeka: osvita, nauka, tekhnika – Electronic scientific journal “Cyber Security: Education, Science, Technology”*, 23(3), 6–16. <https://doi.org/10.28925/2663-4023.2024.23.616> [in Ukrainian].
14. Rudnytskyi, V., et al. (2023). Dyskretno-kazualne predstavlennia modelei elementarnykh funktsii i SET-operatsii [Discrete-casual presentation of models of elementary functions and set operations]. *Systemy upravlinnia, navihatsii ta zviazku – Academic Journal. Control, Navigation and Communication Systems*, 74(4), 96–101. <https://doi.org/10.26906/SUNZ.2023.4.096> [in Ukrainian].
15. Rudnytskyi, V., et al. (2014). Kryptohrafichne koduvannia: metody i zasoby realizatsii (chastyna 2): monohrafia [Cryptographic encoding: methods and implementation tools (Part 2): monograph]. Publisher “Shchedra Sadyba Plus” [in Ukrainian].

Дата надходження статті: 27.10.2025

Дата прийняття статті: 17.11.2025

Опубліковано: 30.12.2025

МЕНЕДЖМЕНТ

УДК 656.078

DOI <https://doi.org/10.32782/2521-6643-2025-2-70.30>

Кузьменко А. І., кандидат технічних наук, доцент,
завідувач кафедри транспортних технологій та міжнародної
логістики Університету митної справи та фінансів
ORCID: 0000-0001-7278-3647

Черчата А. О., кандидат економічних наук, доцент,
завідувач кафедри економіки, менеджменту та підприємництва
Українського державного університету науки і технологій,
ННІ «Придніпровська державна академія будівництва
та архітектури»
ORCID: 0000-0002-6753-2891

Халіпова Н. В., кандидат технічних наук, доцент,
доцент кафедри транспортних технологій та міжнародної логістики
Університету митної справи та фінансів
ORCID: 0000-0001-5605-6781

Леснікова І. Ю., кандидат технічних наук, доцент,
доцент кафедри транспортних технологій та міжнародної логістики
Університету митної справи та фінансів
ORCID: 0000-0001-5605-6781

ОПТИМІЗАЦІЯ МІЖНАРОДНИХ ЛОГІСТИЧНИХ ПОТОКІВ НА ПІДСТАВІ МЕТОДІВ ТЕОРІЇ ПРИЙНЯТТЯ РІШЕНЬ В УМОВАХ НЕВИЗНАЧЕНОСТІ

У статті розглянуто теоретичні та практичні аспекти оптимізації міжнародних логістичних потоків в умовах невизначеності зовнішнього середовища. Визначено основні чинники, що впливають на ефективність функціонування логістичних систем у міжнародному масштабі, серед яких – геополітична нестабільність, коливання валютних курсів, зміни у митному законодавстві, нестача транспортних ресурсів та інформаційна непрозорість. Особливу увагу приділено застосуванню методів теорії прийняття рішень, які дозволяють обґрунтовано обирати оптимальні варіанти дій за умов обмеженої або неповної інформації.

У роботі проаналізовано можливості використання критеріїв Лапласа, Вальда, Севіджа та Гурвіца для оцінки ризиків та ефективності логістичних рішень. Продемонстровано, як кожен із цих критеріїв відображає різний рівень ставлення до ризику (від песимістичного до оптимістичного) і може бути застосований для вибору маршруту доставки, постачальника, перевізника або схеми транспортування. Розв'язано практичну задачу оптимізації міжнародних логістичних потоків у автомобільному сполученні за напрямом Дрезден – Ужгород, як одному з актуальних логістичних коридорів між Європейським Союзом та Україною. Цією метою розроблено мережевий графік, виконано відповідні розрахунки та побудовано порівняльні діаграми.

Розроблено узагальнену матрицю ризиків, що дозволяє оцінити вплив зовнішніх факторів на стабільність міжнародних поставок і визначити пріоритетні напрями управлінських дій.

Результати дослідження підтверджують, що інтеграція методів прийняття рішень у процес управління міжнародною логістикою підвищує її адаптивність, сприяє зменшенню ризиків та забезпечує збалансованість між витратами, часом і надійністю постачань. Запропонований підхід може бути використаний підприємствами-експортерами, транспортно-експедиторськими компаніями та логістичними операторами для стратегічного планування та підвищення конкурентоспроможності на світовому ринку.

Ключові слова: міжнародна логістика, логістичні потоки, теорія прийняття рішень, критерії невизначеності, оптимальні маршрути.

© А. І. Кузьменко, А. О. Черчата, Н. В. Халіпова, І. Ю. Леснікова, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

Kuzmenko A. I., Cherchata A. A., Khalipova N. V., Lesnikova I. Yu. Optimization of international logistics flows based on methods of the theory of decision-making in conditions of uncertainty

The article examines the theoretical and practical aspects of optimizing international logistics flows in conditions of uncertainty in the external environment. The main factors affecting the efficiency of logistics systems on an international scale are identified, including geopolitical instability, exchange rate fluctuations, changes in customs legislation, lack of transport resources and information opacity. Special attention is paid to the application of decision-making theory methods that allow you to reasonably choose the optimal options for actions under conditions of limited or incomplete information.

The paper analyzes the possibilities of using the Laplace, Wald, Savage, and Hurwitz criteria to assess the risks and effectiveness of logistics solutions. It is demonstrated how each of these criteria reflects a different level of risk attitude (from pessimistic to optimistic) and can be applied to select a delivery route, supplier, carrier or transportation scheme. The practical task of optimizing international logistics flows in the Dresden-Uzhgorod direction by road, as one of the relevant logistics corridors between the European Union and Ukraine, has been solved. For this purpose, a network graph was developed, corresponding calculations were performed and comparative diagrams were constructed. A generalized risk matrix has been developed, which allows to assess the impact of external factors on the stability of international supplies and to determine the priority directions of management actions.

The results of the study confirm that the integration of decision-making methods into the international logistics management process increases its adaptability, helps reduce risks and ensures a balance between costs, time and reliability of supplies. The proposed approach can be used by exporting enterprises, transport and forwarding companies and logistics operators for strategic planning and increasing competitiveness in the world market.

Key words: international logistics, logistics flows, decision-making theory, uncertainty criteria, optimal routes.

Постановка проблеми. У світовому господарстві здійснення міжнародних перевезень неможливе без розвинутої логістичної інфраструктури, як на національному рівні, так і на міжнародному. Особливе місце в логістичній інфраструктурі займає транспортна логістика, яка утворює світову транспортну систему [1].

Сучасний етап розвитку світової економіки супроводжується зростанням рівня глобальної невизначеності та ускладненням функціонування міжнародних транспортно-логістичних систем. Воєнні конфлікти, пандемія COVID-19, різкі коливання цін на енергоносії, торговельні війни та кліматичні катаклізми дедалі частіше призводять до збоїв у ланцюгах постачання та обмежень міжнародної торгівлі. Логістичні компанії та національні економіки стикаються з необхідністю швидкого реагування на непередбачувані зміни, що проявляються у закритті кордонів, блокуванні портів, порушенні авіаційних і морських маршрутів, дефіциті транспортних ресурсів і зростанні витрат на перевезення [2]. Повномасштабна війна в Україні спричинила глибоку трансформацію національної економіки, істотно ускладнила функціонування логістичної інфраструктури та вплинула на стабільність глобальних ланцюгів постачання. В умовах підвищеної невизначеності особливого значення набуває ефективне управління експортною логістикою як ключовим чинником забезпечення економічної стійкості країни та збереження її інтеграції у світову економіку.

Традиційні підходи до організації міжнародних перевезень, що базувалися на принципах мінімізації витрат та оптимізації маршрутів у стабільних умовах, виявилися малоефективними в період глобальних криз. Виникає потреба у формуванні нових адаптивних стратегій управління, які враховують багатовекторні ризики, забезпечують гнучкість і здатність швидкого відновлення функціонування логістичних систем [2].

Оптимізація міжнародних логістичних потоків – це складне завдання, що включає управління переміщенням товарів, інформації та фінансових ресурсів між країнами з мінімальними витратами та максимальним рівнем ефективності. Проте на практиці існує низка проблем, які ускладнюють цей процес. Основні з них, згруповані за ключовими напрямками, наступні.

1. Економічні та фінансові проблеми: коливання валютних курсів, що впливають на вартість перевезень, митних зборів і закупівель; висока вартість транспортування, особливо морських та авіаперевезень після пандемії COVID-19 і воєнних конфліктів; нерівномірність розвитку інфраструктури (у різних країнах транспортні системи розвинені нерівномірно, що підвищує витрати на логістику); високі митні та податкові бар'єри, особливо у торгівлі між країнами, які не мають угод про вільну торгівлю.

2. Організаційно-правові проблеми: складні митні процедури (затримки на кордоні, бюрократія, відсутність електронного документообігу); відмінності у правових системах – різні стандарти безпеки, сертифікації, маркування продукції; нестача єдиних міжнародних логістичних стандартів, що ускладнює координацію між перевізниками, експедиторами та митницями; обмеження, пов'язані з санкціями або воєнними діями, наприклад, проти РФ або в зонах конфлікту.

3. Технологічні проблеми: низький рівень цифровізації логістичних процесів (у багатьох країнах досі застосовуються паперові документи); недостатня інтеграція ІТ-систем учасників ланцюга постачання, яка призводить до втрат даних, помилок, дублювання операцій; відсутність аналітичних інструментів для прогнозування затримок чи оптимізації маршрутів та кібербезпека (у даному випадку – ризики витоку даних або саботажу в системах управління логістикою).

4. Геополітичні та безпекові чинники: військові конфлікти (зокрема, російська агресія проти України), які руйнують транспортну інфраструктуру та змінюють маршрути постачання; закриття або блокування

транспортних коридорів (наприклад, Чорноморських портів); політична нестабільність у транзитних країнах, що підвищує ризики зриву поставок.

5. Екологічні та соціальні проблеми: вимоги до зменшення викидів CO₂ (перехід на “зелену логістику” потребує значних інвестицій); нестача кваліфікованих кадрів, особливо водіїв міжнародних перевезень; зміна споживчих очікувань (клієнти вимагають швидких і прозорих поставок, що тисне на логістичні системи).

6. Проблеми оптимізації управління: складність координації між учасниками ланцюга постачання (виробник – перевізник – митниця – дистриб’ютор – споживач); відсутність гнучких стратегій ризик-менеджменту (логістичні ланцюги часто не готові до форс-мажорів); недостатня аналітика і прогнозування попиту, що призводить до надлишкових або дефіцитних запасів.

Аналіз останніх досліджень і публікацій. Процеси управління міжнародною логістикою розглядало багато вітчизняних та зарубіжних науковців. Серед них: Верещагіна Г. В. та Сушко К. В. [1]; Малиновський Ю. В., Ремез О. С., Сорока В. В., Гринчук Т. В. Гриценко С. І. та Дейкун В. О. [2]; Сколозdra Ю. І. [3]; Гриценко С. І. та Дейкун В. О. [4]; Вакуленко В.Л., Лю Юнтао та Лю Сяовой [5] і т.д.

Так, у статті [2] досліджено сучасні проблеми функціонування міжнародної логістики в умовах глобальних криз та обґрунтовано необхідність формування адаптивних стратегій управління транспортно-логістичними системами. Показано, що традиційні моделі організації перевезень, орієнтовані переважно на мінімізацію витрат і стабільність операційних процесів, є недостатньо ефективними за умов високої невизначеності, що зумовлюється воєнними конфліктами, пандеміями, енергетичними коливаннями та кліматичними катаклізмами. Встановлено, що ключовими викликами для міжнародних перевезень є геополітичні загрози, економічні ризики, екологічні та технологічні чинники, а також соціальні аспекти, пов’язані з дефіцитом кадрів і міграційними процесами.

Автор статті [3] розглядає роль логістики в імпорتنій діяльності підприємства. Основну увагу приділено оптимізації транспортних потоків, поліпшенню складських операцій та мінімізації витрат на логістику. Визначено основні фактори, які найбільше враховується у міжнародному логістичному ланцюжку для імпортерів. Розглянуто практичний кейс оптимізації логістичних процесів в імпорتنій діяльності підприємства.

У науковій роботі [4] розглянуто логістику як ключовий компонент конкурентоспроможності підприємства в умовах стрімкого розвитку цифрових технологій і глобалізації бізнесу. Автори зазначають, що успішне функціонування ланцюгів постачання більше не може базуватися на традиційних методах управління, тому необхідно запустити процеси цифрової трансформації, яка відкриває нові можливості для оптимізації логістичних процесів, підвищення ефективності, зниження витрат та посилення прозорості.

Метою дослідження [5] є системний аналіз викликів у сфері логістичного забезпечення експорту в умовах воєнного стану, а також визначення стратегічних підходів до підвищення гнучкості та надійності ланцюгів постачання. Авторами зацентовано увагу на необхідності переосмислення традиційних моделей системи ризик-менеджменту через призму концепцій вразливості та стійкості, які в сучасних умовах виступають пріоритетними в теорії та практиці управління ризиками. У межах вказаного дослідження ідентифіковано форми реалізації ризик-орієнтованого управління в логістиці з метою мінімізації наслідків потенційних збоїв, а також обґрунтовано інтеграцію інструментів адаптивного управління для забезпечення швидкого відновлення функціональних логістичних систем при цифровізації логістичних процесів та використанні сучасного програмного забезпечення.

Не зважаючи на значну кількість досліджень у сфері міжнародної логістики, недостатньо розробленими залишаються комплексні підходи до розподілу міжнародних логістичних підходів. Потребує вирішення завдання формування адаптивних стратегій, здатних забезпечити безперервність і стійкість транспортно-логістичних систем, інтегруючи інноваційні та цифрові рішення [2].

Метою дослідження є визначення інструментів оптимізації міжнародних логістичних потоків у автомобільному сполученні за напрямом Дрезден – Ужгород, як одному з актуальних логістичних коридорів між Європейським Союзом та Україною. Поставлена задача полягає в тому, щоб спроектувати оптимальну транспортно-логістичну схему доставки вантажу за вказаним маршрутом на підставі методів теорії прийняття рішень в умовах невизначеності. Для цього пропонується використовувати мережеве планування і критерії Лапласа, Вальда, Севіджа та Гурвіца, що належать до класичних і в транспортній логістиці допомагають оцінювати ризики, порівнювати варіанти маршрутів, постачальників чи перевізників та вибирати оптимальну стратегію при відсутності точних прогнозів. На основі цих критеріїв виконано відповідні розрахунки для порівняння чотирьох можливих маршрутів та обрано оптимальний.

Виклад основного матеріалу дослідження. Комплексний аналіз чотирьох варіантів маршрутів з метою вибору оптимального рішення для організації міжнародної логістичної доставки передбачає складання можливих схем перевезень (див. рис. 1-4) з детальним описом кожного маршруту.

Формування маршрутів здійснювалося за допомогою спеціалізованого логістичного сервісу Flagma. UA [6], що дозволило врахувати реальні дорожні умови, відстані, можливі затори та швидкісні обмеження. У процесі аналізу було також враховано такі часові витрати, як зміна водіїв, перерви на відпочинок, проходження митного контролю тощо [7].

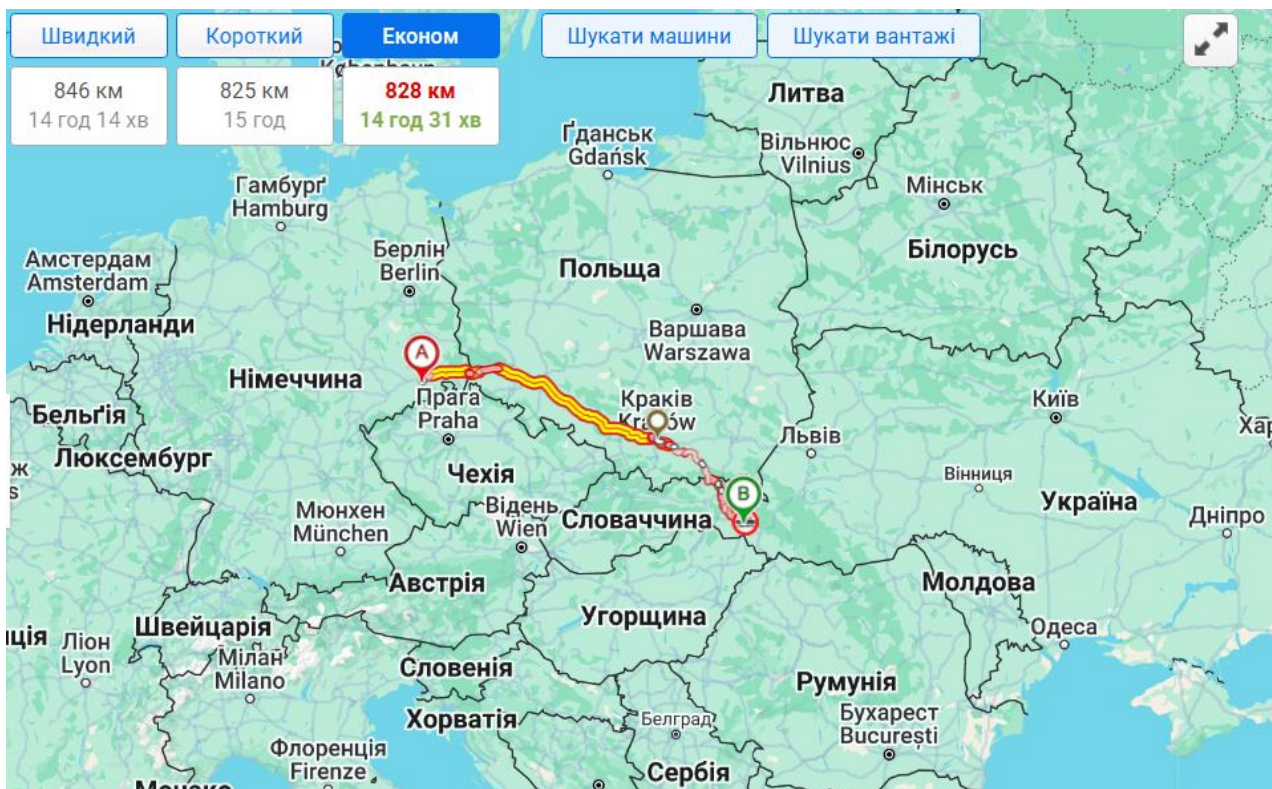


Рис. 1. Маршрут № 1 Німеччина – Україна через території Польщі та Словаччини

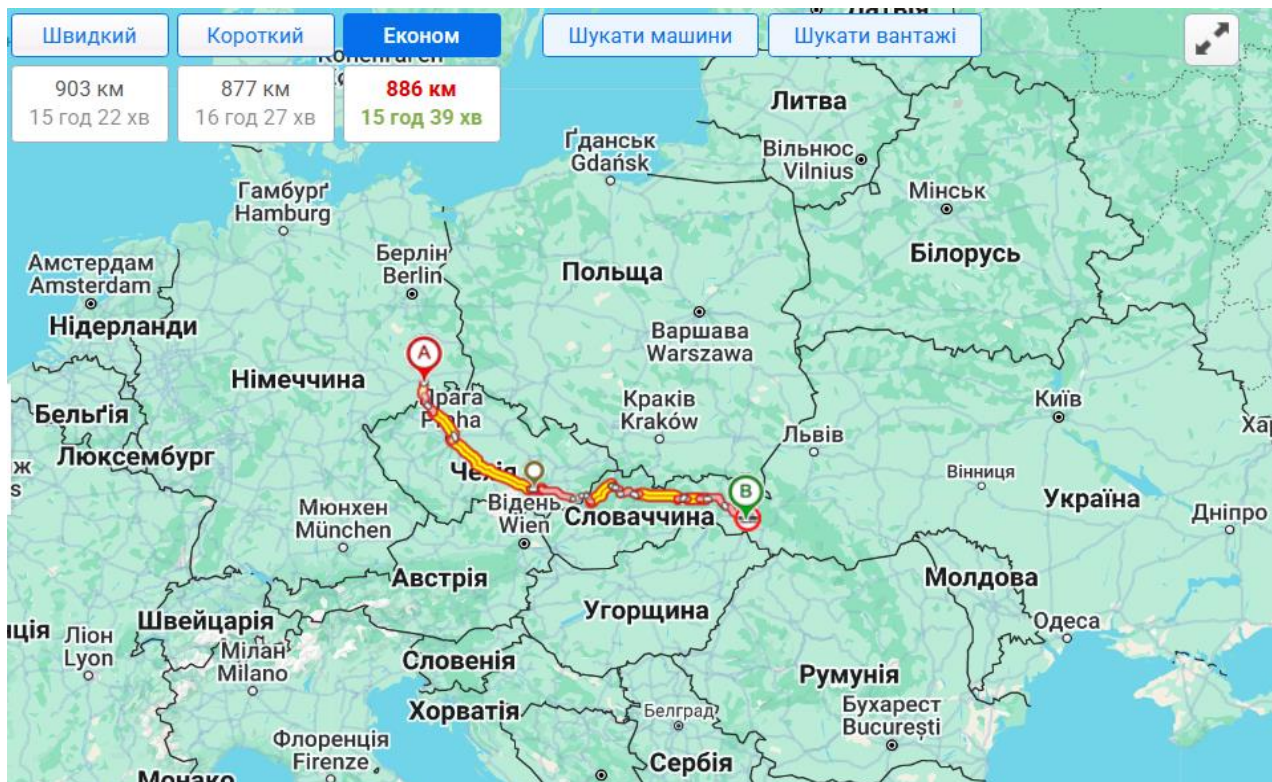


Рис. 2. Маршрут № 2 Німеччина – Україна через території Чехії та Словаччини

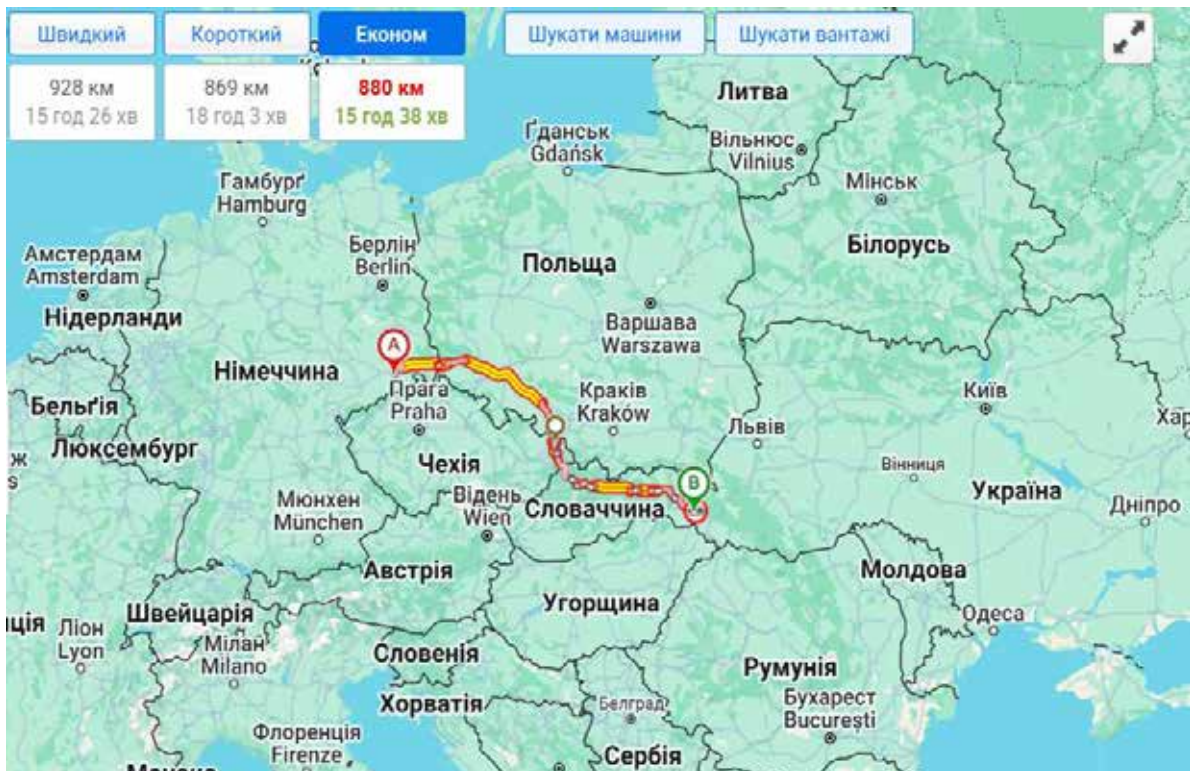


Рис. 3. Маршрут № 3 Німеччина – Україна через території Польщі, Чехії та Словаччини

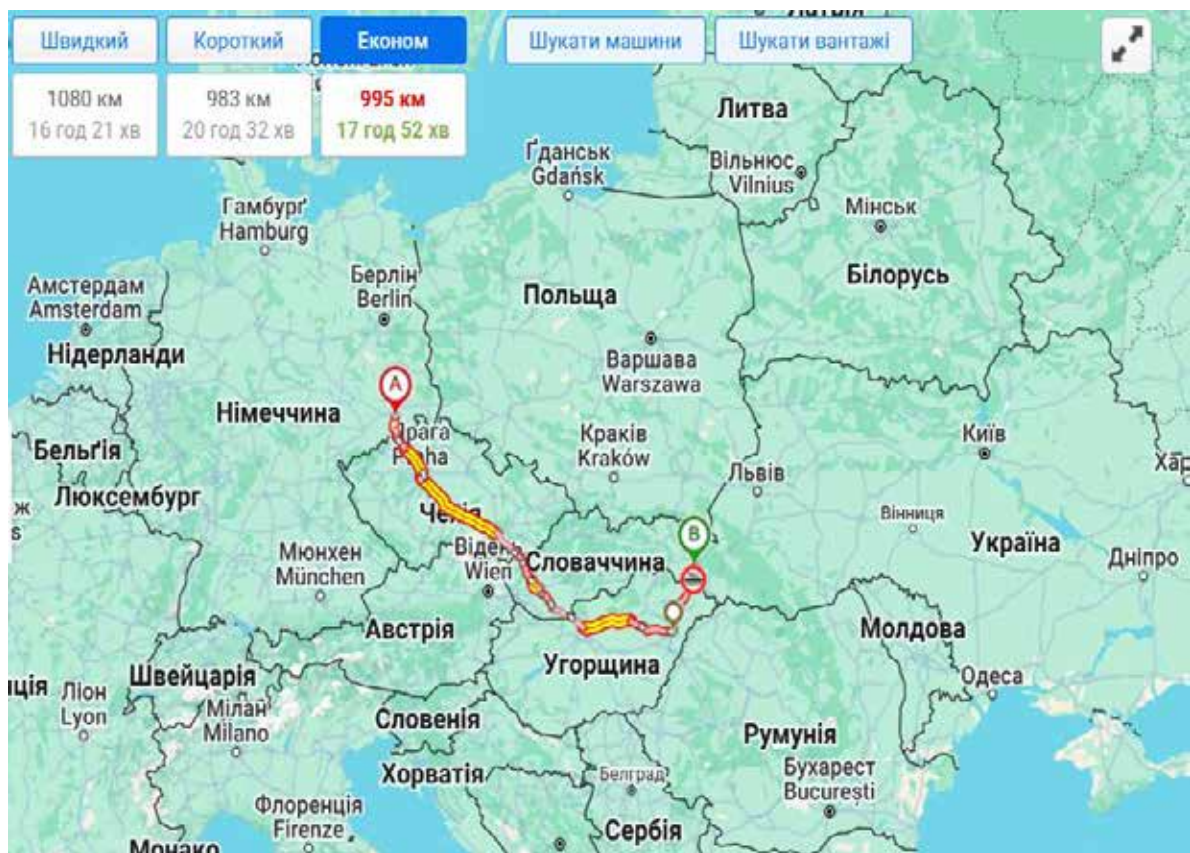


Рис. 4. Маршрут № 4 Німеччина – Україна через території Чехії, Словаччини та Угорщини

Загальна протяжність шляху прямування за маршрутом № 1 становить 828 км, а орієнтований час у дорозі – 27 годин 16 хвилин. Враховано нормативні зупинки, відпочинок водіїв та час на проходження митного контролю. Маршрут проходить через міста Болеславець і Катовіце (PL), Першов (SK) та передбачає в'їзд на територію України через пункт пропуску «Ужгород». Для другого маршруту загальна протяжність становить 872 км, орієнтований час у дорозі – близько 28 годин 40 хвилин. Перевезення здійснюється через міста Прага, Брно та Трнава з вїздом на територію України через пункт пропуску «Ужгород» [7].

Для маршруту № 3 загальна протяжність становить 935 км, орієнтований час у дорозі – близько 30 годин 10 хвилин. Рух здійснюється через великі транзитні міста – Вроцлав, Острава, Жиліна – з вїздом на українську територію через пункт «Ужгород» за рахунок інфраструктури в межах Чехії та Словаччини. Загальна протяжність за четвертим маршрутом становить 995 км, орієнтовний час у дорозі – приблизно 30 годин 55 хвилин. Рух здійснюється через великі транспортні коридори: Прага, Жиліна, Будапешт, з подальшим вїздом до України через пункт пропуску «Захонь – Чоп» [7].

У процесі організації міжнародних автомобільних перевезень автозапчастин важливо забезпечити узгодженість усіх етапів логістичного ланцюга – від формування вантажного місця на складі до остаточної доставки на склад одержувача. Для цього застосовуються методи мережевого планування, які дозволяють ефективно моделювати послідовність робіт, враховувати часові витрати на кожен етап транспортування та виявляти критичні ділянки маршруту [8]. Мережеве планування має ряд переваг: забезпечує наочність технологічної послідовності робіт; дозволяє скласти оперативні та поточні плани, а також прогнозувати складні процеси; дозволяє виявити приховані ресурси часу і матеріальних засобів при виконанні виробничих процесів. Процес мережевого планування і управління (МПУ) містить чотири взаємопов'язаних етапи [8]:

1. Опис комплексу робіт, визначення їх тривалості та послідовності.

2. Побудова мережевого графіка.

3. Розрахунок і аналіз параметрів мережевого графіка.

4. Оптимізація мережевого графіка, контроль і оперативне управління ходом виконання комплексом робіт.

У рамках даного дослідження було розроблено мережевий графік просування логістичних потоків за чотирма маршрутами: Німеччина – Польща – Словаччина – Україна; Німеччина – Чехія – Словаччина – Україна; Німеччина – Польща – Чехія – Словаччина – Україна та Німеччина – Чехія – Словаччина – Угорщина – Україна. Для кожного з маршрутів встановлено послідовність робіт: оформлення документів, завантаження на складі відправника, перетин міжнародних пунктів пропуску, відпочинок водія, технічні зупинки, митне оформлення, доставка на склад в Ужгороді. Для кожної роботи задається очікувана тривалість у годинах (або днях), що дозволяє проаналізувати загальну ефективність маршруту та визначити резерви часу.

На (рис. 5) представлено мережевий графік, який охоплює всі чотири розглянуті маршрути.

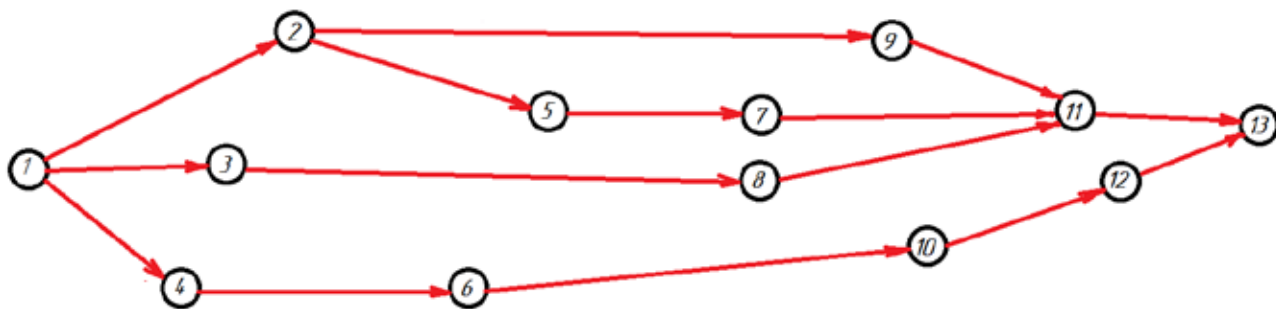


Рис. 5. Мережевий графік просування логістичних потоків за альтернативними маршрутами

Умовні позначення в кружечках [7]:

1 – м. Дрезден, Німеччина;

2 – МАПП Świecko – Frankfurt (Oder) (Німеччина-Польща);

3 – МАПП Breitenau – Petrovice (Німеччина-Чехія);

4 – м. Вроцлав, Польща;;

5 – Катовіце / Краків, Польща;

6 – м. Прага / Пльзень, Чехія;

7 – МАПП Barwinek – Vyšný Komárnik (Польща-Словаччина);

8 – МАПП Břeclav – Kúty (Чехія-Словаччина);

9 – м. Жиліна / Банська Бистриця, Словаччина;

10 – МАПП Rajka – Hegyeshalom (Словаччина-Угорщина);

11 – МАПП Vyšné Nemecké – Ужгород (Словаччина-Україна);

12 – МАПП Захонь – Чоп (Тиса) (Угорщина–Україна);

13 – м. Ужгород, Україна.

Для реалізації сітьового планування доставки автозапчастин було складено повний перелік логістичних робіт, які виконуються на різних етапах маршруту. Час доставки за різними маршрутами представлено на (рис. 6).



Рис. 6. Порівняльна діаграма часу логістичної доставки за різними міжнародними маршрутами

Прийняття рішень в умовах невизначеності є ключовим аспектом аналізу транспортної логістики. Вибір оптимального маршруту міжнародних перевезень залежить від численних економічних, технічних та логістичних факторів. У транспортній галузі широко застосовуються методи теорії прийняття рішень, які дозволяють оцінити ефективність варіантів перевезень за різними критеріями. Серед найбільш поширених підходів, які використовуються для ухвалення рішень в умовах невизначеності, виділяють критерії Лапласа, Вальда, Севіджа та Гурвіца. Вони дозволяють оцінити ризики та обрати оптимальну стратегію перевезень [9].

Коротко розглянемо сутність кожного з означених критеріїв.

Суть критерію Лапласа (принцип рівномірності) полягає у наступному. Якщо немає інформації про ймовірність можливих станів середовища (наприклад, ситуації на митниці, затримки, попит), вважають, що усі варіанти рівномірні. Обирають рішення з найвищим середнім виграшем (або найменшими середніми витратами). У транспортній логістиці даний критерій застосовується, коли немає статистики щодо затримок або часу доставки новими маршрутами. Також він використовується для оцінки альтернативних перевізників чи напрямків за середньою ефективністю [10]. У нашому випадку ймовірності затримок невідомі, тому пропонується обчислювати середній час доставки або середні витрати для кожного маршруту та обирати найкращий за критерієм Лапласа.

Суть критерію Вальда (максимінного, або критерію песимізму) полягає у тому, що приймається найбільш обережне (песимістичне) рішення. Для кожної альтернативи обирають найгірший можливий результат, а потім – варіант з найкращим серед найгірших. У транспортній логістиці він застосовується при високому рівні ризику або нестабільності (наприклад, воєнні дії, страйки, санкції) та допомагає мінімізувати потенційні втрати, навіть якщо прибуток буде менший [10].

Суть критерію Севіджа (мінімаксного жалю, або мінімізації втрат) полягає у тому, що він розглядає не самі результати, а «втрати можливостей» (жалю), які виникають, якщо рішення виявиться не найкращим. Обирають рішення, для якого максимальне можливе жалю мінімальне. У транспортній логістиці цей критерій підходить, коли потрібно зменшити ризик неправильного вибору маршруту чи постачальника [10]. Також він може бути корисним у ситуаціях, коли можна порівняти потенційні упущені вигоди (наприклад, через запізнення, штрафи, або втрату клієнтів).

Критерій Гурвіца (оптимізму–песимізму) виражає компроміс між оптимістичним і песимістичним підходами. Використовується коефіцієнт оптимізму α ($0 \leq \alpha \leq 1$). У транспортній логістиці він підходить для збалансованого прийняття рішень за умов часткової невизначеності та може бути корисним, коли менеджер впевнений частково, але також враховує і ризики [10]. Підсумкове порівняння критеріїв зведено в (табл. 1).

Таблиця 1

Порівняння критеріїв Лапласа, Вальда, Севіджа та Гурвіца

Критерій	Ставлення до ризику	Тип рішень	Приклад у логістиці
Лапласа	Нейтральне	Середнє значення результатів	Вибір маршруту без статистики затримок
Вальда	Песимістичне	Мінімізація втрат у найгіршому випадку	Перевезення через нестабільний регіон
Севіджа	Обережне	Мінімізація упущених можливостей	Вибір перевізника з мінімальним «жалем»
Гурвіца	Збалансоване	Компроміс оптимізму та песимізму	Вибір постачальника за часткової невизначеності

Вибір оптимального маршруту розподілу логістичних потоків міжнародної перевезення доставки вантажів є складним процесом. Оцінка маршрутів за критеріями Лапласа, Вальда, Севіджа та Гурвіца дозволяє логістичним операторам приймати обґрунтовані рішення щодо ефективності логістичної доставки.

Для подальшого аналізу було визначено орієнтовні логістичні витрати за кожним із чотирьох маршрутів, що дозволило застосовувати методи теорії прийняття рішень для оцінки вибору оптимального маршруту (див. табл. 2).

Таблиця 2

Оцінка вартості за критеріями прийняття рішень

Маршрути	Вартість, €	Відносна оцінка (вартість, мін.)
Маршрут № 1	1250,67	1,00
Маршрут № 2	1308,67	1,05
Маршрут № 3	1368,67	1,09
Маршрут № 4	1450,50	1,16

Усі вхідні дані – орієнтовні вартості перевезень – були нормалізовані шляхом поділу на найменше значення серед них. Це дозволило уніфікувати оцінки та зробити результати порівнюваними між маршрутами. Розрахунок для маршруту № 1.

$$\text{Критерій Лапласа: } L_1 = \frac{1+1+1}{3} = 1,00$$

$$\text{Критерій Вальда: } W_1 = \max(1) = 1,00$$

$$\text{Критерій Севіджа: } S_1 = \max(1) = 1,00$$

$$\text{Критерій Гурвіца: } H_1 = 0,6 \cdot 1 + 0,4 \cdot 1 = 1,00$$

Розрахунок для маршруту № 2.

$$\text{Критерій Лапласа: } L_2 = \frac{1,05 + 1,05 + 1,05}{3} = 1,05$$

$$\text{Критерій Вальда: } W_2 = \max(1,05) = 1,05$$

$$\text{Критерій Севіджа: } S_2 = \max(1,16 - 1,05) = 0,11$$

$$\text{Критерій Гурвіца: } H_2 = 0,6 \cdot 1,05 + 0,4 \cdot 1 = 1,03$$

Розрахунок для маршруту № 3.

$$\text{Критерій Лапласа: } L_3 = \frac{1,09 + 1,09 + 1,09}{3} = 1,09$$

$$\text{Критерій Вальда: } W_3 = \max(1,09) = 1,09$$

$$\text{Критерій Севіджа: } S_3 = \max(1,16 - 1,09) = 0,07$$

$$\text{Критерій Гурвіца: } H_3 = 0,9 \cdot 1,05 + 0,4 \cdot 1 = 1,05$$

Розрахунок для маршруту № 4.

$$\text{Критерій Лапласа: } L_4 = \frac{1,16 + 1,16 + 1,16}{3} = 1,16$$

$$\text{Критерій Вальда: } W_4 = \max(1,16) = 1,16$$

$$\text{Критерій Севіджа: } S_4 = \max(1,16 - 1,16) = 0,00$$

$$\text{Критерій Гурвіца: } H_4 = 0,6 \cdot 1,16 + 0,4 \cdot 1 = 1,10$$

У таблиці 3 наведено узагальнені результати розрахунків, що дозволяють наочно порівняти маршрути та зробити остаточні висновки щодо оптимізації міжнародних логістичних потоків на підставі методів теорії прийняття рішень в умовах невизначеності.

Таблиця 3

Порівняльна оцінка маршрутів розподілу логістичних потоків за критеріями прийняття рішень

Критерій	Маршрут № 1	Маршрут № 2	Маршрут № 3	Маршрут № 4
Лапласа	1,00	1,05	1,09	1,16
Вальда	1,00	1,05	1,09	1,16
Севіджа	0,16	0,11	0,07	0,00
Гурвіца $\alpha = 0,6$	1,00	1,03	1,05	1,10

У результаті проведених розрахунків було отримано кількісні оцінки ефективності кожного з чотирьох розглянутих маршрутів та побудовано наочні діаграми для критеріїв Лапласа (див. рис. 7), Вальда (див. рис. 8), Севіджа (див. рис. 9) та Гурвіца (див. рис. 10). Це дозволило комплексно оцінити маршрути з позицій середніх витрат, максимально можливих втрат, рівня ризику та компромісного підходу між песимістичним і оптимістичним сценаріями.

Для оцінки маршрутів за критерієм Лапласа було враховано середнє значення виграшу для кожного маршруту, виходячи з припущення рівноймовірності можливих сценаріїв. Як видно з рисунка 7, найкращий показник отримав маршрут № 1 з оцінкою 1,00, що свідчить про його стабільну ефективність у середньому випадку. Зі збільшенням кількості транзитних країн середнє значення витрат також зростає, що негативно впливає на загальну привабливість інших маршрутів.

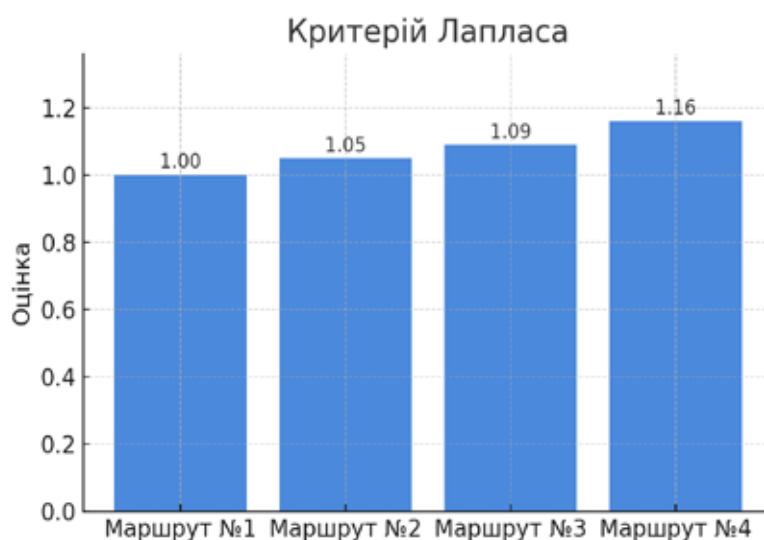


Рис. 7. Порівняльна діаграма розрахунку міжнародних маршрутів за критерієм Лапласа

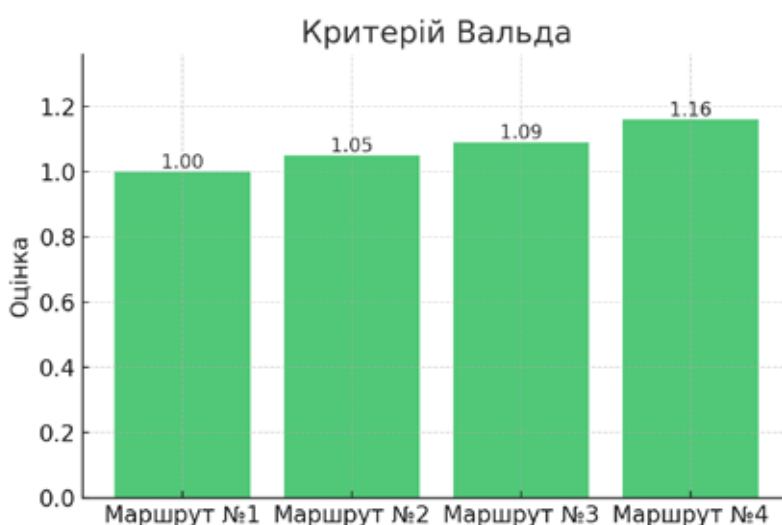


Рис. 8. Порівняльна діаграма розрахунку міжнародних маршрутів за критерієм Вальда



Рис. 9. Порівняльна діаграма розрахунку міжнародних маршрутів за критерієм Севіджа



Рис. 10. Порівняльна діаграма розрахунку міжнародних маршрутів за критерієм Гурвіца

Критерій Вальда демонструє найгірший можливий варіант для кожного маршруту. Цей підхід особливо доцільний у випадках, коли логістичний оператор прагне уникнути максимально негативного сценарію. Як показано на рисунку 8, найменші витрати у найгіршому сценарії знову має маршрут № 1. Найвищий показник характерний для маршруту № 4, що робить його менш привабливим з точки зору ризиків.

Аналіз за критерієм Севіджа базується на обчисленні максимального «розчарування» – тобто різниці між оптимальним варіантом і поточним. На рисунку 9 видно, що мінімального розчарування зазнає маршрут № 4 ($S = 0$), тобто незалежно від обставин цей маршрут не буде найгіршим. Найбільше значення спостерігається для маршруту № 1, оскільки його найкращий результат не гарантує перевагу в усіх можливих сценаріях.

Рисунок 10 ілюструє динаміку змін критеріїв залежно від рівня ризику або оптимізму особи, яка приймає рішення (зокрема, при зміні коефіцієнта α у критерії Гурвіца). Це дозволяє адаптувати модель вибору маршруту під реальні умови транспортної ситуації, наприклад, у разі прогнозування загострень на кордоні чи погіршення дорожньої інфраструктури. Критерій Гурвіца поєднує в собі елементи оптимістичного та песимістичного підходів. При заданому коефіцієнті $\alpha = 0.6$, що відповідає середньому рівню оптимізму, найбільш збалансованим маршрутом знову виступає маршрут № 1 ($H = 1.00$). Водночас маршрут № 4 отримав найвищу оцінку, що свідчить про його підвищені ризики у поєднанні з вартістю доставки.

Результати даної роботи можуть бути корисними за умов управлінської невизначеності в логістиці: неповної інформації для прийняття рішень, коли менеджери не мають достовірних прогнозів щодо попиту,

витрат чи часу доставки; невизначеності у виборі оптимального маршруту або постачальника, коли важко оцінити, який варіант буде найвигіднішим у конкретних умовах.

Невизначеність у міжнародній логістичній доставці проявляється у коливаннях часу, вартості, надійності та безпеки перевезень. Для її зменшення пропонується використовувати: цифрові системи моніторингу (GPS, IoT, блокчейн); страхування ризиків; сценарне планування та критерії прийняття рішень (Лапласа, Вальда, Севіджа, Гурвіца); диверсифікацію маршрутів і партнерів; створення резервних планів постачання (contingency plans).

Створімо для нашого прикладу матрицю ризиків, яка показує, які ризики можуть виникати під час для міжнародних логістичних перевезень вантажів, наскільки вони ймовірні та які можуть мати наслідки (див. табл. 4).

Таблиця 4

Матриця ризиків у міжнародній транспортній логістиці

№	Тип ризику	Ймовірність (Р)	Вплив (І)	Рівень ризику (R = P × I)	Можливі наслідки	Заходи мінімізації
1	Затримка на митниці або кордоні	Висока (4)	Високий (4)	16 (критичний)	Порушення строків поставки, штрафи	Попередня митна декларація, використання “єдиного вікна”, вибір менш завантажених пунктів пропуску
2	Зміна митних або торговельних правил	Середня (3)	Високий (4)	12 (високий)	Збільшення витрат, простої, переоформлення документів	Постійний моніторинг законодавства, співпраця з брокером
3	Колівання вартості фрахту або пального	Висока (4)	Середній (3)	12 (високий)	Зростання транспортних витрат, зниження прибутковості	Фіксація тарифів у контрактах, використання паливних хеджів
4	Пошкодження або втрата вантажу	Низька (2)	Високий (4)	8 (помірний)	Фінансові втрати, зниження довіри клієнта	Страхування вантажів, контроль упаковки, GPS-моніторинг
5	Політична або воєнна нестабільність у транзитній країні	Середня (3)	Дуже високий (5)	15 (критичний)	Блокування маршруту, затримка або втрата товару	Диверсифікація маршрутів, альтернативні перевізники
6	Погодні умови (буря, ожеледиця)	Середня (3)	Середній (3)	9 (помірний)	Затримки у доставці	Гнучке планування, врахування сезонності
7	Збій IT-систем або кібератака	Низька (2)	Високий (4)	8 (помірний)	Втрата даних, порушення координації	Резервне копіювання, кіберзахист
8	Брак транспорту або водіїв	Середня (3)	Середній (3)	9 (помірний)	Втрата часу, порушення графіку	Контракти з кількома перевізниками, кадровий резерв
9	Неточність прогнозу попиту	Висока (4)	Середній (3)	12 (високий)	Надлишкові або дефіцитні запаси	Використання аналітики, регулярне оновлення прогнозів
10	Помилка персоналу при оформленні документів	Середня (3)	Середній (3)	9 (помірний)	Затримка митного оформлення	Навчання персоналу, подвійна перевірка документів

Отримана для оптимального маршруту матриця ризиків допомагає візуально визначити найнебезпечніші чинники (червона зона); розставити пріоритети у роботі з ризиками; розробити план дій для мінімізації впливу кожного ризику.

Висновки. Оптимізація міжнародних логістичних потоків в умовах невизначеності є складним, але необхідним завданням для підвищення ефективності зовнішньоекономічної діяльності. Застосування методів теорії прийняття рішень, зокрема критеріїв Лапласа, Вальда, Севіджа та Гурвіца, дає можливість оцінити ризики, зважити ймовірні наслідки та обрати оптимальний варіант логістичної стратегії. Використання цих підходів сприяє зменшенню впливу невизначених факторів (політичних, економічних, технологічних) і підвищує стійкість ланцюгів постачання.

Таким чином, інтеграція методів прийняття рішень у систему управління міжнародною логістикою забезпечує збалансоване поєднання ефективності, надійності та гнучкості логістичних процесів.

Отже, отримані в даній роботі результати дозволяють не лише визначити оптимальний маршрут у межах кожного окремого критерію, а й виявити загальні тенденції, які можуть бути використані при прийнятті логістичних рішень. Наприклад, завдяки візуалізації даних логістичні оператори можуть обґрунтовано обрати маршрут, який забезпечить мінімальні витрати та найменші логістичні ризики в довгостроковій перспективі.

Список використаних джерел:

1. Верещагіна Г. В., Сушко К. В. Оптимізація транспортно-логістичної діяльності підприємства. *Матеріали конференції «Конкурентоспроможність та інновації: проблеми науки та практики»*. Харків, 2018. С. 126–129. URL: <https://ndc-ipr.org/media/publications/files>
2. Малиновський Ю. В., Ремез О. С., Сорока В. В., Гринчук Т. В. Адаптивні стратегії управління міжнародною логістикою в умовах глобальних криз. *European scientific journal of Economic and Financial innovation*. № 4 (18), 2025. DOI: <http://doi.org/10.32750/2025-0414>
3. Сколоздра Ю. І. Оптимізація логістичних процесів в імпорتنій діяльності. *Економіка та суспільство*. Вип. № 73. 2025. DOI: <https://doi.org/10.32782/2524-0072/2025-73-128>
4. Гриценко С. І., Дейкун В. О. Оптимізація логістичних процесів підприємства в умовах цифрової трансформації (на прикладі Kuehne + Nagel). *Вісник економічної науки України*. № 1 (48), 2025. С. 134–139. DOI: [https://doi.org/10.37405/1729-7206.2025.1\(48\).134-139](https://doi.org/10.37405/1729-7206.2025.1(48).134-139)
5. Вакуленко В. Л., Лю Юнтао, Лю Сяовей. Система ризик-менеджменту логістичних систем в умовах воєнного стану України. *Сталий розвиток економіки*. № 3 (54), 2025. С. 296–300. DOI: <https://doi.org/10.32782/2308-1988/2025-54-45>
6. Flagma.ua. Розрахунок відстаней маршрутів URL: <https://flagma.ua/uk/rozrahunok-vidstaney.html>
7. Азарян Б. С. Проектування транспортно-логістичної схеми доставки автозапчастин у міжнародному сполученні : випускна кваліфікаційна робота на здобуття освітнього ступеня «бакалавр» за спеціальністю 275 Транспортні технології (на автомобільному транспорті). Науковий керівник – к.т.н., доц. Кузьменко А. І. Університет митної справи та фінансів, Дніпро, 2025. 77 с.
8. Кужель В. Біліченко В. Моделювання технологічних процесів підприємств автомобільного транспорту : навчальний посібник. Вінниця : ВНТУ, 2017. 163 с. URL: http://pdf.lib.vntu.edu.ua/books/2022/Bilichenko_2017_163.pdf
9. Savidge, Gurvitz et al. Methods for evaluating international transport routes. 2021. URL: <https://studfile.net/preview/2798320/page:27/>
10. Український Чат GPT 5. URL: <https://chatgpt.com/g/g-S8ztDuUKM-ukrainskii-chat-gpt/c/68fddb20-9ee4-832b-a860-10e6dfc0a599>

References:

1. Vereshchahina, H. V., & Sushko, K. V. (2018). Optymizatsiia transportno-lohistychnoi diialnosti pidpriemstva [Optimization of transport and logistics activity of the enterprise]. In *Materialy konferentsii "Konkurentospromozhnist ta innovatsii: problemy nauky ta praktyky"* (pp. 126–129). Kharkiv. Retrieved from: <https://ndc-ipr.org/media/publications/files>
2. Malynovskyi, Yu. V., Remez, O. S., Soroka, V. V., & Hrynychuk, T. V. (2025). Adaptivni stratehii upravlinnia mizhnarodnoiu lohistykoiu v umovakh hlobalnykh kryz [Adaptive strategies for managing international logistics under global crises]. *European Scientific Journal of Economic and Financial Innovation*, 4(18). <https://doi.org/10.32750/2025-0414>
3. Skolozdra, Yu. I. (2025). Optymizatsiia lohistychnykh protsesiv v importnii diialnosti [Optimization of logistics processes in import activities]. *Ekonomika ta suspilstvo*, 73. <https://doi.org/10.32782/2524-0072/2025-73-128>
4. Hrytsenko, S. I., & Deikun, V. O. (2025). Optymizatsiia lohistychnykh protsesiv pidpriemstva v umovakh tsyfrovoi transformatsii (na prykladi Kuehne + Nagel) [Optimization of enterprise logistics processes in the context of digital transformation (on the example of Kuehne + Nagel)]. *Visnyk Ekonomichnoi Nauky Ukrainy*, 1(48), 134–139. [https://doi.org/10.37405/1729-7206.2025.1\(48\).134-139](https://doi.org/10.37405/1729-7206.2025.1(48).134-139)
5. Vakulenko, V. L., Liu, Yuntao, & Liu, Xiaowei. (2025). Systema ryzyk-menedzhmentu lohistychnykh system v umovakh voiennoho stanu Ukrainy [Risk management system of logistics systems under martial law in Ukraine]. *Stalyi Rozvytok Ekonomiky*, 3(54), 296–300. <https://doi.org/10.32782/2308-1988/2025-54-45>
6. Flagma.ua. (n.d.). Rozrahunok vidstanei marshrutiv [Distance calculation of routes]. Retrieved from: <https://flagma.ua/uk/rozrahunok-vidstaney.html>
7. Azarian, B. S. (2025). Proiektuvannia transportno-lohistychnoi skhemy dostavky avtozapchastyn u mizhnarodnomu spoluchenni [Design of a transport and logistics delivery scheme for auto parts in international connection] (Bachelor's thesis, University of Customs and Finance). Dnipro: University of Customs and Finance.

-
8. Kuzhel, V., & Bilichenko, V. (2017). Modeliuvannia tekhnolohichnykh protsesiv pidpriemstv avtomobilnoho transportu: Navchalnyi posibnyk [Modeling of technological processes of road transport enterprises: Textbook]. Vinnytsia: VNTU. Retrieved from: http://pdf.lib.vntu.edu.ua/books/2022/Bilichenko_2017_163.pdf
9. Savidge, Gurvitz, et al. (2021). Methods for evaluating international transport routes. Retrieved from: <https://studfile.net/preview/2798320/page:27/>
10. Ukrainian Chat GPT 5. (n.d.). Retrieved from: <https://chatgpt.com/g/g-S8ztDuUKM-ukrainskii-chat-gpt/c/68fddb20-9ee4-832b-a860-10e6dfc0a599>

Дата надходження статті: 26.10.2025

Дата прийняття статті: 17.11.2025

Опубліковано: 30.12.2025

UDC 656.078:656.1

DOI <https://doi.org/10.32782/2521-6643-2025-2-70.31>

Muzykin M. I., PhD in Technical Sciences, Associate Professor,
Associate Professor at the Department of Transport Technologies
and International Logistics
University of Customs and Finance
ORCID: 0000-0003-2938-7061

Nesterenko H. I., PhD in Technical Sciences, Associate Professor,
Associate Professor at the Department of Transport Service
and Logistics
Ukrainian State University of Science and Technologies
ORCID: 0000-0003-1629-0201

Bibik S. I., PhD in Technical Sciences, Associate Professor,
Associate Professor at the Department of Transport Technologies
and Transportation Process Management
National Transport University
ORCID: 0000-0002-5832-6949

Chubenko O. I., Postgraduate Student
Ukrainian State University of Science and Technologies
ORCID: 0009-0007-5157-8020

LOGISTICS METHODS FOR MANAGING THE TRANSPORTATION OF HUMANITARIAN CARGO NEAR THE WAR ZONE

This article analyzes existing logistics systems, identifies problematic aspects of their functioning, and develops practical recommendations for optimizing the transportation of humanitarian cargo. The logistical delivery of humanitarian cargo faces many challenges near the war zone. The authors research the impact of various factors, including security risks for carriers, limited access to transport and infrastructure, adverse weather conditions, and insufficient coordination between different humanitarian organizations and local authorities. The characteristics of cargoes provided to citizens as humanitarian aid are considered, and a means of transport for transporting goods from large warehouses to intermediate storage facilities is selected, and transport routes are developed, taking into account the uncertainty of route selection due to shelling. Simulating the routes for delivering humanitarian aid to frontline settlements has been carried out. The research emphasizes that efficient transportation of humanitarian cargo from central warehouses to frontline areas is a complex but solvable challenge. Transporting cargoes from central warehouses to intermediate ones, developing routes for further transportation to humanitarian aid distribution points, and optimising the transport fleet are the main stages in developing an effective transport and logistics system. Specialised algorithms of linear programming tasks were used to allocate vehicles and minimise costs. In the event of a danger associated with hostilities, the availability of developed alternative routes ensures continuity of supply. Risk probability analysis makes it possible to choose the safest routes.

Key words: humanitarian logistics, frontline areas, alternative logistics routes, humanitarian cargo.

Музикін М. І., Нестеренко Г. І., Бібік С. І., Чубенко О. І. Логістичні методи управління перевезеннями гуманітарних вантажів поблизу зони бойових дій

Статтю присвячено аналізу існуючих логістичних систем, виявленню проблемних аспектів їх функціонування, а також розробці практичних рекомендацій для оптимізації процесів транспортування гуманітарних вантажів. Поблизу зони бойових дій логістична доставка гуманітарних вантажів стикається з багатьма труднощами. Досліджується вплив різних факторів, зокрема безпекові ризики для перевізників, обмежений доступ до транспорту та інфраструктури, складні погодні умови, а також недостатня координація між різними гуманітарними організаціями та місцевими органами влади. Розглянуто характеристику вантажів, які надаються громадянам у якості гуманітарної допомоги, обрано транспортний засіб для перевезення вантажів з великих складів до проміжних складів та розроблено маршрути перевезення, враховуючи невизначеність вибору маршруту через обстріли. Виконано моделювання маршрутів розвезення гуманітарної допомоги в прифронтові населені пункти. В дослідженні підкреслюється, що ефективне перевезення

© M. I. Muzykin, H. I. Nesterenko, S. I. Bibik, O. I. Chubenko, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

гуманітарних вантажів від центральних складів до прифронтових районів є складною, але вирішуваною задачею. Перевезення вантажів від центральних складів до проміжних, розробка маршрутів для подальшого транспортування до місць видачі гуманітарної допомоги та оптимізація транспортного парку є основними етапами розробки ефективної транспортно-логістичної системи. Під час виконання розподілу транспортних засобів та мінімізації витрат використано спеціалізовані алгоритми задач лінійного програмування. У разі наявності небезпеки пов'язаної з бойовими діями, наявність розроблених альтернативних маршрутів забезпечує безперервність постачання. Аналіз ймовірності ризиків надає можливість обрати найбезпечніші маршрути.

Ключові слова: гуманітарна логістика, прифронтові райони, альтернативні логістичні маршрути, гуманітарні вантажі.

Problem statement. The modern world faces numerous challenges, one of which is conflicts and armed confrontations, which create a need for humanitarian assistance. In situations where the lives and well-being of millions of people are at risk, it is especially crucial to ensure the efficient transport of humanitarian cargoes. The timely delivery of such cargoes can be vital for saving human lives, supporting the survival of affected communities, and providing basic needs [1].

Logistics strategies for handling the transportation of humanitarian cargoes near war zones are critically important in today's world, where armed conflicts leave millions without livelihoods and essential protections. That is why this article is relevant.

Analysis of recent research and publications. Humanitarian cargo includes a wide range of vital resources, such as medicines, food, clean drinking water, clothing, and other essential items [3]. However, the logistics of these cargoes near the combat zone are often accompanied by numerous difficulties and challenges [12]. Among them, it is worth highlighting security risks for carriers, limited access to transport and infrastructure, difficult weather conditions, as well as insufficient coordination between various humanitarian organizations and local authorities [8].

We have witnessed and participated in another wave of unprecedented mobilization not only of the Ukrainian people, but also of people around the world, to assist both the Armed Forces of Ukraine and the Territorial Defense Forces, as well as to provide humanitarian assistance to those affected by russian military aggression. At the same time, we are forced to admit that the network potential of Ukrainians for cooperation, which ensured the supply chains of necessary assistance, although it played a critically important role, remained untapped due to the lack of effective coordination mechanisms. Logistics specialists in the global humanitarian community are making increasing efforts to formulate principles, protocols, and a structure for effective interaction in the field of humanitarian aid delivery. [9]

The team of authors in the work [11] proposes to use the method of hierarchy analysis to determine the criteria for optimization and selection of the optimal transportation route in conditions of military conflict.

The article states [10] that volunteer organizations play a crucial role in meeting the humanitarian needs of the population in Ukraine. The research analyzes in detail the volumes and types of assistance provided by various organizations, and also emphasizes the importance of coordinated logistical processes and effective organization. Much attention is paid to humanitarian logistics, which needs further development.

The purpose of the article is to comprehensively study methods and strategies that can increase the efficiency of humanitarian cargo transportation near the combat zone. The research aims to conduct an in-depth analysis of existing logistics systems, identify problematic aspects of their functioning, and develop practical recommendations for optimizing transportation processes. The main focus is on studying the best international practices, implementing innovative approaches, and utilizing modern technologies to enhance transportation efficiency.

Presentation of the main material. More than three years have passed since the full-scale invasion of Ukraine began, and the humanitarian needs of the affected population remain high. An estimated 14,6 million people are in urgent need of humanitarian assistance and social and legal protection services in 2024. Since the start of the year, people across Ukraine have been suffering from widespread rocket attacks, which have killed and injured civilians and damaged homes and critical infrastructure. The ongoing fighting in frontline and border communities in the north, south, and east of the country is negatively impacting the population and forcing people to evacuate. From January to March 2024, a total of 9,700 attacks were recorded across Ukraine. [1]

As of the end of December 2023, almost 907,000 people in Ukraine had received assistance to get through the autumn-winter period from the United Nations High Commissioner for Refugees (UNHCR) and partners. It provides cash assistance of UAH 6,660 per person. This aid is planned to reach over 617,900 people with special vulnerability criteria to help them cover increased expenses during the winter months. [1]

Almost 249,000 people received essential items to get through the winter period (namely thermal blankets and regular blankets), of which 77,000 people received winter clothing.

Over 11,800 people received housing assistance during the autumn-winter season. This includes nearly 8,500 individuals who received seasonal home insulation kits to keep their homes warm, such as materials traditionally used in Ukraine to improve thermal insulation: heat-reflecting screens, transparent polyethylene film for windows, foam sealants, and construction tape.

At the same time, in close coordination with the authorities, the distribution of generators to communities continues. As of today, more than 80 of the 147 planned generators have already been provided to communities. This support will help people and communities be prepared for power outages and keep their homes warm.

To support internally displaced persons in temporary housing nationwide, UNHCR also helped repair and insulate premises, as well as maintain heating and electricity systems. This support improved the living conditions of nearly 28,200 IDPs in temporary accommodations.

In 2023, UNHCR provided multi-purpose cash assistance to 899,039 internally displaced families, returnees, and individuals with special vulnerabilities to help cover basic needs such as food, medicine, clothing, shelter, utilities, and energy costs during the autumn-winter season. This included 565,236 people who received three months of multi-purpose cash assistance and 617,926 individuals who received cash support to help offset additional costs related to the autumn-winter period. Many families with special vulnerability criteria received payments from both programs. Over \$223.6 million was allocated in 2023 to assist internally displaced and other war-affected populations through multi-purpose cash assistance and energy cost support during the colder months.

UNHCR signed a Memorandum of Understanding with the Ministry of Social Policy of Ukraine and the Pension Fund of Ukraine to provide additional support to the Ukrainian population during the winter period of 2023-2024. According to the Memorandum, UNHCR will provide cash assistance for the purchase of solid fuel for more than 85,000 families with special vulnerabilities, including households of older people and families with people with disabilities.

In December, UNHCR also provided cash assistance to 120 refugee families remaining in Ukraine to cover their basic needs and rent.

UNHCR supplies standard essential items, such as blankets, kitchen sets, and solar lamps. Non-standard items, like feminine hygiene products, mattresses, and clothing, are also being provided (Fig. 1). All this support is distributed to those in need, mainly in frontline areas and among newly displaced people, as well as in areas where access has recently been restored and which have been affected by shelling for months. The distribution of non-food items is carried out with the help of implementing partners and local authorities. Assistance is delivered directly to those in need or through participation in inter-agency responses.

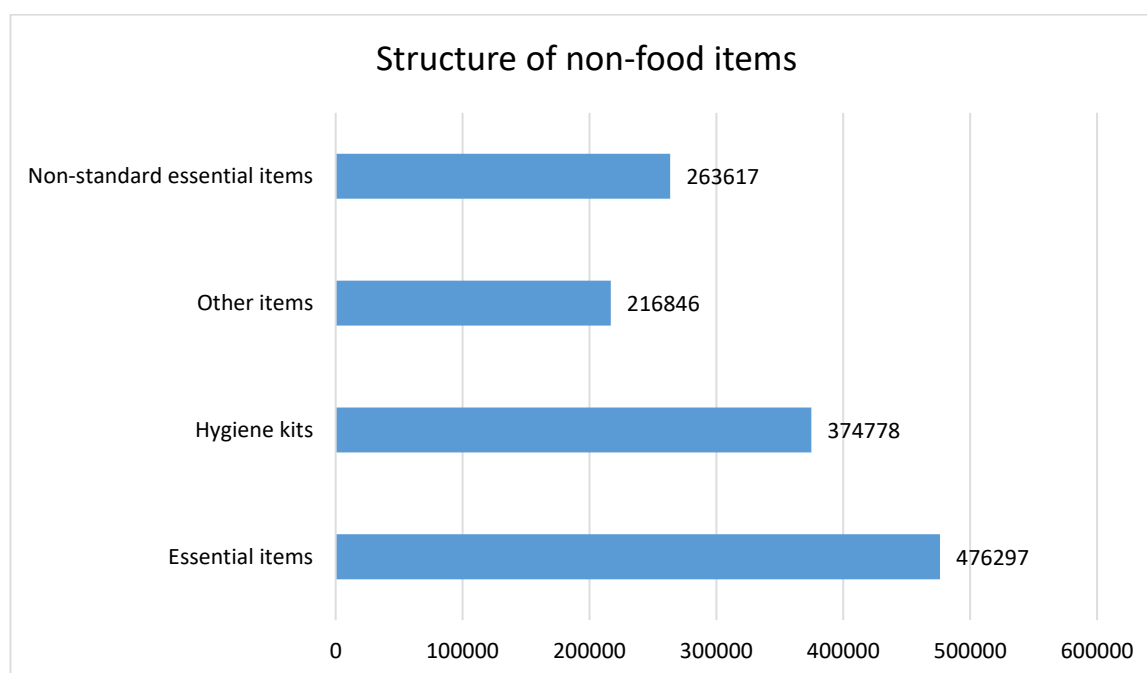


Fig. 1. Structure of non-food items

Helping war-affected families living in frontline communities meet their basic needs remains a priority. UNHCR joined and provided essential items, including tarpaulins, plastic sheeting, blankets, solar lamps, and bedding sets, for six inter-agency convoys in December, including three to Donetsk, two to Kherson, and one to Kharkiv regions.

In total, UNHCR provided 3,120 tarpaulin sheets, 1,374 solar lamps, 30 rolls of plastic sheeting, 300 blankets, 300 bedding sets, 1,500 sleeping bags, and clothing for the six convoys. These items will help families in frontline communities, including vulnerable people and people with disabilities, stay warm this winter.

Overall, more than 575,270 internally displaced and war victims have received essential items this year, including people affected by renewed shelling across the country and people living in hard-to-reach frontline areas.

A summary of the assistance received by item is shown in (Fig. 2).

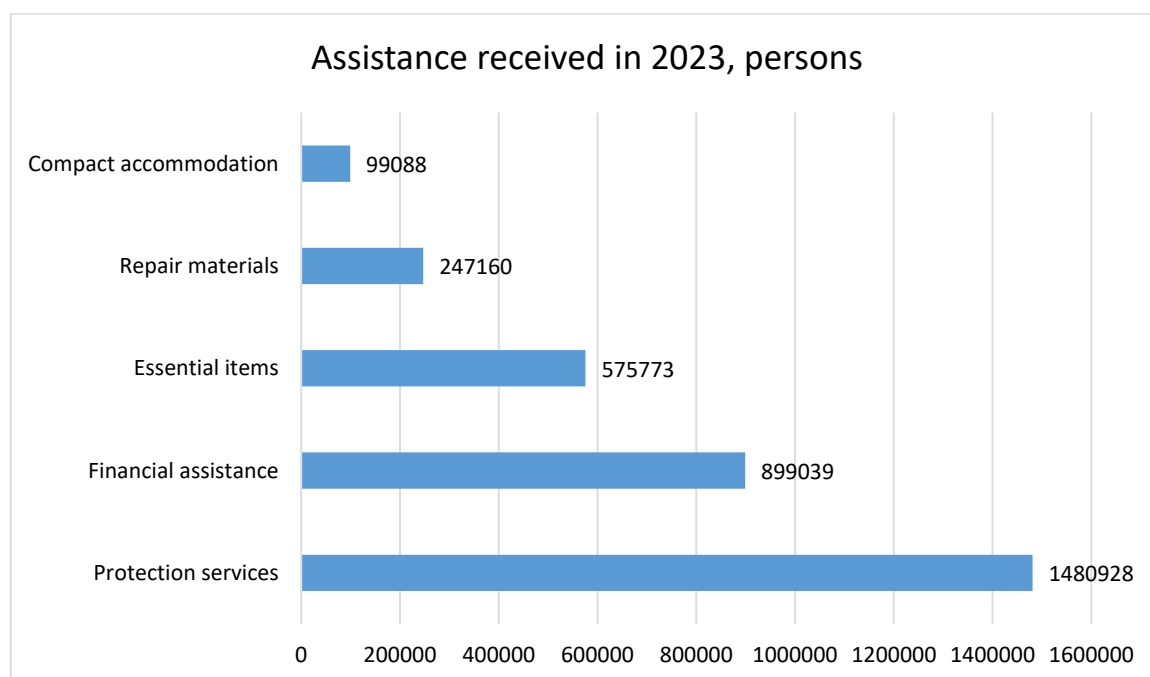


Fig. 2. Aid received in 2023, persons

UNHCR's shelter programme supports people affected by the war in three ways:

- providing emergency shelter to those in urgent need in areas directly affected by shelling;
 - repairs and housing solutions to help people return to their homes, where possible;
 - improving conditions or creating new places in collective accommodation for IDPs who cannot return home.
- The number of people who received assistance in 2024 by month is shown in (Fig. 3).

Almost 2.9 million people were provided with food and agricultural supplies, mostly in frontline communities. Almost 1.7 million people received health assistance.

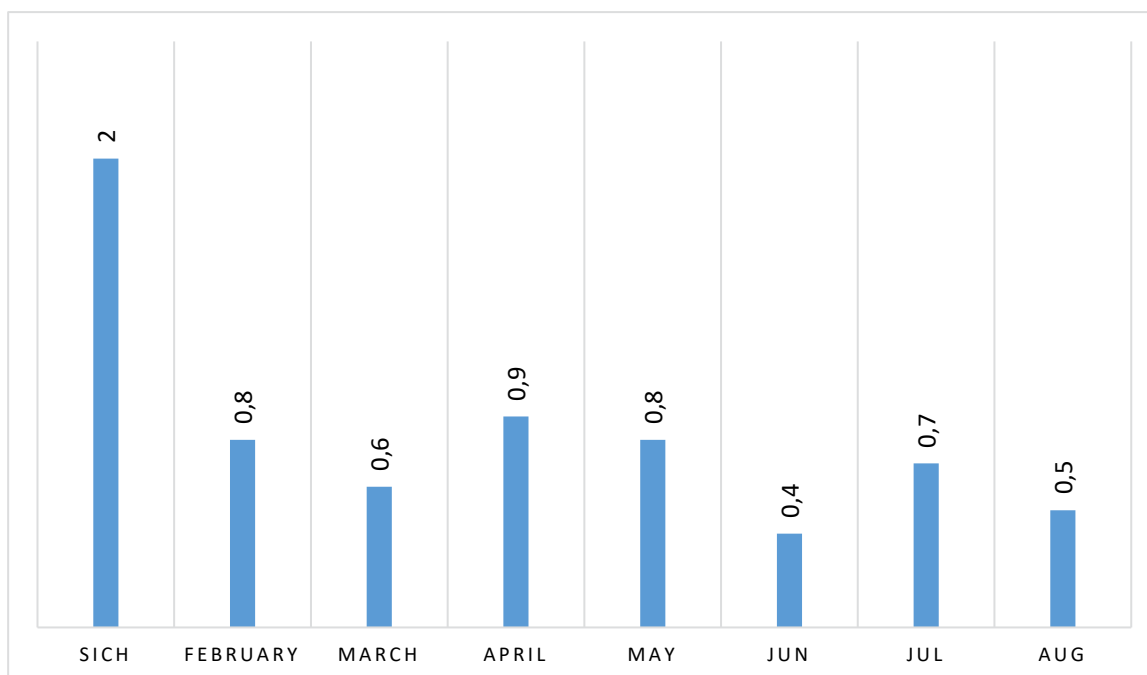


Fig. 3. Number of people who receive assistance in 2024 by months, in millions of people

To prepare for the autumn-winter season, over 1.2 million people received fuel and supplies to repair damaged homes, along with other non-food items. Humanitarian workers also offered protection services, including social and legal support, child protection, gender-based violence prevention and survivor assistance, and mine action, such as explosive ordnance risk education.

To reach frontline areas where access is difficult due to the security situation, humanitarian partners delivered 28 inter-agency convoys with assistance for almost 45,000 people in Donetsk, Zaporizhia, Kharkiv, and Kherson regions, in addition to regular humanitarian assistance activities. The distribution of people assisted by region is shown in (Fig. 4).



Fig. 4. Distribution of people by region [5]

Among the statuses of citizens, assistance is provided to citizens of the following statuses: internally displaced persons (IDPs), people who have returned to their places of residence, and non-displaced (Fig. 5) [5].

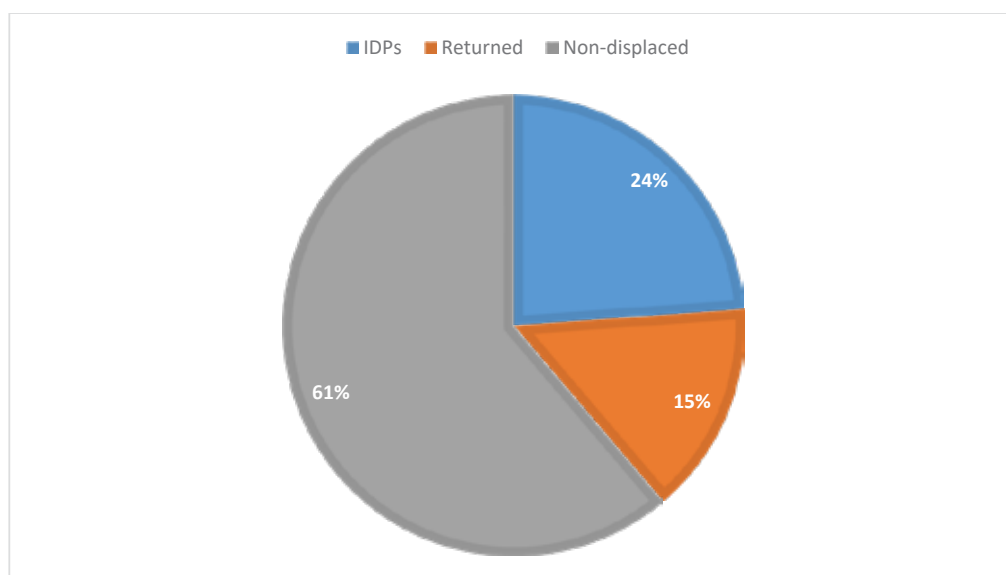


Fig. 5. Population groups provided with assistance for January-September 2024

Humanitarian aid is provided by many Ukrainian manufacturers, which is delivered to central warehouses in the cities of Dnipro, Poltava, Kropyvnytskyi. Then, after the formation of sets, the goods are transported to an intermediate warehouse in the cities of Kramatorsk, Chuhuiv, Izium, Zaporizhzhia, after which, for safety reasons, they are reloaded onto smaller cars (pick-ups) and delivered to settlements in the frontline area.

It is essential to consider the characteristics of goods provided to citizens as humanitarian aid, select a vehicle for transporting goods from large warehouses to intermediate warehouses, develop transportation routes from large to intermediate warehouses, and simulate these routes for delivering humanitarian aid to frontline settlements from Kramatorsk. In the initial phase of transportation, cargo should be moved from central warehouses with humanitarian aid in cities such as Dnipro, Poltava, and Kropyvnytskyi to intermediate warehouses in Kramatorsk, Chuhuiv, Izium, and Zaporizhzhia. In the second phase, delivery routes should be planned from Kramatorsk to frontline cities (Kostiantynivskyi district).

The mathematical model of the transport process of transporting humanitarian cargoes looks like this. If x_{ij} is the volume of humanitarian cargo transported from warehouse i to intermediate warehouse j , then the transport work for such transportation is equal to $c_{ij}x_{ij}$. The total transport work for the transportation of all cargo will be $\sum_{i=1}^m \sum_{j=1}^n c_{ij}x_{ij}$. Transport work must be minimised, so the target function of the transport task will be as follows [4]:

$$Z(X) = \sum_{i=1}^m \sum_{j=1}^n c_{ij}x_{ij} \rightarrow \text{Min}. \quad (1)$$

The system of constraints contains two groups of equations. The first group of m equations indicates that the stocks of all m warehouses are transported in full, i.e.

$$\sum_{j=1}^n x_{ij} = a_i, \quad i=1,2,\dots,m. \quad (2)$$

The second group of constraints from n equations indicates that the needs of all consumers are fully satisfied, i.e. $\sum_{i=1}^m x_{ij} = b_j, \quad j=1,2,\dots,n$. Considering that $x_{ij} \geq 0$:

$$\sum_{i=1}^m x_{ij} = b_j, \quad j=1,2,\dots,n \quad (3)$$

$$x_{ij} \geq 0, \quad i=1,2,\dots,m, \quad j=1,2,\dots,n. \quad (4)$$

In the transport problem model under consideration, it is assumed that the total warehouse stocks equal the total requirements of intermediate warehouses, i.e., the condition is met

$$\sum_{i=1}^m a_i = \sum_{j=1}^n b_j \quad (5)$$

Such a problem is called a problem with a correct balance, and its model is called closed. If equality (5) is not satisfied, then the problem is called a problem with an incorrect balance, and its model is open.

When solving a transport problem, it is necessary to find a transportation plan, a matrix

$$X = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \dots & \vdots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{pmatrix} = \{x_{ij}\}_{i=1; j=1}^{m \quad n}, \quad (6)$$

that satisfies the system of constraints (2), (3), (4) and minimises the objective function (1).

A transportation plan that satisfies the system of constraints (2), (3), (4) is called admissible.

A transportation plan in which the objective function is minimal is called optimal.

The transportation problem is a linear programming problem and can be solved by the simplex method algorithm. Since this problem has several features, namely:

- 1) constraints are given in the form of equations;
- 2) each unknown appears in only two equations;
- 3) the coefficients for unknowns are units.

Special algorithms have been created for its solution [5].

There are 3 types of vehicles and 9 settlements where humanitarian aid needs to be delivered. The requirements for the number of food packages and their weight are given in (Tab. 1).

Table 1

Requirements for food packages by settlements

No.	Settlement	Number of food parcels, pcs	Total weight, kg
1.	Oleksiiievo-Druzhkivka	7000	35000
2.	Stepanivka	400	2000
3.	Predtechyne	300	1500
4.	Ivanopillia	700	3500
5.	Illinivka	1000	5000
6.	Kostiantynivnka 1	1100	5500
7.	Kostiantynivnka 2	1150	5750
8.	Kostiantynivnka 3	1300	6500
9.	Kostiantynivnka 4	800	4000

Considering the capacity of vehicles in terms of load capacity, the cost of transportation to the i -th settlement relative to the j -th type of vehicles, the cost of renting each type of vehicle, and the distribution of vehicles between intermediate warehouses, we obtain the optimal transportation route based on linear programming tasks [2]. Thus, the objective function for the first stage of transportation will be as follows:

$$F(x) = 349.28x_{11} + 376.56x_{12} + 390.51x_{13} + 220.06x_{14} + 161.25x_{21} + 215.87x_{22} + 250.80x_{23} + 205.00x_{24} + 189.01x_{31} + 183.37x_{32} + 187.42x_{33} + 70.00x_{34} \rightarrow \min \quad (7)$$

With the following constraints:

By stocks:

$$x_{11} + x_{12} + x_{13} + x_{14} \leq 500$$

$$x_{21} + x_{22} + x_{23} + x_{24} \leq 450$$

$$x_{31} + x_{32} + x_{33} + x_{34} \leq 500$$

By needs:

$$x_{11} + x_{21} + x_{31} = 300$$

$$x_{12} + x_{22} + x_{32} = 250$$

$$x_{13} + x_{23} + x_{33} = 450$$

$$x_{14} + x_{24} + x_{34} = 450$$

Food packages provided as humanitarian aid to war-affected people include canned goods, cereals, pasta, flour, yeast, oil, salt, sugar, tea, bottled drinking water, and beverages. These goods do not have special storage temperature requirements, but some of them require storage and transportation in dry places.

Choosing the right vehicle for transporting goods is of great importance for the success and efficiency of any logistics process. The right solution in this area can reduce costs, improve transportation safety, and ensure the timely delivery of goods.

Key factors to consider when choosing a vehicle:

Type of cargo: The characteristics of the cargo, such as its size, weight, and sensitivity to transport conditions, influence the choice of transport vehicle. Considering the transportation of goods that only need to be protected from moisture, special vehicles (SV) are not required for transportation.

Choosing the right vehicle can significantly reduce transportation costs. Optimizing fuel, maintenance, and other operating costs helps to increase the profitability of logistics operations.

Choosing a vehicle that meets the needs of fast and reliable delivery guarantees the timely arrival of the cargo at its destination. This is especially important for companies working with urgent orders or perishable goods. Given that the delivery takes place in an area close to the front line, speed is also of great importance.

Vehicles that meet modern environmental standards help reduce environmental impact. Choosing a vehicle with low CO₂ emissions supports corporate responsibility and contributes to a positive company image.

Reliable vehicles with special safety systems ensure the safety of cargo during transportation. This reduces the risk of damage to goods and losses associated with accidents or other incidents on the road.

The correct choice of a motor vehicle allows you to optimize transportation processes, reduce costs, and increase the overall efficiency of logistics. This is a strategic decision that requires careful analysis and consideration of all aspects of transportation.

The selection of the optimal vehicle model from three competitive ones is based on a comparison of the results of technical and operational calculations, as well as technical and economic calculations. Three types of rolling stock were selected for comparison: DAF FX105, Mercedes-Benz Actros, and Renault Magnum [6]. A comparison

of vehicles is given in Table 2. In the future, the selected model will be used for cargo transportation along the Dnipro-Kramatorsk route.

Table 2

Comparative table of rolling stock of vehicles

No.	Indicator	Option 1	Option 2	Option 3
1.	Car brend	DAF FX 105	Renault Magnum	Mercedes-Benz Actros
2.	Euro compliance	Euro 5	Euro 5	Euro 6
3.	Transmission	Manual	Manual	Automatic
4.	Power (horsepower)	410	480	448
5.	Fuel type	diesel	diesel	diesel
6.	Fuel consumption (litres per 100 km)	30	33	28
7.	Load capacity (kg)	45000	45 000	45000
8.	Year of manufacture	2017	2017	2018
9.	Car price	24000 dollars	23000 euros	36000 euros
10.	Car tyre cost	10500 UAH	10500 UAH	10500 UAH
11.	Technical speed (km/h)	85	75	80
12.	Productivity (t)	1.55	1.38	1.45
13.	Productivity (t/km)	1074.18	957.29	1050.76

According to the calculations, the first option (DAF FX 105) showed higher hourly productivity, so we chose it for transportation.

Considering the amount of cargo for different settlements that transport aid directly to the front line, we suggest choosing vehicles of different load capacities.

The vehicle with a load capacity of up to 1.5 tons.

We will prepare the initial data in Excel and provide information on distances, stocks, and requirements (Tab. 3).

Table 3

Initial data for calculating the transport task

	Chuhuiv	Izium	Kramatorsk	Zaporizhzhia	Stocks
Kropyvnytskyi	349.28	376.56	390.51	220.06	500
Poltava	161.25	215.87	250.80	205.00	450
Dnipro	189.01	183.37	187.42	70.00	500
Needs	300	250	450	450	

The obtained results are shown in (Fig. 6).

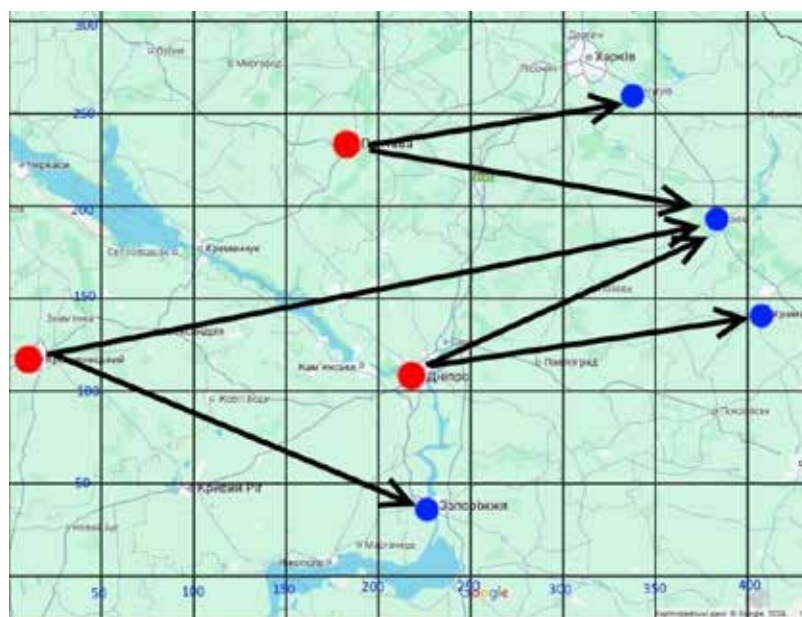


Fig. 6. Transportation scheme based on the solution of the transportation task

Thus, the minimum transport work will be 292114 t-km. From Kropyvnytskyi to Izium, 50 tonnes of cargo are delivered, and to Zaporizhzhia, 450 tonnes. From Poltava, 300 tonnes to Chuhuiv and 150 tonnes to Izium. From Dnipro 50 tonnes to Izium and 450 tonnes to Kramatorsk.

To solve the transport task, distances were calculated using the criterion of minimum distance; however, due to the proximity of the front, it is necessary to have alternative routes in case of shelling.

In case of shelling of this territory, alternative routes can be used, one through Synelnykovo and Vasytkivka, and the second through Pokrovsk.

Since it is essential to deliver food and hygiene products to people in settlements as quickly as possible, we choose the route that is both the shortest and fastest in terms of delivery time. We also have two alternative routes passing through the cities of Pokrovsk and Vasytkivka. Let us analyze the probabilities associated with these alternative routes. Along a similar route, 60% of vehicles carrying humanitarian cargo pass through Pokrovsk, while 40% go through Vasytkivka. Additionally, 5% of vehicles near Pokrovsk come under fire, compared to only 2% near Vasytkivka. We need to calculate these probability indicators.

Let us formulate two hypotheses: H_1 – the vehicle passed through Pokrovsk, H_2 – the vehicle passed through Vasytkivka. Let event A – the vehicle was waiting for the end of the shelling.

To determine the probability that a random vehicle with humanitarian cargo was waiting for the end of the shelling near the city of Pokrovsk, we need to use the Bayes formula, according to which [8]:

$$P(H_1 / A) = \frac{P(H_1) \cdot P(A / H_1)}{P(H_1) \cdot P(A / H_1) + P(H_2) \cdot P(A / H_2)} = \frac{0,6 \cdot 0,05}{0,6 \cdot 0,05 + 0,4 \cdot 0,02} = \frac{0,03}{0,038} = 0,79 = 79\%$$

Then, by the property of probabilities that their sum is equal to 1, we have that the probability of downtime in Vasytkivka will be as follows:

$$P(H_2) = 1 - 0,79 = 0,21$$

Thus, for the alternative transportation route, we choose route No. 2, via Vasytkivka, because this route has a lower probability of coming under fire.

There are 3 types of vehicles and 9 settlements where humanitarian aid needs to be delivered. The requirements for the quantities of food packages and their weights are given in (Tab. 4).

Table 4

Requirements for food packages by settlements

No.	Settlement	Number of food parcels, pcs	Total weight, kg
1.	Oleksiievo-Druzhkivka	7000	35000
2.	Stepanivka	400	2000
3.	Predtechyne	300	1500
4.	Ivanopillia	700	3500
5.	Illinivka	1000	5000
6.	Kostiantynivka 1	1100	5500
7.	Kostiantynivka 2	1150	5750
8.	Kostiantynivka 3	1300	6500
9.	Kostiantynivka 4	800	4000

So, according to the terms of the task, the optimal transportation option should be as follows: four vehicles with a load capacity of 8 tonnes (the third type) and one vehicle with a load capacity of 3 tonnes (the second type) go to the urban village of Oleksiievo-Druzhkivka; two vehicles with a load capacity of up to 1.5 tonnes (the first type) go to the village of Stepanivka; one vehicle of the first type goes to the village of Predtechyne; one vehicle of the first type and one of the second type go to the village of Ivanopillia; one vehicle of the third type goes to the village of Illinivka; one vehicle of the third type goes to the city of Kostiantynivka (the first district); another of the third type to Kostiantynivka (the second district); one vehicle of the third type to Kostiantynivka (the third district); and one vehicle of the first type and one of the second type go to Kostiantynivka (the fourth district). This allocation and distribution of the vehicle fleet ensures minimal costs for transporting humanitarian cargo.

After three years of a full-scale invasion of Ukraine, the humanitarian crisis has not eased; rather, the needs of the affected population remain high. Massive rocket attacks and ongoing hostilities have caused significant civilian casualties and destroyed infrastructure, forcing thousands of people to leave their homes. Organizations such as UNHCR are working to provide a variety of assistance – from monetary to material support – meeting the basic needs of those affected and helping them survive the autumn and winter months.

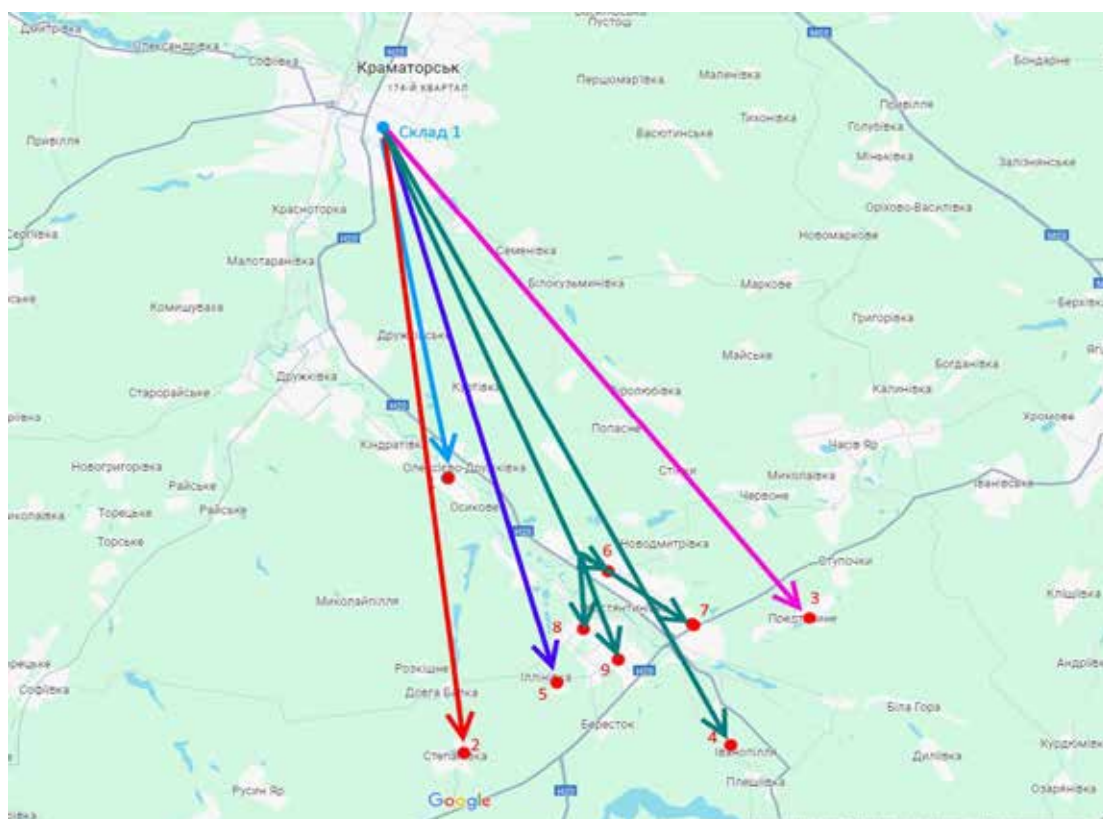


Fig. 7. Transportation scheme from the warehouse to issuing points

Overall, humanitarian assistance in Ukraine as of the end of August 2024 has achieved significant results thanks to the efforts of more than 580 organizations. It was able to support 6.7 million people by providing water supply, food, healthcare, and non-food items. Successful preparations for the autumn-winter season enabled support for over 1.2 million people with fuel and materials for house repairs.

It is worth noting the important role of multi-purpose cash assistance, which helped more than 600,000 people meet their basic needs. Also, providing education to almost 530,000 people is a significant achievement.

Thanks to inter-agency convoys, it was possible to reach hard-to-reach frontline areas and provide assistance to nearly 45,000 people, which is an important step in overcoming the consequences of the conflict.

Humanitarian partners continue to work to improve the living conditions of internally displaced persons and provide necessary support to citizens of various statuses, which is crucial to a further successful response to humanitarian challenges in Ukraine.

Conclusions. Despite numerous challenges, both international and local efforts focus on supporting the most vulnerable groups, improving living conditions for internally displaced persons, and providing them with housing. The success of these initiatives is vital for ensuring the survival and dignity of those facing hardship due to the war. The daily work of coordinating humanitarian aid and responding to emergencies continues to be crucial in overcoming the war's impacts and maintaining stability in the country.

The effective transportation of humanitarian cargo from central warehouses to front-line areas is a challenging but solvable task. Developing an efficient delivery system based on logistics management methods involves several key stages: moving cargo from central warehouses to intermediate depots, creating routes for further transportation to distribution points, and optimizing the transport fleet. Using specialized algorithms for linear programming tasks helps find the best solutions for vehicle allocation and cost reduction.

Choosing the right vehicle for cargo transport is crucial for the success of logistics operations. The proper choice can significantly cut costs, enhance safety, and ensure on-time delivery. Optimizing delivery routes and selecting rolling stock with different capacities helps reduce transportation expenses and improve transport efficiency. The minimum transport work for the first stage will be 292,114 t-km, with 50 t delivered from Kropyvnytskyi to Izium, 450 t to Zaporizhzhia, 300 t from Poltava to Chuhuiev, and 150 t to Izium. Additionally, 50 t move from Dnipro to Izium, and 450 t to Kramatorsk. In the second stage, selecting and distributing the vehicle fleet minimizes transportation costs for humanitarian cargo.

During threats like shelling, having alternative routes ensures supply continuity. Risk analysis helps identify the safest routes. Overall, careful planning, modern vehicles, and route optimization are key to successfully delivering humanitarian aid, especially during military operations.

Bibliography:

1. Публікації – UNHCR Україна. UNHCR Україна. URL: <https://www.unhcr.org/ua/publications> (дата звернення: 01.10.2024).
2. Самойленко М., Кобець А. Інформаційні технології в розв'язанні транспортних задач : монографія. Харків : ХНАМГ, 2011. 256 с.
3. Тюріна Н. М. Логістика: навч. посіб. Київ : «Центр навчальної літератури», 2015. 392 с.
4. Транспортна задача. навч. посіб. для студ. спеціальності 121 «Інженерія програмного забезпечення», спеціалізації «Програмне забезпечення розподілених систем», «Програмне забезпечення Web-технологій та мобільних пристроїв». КПІ ім. Ігоря Сікорського ; уклад.: О. К. Молодід. Київ : КПІ ім. Ігоря Сікорського, 2018. 37 с.
5. Україна: гуманітарне реагування у 2024 році. Situation Reports. URL: <https://reports.unocha.org/uk/country/ukraine> (дата звернення: 30.10.2024).
6. Renault Magnum 460.44T : фото, технічні характеристики Автомобілі. ONLINE.UA. URL: <https://auto.online.ua/ukr/catalog/articles/renault/magnum/magnum/58732/> (дата звернення: 30.11.2024).
7. Музикін М. І., Попов Г. В., Біляченко М. В. Дослідження аварій з транспортними одиницями з використанням ризико-орієнтованого підходу. *Інформаційно-керуючі системи на залізничному транспорті* : тези стендових доповідей та виступів учасників 34-ї Міжнародної науково-практичної конференції. Харків : УкрДУЗТ, 2021. С. 45–46
8. Бех П. В., Нестеренко Г. І., Стрелко О. Г., Музикін М. І. Управління вантажними перевезеннями в умовах ризиків конкурентного середовища. *Системи та технології*. 2021. № 1 (61). С. 85–97.
9. Музикін М. І., Буряк А. С. Шляхи підвищення ефективності перевезень гуманітарних вантажів автомобільним транспортом. *Економіко-правові та управлінсько-технологічні виміри сьогодення: молодіжний погляд : матеріали Міжнародної науково-практичної конференції* : у 2 т. Т. 2. Дніпро : УМСФ, 2024. С. 440–442.
10. Nesterenko H. I., Muzykin M. I., Bibik S. I., Strelko O.H., Aleksieieva A. O. Analysis of organizing the delivery of humanitarian aid in crisis situations. *Системи та технології*. № 2 (68). 2024. С. 130–139.
11. Помазков М. В. Моделювання інтелектуальних маршрутів перевезення гуманітарних вантажів в умовах воєнного часу. *Подільський вісник: сільське господарство, техніка, економіка*. 2023. № 41 С. 46–53.
12. Гринів Н. Гуманітарна логістика як інструмент трансформації логістичних потоків в умовах воєнного часу. *Економіка та суспільство*. 2023. № 56. URL: <https://doi.org/10.32782/2524-0072/2023-56-4> (дата звернення: 01.10.2024).

References:

1. Publications – UNHCR Ukraine. UNHCR Ukraine. (2024). Retrieved from: <https://www.unhcr.org/ua/publications> (Accessed on 01.10.2024).
2. Samoilenko, M., & Kobets, A. (2011). Information technologies in solving transport problems: monograph. Kharkiv: KhNAUE. 256 p.
3. Tyurina, N. M. (2015). Logistics: teaching guide. Kyiv: Center for Educational Literature. 392 p.
4. Molodid O. K. (2018). Transport task: teaching guide for students of specialty 121 "Software Engineering", specializations "Software of distributed systems", "Software of Web technologies and mobile devices". Kyiv: Igor Sikorsky Kyiv Polytechnic Institute. 37 p.
5. Ukraine: humanitarian response in 2024. Situation Reports. (2024). Retrieved from: <https://reports.unocha.org/uk/country/ukraine> (Accessed on: 30.10.2024).
6. Renault Magnum 460.44T : photos, technical specifications of Cars. ONLINE.UA. Retrieved from: <https://auto.online.ua/ukr/catalog/articles/renault/magnum/magnum/58732/> (Accessed on: 30.11.2024).
7. Muzykin, M. I., Popov, G. V., & Bilyachenko, M. V. (2021). Research of accidents with transport units using a risk-oriented approach. *Information and control systems in railway transport: abstracts of poster presentations and speeches of participants of the 34th International Scientific-Practical Conference*. Kharkiv: UkrSURT, pp. 45–46.
8. Bekh, P. V., Nesterenko, H. I., Strelko, O. H., & Muzykin, M. I. (2021). Freight transportation management in conditions of competitive environment risks. *Systems and Technologies*, 1 (61), 85–97.
9. Muzykin, M. I., & Buryak, A. S. (2024). Ways to improve the efficiency of transportation of humanitarian cargo by road. *Economic, legal and managerial and technological dimensions of the present: a youth perspective: Proceedings of the International Scientific-Practical Conference*: in 2 volumes. Vol. 2. Dnipro: UMSF, pp. 440–442.
10. Nesterenko, H. I., Muzykin, M. I., Bibik, S. I., Strelko O.H., & Aleksieieva A. O. (2024). Analysis of organizing the delivery of humanitarian aid in crisis situations. *Systems and Technologies*, 2 (68), 130–139.
11. Pomazkov, M. V. (2023). Modeling of intellectual routes for transporting humanitarian cargo in wartime. *Podolskyi visnyk: agriculture, technology, economics*. (41), 46–53.
12. Hryniv, N. (2023). Humanitarian logistics as a tool for transforming logistics flows in wartime conditions. *Economy and Society*, (56). Retrieved from: <https://doi.org/10.32782/2524-0072/2023-56-4> (Accessed on: 01.10.2024).

Дата надходження статті: 30.10.2025

Дата прийняття статті: 17.11.2025

Опубліковано: 30.12.2025

Погребняк В. Г., доктор технічних наук, професор,
професор кафедри технології захисту навколишнього середовища
та безпеки праці
Івано-Франківського національного технічного університету
нафти і газу
ORCID: 0000-0002-7735-3408

Перкун І. В., кандидат технічних наук,
доцент, здобувач наукового ступеня доктора технічних наук
Івано-Франківського національного технічного університету
нафти і газу
ORCID: 0000-0002-6523-1606

Вишнікіна О. В., кандидат хімічних наук, доцент,
доцент кафедри міжнародного туризму
та готельно-ресторанного бізнесу
Університету митної справи та фінансів
ORCID: 0000-0002-3747-2904

Погребняк А. В., доктор технічних наук, професор,
професор кафедри міжнародного туризму
та готельно-ресторанного бізнесу
Університету митної справи та фінансів
ORCID: 0000-0003-3214-6410

ВІД ОХОРОНИ ПРАЦІ ДО КОМПЛЕКСНОЇ БЕЗПЕКИ: СТРАТЕГІЇ ІНТЕГРАЦІЇ ЦИВІЛЬНОГО ЗАХИСТУ В КОРПОРАТИВНІ СИСТЕМИ УПРАВЛІННЯ

В статті обґрунтовано нагальну науково-практичну проблему, що виникла внаслідок повномасштабної війни в Україні – фундаментальну невідповідність традиційних корпоративних систем управління безпекою новим гібридним загрозам. Доведено, що класична модель, де охорона праці (ОП) та цивільний захист (ЦЗ) функціонують як дві паралельні, слабо інтегровані системи, є застарілою та нездатною забезпечити адекватний рівень захисту персоналу та активів підприємств, які перетворилися на потенційні цілі для воєнних атак. Існуючі системи управління охороною праці, побудовані на стандартах мирного часу (напр., ISO 45001), орієнтовані на запобігання ненавмисним інцидентам і не мають інструментарію для протидії умисним зовнішнім впливам. Метою статті є розробка науково обґрунтованих стратегій інтеграції функцій цивільного захисту в корпоративні системи управління охороною праці для створення єдиної, комплексної моделі безпеки підприємства, адаптованої до умов гібридних загроз. На основі аналізу класичних концепцій (Safety-I, Safety-II, Resilience Engineering) та нормативно-правової бази України виявлено ключову прогалину – відсутність цілісної моделі управління всіма видами безпекових ризиків на корпоративному рівні. Наукова новизна роботи полягає у тому, що вперше запропоновано концептуальну модель інтегрованої системи управління комплексною безпекою (СУКБ). На відміну від існуючих підходів, дана модель синтезує ОП та ЦЗ в єдиний управлінський цикл («аналіз – превенція – реагування – відновлення»), що базується на п'яти ключових принципах: єдності управління, проактивності, тотальної ризик-орієнтованості (з урахуванням воєнних загроз), людиноцентризму та безперервності. Цей підхід дозволяє перейти від застарілої дихотомії «внутрішні/зовнішні загрози» до єдиного контуру управління всіма ідентифікованими ризиками, незалежно від їхнього походження. Практичне значення полягає у розробці чотирьох груп конкретних стратегій інтеграції: нормативно-правової (створення єдиного Положення про комплексну безпеку), організаційно-управлінської (формування єдиного центру безпеки), техніко-технологічної (інтеграція систем оповіщення та моніторингу) та соціально-психологічної (провадження комплексних навчань, психологічна підтримка). Запропоновані стратегії можуть слугувати дорожньою картою для керівників українських підприємств для реформування існуючих систем, підвищення рівня захищеності персоналу, забезпечення стійкості бізнес-процесів та збереження людських життів в умовах воєнного стану.

Ключові слова: комплексна безпека, інтеграція систем, цивільний захист, охорона праці, воєнні ризики, гібридні загрози, система управління безпекою, стійкість підприємства, ризик-орієнтований підхід.

© В. Г. Погребняк, І. В. Перкун, О. В. Вишнікіна, А. В. Погребняк, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

Pogrebnyak V. G., Perkun I. V., Vyshnikina O. V., Pogrebnyak A. V. From Occupational Safety to Comprehensive Security: Strategies for Integrating Civil Defense into Corporate Management Systems

The article substantiates an urgent scientific and practical problem that has arisen due to the full-scale war in Ukraine – the fundamental inadequacy of traditional corporate safety management systems in the face of new hybrid threats. It is proven that the classic model, where occupational safety (OS) and civil defense (CD) function as two parallel, weakly integrated systems, is outdated and incapable of providing an adequate level of protection for personnel and assets of enterprises, which have become potential targets for military attacks. Existing occupational safety management systems, built on peacetime standards (e.g., ISO 45001), are focused on preventing unintentional incidents and lack the tools to counter deliberate external impacts. The purpose of the article is to develop scientifically-backed strategies for integrating civil defense functions into corporate occupational safety management systems to create a unified, comprehensive enterprise security model adapted to the conditions of hybrid threats. Based on an analysis of classical concepts (Safety-I, Safety-II, Resilience Engineering) and the legal framework of Ukraine, a key gap has been identified – the absence of a holistic model for managing all types of security risks at the corporate level. The scientific novelty of the work lies in the proposal of a conceptual model for an integrated comprehensive security management system (CSMS) for the first time. Unlike existing approaches, this model synthesizes OS and CD into a single management cycle ("analysis – prevention – response – recovery"), based on five key principles: unity of command, proactivity, total risk orientation (including military risks), human-centricity, and continuity. This approach allows for a transition from the outdated dichotomy of "internal/external threats" to a single management loop for all identified risks, regardless of their origin. The practical significance lies in the development of four groups of specific integration strategies: legal-regulatory (creating a unified Comprehensive Security Regulation), organizational-managerial (forming a single security center), technical-technological (integrating alert and monitoring systems), and socio-psychological (conducting comprehensive training and providing psychological support). The proposed strategies can serve as a roadmap for managers of Ukrainian enterprises to reform existing systems, enhance personnel protection, ensure the resilience of business processes, and save human lives under martial law.

Key words: comprehensive security, systems integration, civil defense, occupational safety, military risks, hybrid threats, safety management system, enterprise resilience, risk-based approach.

Постановка проблеми. Війна в Україні спричинила тектонічний зсув у безпековому середовищі, перетворивши всю територію країни на зону потенційного ураження. Для українських підприємств це означало кардинальну зміну парадигми ризиків. Якщо раніше система управління безпекою фокусувалася переважно на внутрішніх загрозах – виробничому травматизмі, професійних захворюваннях, технологічних аваріях та пожежах [1–4], – то сьогодні на перший план вийшли прямі зовнішні воєнні загрози [5, 6]. Ракетні обстріли, атаки безпілотних літальних апаратів (БпЛА), артилерійські удари, кібератаки та диверсійна діяльність перетворили промислові, енергетичні, логістичні та навіть аграрні об'єкти з виробничих майданчиків на потенційні цілі.

Ця нова реальність оголила фундаментальну вразливість існуючих корпоративних систем управління. Традиційний підхід, де охорона праці (ОП) та цивільний захист (ЦЗ) існують як дві паралельні, слабо пов'язані між собою системи, виявився нездатним забезпечити адекватний рівень захисту персоналу та активів [1, 2, 6]. Безпека працівника на робочому місці більше не залежить лише від справності обладнання чи дотримання інструкцій; вона прямо залежить від наявності та стану укриття, ефективності системи оповіщення та готовності до дій в умовах бойової загрози. Таким чином, проблема виходить за межі класичної охорони праці і набуває рис комплексної проблеми виживання та стійкості бізнесу.

Постановка завдання. Існуючі корпоративні системи управління охороною праці (СУОП), здебільшого побудовані на філософії стандартів мирного часу, таких як міжнародний стандарт ISO 45001:2018, орієнтовані на ідентифікацію, оцінку та контроль ризиків *ненавмисного* характеру [1–3, 7]. Вони ефективні для запобігання нещасним випадкам, але їх інструментарій та методологія виявляються недостатніми для протидії *умисним*, спланованим зовнішнім загрозам воєнного характеру. Ризик влучання ракети неможливо «усунути» чи «замінити» інженерними засобами контролю на робочому місці [5, 6].

З іншого боку, функції цивільного захисту на підприємствах, регламентовані Кодексом цивільного захисту України [8], до 2022 року часто виконувались формально. Плани реагування, тренування та матеріально-технічне забезпечення не завжди відповідали реальним потребам. Найголовніше – ці функції існували в організаційному вакуумі, окремо від щоденних процесів управління охороною праці та виробничою діяльністю. Виник критичний функціональний розрив між системою, що відповідає за безпеку «всередині» (СУОП), і системою, що має відповідати за безпеку «зовні» (ЦЗ).

Відтак, виникає гостра наукова та практична проблема: відсутність цілісної, теоретично обґрунтованої концепції та практичних стратегій для повної, а не часткової, інтеграції функцій цивільного захисту в корпоративні системи управління. Необхідно подолати застарілу дихотомію «ОП проти ЦЗ» [1, 8] і перейти до єдиної моделі управління комплексною безпекою.

Мета статті – розробка науково обґрунтованих стратегій інтеграції функцій цивільного захисту в корпоративні системи управління охороною праці для створення єдиної, комплексної моделі безпеки підприємства, адаптованої до умов гібридних загроз. Досягнення мети передбачає вирішення наступних завдань: аргументувати необхідність парадигмального зсуву від ізольованих систем ОП та ЦЗ до єдиного контуру

комплексної безпеки; сформулювати ключові принципи та розробити концептуальну модель інтегрованої системи управління комплексною безпекою (СУКБ); запропонувати конкретні нормативно-правові, організаційні, технічні та соціально-психологічні стратегії практичної інтеграції.

Аналіз останніх досліджень і публікацій. *Огляд класичних моделей управління охороною праці та їхньої обмеженості.* Традиційною парадигмою в управлінні безпекою є Safety-I, яка фокусується на аналізі того, що йде не так: розслідування нещасних випадків, виявлення помилок та усунення причин збоїв [9–11]. Цей реактивний підхід, хоч і є базовим, абсолютно недостатній в умовах війни, оскільки він не може запобігти подіям типу «чорний лебідь», якими є військові атаки.

Новітня концепція Safety-II пропонує змістити фокус на вивчення того, чому системи працюють успішно переважну більшість часу, навіть за наявності варіативності та складнощів [10, 11]. Вона акцентує на гнучкості та адаптивності персоналу. Цей підхід є більш корисним, оскільки підкреслює важливість людського фактору в адаптації до несподіваних обставин. Проте, він не надає чітких інструментів для протидії цілеспрямованим руйнівним діям.

Найбільш релевантною є концепція Resilience Engineering (Інженерія стійкості), що визначає безпеку як здатність системи адаптуватися та відновлюватися перед лицем непередбачуваних подій. Вона оперує чотирма ключовими здібностями: моніторинг, реагування, передбачення та навчання. Ця концепція є чудовою філософською основою для побудови комплексної безпеки, але її класичні моделі розроблялися для складних соціотехнічних систем (авіація, атомна енергетика) і не враховували специфіку прямого військового впливу на рядові промислові підприємства.

Аналіз нормативно-правової бази України. Аналіз ключових законодавчих актів – Кодексу цивільного захисту України [8] та Закону України «Про охорону праці» [1] – виявляє їхню паралельність. Кожен документ чітко окреслює свою сферу відповідальності, проте механізми їхньої обов'язкової інтеграції на рівні суб'єкта господарювання прописані недостатньо. Це призводить до ситуації, коли на підприємстві є інженер з охорони праці, відповідальний за СУОП, і начальник штабу ЦЗ (або інша відповідальна особа), які звітують різним керівникам, керуються різними планами та проводять окремі навчання. Відсутній єдиний нормативний документ, який би вимагав створення інтегрованої системи управління всіма видами безпекових ризиків.

Огляд сучасних публікацій (після 2022 року). З початком повномасштабного вторгнення з'явилася значна кількість публікацій, присвячених стійкості критичної інфраструктури, енергетичній безпеці та безперервності бізнесу (business continuity) [5, 6, 12–15]. Українські та закордонні автори активно досліджують досвід функціонування економіки в умовах війни. Проте, більшість цих робіт фокусується на макrorівні (держава, галузь) або на аспектах фізичного захисту об'єктів (інженерні укріплення, ППО). Праці, що пропонують комплексну модель управління безпекою саме на корпоративному рівні, поєднуючи філософію ОП та практику ЦЗ, залишаються поодинокими [16].

Виявлення невирішених раніше частин загальної проблеми. На основі проведеного аналізу можна констатувати, що ключовою невирішеною проблемою є відсутність цілісної концептуальної моделі управління комплексною безпекою на рівні підприємства. Більшість існуючих робіт та практичних підходів продовжують розглядати охорону праці та цивільний захист як паралельні, а не глибоко інтегровані системи. Не розроблено єдиної методології оцінки ризиків, яка б органічно поєднувала аналіз виробничих небезпек (напр., за методом FMEA) та воєнних загроз (напр., за методом CARVER). Відсутній системний підхід до інтеграції планів, ресурсів, навчання та управлінських структур, що відповідають за різні аспекти безпеки. Ця стаття покликана заповнити саме цю концептуальну прогалину.

Виклад основного матеріалу дослідження. *Парадигмальний зсув: Від ізольованих систем до єдиного контуру безпеки.* Традиційне розділення безпеки на «виробничу» (охорона праці) та «цивільну» (цивільний захист) є спадком індустріальної епохи мирного часу. Воно ґрунтується на хибному припущенні, що джерела загроз можна чітко розмежувати: внутрішні (технологічні) та зовнішні (природні, військові). В умовах сучасної гібридної війни це розділення не просто застаріло, а стало небезпечним.

Аргументація: Наслідки інцидентів різного походження можуть бути ідентичними. Наприклад, пожежа на підприємстві може бути спричинена як коротким замиканням (проблема ОП), так і влучанням уламків БпЛА (проблема ЦЗ). Руйнування будівлі може статися через вибух газового балону (ОП) або через влучання ракети (ЦЗ). В обох випадках персоналу потрібно знати шляхи евакуації, вміти надавати першу домедичну допомогу та діяти скоординовано. Розділені системи реагування створюють плутанину, дублювання функцій та, як наслідок, втрату дорогоцінного часу.

Тому перехід до концепції *комплексної безпеки* – це стратегічний підхід, що визначає стан захищеності персоналу, матеріальних та нематеріальних активів, а також виробничих процесів підприємства від усього спектру ідентифікованих загроз, незалежно від їхнього походження: виробничих, природних, техногенних, воєнних, соціальних, інформаційних та кібернетичних. Єдиний контур безпеки означає, що всі загрози аналізуються в рамках єдиної системи оцінки ризиків, а заходи з контролю та реагування плануються та управляються з єдиного центру.

Принципи та модель інтегрованої системи управління комплексною безпекою (СУКБ). Для побудови ефективної інтегрованої системи пропонується базуватися на наступних ключових принципах: 1. *Принцип*

єдності управління: Створення єдиного центру відповідальності та прийняття рішень з усіх питань безпеки. Запобігання дублюванню функцій та конфлікту повноважень. 2. *Принцип проактивності:* Перехід від реактивного реагування на інциденти до проактивного прогнозування та запобігання загрозам. Постійний моніторинг безпекового середовища. 3. *Принцип тотальної ризик-орієнтованості:* Інтеграція аналізу воєнних та терористичних ризиків у стандартні процедури оцінки ризиків на робочих місцях. Кожен ризик розглядається з точки зору ймовірності та потенційних наслідків для життя, здоров'я, активів та безперервності діяльності. 4. *Принцип людиноцентризму:* Визнання життя та здоров'я працівника найвищою цінністю. Це включає не лише фізичний захист (укриття, ЗІЗ), але й психологічну підтримку, навчання та формування культури безпеки. 5. *Принцип безперервності та постійного вдосконалення:* СУКБ не є статичним документом, а циклічним процесом (цикл Демінга: Plan-Do-Check-Act), який регулярно переглядається та вдосконалюється на основі отриманого досвіду, навчань та змін у зовнішньому середовищі.

Концептуальна модель СУКБ може бути візуалізована як цикл, що складається з чотирьох взаємопов'язаних підсистем: 1. *Підсистема аналізу та оцінки ризиків:* Єдиний реєстр ризиків, що включає як виробничі небезпеки (падіння, ураження струмом), так і воєнні загрози (ракетний удар, хімічна атака, окупація). Для кожного ризику проводиться оцінка ймовірності та наслідків. 2. *Підсистема превентивних заходів:* Розробка та впровадження заходів для зниження ризиків. *Інтегрований приклад:* укріплення віконних прорізів (ЦЗ) одночасно знижує ризик травмування уламками скла як від вибухової хвилі, так і від внутрішнього технологічного вибуху (ОП). 3. *Підсистема готовності та реагування на інциденти:* Розробка єдиних, комплексних планів реагування, що охоплюють різні сценарії (напр., "повітряна тривога з подальшим загорянням"). Створення універсальних рятувальних формувань, проведення інтегрованих тренувань. 4. *Підсистема відновлення та аналізу досвіду:* Дії після інциденту, спрямовані на відновлення роботи, надання допомоги постраждалим, розслідування причин (незалежно від їх природи) та внесення змін у всі попередні підсистеми для запобігання повторенню.

Стратегії практичної інтеграції цивільного захисту в корпоративне управління.

Нормативно-правова стратегія: розробка та затвердження єдиного корпоративного стандарту – *Положення про комплексну безпеку підприємства*, який замінює собою розрізнені положення про СУОП та плани ЦЗ; оновлення посадових інструкцій: введення обов'язків з ЦЗ для всіх керівників підрозділів, а не лише для окремої особи; інтеграція вимог ЦЗ (дії під час тривоги, правила перебування в укритті) в програми інструктажів з охорони праці; створення єдиної карти (матриці) ризиків підприємства;

Організаційно-управлінська стратегія: реорганізація структур: створення єдиного *Департаменту (або Центру) комплексної безпеки*, який об'єднує фахівців з ОП, пожежної безпеки, ЦЗ, екологічної та кібербезпеки; введення посади *менеджера з комплексної безпеки (Chief Security Officer)*, який підпорядковується безпосередньо першому керівнику підприємства; розробка *інтегрованих планів реагування на надзвичайні ситуації* (напр., план дій у разі хімічного викиду повинен враховувати сценарій, коли викид стався внаслідок ворожого удару, що унеможливило евакуацію за стандартними маршрутами); формування єдиних добровільних формувань (аварійно-рятувальних, пожежних, санітарних), підготовлених до дій у різних умовах.

Техніко-технологічна стратегія: створення *інтегрованої системи оповіщення та зв'язку*, що поєднує сигнали повітряної тривоги, пожежної сигналізації та оповіщення про технологічні аварії; модернізація та обладнання укриттів з урахуванням вимог як ЦЗ (захист від уламків), так і ОП (вентиляція, освітлення, наявність аптечок, запасних виходів); активне застосування цифрових інструментів – GIS-систем для моделювання зон ураження, дронів для моніторингу периметра та оцінки збитків після атак, інтегрованих систем відеоспостереження та контролю доступу.

Соціально-психологічна стратегія: проведення *комплексних, сценарних навчань*, що максимально наближені до реальних умов і поєднують елементи ОП та ЦЗ (напр., евакуація під час "повітряної тривоги" з імітацією задимлення на шляху до укриття); впровадження *програм психологічної підтримки персоналу*, навчання технік управління стресом, тренінги з підвищення стресостійкості (resilience); системне формування *корпоративної культури всеохоплюючої безпеки*, де кожен працівник усвідомлює свою відповідальність не лише за дотримання правил ОП, а й за виконання вимог ЦЗ, розглядаючи це як єдине ціле.

Висновки

1. У ході дослідження було доведено, що традиційна модель роздільного управління охороною праці та цивільним захистом на підприємствах є застарілою та нездатною забезпечити адекватний рівень захисту в умовах повномасштабної війни. Ефективна протидія сучасним гібридним загрозам вимагає фундаментального парадигмального зсуву – переходу до інтегрованої системи управління комплексною безпекою (СУКБ), яка розглядає всі загрози (виробничі, воєнні, природні тощо) в єдиному контурі аналізу, запобігання та реагування.

2. Наукова новизна роботи полягає у тому, що вперше запропоновано *цілісну концептуальну модель інтегрованої системи управління комплексною безпекою (СУКБ) на рівні підприємства*. На відміну від існуючих підходів, які розглядають ОП та ЦЗ як окремі елементи, дана модель синтезує їх в єдиний управлінський цикл (аналіз – превенція – реагування – відновлення), що базується на принципах єдності управління,

проактивності, тотальної ризик-орієнтованості, людиноцентризму та безперервності. Модель адаптована до специфіки функціонування підприємств в умовах гібридних загроз.

3. Розроблені у статті *чотири групи стратегій (нормативно-правова, організаційно-управлінська, техніко-технологічна, соціально-психологічна)* можуть бути використані керівниками та фахівцями з безпеки українських підприємств як методичні вказівки та дорожня карта для реформування існуючих систем. Впровадження запропонованих стратегій дозволить не формально, а реально підвищити рівень захищеності персоналу, мінімізувати потенційні збитки, забезпечити стійкість бізнес-процесів та, найголовніше, зберегти людські життя в умовах воєнного стану.

Перспективи подальших досліджень. Подальший розвиток цієї теми може відбуватися у кількох напрямках. Перспективним є *розробка галузевих моделей СУКБ*, які б враховували специфічні ризики для підприємств енергетики, хімічної промисловості, агросектору, транспорту та ІТ. Іншим важливим напрямком є *створення методики кількісної оцінки рівня комплексної безпеки підприємства*, що дозволить проводити аудит, порівнювати ефективність різних заходів та обґрунтовувати інвестиції в безпеку.

Список використаних джерел:

1. Про охорону праці : Закон України від 14.10.1992 р. № 2694-XII. *Відомості Верховної Ради України*. 1992. № 49. Ст. 668. URL: <https://zakon.rada.gov.ua/laws/show/2694-12>
2. Левченко О. Г., Бахрушин В. Е., Соколов М. М. Охорона праці та цивільний захист: підручник. Київ: КПІ ім. І. Сікорського. 2019. 420 с. URL: https://ela.kpi.ua/bitstream/123456789/26895/1/OP_ta_TsZ_pidruchnyk.pdf
3. Желібо Є. П., Заверуха Н. М., Зацарний В. В. Безпека життєдіяльності : підручник. Київ : Каравела, 2017. 320 с. URL: <https://xn--elajqk.kiev.ua/wp-content/uploads/2019/12/ZHelibo-Bezpeka-ZHitt-diyalnosti-2008.pdf>
4. Пащенко О. В., Перкун І. В., Погребняк В. Г. Підвищення безпеки праці при проведенні гідроструминної водополімерної перфوراції свердловин. *Системи та технології*. 2025. Vol. 69, No.1, 176–186. <https://doi.org/10.32782/2521-6643-2025-1-69.21>
5. Вплив війни на соціально-економічні системи: безпековий аспект: колективна монографія за заг. ред. д.е.н., проф. М.І. Копитко., д.ю.н., проф. З.Р. Кісіль. Львів: ЛьвДУВС, 2024. 394 с. URL: https://dspace.lvduvs.edu.ua/bitstream/1234567890/7280/1/%D0%9C%D0%BE%D0%BD%D0%B%D0%B3%D1%80%D0%B0%D1%84%D1%96%D1%8F_%D0%92%D0%9B%D0%98%D0%92%D0%92%D0%86%D0%99%D0%9D%D0%98.pdf
6. Про організацію трудових відносин в умовах воєнного стану: Закон України від 15.03.2022 р. № 2136-IX. URL: <https://zakon.rada.gov.ua/laws/show/2136-20#Text>
7. Радецький І. І., Горобинський С. В., Костанян В. Р. і ін. Удосконалення системи управління безпеки виробництва підприємств нафтогазового комплексу. *Нафтова і газова промисловість*, 2012. № 6. С. 16–20.
8. Кодекс цивільного захисту України: Закон України від 02.10.2012 р. № 5403-VI. URL: <https://zakon.rada.gov.ua/laws/show/5403-17#Text>
9. Ібрагім Умар, Тімоті О. І., Адеремі Адехумо, Аввал М. Г. Assessing The Use of AI for Improving Safety and Performance for Building Construction Workers. *Nile Journal of Engineering and Applied Science*. 2024. No. 2(1). P. 1–10. <https://doi.org/10.5455/NJEAS.189076>
10. Badri A., Boudreau-Trudel B., Souissi A. S. Occupational Health and Safety in the Industry 4.0 Era: A Cause for Major Concern? *Safety science*. 2018. Vol. 109. P. 403–411. <https://doi.org/10.1016/j.ssci.2018.06.012>
11. Hale A., Guldenmund F. Safety 4.0: The role of new digital technologies in occupational safety. *Safety Science*. 2022. Vol. 150. Art. no. 105697.
12. Суходоля О. М., Харазішвілі Ю. М., Рябцев Г. Л. Енергетична безпека України: перспективна модель управління ризиками: монографія за ред. О. М. Суходолі. Київ: НІСД, 2023. 152 с. <https://doi.org/10.53679/NISS-book.2023.01>
13. Енергетична безпека України: стан та шляхи зміцнення: аналітична доповідь / Центр Разумкова. 2024. URL: <https://razumkov.org.ua/novyny/energetychna-bezpeka-ukrayiny-standa-shliakhy-zmitsnennya>
14. Робота під тиском війни: важливі інсайти з дослідження психічного здоров'я працівників в Україні / Robota.ua. 2025. URL: https://budni.robota.ua/hr/robota-pid-tiskom-viyni-vazhliivi-insayti-z-doslidzhennya-psihichnogo-zdorov_ya-pratsivnikiv-v-u
15. Safety and Health at Work [Electronic resource] / International Labour Organization. 2024. URL: <https://www.ilo.org/media/67566/download>
16. Погребняк В. Г., Кондрат О. Р. Перкун І. В., Погребняк А. В. Моделювання професійних і воєнних ризиків в системі безпеки праці та цивільного захисту на об'єктах нафтогазової галузі. *Прикарпатський вісник НТШ. Число*. 2025. № 20 (77). 9 с.

References:

1. Pro okhoronu pratsi [On labor protection]. (1992). *Vidomosti Verkhovnoi Rady Ukrainy*, 49, 668. Retrieved from: <https://zakon.rada.gov.ua/laws/show/2694-12>
2. Levchenko, O. H., Bakhrushyn, V. E., & Sokolov, M. M. (2019). *Okhorona pratsi ta tsyvilnyi zakhyst: pidruchnyk* [Occupational safety and civil defense: textbook]. KPI im. I. Sikorskoho. Retrieved from: https://ela.kpi.ua/bitstream/123456789/26895/1/OP_ta_TsZ_pidruchnyk.pdf
3. Zhelibo, Ye. P., Zaverukha, N. M., & Zatsarnyi, V. V. (2017). *Bezpeka zhyttiediialnosti: pidruchnyk* [Life safety: textbook]. Karavela. Retrieved from: <https://xn--elajqk.kiev.ua/wp-content/uploads/2019/12/ZHelibo-Bezpeka-ZHitt-diyalnosti-2008.pdf>
4. Pashchenko, O. V., Perkun, I. V., & Pohrebniak, V. H. (2025). Pidvyshchennia bezpeky pratsi pry provedenni hidrostrumynnoi vodopolimernoi perforatsii sverdllovyn [Improving occupational safety during hydro-jet water-polymer perforation of wells]. *Systemy ta tekhnolohii*, 69(1), 176–186. <https://doi.org/10.32782/2521-6643-2025-1-69.21>
5. Kopytko, M. I., & Kisil, Z. R. (Eds.). (2024). Vplyv viiny na sotsialno-ekonomichni systemy: bezpekovi aspekt: kolektyvna monohrafiia [The impact of war on socio-economic systems: the security aspect: a collective monograph]. LvDUVS. Retrieved from: https://dspace.lvduvs.edu.ua/bitstream/1234567890/7280/1/%D0%9C%D0%BE%D0%BD%D0%BE%D0%B3%D1%80%D0%B0%D1%84%D1%96%D1%8F_%D0%92%D0%9F%D0%9B%D0%98%D0%92%20%D0%92%D0%86%D0%99%D0%9D%D0%98.pdf
6. Pro orhanizatsiiu trudovykh vidnosyn v umovakh voiennoho stanu [On the organization of labor relations under martial law] (Zakon Ukrainy No. 2136-IX). (2022, March 15). Verkhovna Rada Ukrainy. Retrieved from: <https://zakon.rada.gov.ua/laws/show/2136-20#Text>
7. Radetskyi, I. I., Horobynskyi, S. V., Kostanian, V. R., et al. (2012). Udoskonalennia systemy upravlinnia bezpeky vyrobnytstva pidpriemstv naftohazovoho kompleksu [Improving the safety management system for production at oil and gas complex enterprises]. *Naftova i hazova promyslovist*, (6), 16–20. [in Ukrainian].
8. Kodeks tsyvilnoho zakhystu Ukrainy [Civil Protection Code of Ukraine] (Zakon Ukrainy No. 5403-VI). (2012, October 2). Verkhovna Rada Ukrainy. Retrieved from: <https://zakon.rada.gov.ua/laws/show/5403-17#Text>
9. Umar, I., O. I., T., Adehumu, A., & M. G., A. (2024). Assessing The Use of AI for Improving Safety and Performance for Building Construction Workers. *Nile Journal of Engineering and Applied Science*, 2(1), 1–10. <https://doi.org/10.5455/NJEAS.189076>
10. Badri, A., Boudreau-Trudel, B., & Souissi, A. S. (2018). Occupational Health and Safety in the Industry 4.0 Era: A Cause for Major Concern? *Safety Science*, 109, 403–411. <https://doi.org/10.1016/j.ssci.2018.06.012>
11. Hale, A., & Guldenmund, F. (2022). Safety 4.0: The role of new digital technologies in occupational safety. *Safety Science*, 150, Article 105697. <https://doi.org/10.1016/j.ssci.2022.105697>
12. Sukhodolia, O. M., Kharazishvili, Yu. M., & Riabtsev, H. L. (2023). *Enerhetychna bezpeka Ukrainy: perspektyvna model upravlinnia ryzykamy: monohrafiia* [Energy security of Ukraine: a promising risk management model: monograph] (O. M. Sukhodolia, Ed.). NISD. <https://doi.org/10.53679/NISS-book.2023.01>
13. Tsentrazumkova. (2024). Enerhetychna bezpeka Ukrainy: stan ta shliakhy zmitsnennia: analitychna dopovid [Energy security of Ukraine: state and ways to strengthen: analytical report]. Retrieved from: <https://razumkov.org.ua/novyny/energetychna-bezpeka-ukrayiny-stan-ta-shliakhy-zmitsnennia>
14. Robota.ua. (2025). Robota pid tyskom viiny: vazhlyvi insaity z doslidzhennia psykhichnoho zdorovia pratsivnykiv v Ukraini [Work under the pressure of war: important insights from a study of the mental health of workers in Ukraine]. Retrieved from: <https://budni.robota.ua/hr/robota-pid-tyskom-viyni-vazhlyvi-insayti-z-doslidzhennia-psyhichnogo-zdorov-ya-pratsivnykiv-v-u>
15. International Labour Organization. (2024). *Safety and Health at Work*. Retrieved from: <https://www.ilo.org/media/67566/download>
16. Pohrebniak, V. H., Kondrat, O. R., Perkun, I. V., & Pohrebniak, A. V. (2025). Modeliuvannia profesiinykh i voiennykh ryzykiv v systemi bezpeky pratsi ta tsyvilnoho zakhystu na ob'ektakh naftohazovoi haluzi [Modeling of professional and military risks in the system of labor safety and civil protection at oil and gas facilities]. *Prycarpatskyi visnyk NTSh. Chyslo*, 20(77), 9. [in Ukrainian].

Дата надходження статті: 03.10.2022

Дата прийняття статті: 10.11.2025

Опубліковано: 30.12.2025

МАШИНОБУДУВАННЯ (ЗА СПЕЦІАЛІЗАЦІЯМИ)

УДК 621.791:621.785

DOI <https://doi.org/10.32782/2521-6643-2025-2-70.33>

Кирилах С. В., здобувачка вищої освіти кафедри
«Інтегровані технології зварювання та моделювання конструкцій»
Національного університету «Запорізька політехніка»
ORCID: 0009-0001-5688-5616

ВПЛИВ ФРИКЦІЙНО-ЗМІШУВАЛЬНОЇ ОБРОБКИ (FRICTION STIR PROCESSING) НА МІКРОСТРУКТУРУ ТА МЕХАНІЧНІ ВЛАСТИВОСТІ АДИТИВНО ВИГОТОВЛЕНОГО Ti-6Al-4V

У роботі досліджено вплив фрикційно-змішувальної обробки (Friction Stir Processing, FSP) на мікроструктуру та механічні властивості адитивно виготовленого титаново-алюмінієвого сплаву Ti-6Al-4V. Розглянуто гібридний технологічний маршрут «адитивне виробництво + FSW», який включає процес електронно-променевого плавлення (EBM) для виготовлення вихідних зразків та подальше фрикційно-змішувальне зварювання (FSW) з детальним аналізом зон шва: зона перемішування (SZ), термомеханічно впливова зона (TMAZ) та зона, що не зазнала впливу тепла (HAZ). Проведено мікроструктурну характеристику із визначенням орієнтації зерен, рекристалізації β -фази та градієнту деформації, а також оцінено механічні властивості швів, включно з межею міцності на розрив і пластичністю. Порівняльний аналіз з результатами FSW чистого титану дозволив виділити специфічні особливості адитивно виготовленого Ti-6Al-4V, такі як вплив початкової колонної структури та дефектів плавленої мікроструктури на формування дрібнозернистої SZ і механічну стабільність шва. Висновки підкреслюють, що оптимізація параметрів FSW критично впливає на мікроструктурну однорідність та підвищення експлуатаційних властивостей, що робить технологію перспективною для авіаційної та транспортної галузей. Результати дослідження створюють наукову основу для подальшого вдосконалення гібридних маршрутів AM + FSW і прогнозування механічної поведінки адитивно виготовлених деталей під високим навантаженням.

Ключові слова: Friction Stir Processing, адитивне виробництво, Ti-6Al-4V, мікроструктура, зона перемішування (SZ), термомеханічно впливова зона (TMAZ), зона термічного впливу (HAZ), β -фаза, механічні властивості, гібридні технологічні маршрути, електронно-променеве плавлення (EBM), фрикційно-змішувальне зварювання (FSW), авіаційні конструкційні деталі.

Kyrylakha S. V. Effect of Friction Stir Processing (FSP) on the Microstructure and Mechanical Properties of Additively Manufactured Ti-6Al-4V

This study investigates the influence of Friction Stir Processing (FSP) on the microstructure and mechanical properties of additively manufactured Ti-6Al-4V alloy. A hybrid technological route, “additive manufacturing + friction stir welding (AM + FSW),” was employed, involving electron beam melting (EBM) for the fabrication of initial samples, followed by FSP to enhance structural integrity. The research focuses on the detailed characterization of weld zones, including the stir zone (SZ), thermo-mechanically affected zone (TMAZ), and heat-affected zone (HAZ). Microstructural analysis was conducted to evaluate grain orientation, β -phase recrystallization, and deformation gradients, while mechanical testing assessed ultimate tensile strength and plasticity. Comparative analysis with conventional FSW of pure titanium highlighted specific features of additively manufactured Ti-6Al-4V, such as the influence of the initial columnar grain structure and intrinsic melting-induced defects on the formation of a fine-grained SZ and the overall mechanical performance of the joints. Results indicate that optimization of FSP parameters, particularly tool rotational speed and traverse rate, critically affects microstructural homogeneity, defect elimination, and enhancement of mechanical properties. The findings demonstrate that the hybrid AM + FSW route can effectively compensate for typical additive manufacturing defects, improving isotropy, tensile strength, and ductility of structural components. This research provides a scientific basis for the further development and optimization of hybrid manufacturing routes, allowing for the prediction and control of mechanical behavior in high-performance Ti-6Al-4V components intended for aerospace and transport applications. The integration of additive manufacturing with friction stir processing represents a promising strategy to produce reliable, high-quality titanium alloy components with enhanced durability and mechanical resilience.

Key words: Friction Stir Processing, additive manufacturing, Ti-6Al-4V, microstructure, stir zone (SZ), thermo-mechanically affected zone (TMAZ), heat-affected zone (HAZ), β -phase, mechanical properties, hybrid manufacturing routes, electron beam melting (EBM), friction stir welding (FSW), aerospace structural components.

© С. В. Кирилах, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

Постановка проблеми. У сучасному машинобудуванні та авіаційній галузі зростає потреба у високоміцних, легких та довговічних конструкційних матеріалах. Титановий сплав Ti-6Al-4V завдяки поєднанню високої міцності, корозійної стійкості та низької щільності є одним із пріоритетних матеріалів для виготовлення відповідальних деталей літаків, космічної та транспортної техніки. Однак традиційні методи виготовлення та зварювання цього сплаву обмежені низькою продуктивністю та високою ймовірністю дефектів, таких як пористість, тріщини та неоднорідність мікроструктури, що негативно впливає на механічні властивості виробів.

Адитивні технології (AM), зокрема електронно-променеве плавлення (EBM), дозволяють виготовляти складні геометричні конструкції, але створюють специфічні дефекти та колонну морфологію зерен, які зумовлюють анізотропію механічних властивостей. У цьому контексті фрикційно-змішувальне зварювання (FSW) виступає ефективним методом модифікації мікроструктури, усунення дефектів AM-деталей і підвищення механічної продуктивності.

Аналіз останніх досліджень (Singh, 2020; Rubino, 2019) показує, що поєднання адитивного виробництва з FSW (гібридний технологічний маршрут «AM + FSW») дозволяє значно поліпшити мікроструктурні та механічні характеристики деталей Ti-6Al-4V. Проте досі залишаються невирішеними питання:

- точне визначення впливу параметрів FSW на градієнт деформації та β -фазу в зоні перемішування (SZ);
- механізми формування термомеханічно впливової зони (TMAZ) і її взаємозв'язок із вихідною колонною структурою AM-матеріалу;
- оптимізація швидкості обертання та переміщення інструмента для досягнення балансу між міцністю та пластичністю без утворення дефектів.

Таким чином, існує потреба у комплексному дослідженні мікроструктурної еволюції та механічних властивостей Ti-6Al-4V, виготовленого AM і обробленого FSW, що дозволить сформулювати рекомендації щодо оптимізації гібридних технологічних маршрутів для авіаційних та транспортних застосувань.

Аналіз останніх досліджень і публікацій. Упродовж останнього десятиліття вивчення мікроструктурних та механічних характеристик титанових сплавів, зокрема Ti-6Al-4V, виготовлених за допомогою адитивних технологій та оброблених методом фрикційно-змішувального зварювання (FSW), стало предметом численних досліджень. Ці роботи спрямовані на вдосконалення технологічних процесів, підвищення якості з'єднань та розширення застосування таких матеріалів у високонавантажених конструкціях.

1. Вплив параметрів FSW на мікроструктуру та механічні властивості Ti-6Al-4V

У статті автори Rajath et al. (2024) [1] здійснили всебічний огляд параметрів FSW, що впливають на мікроструктуру та механічні властивості Ti-6Al-4V. Автори відзначають, що оптимізація швидкості обертання інструмента та швидкості переміщення може значно покращити механічні характеристики, зокрема міцність на розрив та втомну міцність, завдяки утворенню дрібнозернистої структури в зоні перемішування (SZ) [1].

2. Мікроструктурні зміни в адитивно виготовлених зразках Ti-6Al-4V після FSW

Досліджено Singh et al. (2020) [2] мікроструктурні зміни в адитивно виготовлених зразках Ti-6Al-4V після FSW. Вони виявили, що колонна структура, характерна для зразків, виготовлених методом електронно-променевого плавлення (EBM), руйнується під час FSW, що призводить до формування дрібнозернистої структури в SZ та часткової рекристалізації в термомеханічно впливовій зоні (TMAZ) [2].

3. Використання віртуального FSW для оптимізації процесу

У статті, [3] опублікованій у 2024 році, розглянуто використання віртуального FSW (VFSW) для оптимізації параметрів процесу при зварюванні Ti-6Al-4V. Цей підхід дозволяє моделювати процес зварювання та прогнозувати мікроструктурні зміни, що сприяє більш точному налаштуванню параметрів процесу.

4. Вплив швидкості обертання інструмента на мікроструктуру

Дослідження, проведене Brayden Terry et al. (2024), [4] встановило, що швидкість обертання інструмента має значний вплив на мікроструктуру та механічні властивості Ti-6Al-4V при FSW. Збільшення швидкості обертання призводить до підвищення температури в зоні зварювання, що сприяє кращій пластичній деформації та утворенню дрібнозернистої структури.

5. Використання нових технологій FSW

У роботі, опублікованій у 2023 році, розглянуто використання нових технологій FSW, таких як Additive Friction Stir Deposition (AFSD), для виготовлення частин у твердому стані без досягнення температури плавлення. Цей підхід дозволяє створювати складні геометричні форми з високою точністю та мінімізацією теплових деформацій [5].

6. Дослідниками Liu H та інші було продемонстровано [6] вплив параметрів фрикційно-змішувального зварювання (FSW) на мікроструктуру та механічні властивості чистого титану. Було встановлено:

Мікроструктура: утворюється чітка зональна структура шва – SZ, TMAZ, HAZ. У SZ спостерігається дрібнозерниста структура через інтенсивну пластичну деформацію; TMAZ характеризується частковою деформацією та нагрівом без повного перемішування; HAZ зберігає початкову структуру.

Механічні властивості: підвищення швидкості обертання інструмента покращує міцність на розрив та пластичність шва.

Текстура та орієнтація зерен: FSW руйнує колонну орієнтацію зерен, сприяючи більш ізотропній структурі.

Мета статті – метою цієї роботи є аналіз впливу фрикційно-змішувальної обробки (Friction Stir Processing, FSP) на мікроструктуру та механічні властивості адитивно виготовленого титаново-алюмінієвого сплаву Ti-6Al-4V, а також оцінка ефективності гібридного технологічного маршруту «адитивне виробництво + FSW» для підвищення механічної стабільності та якості зварних з'єднань у конструкційних матеріалах авіаційного та транспортного призначення.

Виклад основного матеріалу. Фрикційно-змішувальна обробка (FSP) є перспективною технологією для удосконалення механічних властивостей матеріалів, особливо в контексті адитивного виробництва. Цей метод дозволяє значно покращити характеристики матеріалу, зокрема його міцність, твердість та витривалість, шляхом локального змішування та рекристалізації мікроструктури. Особливо актуальним є застосування FSP до титанових сплавів, таких як Ti-6Al-4V, які широко використовуються в аерокосмічній та медичній промисловості завдяки своїй високій міцності та корозійній стійкості [7, 8, 9, 10, 11].

Фрикційно-змішувальна обробка (FSP) – це технологія обробки металів у твердому стані, яка була розроблена на основі методу фрикційного зварювання. Вона полягає у зануренні обертового інструмента в заготовку та його переміщенні уздовж неї при значущому осьовому навантаженні. У результаті генерується локальний нагрів, матеріал пластично деформується, відбувається перемішування та рекристалізація мікроструктури, що сприяє тоншому зерну і покращеним механічним властивостям (міцність, твердість, витривалість) [10, 11].

Сучасні дослідження у галузі обробки титанових сплавів, зокрема Ti-6Al-4V, демонструють ефективність гібридних технологічних маршрутів «AM + FSW», що поєднують переваги адитивного виробництва (Additive Manufacturing, AM) та фрикційного зварювання з перемішуванням (Friction Stir Welding, FSW). Такий підхід дає змогу не лише усунути дефекти, характерні для адитивних процесів (пори, неоднорідна мікроструктура, залишкові напруження), а й сформувати дрібнозернисту рекристалізовану структуру з підвищеною міцністю та пластичністю. У результаті гібридна схема «AM + FSW» розглядається як перспективна технологічна стратегія для виготовлення високоміцних з'єднань і зміцнених поверхонь титанових деталей, призначених для авіаційного та транспортного машинобудування.

Відповідно до дослідження Singh et al. (2020) [2, с. 392–397], мікроструктура адитивно виготовленого сплаву Ti-6Al-4V, обробленого методом фрикційно-змішувального зварювання, характеризується чітко вираженою зональною будовою – зоною перемішування (SZ), термомеханічно деформованою зоною (TMAZ) та зоною термічного впливу (HAZ). У вихідному адитивному матеріалі спостерігалася колонна морфологія зерен, орієнтованих уздовж напрямку нарощування шарів, що зумовлювало значну анізотропію властивостей. Після FSW така орієнтація повністю руйнується внаслідок інтенсивної динамічної рекристалізації у SZ, де формується дрібнозерниста $\alpha+\beta$ -структура з розміром зерен 3–5 мкм. У межах α -пластин фіксується залишкова β -фаза, яка утворюється внаслідок швидкого охолодження та сприяє підвищенню в'язкості й зниженню схильності до крихкого руйнування. На межі SZ і TMAZ спостерігається градієнт деформації, що виявляється у поступовій зміні морфології зерен від рівномірної в зоні перемішування до витягнутої в TMAZ. У TMAZ α -пластини частково орієнтовані за напрямком обертання інструмента, що зумовлює локальну текстуру. Структура HAZ при цьому залишається близькою до вихідної $\alpha+\beta$ -морфології, зазнаючи лише незначного видовження пластин під дією теплового впливу. Для всіх зон характерна відсутність пор і тріщин, що свідчить про високу ефективність FSW як методу ущільнення адитивних матеріалів. Підвищення швидкості обертання інструмента (від 400 до 700 об/хв) інтенсифікує процес рекристалізації, сприяючи утворенню дрібнозернистої рівномірної структури та покращенню поєднання міцності й пластичності матеріалу [2].

У порівнянні з традиційними плавленими процесами, FSP дозволяє зменшити дефекти, властиві металургії плавлення (пори, тріщини, крупні зерна), що робить її особливо актуальною для адитивного виробництва, де ці дефекти часто присутні [1, 2].

Згідно з роботою [2, с. 392–401], на рисунку 1 представлено інтегрований процес виготовлення та з'єднання зразків із Ti-6Al-4V, а також подальшу механічну характеристику:

1. Технологічна схема (ліворуч):

Показано процес електронно-променевого плавлення (EBM) для виготовлення пластин із Ti-6Al-4V. Заготовки формуються у вигляді порошкової подушки, яка під дією електронного променя плавиться і ущільнюється шар за шаром. Цей процес дозволяє отримати деталі зі складною геометрією, однак у матеріалі можуть залишатися дефекти, такі як пористість та неоднорідність мікроструктури.

2. Мікроструктурна характеристика (центр):

Далі зразки піддаються фрикційному зварюванню з перемішуванням (FSW). Мікрофотографії показують три основні зони зварного шва:

– Зона перемішування (SZ) – зона інтенсивної пластичної деформації та рекристалізації, де формуються дрібні, пластинчасті зерна.

- Термомеханічно впливова зона (TMAZ) – область, де відбувається часткова деформація та нагрівання без повного перемішування.
- Зона, що не зазнала впливу тепла (HAZ) – структура залишається аналогічною вихідному матеріалу.

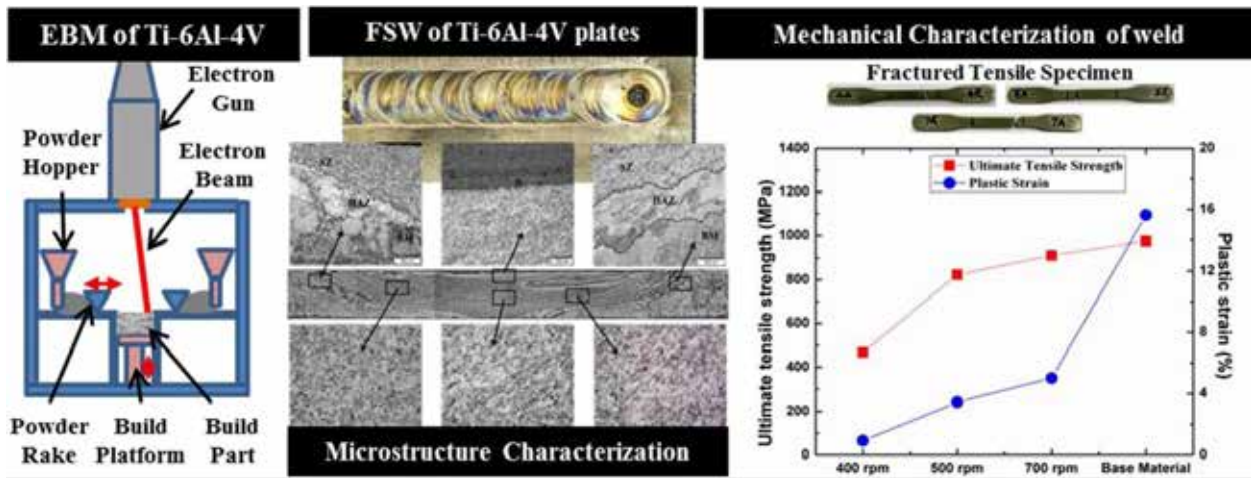


Рис. 1. Схема експериментальної методики: (а) процес електронно-променевого плавлення (EBM) Ti-6Al-4V; (б) фрикційно-змішувальне зварювання (FSW) та мікроструктурна характеристика; (с) результати механічних випробувань зварного шва

Джерело: [2]

Мікроструктурні зони після FSW/FSP

Зона перемішування (SZ) – зона інтенсивної пластичної деформації та рекристалізації, де формуються дрібні, пластинчасті зерна. У цій зоні відбувається повна динамічна рекристалізація, структура стає однорідною, дрібнозернистою, що сприяє покращенню механічних властивостей, зокрема межі міцності на розрив і пластичності.

Термомеханічно впливова зона (TMAZ) – область, де відбувається часткова деформація та нагрівання без повного перемішування. Характеризується частковою пластичною деформацією та частковою рекристалізацією зерен, розташована між SZ та HAZ. Структура зерен змінюється, але зберігає деякі особливості вихідного матеріалу, що визначає проміжні механічні властивості.

Зона, що не зазнала впливу тепла (HAZ) – структура залишається аналогічною вихідному матеріалу, проте можуть спостерігатися незначні термічні зміни, які практично не впливають на механічні властивості.

Базовий матеріал (BM, Base Material) – область, яка не піддавалася жодному термічному або механічному впливу під час FSW/FSP. Зберігає початкову мікроструктуру після адитивного виробництва або лиття, з властивостями, характерними для вихідного матеріалу.

Відзначається еволюція мікроструктури від грубих колонних зерен у вихідному EBM-матеріалі до дрібнозернистої, більш однорідної структури в SZ. Це підтверджує ефективність FSW у покращенні матеріальних характеристик шляхом змішування та динамічної рекристалізації.

3. Механічні результати (праворуч):

Механічне випробування демонструє вплив швидкості обертання інструмента на межу міцності на розрив (UTS) та пластичну деформацію. Результати такі:

400 rpm: UTS $\approx$ 900–950 MPa, пластична деформація $\approx$ 6–7 %.

500 rpm: UTS $\approx$ 1100 MPa, пластичність $\approx$ 10–12 %.

700 rpm: UTS наближається до міцності базового матеріалу ($\sim$ 1200 MPa), пластичність $>$ 15 %.

Зі збільшенням швидкості обертання інструмента спостерігається покращення механічних властивостей, що пояснюється більш інтенсивним перемішуванням матеріалу та дрібнозернистою рекристалізацією у зоні перемішування. Це дозволяє зменшити дефекти, характерні для EBM-процесу, і підвищити стійкість зварного шва до навантажень.

Таким чином, представлений на рисунку процес демонструє повний ланцюг «виробництво – зварювання – механічна характеристика» для Ti-6Al-4V. Це підкреслює важливість інтегрованого підходу до моделювання залишкових напружень, еволюції мікроструктури та механічної відповіді матеріалу для підвищення експлуатаційних характеристик компонентів авіаційного транспорту [2, с. 392–401].

Аналіз мікроструктури та механічних властивостей Ti-6Al-4V після FSW

Згідно з роботою [2, с. 392–401], застосування фрикційного зварювання з перемішуванням (FSW) до адитивно виготовлених заготовок Ti-6Al-4V суттєво впливає на мікроструктуру та механічні властивості

матеріалу. До обробки, матеріал, виготовлений методом EBM, характеризується колонною зернистою структурою та наявністю пористості й неоднорідного розподілу фаз α та β . Після FSW спостерігається чітка зона градієнту властивостей:

Зона перемішування (SZ) демонструє дрібнозернисту структуру з пластинчастими елементами α -фази та ознаками динамічної рекристалізації, що підвищує однорідність та пластичну деформацію.

Термомеханічно впливова зона (TMAZ) характеризується частковим змішуванням та деформацією зерен, із збереженням деяких особливостей вихідного матеріалу.

Зона, що не зазнала впливу тепла (HAZ) практично зберігає структуру базового матеріалу, але може мати незначні зміни через тепловий вплив.

Механічні випробування показують, що збільшення швидкості обертання інструмента від 400 до 700 г/м сприяє підвищенню межі міцності на розрив від $\approx 900\text{--}950$ МПа до ≈ 1200 МПа та збільшенню пластичної деформації від 6–7 % до понад 15 %. Це свідчить про ефективність FSW у зменшенні дефектів, характерних для EBM, та покращенні механічних властивостей через інтенсивніше перемішування матеріалу та дрібнозернисту рекристалізацію.

Додатково, Rubino et al. (2019) [7, с. 321–331] підтверджує ці тенденції, водночас надаючи більш детальну характеристику поверхневих і внутрішніх змін. Автори відзначають, що після FSW з'являється більш однорідна $\alpha+\beta$ структура у SZ, зменшуються розміри колонних зерен та вирівнюється розподіл фаз, що впливає на рівномірність пластичного потоку та зносостійкість. TMAZ і HAZ демонструють аналогічні зміни, але з менш вираженою рекристалізацією та залишковими ознаками вихідної структури.

Порівняння двох робіт дозволяє зробити висновок, що оптимізація параметрів FSW (швидкість обертання, швидкість подачі, температура) критично впливає на кінцеві механічні характеристики та мікроструктурну однорідність (табл. 1).

Таблиця 1

Мікроструктурні зміни після FSW

Зона	Характеристика мікроструктури	Механічні властивості
SZ	Дрібнозерниста $\alpha+\beta$ структура, рекристалізація, однорідність, зменшені колонні зерна	Збільшення UTS до ≈ 1200 МПа, пластичність >15 %
TMAZ	Часткове перемішування, деформація зерен, збереження деяких особливостей базової структури	Помірне збільшення UTS, пластичність 10–12 %
HAZ	Залишкова структура базового матеріалу, незначні термічні зміни	Мінімальні зміни UTS і пластичності, можливі залишкові напруження

Джерело: [2, 7]

На основі порівняння двох робіт можна зробити висновок, що FSW після адитивного виробництва значно покращує мікроструктурну однорідність і механічні характеристики Ti-6Al-4V. Основний ефект досягається у SZ, де дрібнозерниста структура і рекристалізація підвищують міцність і пластичність, а TMAZ і HAZ демонструють менший вплив, що підтверджує важливість точного контролю параметрів обробки. Таке порівняння дозволяє не лише оцінити ефективність FSW, а й виділити ключові напрямки оптимізації процесу для досягнення стабільної якості зварних з'єднань у адитивно виготовлених деталях авіаційного транспорту.

Мікроструктурна еволюція та механічні властивості FSW-з'єднань Ti-6Al-4V. Подальші дослідження повинні бути спрямовані на оптимізацію технологічних параметрів процесу фрикційного зварювання з перемішуванням (FSW) – зокрема, швидкості обертання інструмента, швидкості подачі, осьового зусилля та геометрії штифта. Це дозволить керовано формувати мікроструктуру в зоні перемішування (SZ) шляхом контролю співвідношення між тепловим і механічним внеском енергії, що визначає інтенсивність динамічної рекристалізації та зернової рафінації.

Особливу увагу слід приділяти зниженню залишкових напружень і гетерогенності структури в перехідних зонах (TMAZ і HAZ), оскільки саме вони є потенційними концентраторами напружень і місцями зародження втомних тріщин. Оптимізація параметрів FSW у поєднанні з попереднім моделюванням теплових полів та швидкостей деформації може забезпечити баланс між високою міцністю з'єднання та достатньою пластичністю, що є критично важливим для адитивно виготовлених елементів з Ti-6Al-4V, які працюють під змінними навантаженнями.

Крім того, доцільним є поєднання FSW із локальними термомеханічними або термічними постобробками, спрямованими на усунення пор і текстурних неоднорідностей, притаманних виробам, виготовленим методами електронно-променевого або лазерного плавлення порошкового шару (EBM, SLM). Такі підходи відкривають перспективи для створення гібридних технологічних маршрутів «AM + FSW», що забезпечують підвищену надійність і прогнозовані механічні характеристики з'єднань.

Здійснено детальний аналіз мікроструктурних зон SZ, TMAZ та HAZ адитивно виготовленого Ti-6Al-4V після фрикційно-змішувального зварювання (FSW), із оцінкою морфології зерен, розподілу β -фази та градієнтів деформації. Мікроструктурна еволюція визначається взаємодією локального термомеханічного впливу та попередньої колонної текстури EBM-матеріалу, що включає інтенсивну пластичну деформацію, локальне нагрівання, динамічну рекристалізацію та формування залишкової β -фази. Нижче наведено ключові аспекти трансформації мікроструктури в різних зонах з'єднання [2, с. 392–397].

1. Орієнтація колонних зерен у вихідному адитивному матеріалі

Перед FSW, у вихідному EBM-зразку, спостерігалася чітко виражена колонна структура зерен, орієнтована вздовж напрямку нарощування шару. Така текстура характерна для процесів електронно-променевого плавлення (EBM) і зумовлює анізотропію механічних властивостей.

Важливо зазначити, що FSW повністю руйнує колонну орієнтацію, сприяючи формуванню ізотропної дрібнозернистої структури у SZ.

2. Динамічна рекристалізація у зоні перемішування (SZ)

У зоні SZ відбувається інтенсивна динамічна рекристалізація через локальний термомеханічний вплив. Це призводить до формування рівномірно розподіленої $\alpha + \beta$ структури з розміром зерен 3–5 мкм.

Така структура забезпечує високу міцність (~1150 МПа) і пластичність до 15%, оскільки дрібнозерниста морфологія покращує опір зародженню тріщин.

3. Наявність залишкової β -фази

На межах α -пластин у SZ зафіксовано залишкову β -фазу, яка утворюється внаслідок швидкого охолодження після локального нагрівання.

Ця фаза має кубічну решітку (BCC) і відіграє роль «амортизатора» деформацій – підвищує в'язкість і знижує схильність до крихкого руйнування.

4. Градієнт деформації між SZ і TMAZ

На межі між зонами спостерігається поступовий перехід зернової морфології – від рівномірної в SZ до витягнутої в TMAZ.

Це свідчить про градієнт пластичної деформації, який визначає механічну стабільність з'єднання: TMAZ виконує роль «буферного шару», що знижує напруження між SZ і HAZ.

5. Орієнтація α -пластин у TMAZ

У TMAZ α -пластини частково орієнтовані за напрямком обертання інструмента, що створює локальну текстуру.

Це цікаво з точки зору керування анізотропією, адже така орієнтація може бути корисною для регулювання напрямних властивостей матеріалу.

6. Мінімальні зміни у HAZ

У зоні HAZ структура зберігає $\alpha + \beta$ морфологію, характерну для вихідного EBM-матеріалу, проте α -пластини дещо видовжені через термічний вплив.

Це показує, що тепловий вплив процесу FSW є локалізованим, і структура поза зоною перемішування залишається стабільною.

7. Відсутність пор та дефектів після FSW

На відміну від вихідного AM-матеріалу, FSW-з'єднання практично не містять пор і мікротріщин.

Це ключовий аргумент на користь гібридного підходу «AM + FSW»: зменшення внутрішніх дефектів → підвищення щільності → покращення втомної довговічності.

8. Вплив швидкості обертання інструмента

Автори [2] показали, що зі збільшенням швидкості від 400 до 700 rpm зерна дрібнішають, а механічні властивості покращуються.

Підвищення швидкості обертання сприяє інтенсифікації термомеханічного впливу, збільшенню обсягу динамічної рекристалізації та формуванню дрібнозернистої структури з покращеною комбінацією міцності та пластичності».

Порівняння результатів фрикційно-змішувального зварювання чистого титану та адитивно виготовленого Ti-6Al-4V. У порівнянні з Singh et al. (2020) [2], де досліджувалася адитивно виготовлена Ti-6Al-4V, результати Liu et al. (2010) [6] дозволяють виділити кілька ключових моментів:

1. Колонна структура вихідного матеріалу: у Singh et al. (2020) колонна структура EBM-зразків руйнується під час FSW, аналогічно до ефекту, описаного Liu для чистого титану.

2. Зони шва (SZ, TMAZ, HAZ): обидві роботи підтверджують формування типових зон, однак у адитивно виготовленого Ti-6Al-4V в SZ спостерігається сильніший градієнт деформації та прискорена рекристалізація β -фази.

3. Механічні властивості: як і Liu, Singh відзначає, що підвищення швидкості обертання інструмента сприяє покращенню міцності та пластичності, але для адитивних зразків це особливо важливо для компенсації дефектів плавленої структури (пори, тріщини).

4. Застосування до гібридного маршруту AM + FSW: порівняння демонструє, що адитивне виробництво додає специфічні мікроструктурні особливості, які потрібно враховувати при оптимізації FSW, зокрема вплив на β -фазу та товщину зерна.

З порівняння результатів досліджень [2, 6] видно, що фрикційно-змішувальна обробка ефективно руйнує колонну структуру зерен вихідного матеріалу та формує дрібнозернисту мікроструктуру в SZ, забезпечуючи підвищення механічних властивостей зварних швів. Водночас вплив параметрів процесу, таких як швидкість обертання інструмента, критично впливає на градієнт деформації, рекристалізацію β -фази та міцність на розрив. Ці результати підкреслюють актуальність оптимізації гібридного технологічного маршруту «AM + FSW» для адитивно виготовленого Ti-6Al-4V.

Висновки

1. Було встановлено, що фрикційно-змішувальна обробка (Friction Stir Processing, FSP) ефективно модифікує мікроструктуру адитивно виготовленого Ti-6Al-4V, руйнуючи колонну структуру вихідного матеріалу та формуючи дрібнозернисту, пластинчасту структуру в зоні перемішування (SZ).

2. Термомеханічно впливова зона (TMAZ) характеризується частковою пластичною деформацією та частковою рекристалізацією зерен, тоді як зона, що не зазнала впливу тепла (HAZ), зберігає початкову структуру матеріалу.

3. Оптимізація параметрів FSW, зокрема швидкості обертання інструмента та швидкості переміщення, дозволяє підвищити міцність на розрив та пластичність матеріалу, зменшити дефекти та забезпечити більш однорідну мікроструктуру, що особливо важливо для адитивно виготовлених деталей.

4. Порівняння з результатами FSW чистого титану показало, що адитивно виготовлений Ti-6Al-4V має специфічні мікроструктурні особливості, такі як градієнт деформації та рекристалізація β -фази, які слід враховувати при налаштуванні технологічного процесу.

5. Подальші дослідження повинні зосередитися на:

- детальному вивченні взаємодії параметрів FSP з мікроструктурою та фазовим складом;
- моделюванні термомеханічних процесів у SZ, TMAZ та HAZ для прогнозування механічних властивостей;
- розробці гібридних технологічних маршрутів «AM + FSW» для отримання високоякісних конструкційних деталей з мінімізацією дефектів;
- вивченні впливу FSP на експлуатаційні характеристики деталей у реальних умовах, зокрема для авіаційної та транспортної галузі.

Отримані результати створюють наукову та практичну основу для подальшого вдосконалення технологій адитивного виробництва та фрикційно-змішувальної обробки титанових сплавів. Вони демонструють ефективність гібридного маршруту AM + FSW у формуванні дрібнозернистої, однорідної мікроструктури та підвищенні механічних властивостей Ti-6Al-4V, що дозволяє рекомендувати його для застосування у виготовленні високонавантажених конструкційних компонентів авіаційної та транспортної галузі.

Список використаних джерел:

1. Rajath N., Rao S. S., Rao V. V. A comprehensive review of friction stir welding parameters on Ti-6Al-4V alloy microstructure and mechanical properties. *Advanced Materials Processing and Technology*. 2024. Vol. 10(2). P. 196–237 <https://doi.org/10.1080/2374068X.2024.2306571>
2. Singh A. K., Kumar N., Jha K., Astarita A., Squillace A., Franchitti S., Arora A. Friction stir welding of additively manufactured Ti-6Al-4V: Microstructure and mechanical properties. *Journal of Materials Processing Technology*. 2020. Vol. 283. P. 392–401. <https://doi.org/10.1016/j.jmatprotec.2019.116433>
3. Liu X., Zhang P., Li W., Ye T., Xia C., Zhang T., Shi J., Xie J. Ni Z. Microstructure and Mechanical Properties of TC4 Titanium Alloy Joint Fabricated By Vortex Flow-Based Friction Stir Welding. *Chinese Journal of Mechanical Engineering*. 2024. Vol. 37. P. 159. <https://doi.org/10.1186/s10033-024-01152-z>
4. Terry B., Shelton A. Strauss A. Microstructure and mechanical properties of dissimilar friction stir welding between butted, thin sheet NiTi and Ti-6Al-4V. *Weld World* (2025). <https://doi.org/10.1007/s40194-025-02162-8>
5. Prabhakar D.A.P., Arun Kumar Shettigar, Mervin A. Herbert, Manjunath Patel G C, Danil Yu. Pimenov, Khaled Giasin, Chander Prakash A comprehensive review of friction stir techniques in structural materials and alloys: challenges and trends. *Journal of Materials Research and Technology*. 2022. Vol. 20. P. 3025–3060. <https://doi.org/10.1016/j.jmrt.2022.08.034>
6. Liu H., Nakata K., Liao J. Friction stir welding of pure titanium lap joint. *Science and Technology of Welding and Joining*. 2010. Vol. 15, Issue 5. <https://doi.org/10.1179/136217110X12731414740031>
7. Rubino F., Scherillo F., Franchitti S., Squillace A., Astarita A., Carlone P. Microstructure and surface analysis of friction stir processed Ti-6Al-4V plates manufactured by electron beam melting. *Journal of Manufacturing Processes*. 2019. Vol. 37. P. 392–401 <https://doi.org/10.1016/j.jmapro.2018.12.015>
8. Mishra R. S., Ma Z. Y. Friction stir welding and processing. *Materials Science and Engineering: R: Reports*. 2005. Vol. 50(1–2), P. 1–78. <https://doi.org/10.1016/j.mser.2005.07.001>

-
9. MA Z. Y. Friction Stir Processing Technology: A Review. *METALLURGICAL AND MATERIALS TRANSACTIONS A*. 2008. Vol. 39, P. 642–658. <https://doi.org/10.1007/s11661-007-9459-0>
 10. Wikipedia contributors. (2025, October 21). Friction stir processing. In Wikipedia, The Free Encyclopedia. URL: https://en.wikipedia.org/wiki/Friction_stir_processing
 11. Hovanski Ю., Mishra P., Sato Ю. (ред.). Friction Stir Welding and Processing IX. Cham: Springer Cham, 2017. 324 с. ISBN 978-3-319-52382-8. <https://doi.org/10.1007/978-3-319-52383-5>

References:

1. Rajath, N., Rao, S. S., & Rao, V. V. (2024). A comprehensive review of friction stir welding parameters on Ti-6Al-4V alloy microstructure and mechanical properties. *Advanced Materials Processing and Technology*, no 10(2), S. 196–237 <https://doi.org/10.1080/2374068X.2024.2306571>
2. Singh, A. K., Kumar, N., Jha, K., Astarita, A., Squillace, A., Franchitti, S., & Arora, A. (2020). Friction stir welding of additively manufactured Ti-6Al-4V: Microstructure and mechanical properties. *Journal of Materials Processing Technology*, no 283, S. 392–401. <https://doi.org/10.1016/j.jmatprotec.2019.116433>
3. Liu, X., Zhang, P., Li, W., Ye, T., Xia, C., Zhang, T., Shi, J., Xie, J. & Ni, Z. (2024). Microstructure and Mechanical Properties of TC4 Titanium Alloy Joint Fabricated By Vortex Flow-Based Friction Stir Welding. *Chinese Journal of Mechanical Engineering*. no 37, S.159. <https://doi.org/10.1186/s10033-024-01152-z>
4. Terry, B., Shelton, A. & Strauss, A. (2025). Microstructure and mechanical properties of dissimilar friction stir welding between butted, thin sheet NiTi and Ti-6Al-4V. *Weld World*. <https://doi.org/10.1007/s40194-025-02162-8>
5. Prabhakar D. A. P., Arun Kumar Shettigar, Mervin A. Herbert, Manjunath Patel G C, Danil Yu. Pimenov, Khaled Giasin, Chander Prakash (2022). A comprehensive review of friction stir techniques in structural materials and alloys: challenges and trends. *Journal of Materials Research and Technology*. no 20, S. 3025–3060. <https://doi.org/10.1016/j.jmrt.2022.08.034>
6. Liu, H., Nakata, K., & Liao, J. (2010). Friction stir welding of pure titanium lap joint. *Science and Technology of Welding and Joining*. no 15, Issue 5. <https://doi.org/10.1179/136217110X12731414740031>
7. Rubino, F., Scherillo, F., Franchitti, S., Squillace, A., Astarita, A., & Carlone, P. (2019). Microstructure and surface analysis of friction stir processed Ti-6Al-4V plates manufactured by electron beam melting. *Journal of Manufacturing Processes*. no 37, S. 392–401 <https://doi.org/10.1016/j.jmapro.2018.12.015>
8. Mishra, R. S., & Ma, Z. Y. (2005). Friction stir welding and processing. *Materials Science and Engineering: R: Reports*, no 50(1–2), S. 1–78. <https://doi.org/10.1016/j.mser.2005.07.001>
9. MA, Z. Y. (2008). Friction Stir Processing Technology: A Review. *METALLURGICAL AND MATERIALS TRANSACTIONS A*. no 39, S. 642–658. <https://doi.org/10.1007/s11661-007-9459-0>
10. Wikipedia contributors. (2025, October 21). Friction stir processing. In Wikipedia, The Free Encyclopedia. Retrieved from: https://en.wikipedia.org/wiki/Friction_stir_processing
11. Hovanski, Y., Mishra, R., & Sato, Y. (Eds.). (2017). Friction Stir Welding and Processing IX. Springer Cham. <https://doi.org/10.1007/978-3-319-52383-5>

Дата надходження статті: 21.10.2025

Дата прийняття статті: 10.11.2025

Опубліковано: 30.12.2025

Майбородіна Н. В., кандидат фізико-математичних наук, доцент, доцент кафедри природничо-математичних та загально інженерних дисциплін Відокремленого підрозділу Національного університету біоресурсів і природокористування України «Ніжинський агротехнічний інститут»
ORCID: 0000-0003-1754-6790

Герасименко В. П., кандидат технічних наук, доцент, доцент кафедри електричної інженерії Відокремленого підрозділу Національного університету біоресурсів і природокористування України «Ніжинський агротехнічний інститут»
ORCID: 0000-0002-4017-1141

МАТЕМАТИЧНА МОДЕЛЬ ЗАЛЕЖНОСТІ ОБСЯГУ РЕАЛІЗОВАНОЇ ПРОДУКЦІЇ МАШИНОБУДУВАННЯ ВІД ОБСЯГУ ВИДОБУТКУ ЗАЛІЗНИХ РУД

Машинобудування є ключовою та стратегічно важливою галуззю української економіки, незважаючи на значні виклики, які особливо пов'язані з повномасштабною війною. Ця галузь має значний виробничий і науково-технічний потенціал. У статті розглядається розробка математичної моделі для прогнозування та аналізу обсягу реалізованої продукції підприємств машинобудівного комплексу в залежності від обсягу видобутку залізних руд в Україні. Для створення економічної моделі чи прогнозу немає потреби враховувати всю сукупність факторів, що впливають на досліджувані показники. Ключовим завданням моделювання є виділення обмеженої кількості факторів, які мають найвищий ступінь впливу на досліджувані показники. Саме ці домінуючі фактори є основою для надійного аналізу, тоді як вплив другорядних факторів можна не враховувати. В даній роботі домінуючим фактором, який впливає на обсяг реалізованої продукції машинобудування, обрано обсяг видобутку залізних руд. Для моделювання залежності обсягу реалізованої продукції машинобудування в Україні використано статистичні дані Державної служби статистики України та методи регресійного аналізу. Для автоматизації побудови моделі, аналізу моделі та з метою прогнозування на основі обраної моделі обрано табличний процесор Excel. В роботі розраховані точкові прогнозовані значення обсягу реалізації продукції машинобудівного комплексу України на 2026 рік. Для забезпечення достовірності одержаних прогнозованих значень обрано лінію регресії з найвищим коефіцієнтом апроксимації. Дослідження повністю відповідає вимогам математичного моделювання: від постановки завдання та вибору ключового показника до побудови моделі, аналізу моделі та застосування результатів для прогнозування обсягу реалізації продукції машинобудівного комплексу України на 2026 рік.

Ключові слова: машинобудування, залізна руда, регресійний аналіз, прогноз.

Maiborodina N. V., Gerasymenko V. P. Mathematical modeling of engineering production volume dependence on iron ore extraction

Engineering is a key and strategically important sector of the Ukrainian economy, despite significant challenges, especially those associated with the full-scale war. This sector possesses considerable production and scientific-technical potential. This article examines the development a mathematical model for forecasting and analyzing the volume realized output of the engineering complex enterprises, based on the volume of iron ore mining in Ukraine. When creating an economic model or forecast, it is not necessary to consider the entire multitude of factors influencing the studied indicator. The key task is the filtering and selection a limited number of factors that exert the highest degree of influence. These dominant factors form the basis for reliable analysis, while the impact of secondary factors can be legitimately omitted. In this work, the volume of iron ore mining was chosen as the dominant factor affecting the volume the realized engineering output. For modeling the dependence realized engineering output volume in Ukraine, statistical data from the State Statistics Service of Ukraine and regression analysis methods were utilized. The model was developed using the Excel spreadsheet processor to automate the construction, analysis, and forecasting process based on the chosen model. The study includes the calculation of confidence intervals for the forecasted values the realized output volume of the Ukrainian engineering complex for the year 2026. To ensure the reliability the obtained forecasted values, the regression line with the highest approximation coefficient was selected. The research fully complies with the requirements of mathematical modeling: from setting the problem and selecting the key indicator to constructing, analyzing the model, and applying the results for forecasting the realized output volume of the Ukrainian engineering complex for 2026.

Key words: engineering (machine-building), iron ore, regression analysis, forecast.

© Н. В. Майбородіна, В. П. Герасименко, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

Постановка проблеми. Машинобудування є однією з найнеобхідніших галузей сучасної української економіки в умовах війни. Галузь охоплює такі критичні підгалузі: авіабудування, ракетно-космічна галузь, транспортне машинобудування (залізничне, суднобудування) та важке машинобудування (устаткування для металургії та енергетики). Така структура була традиційно зав'язана на сировинній базі та великих наукових центрах. Однак, з 2022 року галузь зіткнулася з гострою втратою потужностей на окупованих територіях, критичним скороченням експорту та проблемами з логістикою і кадровим дефіцитом. Попри значні втрати, машинобудівний комплекс розглядається як стратегічна точка зростання для відновлення економіки.

Актуальність дослідження зумовлена високою динамічністю ринкового середовища та необхідністю підвищення точності управлінського планування в умовах значної економічної невизначеності.

Аналіз останніх досліджень і публікацій. Дослідженню машинобудування приділяється велика увага вчених, оскільки забезпечення сталого розвитку є одним зі стратегічних завдань національної економіки країни.

М.О. Самофалова в роботі [1] дослідила еволюцію розвитку машинобудування в Україні за період XVII – XIX століття та початок XX століття. Основними чинниками, які сприяли розвитку машинобудування були іноземний капітал, західні технології та нове відношення до виробництва.

В дослідженні А.В. Державської та П.В. Круш [2] проаналізовано загально економічний стан підприємств машинобудування, загальний товарообіг з іноземними країнами, зазначено низький рівень їх конкурентоспроможності, незадовільний фінансовий стан та негативне фінансове сальдо зовнішньоторгового товарообігу промисловою продукцією.

В статті Л.В. Соколової та О.В. Стойки [3] обґрунтовано необхідність поновлення зв'язків машинобудівних підприємств з підприємствами інших галузей, відновлення їх матеріально-технічної бази.

В дослідженні Л.М. Сакун та ін. [4], надано рекомендації для розвитку підприємств, впровадження інноваційних рішень у виробництво та технологічний процес.

О. Ходирева [5] дослідила інноваційну активність машинобудівних підприємств та капітальні інвестиції в основне виробництво. В роботі зазначено, що критичними проблемами машинобудівних підприємств є значне недофінансування та бюрократизованість.

Дослідження А.С. Целікової та Є.М. Місько [6] присвячено сучасному стану та проблемам розвитку машинобудівних підприємств України. В роботі запропоновано шляхи впровадження державної підтримки на різних етапах, використання інноваційної моделі технологічних процесів, відновлення існуючої та будівництво нової інфраструктури, підвищення інвестиційної привабливості машинобудівних підприємств, посилення конкурентоспроможності підприємств та виробів.

Н.П. Карачина [7] дослідила трансформація моделей економічної поведінки машинобудівних підприємств за умов розвитку у воєнний час.

Беручи до уваги існуючі наукові дослідження у вивченні основних аспектів розвитку галузі машинобудування слід зазначити, що недостатньо досліджена залежність обсягу реалізованої продукції машинобудування в Україні від обсягу добування залізних руд.

Мета статті – побудова математичної моделі обсягу реалізованої продукції машинобудування від обсягу видобутку залізних руд в умовах невизначеності на основі наявних статистичних даних за попередні 2010 – 2024 роки із застосуванням методів математичної статистики. Пріоритетним завданням роботи є дослідження, аналіз побудованої моделі та прогнозування обсягу реалізованої продукції машинобудування на 2026 рік.

Виклад основного матеріалу. Дослідження обсягів машинобудування має важливе значення для сучасної економіки України, оскільки дохід промислового та сільськогосподарського виробництва, обороноспроможність України в умовах війни на пряму залежить від його можливостей.

Для успішного прогнозування обсягу реалізованої продукції машинобудування і вибору моделі, що використовуватиметься, необхідно обирати регресійну функцію серед можливих варіантів. Для вибору адекватної моделі будемо застосовувати спеціальні критерії якості моделі [8]. З метою перевірки достовірності запропонованої моделі, скористаємося елементами дисперсійного аналізу та перевіримо значущість отриманих результатів.

Оберемо для даної задачі рівень значущості $\alpha = 0,05$ (або 5%). Тоді можна стверджувати, що всі сформульовані висновки будуть виконуватись з рівнем довіри $p = 1 - \alpha = 0,95$, тобто з ймовірністю 95%.

Для створення математичної моделі залежності обсягу реалізованої продукції машинобудування від обсягу добування залізних руд скористаємося даними, які розміщені на сайті Державної служби статистики України в розділі Статистична інформація / Економічна статистика / Економічна діяльність / Промисловість / Обсяг реалізованої промислової продукції за видами діяльності (2010 – 2024) [9].

Статистичні дані обсягу реалізованої промислової продукції машинобудування та обсягу добування залізних руд за період 2010 – 2024 роки наведено в (табл. 1).

Таблиця 1

**Обсяги реалізованої промислової продукції машинобудування та добування залізних руд
за період 2010 – 2024 роки**

Роки	Обсяг реалізованої промислової продукції машинобудування, тис. грн	Обсяг видобування залізних руд, тис. грн
2010	50678234,2	109095225,8
2011	68760874,3	151196115,5
2012	58206687,3	164920075,0
2013	65089095,2	131379637,4
2014	72532798,4	113842226,4
2015	73580218,0	129652335,6
2016	82787136,5	151536511,3
2017	117368292,6	185347635,6
2018	128674907,1	220466275,4
2019	143118131,9	235742912,0
2020	154395680,0	217771772,7
2021	289199705,5	260029247,1
2022	104389130,3	208266353,0
2023	85101672,2	328156055,4
2024	130977099,9	571526642,5

Задача регресійного аналізу полягає в тому, щоб за наявними статистичними даними:

1. Отримати найкращі оцінки $\hat{a}_0, \hat{a}_1, \dots, \hat{a}_n$ невідомих параметрів $a_0, a_1, \dots, a_n$ шуканої моделі.

2. Перевірити статистичні гіпотези про параметри моделі.

3. Перевірити, чи добре отримана модель узгоджується зі статистичними даними (перевірити адекватність моделі) [8].

Якість отриманих оцінок параметрів $\hat{a}_0, \hat{a}_1, \dots, \hat{a}_n$ можна аналізувати шляхом дослідження відхилень між теоретичними значеннями моделі та фактичними статистичними даними. Такі відхилення дозволяють оцінити ступінь відповідності побудованої моделі реальним спостереженням. Одним із найпоширеніших підходів для оцінювання параметрів є метод найменших квадратів, який передбачає знаходження таких значень оцінок параметрів $\hat{a}_0, \hat{a}_1, \dots, \hat{a}_n$, при яких сума квадратів відхилень теоретичних значень від емпіричних є мінімальною:

$$S(\hat{a}_0, \hat{a}_1, \dots, \hat{a}_n) = \sum_{i=1}^n \left(y_i - f(x_i; \hat{a}_0, \hat{a}_1, \dots, \hat{a}_n) \right)^2 \quad (1)$$

де y_i – емпіричні дані, $f(x_i; \hat{a}_0, \hat{a}_1, \dots, \hat{a}_n)$ – емпірична модель, яка залежить від параметрів. Таким чином, метод найменших квадратів забезпечує знаходження найкращої апроксимації емпіричних даних за заданою моделлю.

З метою автоматизації побудови моделі застосуємо табличний процесор Microsoft Excel [10].

Обробку числових даних розпочнемо з використання Надстройки / Аналіз даних / Описова статистика. Результати наведено в (табл. 2).

Таблиця 2

Результати описової статистики в Excel

Характеристика	Обсяг реалізованої промислової продукції машинобудування, тис.грн	Обсяг добування залізних руд, тис.грн
Середнє значення	1,08E+08	2,12E+08
Стандартна похибка	15433835	30041714
Медіана	85101672	1,85E+08
Мода	-	-
Стандартне відхилення	59774984	1,16E+08
Дисперсія вибірки	3,57E+15	1,35E+16
Експес	5,732335	6,56467
Асиметрія	2,122784	2,329399
Інтервал	2,39E+08	4,62E+08
Мінімум	50678234	1,09E+08
Максимум	2,89E+08	5,72E+08
Сума	1,62E+09	3,18E+09
Кількість	15	15

На основі даних про обсяг реалізованої промислової продукції машинобудування та обсяг добування залізних руд з 2010 року по 2024 роки (таблиця 1):

1. Визначимо вид функції регресії.
2. Побудуємо функцію регресії.
3. Дослідимо достовірність побудованої функції регресії.
4. Виконаємо прогнозування за побудованою функцією регресії: спрогнозуємо обсяг реалізованої промислової продукції машинобудування на 2026 рік.

1. Для визначення виду функції залежності обсягу реалізованої промислової продукції машинобудування від обсягу добування залізних руд будемо кореляційне поле точок (рис. 1).

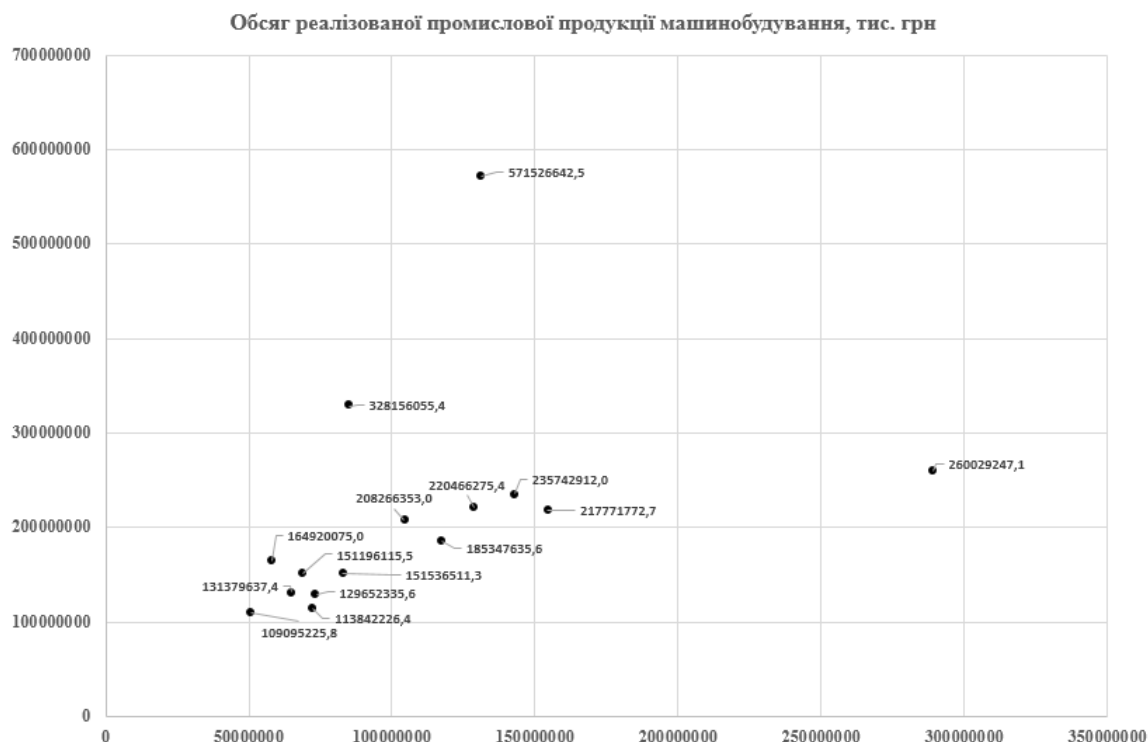


Рис. 1. Кореляційне поле точок

Як видно з приведенного графічного матеріалу на рис. 1 дані за 2023 та 2024 роки виходять за межі загальної тенденції залежності, тобто є аномальними. Таке відхилення можна пояснити військовими діями на території України в період 2023 – 2024 років, які значно вплинули на машинобудівний комплекс.

2. Засобами табличного процесору Microsoft Excel здійснюємо вибір функції регресії. На (рис. 2) зображено вибір лінії тренду.

Як видно з приведенного графічного матеріалу на рис. 2, наявність аномальних даних за 2023 та 2024 роки не дає можливості побудувати якісну модель залежності обсягу реалізованої промислової продукції машинобудування від обсягу добування залізних руд. Лінії тренду, які пропонує табличний процесор Excel мають дуже малий показник апроксимації $R^2 < 0,4$.

Саме тому авторами роботи було прийнято рішення відкинути аномальні дані за 2023 та 2024 роки і надалі досліджувати лише дані за період з 2010 року по 2022 рік. На рис. 3 зображено вибір лінії тренду для статистичних даних з 2010 року по 2022 рік.

Проведений аналіз запропонованих ліній тренду дає підстави стверджувати, що в якості математичної моделі можна обрати логарифмічну або поліноміальну функцію починаючи з другого порядку, оскільки їхні показники апроксимації $R^2 > 0,8$.

Після проведення детального аналізу даних, зображених на рис. 3 автори дослідження зробили припущення, про можливість аномального характеру даних з 2021 року. Відкидання аномальних даних за 2021 рік дало можливість побудувати лінійну модель залежності обсягу реалізованої промислової продукції машинобудування від обсягу добування залізних руд задовільної точності з показником апроксимації $R^2 > 0,8$ (рис. 4).

Згідно рис. 4 лінійна функція регресії має вигляд

$$y = 1,113x + 6 \cdot 10^7.$$

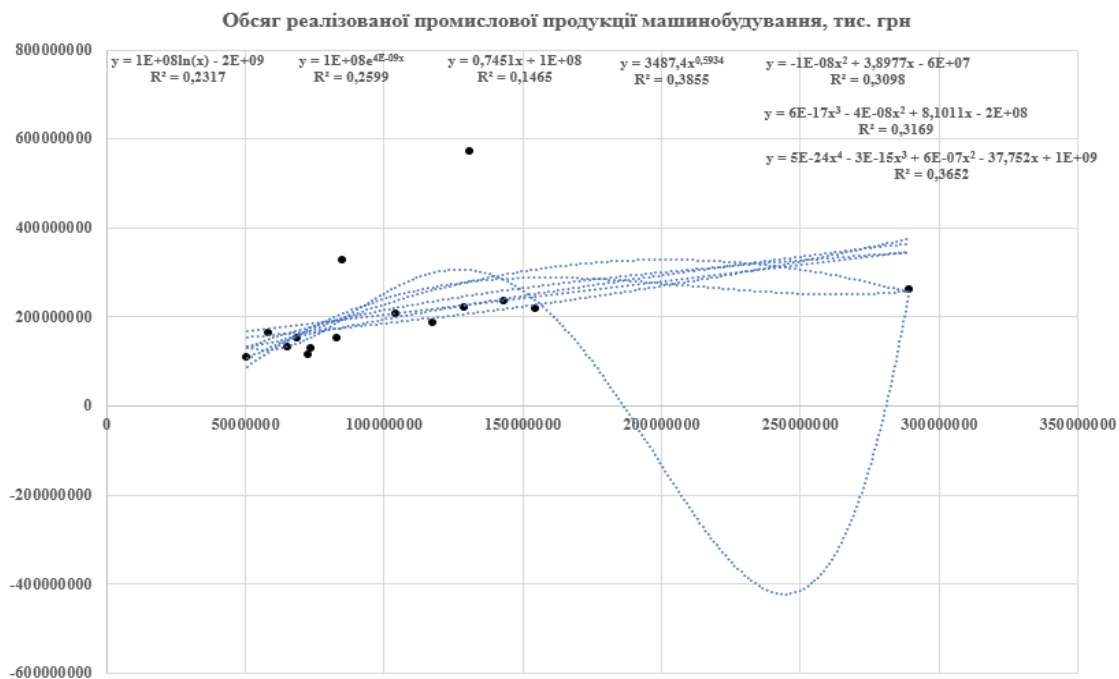


Рис. 2. Графіки залежностей обсягу реалізованої промислової продукції машинобудування від обсягу добування залізних руд в період 2010 – 2024 роки

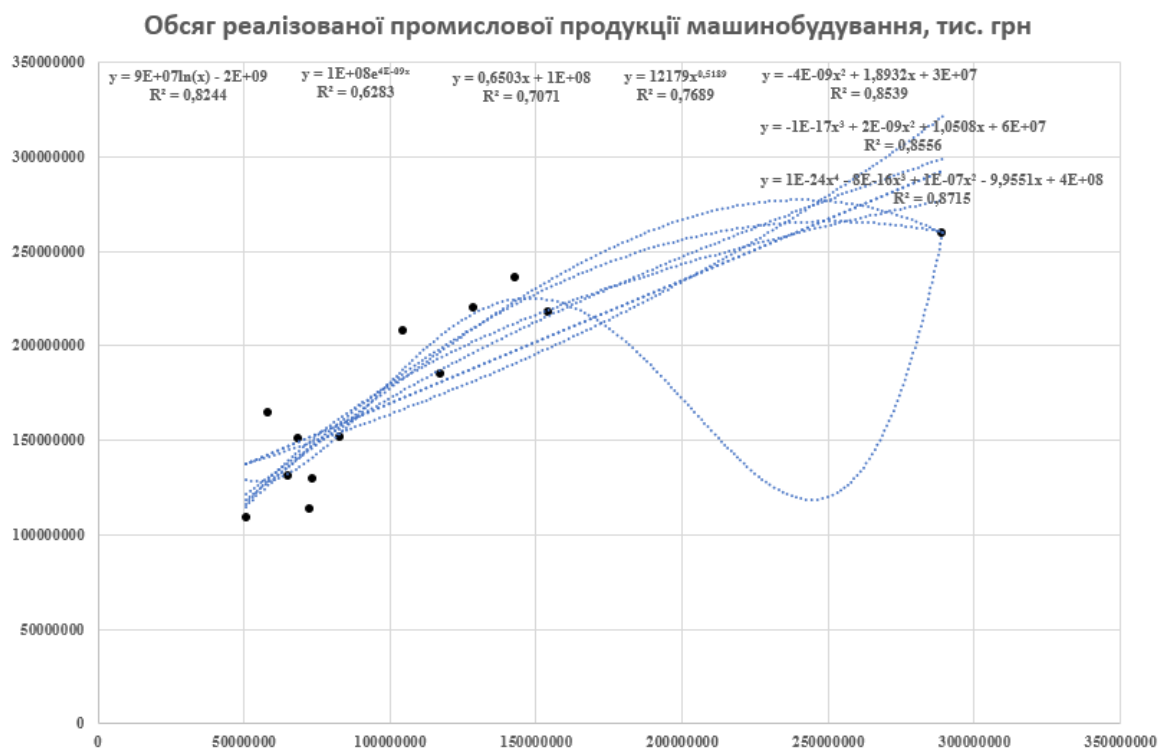


Рис. 3. Графіки залежностей обсягу реалізованої промислової продукції машинобудування від обсягу добування залізних руд в період 2010 – 2022 роки

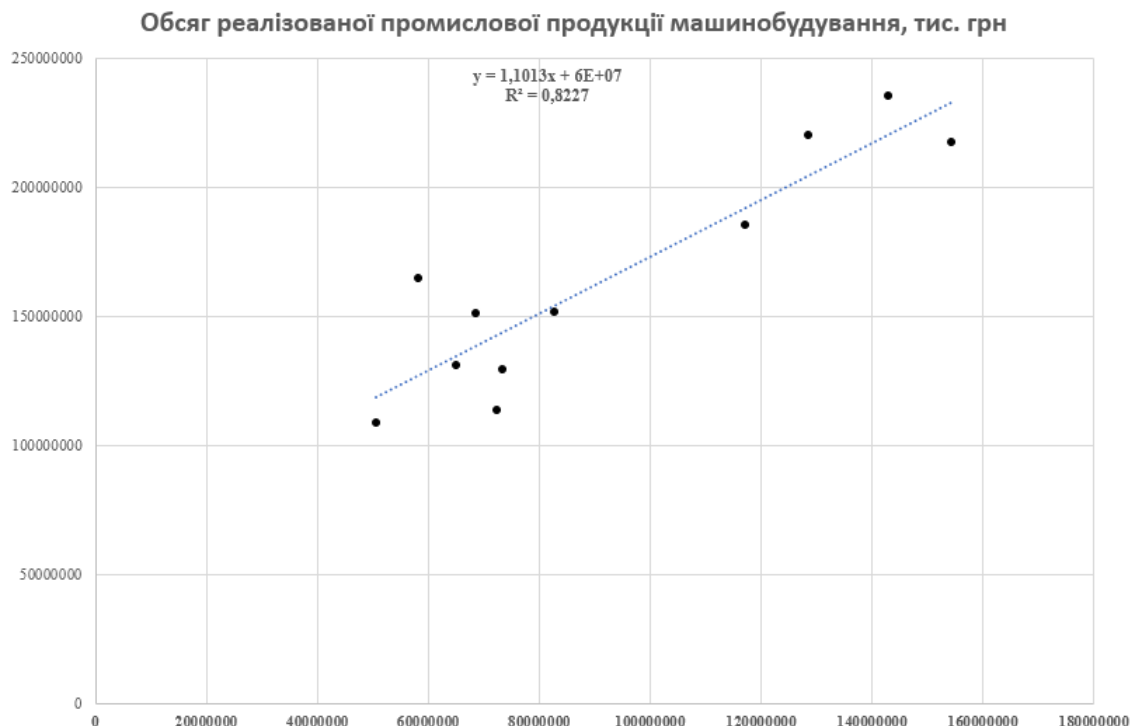


Рис. 4. Графік лінійної залежності обсягу реалізованої промислової продукції машинобудування від обсягу добування залізних руд в період 2010 – 2020 роки

Для вибору адекватної моделі скористаємося відомим методом оцінювання параметрів моделі – методом найменших квадратів (1). Результати порівняння сум квадратів відхилень теоретичних значень обсягу реалізованої промислової продукції машинобудування від емпіричних значень, обрахованих за побудованими моделями, показали, що найменші відхилення у поліноміальної функції третього степеня.

Емпірична функція регресії має вигляд:

$$y = -10^{17}x^3 + 2 \cdot 10^{-9}x^2 + 1,0508x + 6 \cdot 10^7. \quad (2)$$

Теоретична функція регресії має вигляд:

$$y = -10^{17}x^3 + 2 \cdot 10^{-9}x^2 + 1,0508x + 6 \cdot 10^7 + u. \quad (3)$$

де u – випадкова складова.

Причини наявності випадкової складової u в регресійній моделі (3) обумовлені введенням в модель не всіх пояснюючих змінних. Для дослідження було розглянуто залежність обсягу реалізованої промислової продукції машинобудування лише від обсягу добування залізних руд з метою спрощення реальної ситуації, та вважаючи обсяг добування залізних руд – одним із домінуючих факторів впливу. Реальна ситуація є складною комбінацією різних факторів, що призводить до відхилення реальних значень обсягу реалізованої промислової продукції машинобудування від його змодельованих значень. Також похибки збору статистичних даних впливають на розбіжності між модельними та емпіричними даними, що позначається на величині випадкової складової. Обмеженість зібраних статистичних даних з 2010 по 2022 роки також відображена у випадковій складовій u . Поява випадкової похибки у моделях часто обумовлена людським фактором, який є найменш прогнозованим елементом. Незважаючи на коректний вибір специфікації моделі та скрупульозний відбір змінних, непередбачуваність поведінки індивідів залишається причиною, що може суттєво знизити якість навіть найбільш досконалої моделі.

3. Для знайденої функції регресії (2) достовірність апроксимації [8] (коефіцієнт детермінації) $R^2 = 0,8556$ ($0,8 < R^2 < 1$), тому можна вважати, що побудована функція регресії на 85,56% відповідає вхідним статистичним даним обсягу реалізованої промислової продукції машинобудування.

Обчислимо індекс кореляції R за формулою

$$R = \sqrt{R^2}. \quad (4)$$

$$R = \sqrt{0,8556} = 0,9.$$

Оскільки величини R^2 і R наближаються до одиниці, то для побудованої функції регресії (3) це свідчить про її достовірність.

Обчислимо емпіричні значення $\hat{y}_i$, використавши функцію регресії (2).

Обчислимо залишки функції регресії u_i за формулою

$$u_i = y_i - \hat{y}_i. \quad (5)$$

Відносні похибки залишків функції регресії δ_i обчислимо за формулою

$$\delta_i = \frac{u_i}{y_i} \cdot 100\%. \quad (6)$$

Обчислимо середнє значення відносної похибки залишків $\bar{\delta}$ за формулою

$$\bar{\delta} = \frac{1}{n} \cdot \sum_{i=1}^n \delta_i. \quad (7)$$

$$\bar{\delta} = \frac{1}{13} \cdot \sum_{i=1}^n \delta_i = \frac{1}{13} \cdot (-0,1\%) = -0,01\%.$$

Оскільки $\bar{\delta} = -0,01\%$ значно менше 10%, то функцію регресії можна вважати достатньо точною.

4. Обчислимо точковий прогноз y_{np} обсягу реалізованої промислової продукції машинобудування на 2026 рік за побудованою моделлю.

Знайдемо прогнозоване x_{np} для обсягу добування залізних руд з використанням функції FORECAST.ET:

$$x_{np} = 141431921,2 \text{ (тис. грн)}.$$

Для заданого значення $x_{np} = 141431921,2$ (тис. грн) на 2026 рік знайдемо точковий прогноз побудованої емпіричної функції регресії (2):

$$\hat{y}_{np} = -10^{17} x_{np}^3 + 2 \cdot 10^{-9} x_{np}^2 + 1,0508 x_{np} + 6 \cdot 10^7.$$

Отже, значення $\hat{y}_{np}$ обсягу реалізованої промислової продукції машинобудування на 2026 рік за умови обсягу добування залізних руд $x_{np} = 141431921,2$ (тис. грн) згідно одержаної функції регресії (2):

$$\hat{y}_{np} = 220332028,8 \text{ (тис. грн)}.$$

Таким чином, з ймовірністю 95% можна стверджувати, обсяг реалізованої промислової продукції машинобудування на 2026 рік за умови обсягу добування залізних руд $x_{np} = 141431921,2$ (тис. грн) становитиме $\hat{y}_{np} = 220332028,8$ (тис. грн).

Висновки. У даному дослідженні, спираючись на статистичні дані в період з 2010 року по 2024 роки, за допомогою табличного процесора Microsoft Excel побудовано функцію залежності обсягу реалізованої промислової продукції машинобудування від обсягу добування залізних руд. Здійснено регресійний аналіз, в результаті якого підтверджено достовірність отриманої моделі. Статистичні висновки, зроблені у рамках дослідження, мають рівень довіри 95%, що свідчить про надійність одержаних результатів.

Побудована функція регресії є статистично значущою та може бути використана для прогнозування майбутнього обсягу реалізованої промислової продукції машинобудування від обсягу добування залізних руд на 2026 рік. Згідно з розрахунками, при прогнозованому обсязі добування залізних руд 141431921,2 тис. грн у 2026 році, очікуваний обсяг реалізованої промислової продукції машинобудування становитиме приблизно 220332028,8 тис. грн.

Результати дослідження мають важливе практичне значення. Цінність роботи полягає у можливості інтеграції розробленої математичної моделі в систему підтримки прийняття рішень на рівні машинобудівного підприємства для оптимізації виробничого планування, формування реалістичних фінансових прогнозів та для прийняття економічно обґрунтованих управлінських рішень.

У подальших дослідженнях доцільно розширити модель, врахувавши додаткові фактори, які можуть впливати на обсяг реалізованої промислової продукції машинобудування, зокрема: попит споживачів, обсяг інвестицій, платоспроможність та державні замовлення. Також перспективним є використання більш складних моделей та машинного навчання для підвищення точності прогнозування.

Список використаних джерел:

1. Самофалова М. Еволюція розвитку машинобудування в Україні. *Grail of Science*, 2024. № 46, С. 368–376. DOI: <https://doi.org/10.36074/grail-of-science.29.11.2024.042>
2. Державська А. В., Круш П. В. Сучасний стан галузі машинобудування та тенденції її розвитку. *Науковий вісник Ужгородського національного університету*, 2018. Випуск 21. Частина 1. С. 69–73.

3. Соколова Л. В., Стойка О. В. Сучасний стан машинобудування України та тенденції його розвитку за умов незбалансованої економіки. *Ефективна економіка*. 2019. № 11. DOI: <https://doi.org/10.32702/2307-2105-2019.11.5>.

4. Сакун Л. М., Веденіна Ю. Ю., Сухомлин Л. В., Цимбал О. С. Формування стратегії антикризового управління на підприємствах машинобудування в умовах змін. *Вісник Хмельницького національного університету*, 2021. № 5. Том 2. С. 35–40. DOI: [https://doi.org/10.31891/2307-5740-2021-298-5\(2\)-5](https://doi.org/10.31891/2307-5740-2021-298-5(2)-5).

5. Ходирєва О. Сучасний стан та проблеми розвитку машинобудівних підприємств України. *Економічний аналіз*, 2021. Том 31. № 1. С. 227–238. DOI: <https://doi.org/10.35774/econa2021.01.227>.

6. Tselikova A., Misko Y. The current state and problems of the development of machine-building enterprises in Ukraine. *Ukrainian Journal of Applied Economics and Technology*, 2024. Volume 9. № 2, pp. 228–235. DOI: <https://doi.org/10.36887/2415-8453-2024-2-39>

7. Карачина Н. П. Трансформація моделей економічної поведінки машинобудівних підприємств за умов розвитку у воєнний час. *Реальний сектор економіки України в умовах невизначеності: Матеріали Всеукраїнського симпозіуму (04 квітня 2024 року), Київ, Університет «Україна»*, 2024. С. 55–58.

8. Майбородіна Н. В. Економетрика. Ніжин: ПП Лисенко М.М., 2021. 280 с.

9. Державна служба статистики України. URL: <https://www.ukrstat.gov.ua> (дата звернення: 25.10.2025).

10. Лебідь В. В., Мейш Ю. А., Майбородіна Н. В., Герасименко В. П. Застосування математичного моделювання до побудови лінійної багатфакторної моделі перевезень вантажів у міжнародному сполученні. *Вісник Національного транспортного університету. Серія «Технічні науки»*. 2022. № 51. С. 260–266. DOI: <https://doi.org/10.33744/2308-6645-2022-1-51-260-266>.

References:

1. Samofalova, M. (2024). Evoliutsiia rozvytku mashynobuduvannia v Ukraini [The Evolution of Mechanical Engineering Development in Ukraine]. *Grail of Science* (46), 368–376. DOI: <https://doi.org/10.36074/grail-of-science.29.11.2024.042>

2. Derzhavska, A. V., Krush, P. V. (2018). Suchasnyi stan haluzi mashynobuduvannia ta tendentsii yii rozvytku [The current state of the machine-building industry and trends in its development]. *Naukovyi visnyk Uzhhorodskoho natsionalnoho universytetu – Scientific Bulletin of Uzhhorod National University*, (21). 69–73.

3. Sokolova, L. V., Stoika, O. V. (2019). Suchasnyi stan mashynobuduvannia Ukrainy ta tendentsii yoho rozvytku za umov nezbalansovanoi ekonomiky [The current state of machine-building in Ukraine and trends in its development under an unbalanced economy]. *Efektivna ekonomika – Efficient Economy*, (11). DOI: <https://doi.org/10.32702/2307-2105-2019.11.5>.

4. Sakun, L. M., Viedienina, Yu. Iu., Cukhomlyn, L. V., Tsymbal, O. S. (2021). Formuvannia stratehii antykrizovoho upravlinnia na pidpriemstvakh mashynobuduvannia v umovakh zmin [Formation of an anti-crisis management strategy at machine-building enterprises under changing conditions]. *Visnyk Khmelnytskoho natsionalnoho universytetu – Bulletin of Khmelnytskyi National University*, (5). 35–40. DOI: [https://doi.org/10.31891/2307-5740-2021-298-5\(2\)-5](https://doi.org/10.31891/2307-5740-2021-298-5(2)-5).

5. Khodyrieva, O. (2021). Suchasnyi stan ta problemy rozvytku mashynobudivnykh pidpriemstv Ukrainy [The current state and problems of the development of machine-building enterprises in Ukraine]. *Ekonomichniy analiz – Economic Analysis*, (31). 227–238. DOI: <https://doi.org/10.35774/econa2021.01.227>.

6. Tselikova, A., Misko, Y. (2024). The current state and problems of the development of machine-building enterprises in Ukraine. *Ukrainian Journal of Applied Economics and Technology*, (9). 228–235. DOI: <https://doi.org/10.36887/2415-8453-2024-2-39>

7. Karachyna, N. P. (2024). Transformatsiia modelei ekonomichnoi povedinky mashynobudivnykh pidpriemstv za umov rozvytku u voiennyi chas [Transformation of economic behavior models of machine-building enterprises under wartime development conditions]. *Realnyi sektor ekonomiky Ukrainy v umovakh nevyznachenosti: Materialy Vseukrainskoho sympoziumu (04 kvitnia 2024 roku) Kyiv, Universytet «Ukraina» – The Real Sector of Ukraine's Economy under Uncertainty: Proceedings of the All-Ukrainian Symposium (Kyiv, University "Ukraine," April 4, 2024)*, 55–58.

8. Maiborodina, N. V. (2021). Ekonometryka [Econometrics]: Nizhyn: PP Lysenko M.M.

9. Derzhavna sluzhba statystyky Ukrainy [State Statistics Service of Ukraine]. Retrieved from: <https://www.ukrstat.gov.ua>.

10. Lebid, V. V., Meish, Yu. A., Maiborodina, N. V., Herasymenko, V. P. (2022). Zastosuvannia matematychnoho modeliuvannia do pobudovy liniinoi bahatofaktornoї modeli perevezen vantazhiv u mizhnarodnomu spoluchenni [Application of mathematical modeling to the construction of a linear multifactor model of cargo transportation in international communication]. *Visnyk Natsionalnoho transportnoho universytetu. Seriiia «Tekhnichni nauky» – Bulletin of the National Transport University. Series "Technical Sciences"*, (51). 260–266. DOI: <https://doi.org/10.33744/2308-6645-2022-1-51-260-266>.

Дата надходження статті: 27.10.2025

Дата прийняття статті: 17.11.2025

Опубліковано: 30.12.2025

Перкун І. В., кандидат технічних наук, доцент,
докторант Івано-Франківського національного
технічного університету нафти і газу
ORCID: 0000-0002-6523-1606

ГІДРОІМПУЛЬСНА ПЕРФОРАЦІЯ НАФТОГАЗОВИХ СВЕРДЛОВИН СТРУМЕНЕМ РОЗЧИНУ ПОЛІМЕРУ

Аналізується існуючий метод інтенсифікації – гідропіскоструминна перфорація. Попри свою ефективність у створенні надійного гідродинамічного зв'язку між пластом і свердловиною, цей метод має суттєві недоліки. Головними з них є інтенсивне гідроабразивне зношування насосно-компресорного обладнання та арматури, що призводить до погіршення якості перфораційних каналів через ерозію сопел, а також висока ресурсо- та енергоємність процесу. Як перспективну та більш досконалу альтернативу пропонується використання високошвидкісного струменя розчину полімеру (поліетиленоксиду). Ключова інноваційна ідея роботи полягає у застосуванні не стаціонарного (безперервного), а імпульсного режиму подачі водополімерного струменя. Головна гіпотеза дослідження стверджує, що такий підхід, завдяки унікальним реофізичним властивостям полімерів, дозволяє суттєво підвищити руйнівну здатність струменя. Теоретичне обґрунтування та серія експериментів повністю підтвердили цю гіпотезу. Дослідження поведінки розчинів полімерів в умовах коливальної течії, що моделює роботу перфоратора, виявили існування критичної частоти пульсацій. Встановлено, що максимальна ефективність руйнування досягається, коли частота імпульсів відповідає максимуму дисипативної функції. У цьому режимі найбільша частина енергії потоку безповоротно розсіюється і перетворюється на роботу руйнування. Експерименти з перфорації моделі перешкоди (сталь, бетон, гірська порода) наочно продемонстрували, що імпульсний струмінь створює значно глибші канали, ніж стаціонарний, за однакових вихідних тисків. Базуючись на отриманих теоретичних і експериментальних результатах розроблено конструкцію нового пристрою – гідроімпульсного перфоратора. Його новизна полягає у використанні обертового ротора з V-подібними пазами, який приводиться в рух вбудованою гідравлічною турбіною. Обертаючись, ротор механічно перериває потік робочої рідини, що надходить до соплових насадок, генеруючи потужні гідродинамічні імпульси. Важливою особливістю конструкції є можливість осьового зміщення ротора, що дозволяє точно налаштовувати форму, тривалість та частоту імпульсів для досягнення максимального руйнівного ефекту залежно від властивостей породи, що руйнується. Таким чином, робота доводить, що перехід до імпульсного режиму водополімерної перфорації є науково обґрунтованим та високоефективним. Запропонований гідроімпульсний перфоратор є інноваційним технічним рішенням, що відповідає критерію «винахідницький рівень» і відкриває шлях до створення енергоощадної та більш продуктивної технології розкриття нафтогазових пластів.

Ключові слова: гідроімпульсний перфоратор, розчин полімеру, сопло-насадка, обсадна колона, цементне кільце, порода, дисипативна функція, в'язкопружність.

Perkun I. V. Hydro-pulse perforation of oil and gas wells with a polymer solution jet

This paper analyzes the existing intensification method: hydro-abrasive jetting (or hydro-sandblasting perforation). Despite its effectiveness in establishing a reliable hydrodynamic connection between the formation and the wellbore, this method has significant drawbacks. Chief among these are the intense hydro-abrasive wear of pumping equipment and fittings, leading to a decline in the quality of perforation channels due to nozzle erosion, as well as the high resource and energy intensity of the process. As a promising and more advanced alternative, the authors propose the use of a high-velocity polymer solution jet (polyethylene oxide). The key innovative idea of the work is the application of a pulsed, rather than a steady (continuous), mode for delivering the water-polymer jet. The main hypothesis of the research posits that this approach, owing to the unique rheophysical properties of polymers, can significantly enhance the destructive capability of the jet. Theoretical justification and a series of experiments have fully confirmed this hypothesis. Studies of the behavior of polymer solutions under oscillatory flow conditions, which simulate the perforator's operation, revealed the existence of a critical pulsation frequency. It was established that maximum destructive efficiency is achieved when the pulse frequency corresponds to the maximum of the dissipative function. In this mode, the largest portion of the flow's energy is irreversibly dissipated and converted into destructive work. Experiments involving the perforation of a target model (steel, concrete, and rock formation) clearly demonstrated that the pulsed jet creates significantly deeper channels than a steady jet at identical initial pressures. Based on the results obtained, a conceptual solution and design for a new device—the hydro-pulse perforator—were developed. Its novelty lies in the use of a rotating rotor with V-shaped grooves, driven by an integrated hydraulic turbine. As it rotates, the rotor mechanically interrupts the flow of the working fluid to the nozzles, thereby generating powerful hydrodynamic pulses. An important design feature is the capability for

axial displacement of the rotor, which allows for the precise adjustment of the pulse shape, duration, and frequency to achieve maximum destructive effect depending on the properties of the rock formation. Thus, the work demonstrates that the transition to a pulsed mode for water-polymer perforation is scientifically justified and highly effective. The proposed hydro-pulse perforator represents an innovative technical solution that meets the 'inventive step' criterion and paves the way for creating an energy-efficient and more productive technology for developing oil and gas formations.

Key words: *hydro-pulse perforator, polymer solution, nozzle, casing string, cement ring, rock, dissipative function, viscoelasticity.*

Постановка проблеми. У процесі тривалої експлуатації свердловин відбувається кольматація привибійної зони пласту твердими частинками – продуктами руйнування гірських порід, що транспортуються потоками вуглеводнів. Крім того, у цій зоні можуть накопичуватися компоненти бурового та цементного розчинів, пластова вода, вуглеводневий конденсат, солі та технологічні рідини, які використовуються під час спорудження та ремонту свердловин. Це призводить до утворення додаткових фільтраційних опорів, що, своєю чергою, зменшує дебіт свердловин.

До ефективних методів інтенсифікації припливу флюїдів до вибою свердловин належить гідропіско-струминна перфорація. Цей метод дозволяє знизити фільтраційні втрати тиску у привибійній зоні, активізувати неефективно працюючі перфоровані інтервали пласта та залучити до розробки малопотужні й низькопроникні нафто- й газонасичені прошарки, які раніше вважалися нерентабельними [1, 2].

Гідропіскоструминна перфорація забезпечує надійний гідродинамічний зв'язок між продуктивним пластом і стовбуром свердловини, демонструючи істотні переваги відносно інших методів, зокрема вибуховою перфорацією. Проте ця технологія має і суттєві недоліки, серед яких – інтенсивне зношування насосно-компресорного обладнання та арматури під дією піску. Збільшення діаметра сопел і погіршення профілю їх отворів призводить до зниження швидкості гідроабразивного струменя, що погіршує якість формування перфораційних каналів в обсадних колонах, цементному кільці й гірській породі, а також збільшує тривалість процесу. Крім того, глибоке проникнення абразивних частинок у породу знижує ефективність віддачі пласта. Усі ці чинники зумовлюють високу ресурсо- та енергоємність процесу [2].

Альтернативним рішенням є використання високошвидкісного струменя водного розчину поліетеленоксида (ПЕО) замість гідроабразивної суміші. Експериментальні дані свідчать, що малі добавки полімерів здатні істотно підвищити руйнівну здатність водного струменя [3–6]. Втім, сучасні уявлення про поведінку макромолекул у збіжній течії – характерний для вхідних ділянок струменеформуючих сопел гідроперфоратора – залишаються обмеженими.

Становлення теоретичних засад реофізики водополімерної перфорації є необхідною умовою для створення енергоощадної та екологічно безпечної технології. Наразі спостерігається перехід від накопичення експериментальної інформації до розуміння реофізичної сутності явищ, що виникають при збіжній течії розчинів полімерів, і встановлення закономірностей прояву гумоподібної високоеластичності.

Дослідження, присвячені реофізиці перфорації полімерним струменем, а також розробці відповідного обладнання, є інноваційними й актуальними. *Концептуальна гіпотеза роботи* ґрунтується на здатності розчинів полімерів змінювати свої властивості під впливом деформаційної дії потоку. Це відкриває перспективи для створення технології гідроструминної водополімерної перфорації з високою енергоефективністю та мінімальним негативним екологічним впливом.

Аналіз експериментальних даних [5–8] показує, що вдосконалення процесу водополімерної перфорації нафтогазових свердловин пов'язано, в першу чергу, з формуванням гідродинамічних властивостей струменя за умов енергозбереження. Це забезпечує максимальну продуктивність і збільшує ресурс перфораційного обладнання. Підвищення якості водополімерного струменя дає змогу мінімізувати енергетичні витрати за рахунок зниження значень робочого тиску перед насадками при задоволенні технологічних вимог до процесу перфорації.

Теорія деформаційних ефектів при течії розчинів полімерів [9] дозволяє сподіватися, що руйнівна здатність водополімерного струменя буде значно більша при імпульсному протіканні полімерного розчину крізь соплові насадки гідроперфоратора.

Метою статті є: обґрунтувати припущення, що висока руйнівна здатність водополімерного струменя досягається при імпульсній його течії;

– вивчення особливості імпульсної течії (з розтягом) розчину полімеру в модельних умовах соплової насадки гідроперфоратора;

– проведення теоретичного аналізу особливостей імпульсної течії розчинів полімерів для використання отриманих теоретико-експериментальних результатів при розробці гідроімпульсного перфораційного обладнання;

– розробка концептуального рішення щодо створення гідроімпульсного перфоратора.

Методи і матеріали. Дослідження особливостей осцилюючої течії водних розчинів полімеру проводили з використанням гідродинамічної установки, схема якої зображена на рис. 1. Модельний режим формування осцилюючого водополімерного струменя реалізовувався при течії водного розчину ПЕО крізь точну камеру під дією постійного перепаду тиску і впливу на течію електромагнітного вібратора.

Ефективність досліджуваних полімерних розчинів оцінювали за величиною їх гідродинамічного опору η_{ef} . Цей параметр визначали шляхом порівняння об'ємної витрати розчину полімеру та калібрувальної ньютонівської рідини за однакового перепаду тиску.

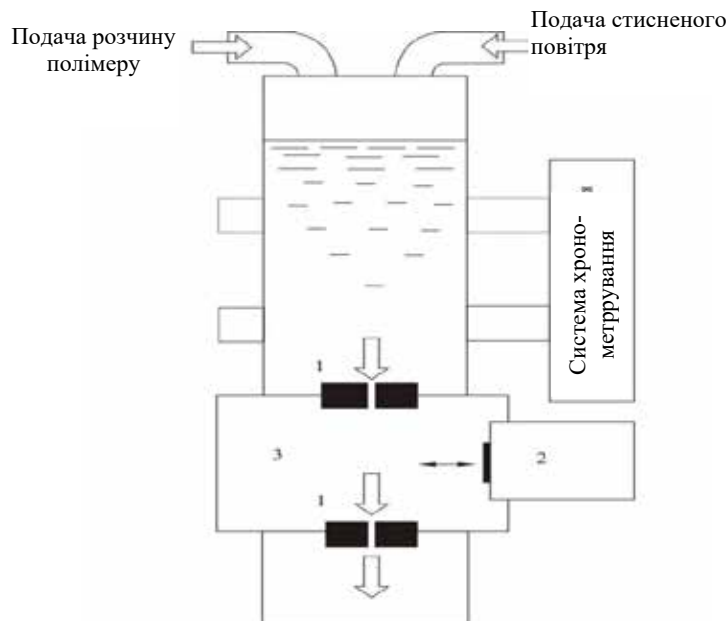


Рис. 1. Схема гідродинамічної установки

Примітка: 1 – пластини з отворами, 2 – електромагнітний вібратор, 3 – зона впливу вібратора

Дослідження впливу частоти переривання струменя водного розчину ПЕО в імпульсному режимі на довжину формованого каналу проводили на гідрорізальній установці [9]. Установа була оснащена електромеханічним клапаном для періодичного переривання потоку. Модель перешкоди, що імітувала обсадну колону, цементне кільце та керн, складалася зі сталевго листа (товщина 10 мм), шару бетону (20 мм) і гірської породи (770 мм) [5].

Робочі водні розчини ПЕО (мол. маси $6 \cdot 10^6$) готували шляхом розведення вихідного концентрованого розчину концентрації 0,1 % [10].

Результати і їх обговорення. Результати експериментального дослідження наведено на рисунку 2,а, де – η_{ef} позначає ефективний гідродинамічний опір за умов стаціонарної течії розчину полімеру, а $\Delta\eta_{\text{ef}}$ – додатковий внесок у опір, що виникає внаслідок імпульсного збурення основного потоку [11].

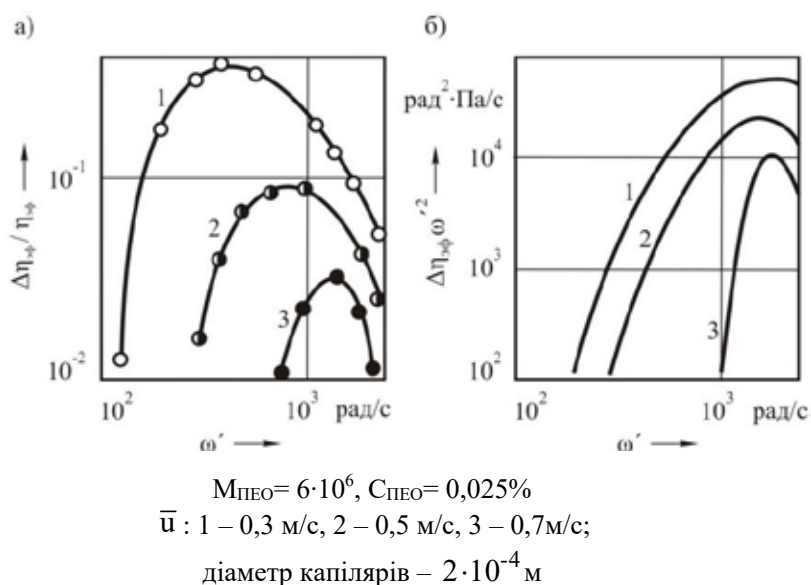


Рис. 2. Вплив частоти зовнішнього поля на гідродинамічний опір (а) і дисипативну функцію (б) при течії водополімерного розчину для різних середніх швидкостей протікання

Аналіз отриманих залежностей показує існування критичної частоти дії $\omega'_{кр}$ на течію водного розчину ПЕО крізь отвори в пластинах. При перевищенні цієї частоти потік характеризується значним зростанням дисипації енергії. Встановлено, що зі збільшенням середньої швидкості потоку значення критичної частоти $\omega'_{кр}$ зростає, тоді як відносний додатковий опір $\Delta\eta_{ef}$ знижується. Окрім того, за умов сталої швидкості потоку, зменшення концентрації полімеру в водному розчині ПЕО призводить до зміщення критичної частоти в область вищих значень.

Відомо, що зміни величини деформації ε і напруження σ при деформації розчинів гнучколанцюгових полімерів відбуваються на різних фазах [12, 13]. Кут δ' характеризує цю різницю і є функцією частоти дії ω' вібратора на потік розчину полімеру. Деформація в умовах осцилюючої течії водополімерного розчину виникає під дією напруження, що змінюється за наступним законом:

$$\sigma(\tau) = \bar{\sigma} + \sigma_0 \exp(i \omega' \tau). \quad (1)$$

де $\bar{\sigma}$ – стаціонарне напруження; τ – час.

Враховуючи, що $\sigma(\tau) = \sigma_0 \cos \omega' \tau$, $\varepsilon(\tau) = \varepsilon_0 \cos(\omega' \tau - \delta')$ і виключаючи з цих рівнянь час, одержимо вираз:

$$\left(\frac{\sigma}{\sigma_0}\right)^2 + \left(\frac{\varepsilon}{\varepsilon_0}\right)^2 = \varepsilon^2 \delta' + \left(\frac{\sigma}{\sigma_0}\right) \cdot \left(\frac{\varepsilon}{\varepsilon_0}\right) \cos \delta'. \quad (2)$$

Вираз (2) описує площу еліпса, яка дорівнює роботі A_0 , що за період гармонійних коливань дисипується при деформації. Тоді функція дисипації W обчислюється як A_0 на частоту осциляції:

$$W = A_0 \frac{\omega'}{2\pi} = \frac{\varepsilon_0 \sigma_0 \omega'}{2} \sin \delta'. \quad (3)$$

З (3) видно, що дисипативна функція залежить від молекулярних властивостей полімеру, що враховується параметром δ' . Якщо напруження змінюється за гармонійним законом, то легко отримати вираз для зміни швидкості деформації $\dot{\varepsilon}$. Остання пов'язана з деформацією таким чином:

$$\dot{\varepsilon} = \frac{d\varepsilon}{d\tau} = \varepsilon_0 i \omega' \exp[i(\omega' \tau - \delta')] = i \omega' \varepsilon. \quad (4)$$

Величина відношення $\frac{\sigma}{\dot{\varepsilon}}$ є комплексною динамічною в'язкістю η^* . Її можна подати, як різницю дійсної та уявної компонент в'язкості [14–16]

$$\eta^* = \eta' - i\eta'', \quad (5)$$

$$\text{де } \eta' = \left(\frac{\sigma_0}{\varepsilon_0 \omega'}\right) \sin \delta', \quad \eta'' = \left(\frac{\sigma_0}{\varepsilon_0 \omega'}\right) \cos \delta'.$$

Введемо у вираз (3) η' і зробивши перетворення, отримаємо

$$W = \eta' \frac{\varepsilon_0^2 \omega'^2}{2}. \quad (6)$$

Динамічна в'язкість розчину полімеру проявляється при його деформації, і приріст гідродинамічного опору $\Delta\eta_{ef}$ в умовах осцилюючого впливу на основну його течію можна віднести до появи η' . Оскільки амплітудне значення величини деформації μ_0 залишалося в експериментах постійним, можна скористатися скейлінгом і переписати (6) у вигляді:

$$W \approx \Delta\eta_{ef} \cdot \omega'^2. \quad (7)$$

Враховуючи експериментальні результати, що представлені на рис. 2,а були отримані з використанням рівняння (7) залежності для дисипативної функції, що наведені на рис. 2,б. Аналіз результатів, наведених на рис. 2,б свідчить, що дисипація енергії при осцилюючій течії зменшується зі зростанням середньої швидкості течії розчину полімеру. Цей результат корелює зі збільшенням частки розгорнутих макромолекул в потоці розчину ПЕО [17] і, як наслідок, зі зростанням наведеної жорсткості полімерних молекул [18].

Відповідно, для систем із більш жорсткими та розгорнутими ланцюгами частоти, за якими режим течії стає високодисипативним, зміщуються в область вищих значень. Таким чином, можна сподіватися, що осцилюючий режим формування водополімерного струменя буде забезпечувати вищу руйнівну здатність порівняно зі стаціонарним струменем.

Для верифікації цього припущення було проведено експериментальне дослідження впливу частоти осцилюючого формування водополімерного струменя на його проникаючу здатність. Досліди виконували на гідрорізальній установці [9], вимірюючи довжину каналу, що утворюється в модельній перешкоді струменем водного розчину ПЕО, генерованим в осцилюючому режимі. Результати, наведені в Таблиці 1,

підтверджують, що реофізичні ефекти, які зумовлюють високодисипативний режим течії, відіграють ключову роль у підвищенні ефективності гідроструминної перфорації нафтогазових свердловин.

Експериментально встановлено, що ефективність процесу гідроструминної водополімерної перфорації нафтогазових свердловин значно підвищується за умов формування струменя в імпульсному режимі, який забезпечує прояв пружних деформацій. При цьому оптимальною є частота пульсаційного впливу, що відповідає максимуму дисипативної функції $W \approx \Delta\eta_{ef} \cdot \omega'^2$.

Таблиця 1

Довжина каналу сформованого стаціонарним та імпульсним водополімерним струменями в моделі перешкоди ($M_{\text{ПЕО}} = 6 \cdot 10^6$, $C_{\text{ПЕО}} = 0,005\%$, $d_{\text{сон}} = 0,35 \cdot 10^{-3}$ м, час перфорації – 10 хв, $l_0 = l_{\text{онт}}$)

Довжина сформованого каналу $h \cdot 10^3$, м		
ΔP_0 , МПа	Стаціонарний струмінь водного розчину ПЕО	Осцилюючий водополімерний струмінь
100	380	445
200	405	474
300	420	491

Таким чином, проведені експерименти не лише підтвердили, але й всебічно обґрунтували фундаментальну гіпотезу про значне підвищення руйнівної здатності водополімерного струменя за умов його імпульсного (осцилюючого) формування. Було доведено, що цей ефект безпосередньо пов'язаний з проявом унікальних в'язкопружних властивостей полімерного розчину в потоці, що дозволяє перетворювати енергію потоку на роботу руйнування з максимальною ефективністю.

Ці фундаментальні результати стали науковою основою для переходу від теорії до практики – розробки вдосконаленого гідроперфоратора, конструкція якого цілеспрямовано використовує відкриті реофізичні явища [7, 9] для підвищення ефективності процесу перфорації нафтогазових свердловин. Запропонований гідроімпульсний перфоратор є інноваційним технічним рішенням, що включає порожнистий корпус із сопловими насадками та ключовий елемент – ротор, встановлений усередині з можливістю обертання. Для забезпечення автономної та стабільної роботи привід обертання ротора реалізовано у вигляді вбудованої багатоступінчастої гідравлічної турбіни, що працює за рахунок енергії самого потоку рідини [19].

Ключові конструктивні особливості та їх обґрунтування. Головна перевага розробленого пристрою полягає в унікальних конструктивних особливостях, що дозволяють не просто створювати пульсації, а й гнучко керувати їх параметрами:

- Ротор виконано у формі порожнистого циліндра, на зовнішній поверхні якого нанесено спеціальні V-подібні пази. Така форма, по-перше, спрощує конструкцію та здешевлює виробництво. По-друге, геометричне розташування пазів (їхній нижній край знаходиться не вище верхнього краю соплових отворів) забезпечує повне періодичне перекриття потоку рідини, що генерує потужні гідродинамічні імпульси.

- Механізм регулювання імпульсів: Перфоратор оснащено прецизійним механізмом для осьового зміщення та надійної фіксації ротора відносно сопел. Цей механізм реалізовано у вигляді штока, що проходить крізь осьовий отвір приводного вала турбіни. На нижньому кінці штока закріплений ротор, а на верхньому – встановлено цанговий затискач для його надійної фіксації.

Функціональні переваги. Таке поєднання конструктивних елементів надає пристрою виняткові експлуатаційні можливості. Наявність V-подібних пазів у поєднанні з можливістю осьового зміщення ротора дозволяє точно регулювати форму, тривалість та частоту гідродинамічних імпульсів. Ця функція є критично важливою, оскільки дає змогу налаштовувати режим роботи перфоратора для досягнення максимального руйнівного ефекту залежно від властивостей породи та реофізичних характеристик полімерного розчину.

Надійна фіксація положення штока за допомогою цангового затискача гарантує стабільність заданих параметрів протягом усього процесу перфорації. При цьому конструкція залишається технологічною у виготовленні та простою в експлуатації. Розміщення затискного механізму на верхньому кінці штока забезпечує зручність та швидкість налаштування пристрою безпосередньо перед його спуском у свердловину.

Таким чином, завдяки цим продуманим конструктивним удосконаленням було створено пристрій, що дозволяє цілеспрямовано генерувати гідродинамічні імпульси з оптимальними, заздалегідь визначеними характеристиками. Заявлена сукупність ознак є інноваційною, відповідає критерію «винахідницький рівень» та демонструє якісно новий технічний результат, відкриваючи шлях до створення більш продуктивної та енергоощадної технології розкриття нафтогазових пластів.

Висновки

1. На основі теоретичного аналізу та експериментальних досліджень обґрунтовано переваги імпульсного режиму формування водополімерного струменя для підвищення ефективності перфорації нафтогазових свердловин. Визначено ключовий параметр оптимізації процесу: частота пульсацій має відповідати максимуму дисипативної функції.

2. Створено нову конструкцію гідроімпульсного перфоратора, що дозволяє ефективніше здійснювати водополімерну перфорацію нафтогазових свердловин. Наукова новизна та практична значущість розробки полягає у досягненні вищих показників процесу завдяки запропонованим конструктивним удосконаленням.

Напрями подальших досліджень. Проведена робота відкриває значні перспективи для подальшого вдосконалення технології водополімерної гідроімпульсної перфорації. Подальші наукові та інженерні пошуки доцільно зосередити на кількох ключових напрямках, що дозволять не лише оптимізувати існуючу розробку, але й розширити сферу її застосування.

По-перше, пріоритетним завданням є розробка математичної моделі для прогнозування оптимальної частоти імпульсів залежно від конкретних геолого-фізичних умов: типу гірської породи, її пористості, проникності та механічних властивостей, а також пластового тиску і температури. Створення такої моделі дозволить ще на етапі планування робіт обирати оптимальні параметри робочої рідини (тип полімеру, його концентрацію та молекулярну масу) і режими роботи перфоратора, що значно підвищить ефективність та передбачуваність результатів.

По-друге, перспективним є дослідження нових полімерних композицій та добавок до робочої рідини. Вивчення впливу наночастинок, поверхнево-активних речовин або сумішей різних полімерів може призвести до створення рідин з унікальними реофізичними властивостями. Це дозволить посилити руйнівну здатність струменя, підвищити його стабільність на великих відстанях від сопла та зменшити екологічне навантаження шляхом використання біорозкладних компонентів.

По-третє, важливим етапом є створення "інтелектуальної" системи керування перфоратором. Це передбачає оснащення пристрою датчиками зворотного зв'язку (наприклад, тиску, вібрації, акустичними сенсорами), які б у режимі реального часу відстежували процес перфорації нафтогазових свердловин. Аналіз цих даних за допомогою бортового мікропроцесора дозволить би автоматично коригувати частоту та форму імпульсів шляхом регулювання положення ротора, адаптуючи роботу перфоратора до змін у властивостях пласта. Це виведе технологію на новий рівень автоматизації та ефективності.

Нарешті, необхідно провести повномасштабні промислові випробування розробленого гідроімпульсного перфоратора на реальних свердловинах. Це дозволить не тільки підтвердити отримані лабораторні результати, але й оцінити довговічність та надійність конструкції, відпрацювати технологічні регламенти його застосування та зібрати дані для подальшої оптимізації.

Комплексна реалізація цих напрямів досліджень сприятиме створенню високоефективної, енергоощадної та екологічно безпечної технології інтенсифікації видобутку вуглеводнів, що має значний потенціал для широкого впровадження в нафтогазовій промисловості.

Список використаних джерел:

1. Чорний М. І., Чорний О. М., Метошоп І. М., Кузів І. М. *Геологічні основи розкриття і випробування продуктивних пластів*. Івано-Франківськ : ІФНТУНГ, 2013. 306 с.
2. Бойко В. С. *Підземний ремонт свердловин* : підручник : у 4 ч. Івано-Франківськ : ІФНТУНГ, 2009. Ч. 3. 809 с.
3. Pogrebnyak V. G., Naumchik N. V. Hydrodynamic activity of polymers in high-velocity flows. *Journal of Engineering Physics and Thermophysics*. 1995. Vol. 68, no. 1. P. 138–140.
4. Погребняк В. Г., Погребняк А. В. Технології високоефективного гідрорізання. *Актуальні напрями розвитку технічного та виробничого потенціалу національної економіки: монографія* за ред. В. О. Пінчук, Г. С. Прокудіна. Дніпро : Пороги, 2021. С. 84–99.
5. Pogrebnyak V. G., Chudyk I. I., Pogrebnyak A. V., Perkun I. V. Perforation of oil and gas wells by a high-velocity jet of polymer solution. *Nafta-Gaz*. 2022. No. 78(1). P. 3–12. DOI:10.18668/NG.2022.01.01
6. Pogrebnyak V. G., Chudyk I. I., Pogrebnyak A. V., Perkun I. V. High-efficiency Casing Perforation Oil and Gas Wells. *SOCAR Proceedings*. 2021. No.2. P. 112–120. DOI: 10.5510/OGP2021SI200578
7. Perkun I., Pogrebnyak V., Szymanskiy V. Perforation of casing strings in oil and gas wells with a polymer solution jet: scientific and practical aspects. *Trends in petroleum engineering*. 2025. Vol. 5, no. 1. P. 1–7. DOI: 10.53902/TPE.2025.05.00045.
8. Спосіб гідроструминної перфорації свердловин: пат. на корисну модель № 150245, Україна. МПК (2022.01) E21B43/00 E21B 43/14 (2006.01) / Погребняк В. Г., Перкун І. В., Погребняк А. В., Шиманський В. Я. № u 2021 04406, заявл. 29.07.2021, опубл. 19.01.2022. Бюл. № 3.
9. Погребняк В. Г., Волошин В. С. Екологічна технологія створення водозахисних екранів: монографія. Донецьк: Ноулідж, 2010. 482 с. URL: <https://scholar.google.com/citations?user=rwvMpTQAAAAJ&hl=en&oi=sra>
10. Pogrebnyak V. G., Naumchuk N. V., Maksyutenko S. N. Use of heat carriers with hydrodynamically active additives. *Izvestiya Vysshikh Uchebnykh Zavedenii, Energetika*. 1982. No. 5. P. 109–111.
11. Pogrebnyak V. G., Perkun I. V., Pogrebnyak A. V. Displacement of Oil from Porous Bed by the Oscillating Flow of Polymer Solution. *American Journal of Science, Engineering and Technology*. 2017. Vol. 2, no. 1. P. 53–57.
12. Ferry J. D. *Viscoelastic Properties of Polymers*. New York : John Wiley & Sons, 1961. 482 p.

-
13. Middleman S. *The Flow of High Polymers: Continuum and Molecular Rheology*. New York : Interscience Publishers, 1968. 245 p.
 14. Vinogradov G. V., Malkin A. Ya. *Rheology of Polymers*. M.: Mir Publishers; Berlin: Springer-Verlag, 1980. 468 p.
 15. Kleiman R. N. Analysis of the oscillating-cup viscometer for the measurement of viscoelastic properties. *Physical Review A*. 1987. Vol. 35, no. 1. P. 261–275.
 16. Ferry J. D. Relaxations in Polymer Solutions, Liquids, and Gels // *Physical Acoustics: Principles and Methods*. ed. by W. P. Mason. New York; London: Academic Press. 1965. Vol. II, part B. P. 1–115.
 17. Pogrebnyak A., Pogrebnyak V. The Elastic Deformations and Anomalous High Cutting Ability of a Polymer Solution Jet. *Scientific Journal of the Ternopil National Technical University*. 2017. No. 1 (85). P. 89–94.
 18. Elyashevich G. K. Thermodynamics and kinetics of orientational crystallization of flexible chain polymers. *Advances in Polymer Science*. 1982. Vol. 43. P. 205–245.
 19. Гідроімпульсний перфоратор : пат. 152383 Україна : МПІК Е21В 43/114 / Погребняк В. Г., Перкун І. В., Погребняк А. В., Шиманський В. Я. № u202202267 ; заявл. 30.06.2022 ; опубл. 18.01.2023, Бюл. № 3.
 20. Pogrebnyak V. G., Chudyk I. I. Solutions of Polymers in the Oil and Gas Technologies. *Prospects for developing resource-saving technologies in mineral mining and processing : multi-authored monograph*. Petrosani : UNIVERSITAS Publishing, 2022. P. 110–294. DOI: 10.31713/m1106.

References:

1. Chornyi, M. I., Chornyi, O. M., Metoshop, I. M., & Kuziv, I. M. (2013). *Heolohichni osnovy rozkryttia i vyprovuvannia produktyvnykh plastiv* [Geological bases of disclosure and testing of productive layers]. Ivano-Frankivsk: IFNTUNH.
2. Boiko, V. S. (2009). *Pidzemnyi remont sverdlovyh: pidruchnyk: u 4 ch.* [Underground well repair: textbook: in 4 parts] (Ch. 3). Ivano-Frankivsk: IFNTUNH.
3. Pogrebnyak, V. G., & Naumchik, N. V. (1995). Hydrodynamic activity of polymers in high-velocity flows. *Journal of Engineering Physics and Thermophysics*, 68(1), 138–140.
4. Pogrebnyak, V. H., & Pogrebnyak, A. V. (2021). Tekhnologii vysokoefektyvnoho hidrorizannia [Technologies of highly effective hydrocutting]. In V. O. Pinchuk & H. S. Prokudina (Eds.), *Aktualni napriamy rozvytku tekhnichnoho ta vyrobnychoho potentsialu natsionalnoi ekonomiky: monohrafiia* [Current directions of development of the technical and production potential of the national economy: monograph] (pp. 84–99). Dnipro: Porohy.
5. Pogrebnyak, V. G., Chudyk, I. I., Pogrebnyak, A. V., & Perkun, I. V. (2022). Perforation of oil and gas wells by a high-velocity jet of polymer solution. *Nafta-Gaz*, 78(1), 3–12. <https://doi.org/10.18668/NG.2022.01.01>
6. Pogrebnyak, V. G., Chudyk, I. I., Pogrebnyak, A. V., & Perkun, I. V. (2021). High-efficiency Casing Perforation Oil and Gas Wells. *SOCAR Proceedings*, (2), 112–120. <https://doi.org/10.5510/OGP2021SI200578>
7. Perkun, I., Pogrebnyak, V., & Szymanskiy, V. (2025). Perforation of casing strings in oil and gas wells with a polymer solution jet: scientific and practical aspects. *Trends in petroleum engineering*, 5(1), 1–7. <https://doi.org/10.53902/TPE.2025.05.00045>
8. Pogrebnyak, V. H., Perkun, I. V., Pogrebnyak, A. V., & Shymanskyi, V. Ya. (2022). *Sposib hidrostrumynnoi perforatsii sverdlovyh* [Method of hydrojet perforation of wells] (Patent No. 150245). Ukrainskyi instytut intelektualnoi vlasnosti (Ukrpatent).
9. Pogrebnyak, V. G., & Voloshyn, V. S. (2010). *Ekolohichna tekhnologhiia stvorennia vodozakhysnykh ekraniv: monohrafiia* [Ecological technology for creating water protection screens: monograph]. Donetsk: Knowledge. <https://scholar.google.com/citations?user=rwvMpTQAAAAJ&hl=en&oi=sra>
10. Pogrebnyak, V. G., Naumchuk, N. V., & Maksyutenko, S. N. (1982). Use of heat carriers with hydrodynamically active additives. *Izvestiya Vysshikh Uchebnykh Zavedenii, Energetika*, (5), 109–111.
11. Pogrebnyak, V. G., Perkun, I. V., & Pogrebnyak, A. V. (2017). Displacement of Oil from Porous Bed by the Oscillating Flow of Polymer Solution. *American Journal of Science, Engineering and Technology*, 2(1), 53–57.
12. Ferry, J. D. (1961). *Viscoelastic Properties of Polymers*. John Wiley & Sons.
13. Middleman, S. (1968). *The Flow of High Polymers: Continuum and Molecular Rheology*. Interscience Publishers.
14. Vinogradov, G. V., & Malkin, A. Ya. (1980). *Rheology of Polymers*. Mir Publishers; Springer-Verlag.
15. Kleiman, R. N. (1987). Analysis of the oscillating-cup viscometer for the measurement of viscoelastic properties. *Physical Review A*, 35(1), 261–275.
16. Ferry, J. D. (1965). Relaxations in Polymer Solutions, Liquids, and Gels. In W. P. Mason (Ed.), *Physical Acoustics: Principles and Methods* (Vol. II, part B, pp. 1–115). Academic Press.
17. Pogrebnyak, A., & Pogrebnyak, V. (2017). The Elastic Deformations and Anomalous High Cutting Ability of a Polymer Solution Jet. *Scientific Journal of the Ternopil National Technical University*, (1(85)), 89–94.
18. Elyashevich, G. K. (1982). Thermodynamics and kinetics of orientational crystallization of flexible chain polymers. *Advances in Polymer Science*, 43, 205–245.

-
19. Pogrebniyak, V. H., Perkun, I. V., Pogrebnyak, A. V., & Shymanskyi, V. Ya. (2023). *Hidroimpulsnyi perforator* [Hydro-impulse perforator] (Patent No. 152383). Ukrainskyi instytut intelektualnoi vlasnosti (Ukrpatent).
20. Pogrebnyak, V. G., & Chudyk, I. I. (2022). Solutions of Polymers in the Oil and Gas Technologies. In *Prospects for developing resource-saving technologies in mineral mining and processing: multi-authored monograph* (pp. 110–294). UNIVERSITAS Publishing. <https://doi.org/10.31713/m1106>

Дата надходження статті: 06.10.2025

Дата прийняття статті: 10.11.2025

Опубліковано: 30.12.2025

Поляков В. О., кандидат технічних наук,
старший науковий співробітник відділу динаміки та міцності нових
видів транспорту Інституту транспортних систем і технологій
Національної академії наук України
ORCID: 0000-0002-4957-8028

Пославський С. Ю., кандидат технічних наук,
старший науковий співробітник відділу динаміки та міцності нових
видів транспорту Інституту транспортних систем і технологій
Національної академії наук України
ORCID: 0009-0007-8972-5823

РОЗРОБКА АЛГОРИТМУ ПОБУДОВИ МОДЕЛІ ДИНАМІКИ ВАГОНА МАГЛЕВ ПОЇЗДА

Основними вимогами до моделі системної динаміки є: актуальність, результативність, достовірність, економічність, системність, оглядність, інформаційність. Найбільшою мірою таким вимогам відповідають моделі, які отримують тензорними методами.

Характер задач, при вирішенні яких передбачається використовувати методику, що розробляється, а також структури їх об'єктів дослідження свідчать про те, що в якості їхніх розрахункових схем можуть бути прийняті агрегати механічних і електромагнітних інерційних елементів, які взаємодіють через елементи податливості.

У якості мети дослідження було обрано подальший розвиток методики тензорного моделювання рухів маглев поїздів, на прикладі їхнього вагона, з урахуванням наявності у ньому детермінуючих руху – механічних і електромагнітних – елементів.

До об'єднання в агрегат, що приймається як розрахункова схема об'єкта розгляду, опорні елементи цієї схеми нічим не з'єднані, їх рухи нічим не обмежені. Після сполучення, на рухи елементів накладаються обмеження, які відображаються рівняннями зв'язів. Структура агрегату описується його структурною матрицею. Розрахункова схема вагона є агрегатом ідеальних інерційних елементів, з'єднаних згідно зі структурною матрицею.

Рухи як відокремленого опорного тіла, так і їх несполученого конгломерату описуються тензорними рівняннями. Для урахування структури розрахункової схеми, рівняння, що описують рухи незв'язаної сукупності опорних тіл, згортаються зі структурною матрицею агрегату. Результуюча модель динаміки вагона – також тензорна і може бути легко перетворена до іншої структури сполучення шляхом згортки з матрицею переходу від попередньої структури сполучення опорних тіл до наступної.

Основою побудови моделі руху вагона є рівняння, що описують динаміку опорних тіл, а також структурна матриця агрегату, яка відображає його структуру. Така модель відповідає вимогам холистичного врахування як параметричної, так і структурної організації об'єкта, що розглядається. Розроблювана методика її побудови може з успіхом використовуватися під час досліджень зазначеного типу поїздів.

Ключові слова: маглев поїзд, модель руху, розрахункова схема, опорне тіло, структурна матриця, тензорні методи, згортка.

Polyakov V. O., Poslavsky S. Yu. Development of a maglev train's carriage's dynamic's model constructing algorithm

The main requirements for a system dynamics model are: relevance, effectiveness, reliability, cost-effectiveness, systematicity, accessibility for review, and informativeness. To the greatest extent, models obtained using tensor methods meet these requirements.

The nature of the problems, in the solution of which it is planned to use the developed methodology, as well as the structure of their object of study, indicate that an aggregate of mechanical and electromagnetic inertial elements that interact through flexible ones can be adopted as its calculation scheme.

The aim of the research was to further develop the method of maglev train's motion's tensor modeling using the example of its carriage, taking into account the presence of mechanical and electromagnetic elements which determine its movement.

Before being combined into an aggregate, which is accepted as the design scheme of the object under consideration, the supporting elements of this scheme are not connected in any way, their movements are not restricted in any way. After coupling, constraints are imposed on the movements of the elements, which are reflected by the bond equations. The structure of an aggregate is described by its structural matrix. The design scheme of the carriage is an aggregate of ideal inertial elements connected according to a structural matrix.

© В. О. Поляков, С. Ю. Пославський, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

The motions of both a separate support body and their non-conjugated conglomerate are described by tensor equations. To take into account the structure of a wagon's calculation scheme, the equations describing the movements of an unconnected set of support bodies are convoluted with the structural matrix of the unit. The resulting carriage's dynamics model is also tensor and can be easily converted to another coupling structure by convolution with the transition matrix from the previous support body coupling structure to the next.

The basis for building a carriage's motion model is the equations describing the dynamics of the supporting bodies, as well as the structural matrix of the unit, which reflects its structure. Such a model meets the requirements of holistic consideration of both the parametric and structural organization of the object under consideration. The developed methodology for its construction can be successfully used when researching such trains dynamics.

Key words: maglev train, motion model, calculation scheme, support body, structural matrix, tensor methods, convolution.

Постановка проблеми. Маглев поїзда, в порівнянні з багатьма іншими видами транспорту, мають безліч переваг: висока швидкість пасажирських і вантажних перевезень, безконтактність рушіїв, високі економічність і керованість двигунів, екологічність, безпека, досить швидка окупність капітальних витрат і низькі витрати експлуатаційні. У той же час, побудова бажаних керованих рухів цих поїздів у різних їх експлуатаційних режимах – проблема вельми не тривіальна. Побудова кожного з таких рухів у своїй основі, безумовно, має модель динаміки не керованої – природної. Тому актуальність розробки та подальшого розвитку алгоритмів побудови моделей, у тому числі, природної динаміки маглев поїздів та їх елементів з часом все підвищується.

Аналіз рівня дослідженості проблеми. Рух маглев поїздів, як складних динамічних систем, піддається багатосторонньому дослідженню [1-6]. Проте, зазвичай, такі дослідження мають або експериментальний, або імітаційний характер. Дослідження ж аналітичного та чисельного характеру для таких систем або не проводяться зовсім, або опис результатів таких досліджень не наводиться у відкритій пресі. У роботах [7, 8] було наведено абрисний опис парадигми тензорного моделювання динаміки системи, що розглядається, але лише з урахуванням в ній механічних елементів і без повного опису елементів розробленої методики моделювання.

Метою дослідження є подальший розвиток методики тензорного моделювання рухів маглев поїзда, на прикладі його вагона, з урахуванням наявності в ньому основних – механічних та електромагнітних – елементів.

Виклад основного матеріалу дослідження. Основними вимогами до моделі системної динаміки є: актуальність, результативність, достовірність, економічність, системність, доступність для огляду, інформаційність [7]. Найбільшою мірою таким вимогам відповідають моделі, які отримують тензорними методами [9].

Урахування характеру задач, при вирішенні яких передбачається використовувати методику, що розробляється, а також структури їхнього об'єкта дослідження призводить до висновку про те, що в якості його розрахункової схеми може бути прийнятий агрегат механічних і електромагнітних інерційних елементів, що взаємодіють через елементи податливі. Ступінь глобальності цих елементів залежить від цілей та умов дослідження і обмежена лише тим, що агрегування можливе аж до імітації усієї системи одним елементом, а декомпозиція – до рівня фізичних фреймів. Під останніми будемо розуміти функціональні елементи, кожен з яких необхідний і достатній для представлення класу якісно однотипних об'єктів. Ідеальні тверді тіла, пружності та демпфери вважатимемо фреймами механічної природи, а ідеальні індуктивності, ємності та омичні опори – електромагнітної.

До об'єднання в агрегат елементи розрахункової схеми нічим не з'єднані, їх рухи нічим не обмежені. Конфігурація сукупності визначається опорними координатами

$$\xi^{\beta} \quad \forall \beta \in [\overline{1, N}], \quad (1)$$

де N – кількість таких координат.

Після сполучення, на рухи елементів накладаються обмеження, які відображаються рівняннями зв'язків. Вважатимемо їх голономними склерономними

$$\xi^{\beta} = \xi^{\beta}(\eta^{\lambda}) \quad \forall \beta \in [\overline{1, N}], \lambda \in [\overline{1, \Lambda}], \quad (2)$$

де $\eta^{\lambda} \quad \forall \lambda \in [\overline{1, \Lambda}]$ – узагальнені координати системи;

Λ – кількість її ступенів волі.

Тобто елементи з'єднані згідно з структурою агрегату, що визначається його структурною матрицею [10]:

$$S = \frac{\partial \xi^{\beta}}{\partial \eta^{\lambda}} \quad \forall \beta \in [\overline{1, N}], \lambda \in [\overline{1, \Lambda}]. \quad (3)$$

Ідеальні (безінерційні) податливі елементи розрахункової схеми можуть бути враховані шляхом запровадження узагальнених сил, що виникають за зміни їх стану. Зазвичай, прийнято вважати, що це сили є пропорційними значенням і швидкостям зміни величин, які характеризують стани зазначених елементів:

$$E_\mu = k_\mu \cdot \Delta_\mu + n_\mu \cdot \dot{\Delta}_\mu \quad \forall \mu \in [\overline{1, D}], \quad (4)$$

де $E_\mu, \Delta_\mu, \dot{\Delta}_\mu$ – узагальнена сила в μ -ому елементі, значення та швидкість зміни величини, що характеризує його стан;

$k_\mu, n_\mu; D$ – коефіцієнти пропорційності, а також кількість податливих елементів у розрахунковій схемі вагона.

Використовуючи будь-яку систему електромеханічних аналогій [11], кожної з величин, що входять до (4), може бути надана фізична інтерпретація. Наприклад, у першій з таких систем:

– для механічного податливого елемента:

E_μ – сила, або момент;

$\Delta_\mu, \dot{\Delta}_\mu$ – значення та швидкість зміни деформації;

k_μ, n_μ – жорсткість та коефіцієнт в'язкого опору;

– для електромагнітного елемента:

E_μ – прикладена до нього електрична напруга;

$\Delta_\mu, \dot{\Delta}_\mu$ – заряд та струм через елемент;

k_μ, n_μ – його ж інверсна ємність та омичний опір.

При такій постановці, розрахункова схема системи являє собою агрегат ідеальних інерційних елементів, з'єднаних згідно структурної матриці S . На кожен із цих елементів через в'язі, у їх взаємодії з податливими елементами, діють узагальнені сили реакцій:

$$R_\gamma = R_\gamma(E_\lambda) \quad \forall \gamma \in [\overline{1, J}], \lambda \in [\overline{1, H}], \quad (5)$$

де J, H – число в'язів, накладених на інерційний елемент, і кількість взаємодіючих із ним елементів податливих.

Ззовні на інерційний елемент діють узагальнені сили:

$$W_\varepsilon = W_\varepsilon(\chi_\tau) \quad \forall \varepsilon \in [\overline{1, V}], \tau \in [\overline{1, A}], \quad (6)$$

де $V, \chi_\tau \quad \forall \tau \in [\overline{1, A}], A$ – число шляхів взаємодії інерційного елемента із середовищем, параметри середовища та кількість таких параметрів.

З використанням принципу звільнення від в'язів, а також узагальненої теореми Пуансо [12], система сил, що діє на j -ий інерційний елемент, може бути приведена до їх головного вектора

$$F_j = F_j(R_\gamma, W_\varepsilon) \quad \forall \gamma \in [\overline{1, J}], \varepsilon \in [\overline{1, V}] \quad (7)$$

та головного моменту

$$M_j = M_j(R_\gamma, W_\varepsilon) \quad \forall \gamma \in [\overline{1, J}], \varepsilon \in [\overline{1, V}]. \quad (8)$$

Тоді рух i -го інерційного елемента розрахункової схеми може бути описано рівнянням [13]

$$f_{i\alpha\beta} \cdot \ddot{\varepsilon}_i^\beta + E_{i\alpha, \beta\gamma} \cdot \dot{\varepsilon}_i^\beta \cdot \dot{\varepsilon}_i^\gamma = K_{i\alpha} \quad \forall \alpha, \beta, \gamma \in [\overline{1, G}], \quad (9)$$

де $f_{i\alpha\beta}, E_{i\alpha, \beta\gamma} \quad \forall \alpha, \beta, \gamma \in [\overline{1, G}]$ – розширений тензор інерції елемента та його триндексний символ Крістоффеля першого роду;

$\varepsilon_i^\beta \quad \forall \beta \in [\overline{1, G}], G, K_{i\alpha} \quad \forall \alpha \in [\overline{1, G}]$ – його ж опорні координати, їх число та відповідні узагальнені сили.

Справедливими є також співвідношення [13]

$$E_{i\alpha, \beta\gamma} = 0,5 \cdot \left(\frac{\partial f_{i\alpha\beta}}{\partial \varepsilon_i^\gamma} + \frac{\partial f_{i\alpha\gamma}}{\partial \varepsilon_i^\beta} - \frac{\partial f_{i\beta\gamma}}{\partial \varepsilon_i^\alpha} \right) \quad \forall \alpha, \beta, \gamma \in [\overline{1, G}]. \quad (10)$$

Крім того

$$K_{i\alpha} = K_{i\alpha}(F_i, M_i) \quad \forall \alpha \in [\overline{1, G}]. \quad (11)$$

Інтерпретація та вид залежностей елементів рівнянь (9) від параметрів інерційного елемента визначається його природою.

Для i -го твердого тіла (механічного фрейму), $G = 6$ та його динаміку зручно описувати з використанням правих ортонормованих трієдрів:

$OX_p \quad \forall p \in [\overline{1, 3}]$ – абсолютного (інерційного); $iCZ_r \quad \forall r \in [\overline{1, 3}]$ – жорстко пов'язаного з розглянутим тілом. Осі останнього трієдра збігаються з головними центральними осями тіла. Взаємна орієнтація цих трієдрів визначима за допомогою напрямних косинусів:

$$\begin{matrix} x_{iAp} \\ z_{iAr} \\ u_{ip} \end{matrix}, \quad (12)$$

де $x_{iAp}, z_{iAr}, \forall p, r \in [1, 3]$ – декартові координати точки А тіла у цих трієдрах;
 $u_{irp}, \forall r, p \in [1, 3]$ – матриця перетворення координат 3-поміж них.
Після введення координат $\varepsilon_i^{\beta}, \forall \beta \in [1, 6]$

$$\varepsilon_i^1 = x_{iC1}; \quad \varepsilon_i^2 = x_{iC2}; \quad \varepsilon_i^3 = x_{iC3};$$

$$\varepsilon_i^4 = \tilde{\varphi}_i; \varepsilon_i^5 = \tilde{\Theta}_i; \varepsilon_i^6 = \tilde{\gamma}_i, \quad (13)$$

отримуємо

$$f_{i\alpha\beta} = \begin{Bmatrix} f_{it} & 0 \\ 0 & f_{ir} \end{Bmatrix} \quad \forall \alpha, \beta \in [1, 6];$$

$$f_{it} = \text{diag}\{m_i, m_i, m_i\}; \quad f_{ir} = \begin{Bmatrix} f_{i44} & f_{i45} & f_{i46} \\ f_{i54} & f_{i55} & 0 \\ f_{i64} & 0 & f_{i66} \end{Bmatrix}; \quad (14)$$

$$f_{i44} = (I_{i11} \cdot \cos^{(2)} \tilde{\gamma}_i + I_{i22} \cdot \sin^{(2)} \tilde{\gamma}_i) \cdot \cos^{(2)} \tilde{\Theta}_i + I_{i33} \cdot \sin^{(2)} \tilde{\Theta}_i;$$

$$f_{i45} = f_{i54} = (I_{i11} - I_{i22}) \cdot \cos \tilde{\Theta}_i \cdot \sin \tilde{\gamma}_i \cdot \cos \tilde{\gamma}_i; \quad f_{i46} = f_{i64} = I_{i33} \cdot \sin \tilde{\Theta}_i;$$

$$f_{i55} = I_{i11} \cdot \sin^{(2)} \tilde{\gamma}_i + I_{i22} \cdot \cos^{(2)} \tilde{\gamma}_i; \quad f_{i66} = I_{i33}; \quad (15)$$

ненульові елементи $E_{i\alpha, \beta\gamma}, \forall \alpha, \beta, \gamma \in [1, 6]$

$$E_{i4,45} = E_{i4,54} = [I_{i33} - (I_{i11} \cdot \cos^{(2)} \tilde{\gamma}_i + I_{i22} \cdot \sin^{(2)} \tilde{\gamma}_i)] \cdot \sin \tilde{\Theta}_i \cdot \cos \tilde{\gamma}_i;$$

$$E_{i4,46} = E_{i4,64} = (I_{i22} - I_{i11}) \cdot \cos^{(2)} \tilde{\Theta}_i \cdot \sin \tilde{\gamma}_i \cdot \cos \tilde{\gamma}_i;$$

$$E_{i4,55} = (I_{i22} - I_{i11}) \cdot \sin \tilde{\Theta}_i \cdot \sin \tilde{\gamma}_i \cdot \cos \tilde{\gamma}_i;$$

$$E_{i4,56} = E_{i4,65} = 0,5 \cdot [I_{i33} + (I_{i11} - I_{i22}) \cdot (\cos^{(2)} \tilde{\gamma}_i - \sin^{(2)} \tilde{\gamma}_i)] \cdot \cos \tilde{\Theta}_i;$$

$$E_{i5,44} = -[I_{i33} - (I_{i11} \cdot \cos^{(2)} \tilde{\gamma}_i + I_{i22} \cdot \sin^{(2)} \tilde{\gamma}_i)] \cdot \sin \tilde{\Theta}_i \cdot \cos \tilde{\gamma}_i;$$

$$E_{i5,46} = E_{i5,64} = 0,5 \cdot [(I_{i11} - I_{i22}) \cdot (\cos^{(2)} \tilde{\gamma}_i - \sin^{(2)} \tilde{\gamma}_i) - I_{i33}] \cdot \cos \tilde{\Theta}_i;$$

$$E_{i5,56} = E_{i5,65} = (I_{i11} - I_{i22}) \cdot \sin \tilde{\gamma}_i \cdot \cos \tilde{\gamma}_i;$$

$$E_{i6,44} = (I_{i11} - I_{i22}) \cdot \cos^{(2)} \tilde{\Theta}_i \cdot \sin \tilde{\gamma}_i \cdot \cos \tilde{\gamma}_i;$$

$$E_{i6,45} = E_{i6,54} = 0,5 \cdot [I_{i33} - (I_{i11} - I_{i22}) \cdot (\cos^{(2)} \tilde{\gamma}_i - \sin^{(2)} \tilde{\gamma}_i)] \cdot \cos \tilde{\Theta}_i;$$

$$E_{i6,55} = (I_{i22} - I_{i11}) \cdot \sin \tilde{\gamma}_i \cdot \cos \tilde{\gamma}_i. \quad (16)$$

$$K_{i\alpha} = \begin{Bmatrix} F_{iX1} & F_{iX2} & F_{iX3} & M_{i1}^* & M_{i2}^* & M_{i3}^* \end{Bmatrix}^T; \quad (17)$$

$$M_{i1}^* = (M_{iZ1} \cdot \cos \tilde{\gamma}_i - M_{iZ2} \cdot \sin \tilde{\gamma}_i) \cdot \cos \tilde{\Theta}_i + M_{iZ3} \cdot \sin \tilde{\Theta}_i;$$

$$M_{i2}^* = M_{iZ1} \cdot \sin \tilde{\gamma}_i + M_{iZ2} \cdot \cos \tilde{\gamma}_i; \quad M_{i3}^* = M_{iZ3}. \quad (18)$$

У виразах (13)-(18) позначено:

$x_{iC\alpha}, \forall \alpha \in [1, 3]$ – координати центру мас тіла в інерційній системі $OX_p, \forall p \in [1, 3]$;

$m_i; I_{iZ\alpha}, \forall \alpha \in [1, 3]; \tilde{\varphi}_i, \tilde{\Theta}_i, \tilde{\gamma}_i$ – його маса, головні центральні моменти інерції, а також ейлерові кути щодо тієї самої системи відліку;

$F_{iX\alpha}, M_{iZ\alpha}, \forall \alpha \in [1, 3]$ – задані в інерційній та зв'язаній системах координат, приведені до центру інерції тіла головний вектор та головний момент доданих до нього зовнішніх сил.

Для електромагнітного інерційного кадру, приймаючи

$$\varepsilon_i^1 = q_i, \quad (19)$$

$$f_{i\alpha\beta} = L_i, \quad \alpha = \beta = 1; \quad (20)$$

$$E_{i\alpha,\beta\gamma} = 0, \quad \alpha = \beta = \gamma = 1; \quad (21)$$

$$K_{i\alpha} = U_i, \quad \alpha = 1. \quad (22)$$

У виразах (19)–(22) позначено:

q_i, L_i, U_i – заряд, що протікає через поперечний переріз фрейму, його індуктивність, а також прикладена до нього напруга.

Використовуючи вираз (9), рух агрегату, що є розрахунковою схемою вагона, можна описати тензорним рівнянням виду:

$$b_{\alpha\beta} \cdot \ddot{\xi}^\beta + B_{\alpha,\beta\gamma} \cdot \dot{\xi}^\beta \cdot \dot{\xi}^\gamma = Q_\alpha \quad \forall \alpha, \beta, \gamma \in [\overline{1, N}], \quad (23)$$

$$b_{\alpha\beta} = \text{diag} \{ \|f_{1\varepsilon v}\|, \|f_{2\varepsilon v}\|, \dots, \|f_{\Phi\varepsilon v}\| \} \quad \forall \alpha, \beta, \gamma \in [\overline{1, N}]; \quad \varepsilon, v \in [\overline{1, G}]; \quad (24)$$

$$B_{\alpha,\beta\gamma} = \text{diag} \{ [E_{1\varepsilon, v\chi}], [E_{2\varepsilon, v\chi}], \dots, [E_{\Phi\varepsilon, v\chi}] \} \\ \forall \alpha, \beta, \gamma \in [\overline{1, N}]; \quad \varepsilon, v, \chi \in [\overline{1, G}]; \quad (25)$$

$$Q_\alpha = \|K_{1\varepsilon} \ K_{2\varepsilon} \ \dots \ K_{\Phi\varepsilon}\|^T \quad \forall \alpha \in [\overline{1, N}]; \quad \varepsilon \in [\overline{1, G}], \quad (26)$$

де $b_{\alpha\beta}, B_{\alpha,\beta\gamma} \forall \alpha, \beta, \gamma \in [\overline{1, N}]$ – тензор інерції агрегату та його триндексний символ Крістоффеля в координатах $\xi^\beta \forall \beta \in [\overline{1, N}]$;

$Q_\alpha \forall \alpha \in [\overline{1, N}]$ – узагальнені сили, що відповідають цим координатам;

Φ – кількість інерційних елементів у розрахунковій схемі вагона.

Рівняння (23) тензорні. Тому вони інваріантні до перетворень координат. Тобто в узагальнених координатах $\eta^\lambda \forall \lambda \in [\overline{1, \Lambda}]$ форма моделі руху вагона матиме вигляд, аналогічний (23). Зміняться, з використанням матриці перетворення S , лише вирази елементів моделі.

З рівнянь зв'язків (2) випливає, що

$$\dot{\xi}^\beta = \frac{\partial \xi^\beta}{\partial \eta^\lambda} \cdot \dot{\eta}^\lambda \quad \forall \beta \in [\overline{1, N}], \quad \lambda \in [\overline{1, \Lambda}]; \quad (27)$$

$$\ddot{\xi}^\beta = \frac{\partial^2 \xi^\beta}{\partial \eta^\lambda \partial \eta^\mu} \cdot \dot{\eta}^\lambda \cdot \dot{\eta}^\mu + \frac{\partial \xi^\beta}{\partial \eta^\lambda} \cdot \ddot{\eta}^\lambda \quad \forall \beta \in [\overline{1, N}]; \quad \lambda, \mu \in [\overline{1, \Lambda}]. \quad (28)$$

Помножуючи обидві частини рівняння (23) на структурну матрицю (3), зі згортокою по α , приходимо до рівнянь:

$$b_{\alpha\beta} \cdot \frac{\partial \xi^\alpha}{\partial \eta^\lambda} \cdot \ddot{\xi}^\beta + \frac{\partial \xi^\alpha}{\partial \eta^\lambda} \cdot B_{\alpha,\beta\gamma} \cdot \dot{\xi}^\beta \cdot \dot{\xi}^\gamma = \frac{\partial \xi^\alpha}{\partial \eta^\lambda} \cdot Q_\alpha \quad \forall \alpha, \beta, \gamma \in [\overline{1, N}]; \quad \lambda \in [\overline{1, \Lambda}]. \quad (29)$$

Врахування ж співвідношень (7), (8), (11) і (26) дозволяє записати:

$$Q_\alpha = Q_\alpha (R_\gamma, W_\varepsilon) \quad \forall \alpha \in [\overline{1, N}]; \quad \eta \in [\overline{1, \tilde{J}}]; \quad \varepsilon \in [\overline{1, \tilde{V}}], \quad (30)$$

де $\tilde{J}, \tilde{V}$ – число в'язів, накладених на елементи розрахункової схеми, а також шляхів їх взаємодії із зовнішнім середовищем. З урахуванням теореми про ідеальність в'язів [12]:

$$\frac{\partial \xi^\alpha}{\partial \eta^\lambda} \cdot Q_\alpha (R_\gamma) \equiv 0 \quad \forall \alpha \in [\overline{1, N}], \quad \lambda \in [\overline{1, \Lambda}], \quad \gamma \in [\overline{1, \tilde{J}}], \quad (31)$$

рівняння (29) набувають вигляду:

$$b_{\alpha\beta} \cdot \frac{\partial \xi^\alpha}{\partial \eta^\lambda} \cdot \ddot{\xi}^\beta + \frac{\partial \xi^\alpha}{\partial \eta^\lambda} \cdot B_{\alpha,\beta\gamma} \cdot \dot{\xi}^\beta \cdot \dot{\xi}^\gamma = \frac{\partial \xi^\alpha}{\partial \eta^\lambda} \cdot Q_\alpha (W_\varepsilon) \\ \forall \alpha, \beta, \gamma \in [\overline{1, N}]; \quad \lambda \in [\overline{1, \Lambda}]; \quad \varepsilon \in [\overline{1, \tilde{V}}]. \quad (32)$$

Підстановка виразів (27) та (28) у рівняння (32), після введення додаткових позначень:

$$c_{\lambda\mu} = b_{\alpha\beta} \cdot \frac{\partial \xi^\alpha}{\partial \eta^\lambda} \cdot \frac{\partial \xi^\beta}{\partial \eta^\mu} \quad \forall \alpha, \beta \in [\overline{1, N}]; \quad \lambda, \mu \in [\overline{1, \Lambda}]; \quad (33)$$

$$C_{\lambda\mu\nu} = b_{\alpha\beta} \cdot \frac{\partial \xi^\alpha}{\partial \eta^\lambda} \cdot \frac{\partial^2 \xi^\beta}{\partial \eta^\lambda \partial \eta^\nu} + \frac{\partial \xi^\alpha}{\partial \eta^\lambda} \cdot \frac{\partial \xi^\beta}{\partial \eta^\mu} \cdot \frac{\partial \xi^\gamma}{\partial \eta^\nu} \cdot B_{\alpha\beta\gamma} \\ \alpha, \beta, \gamma \in [\overline{1, N}]; \lambda, \mu, \nu \in [\overline{1, \Lambda}]; \quad (34)$$

$$Y_\lambda = \frac{\partial \xi^\alpha}{\partial \eta^\lambda} \cdot Q_\alpha(W_\varepsilon) \quad \forall \alpha \in [\overline{1, N}]; \lambda \in [\overline{1, \Lambda}]; \varepsilon \in [\overline{1, \tilde{V}}], \quad (35)$$

дозволяє привести модель природного руху вагона маглев поїзда, записану в узагальнених координатах $\eta^\lambda \forall \lambda \in [\overline{1, \Lambda}]$, до вигляду

$$c_{\lambda\mu} \cdot \ddot{\eta}^\mu + C_{\lambda,\mu\nu} \cdot \dot{\eta}^\mu \cdot \dot{\eta}^\nu = Y_\lambda \quad \forall \lambda, \mu, \nu \in [\overline{1, \Lambda}], \quad (36)$$

де $c_{\lambda\mu}$, $C_{\lambda,\mu\nu}$, $Y_\lambda \forall \lambda, \mu, \nu \in [\overline{1, \Lambda}]$ – коваріантний метричний тензор агрегату, що є розрахунковою схемою вагона, триіндексний символ Крістофеля першого роду зазначеного агрегату в координатах $\eta^\lambda \forall \lambda \in [\overline{1, \Lambda}]$, а також узагальнені сили, що відповідають цим координатам.

Може бути показано [13,14], що, аналогічно (10):

$$C_{\lambda,\mu\nu} = 0,5 \cdot \left(\frac{\partial c_{\lambda\mu}}{\partial \eta^\nu} + \frac{\partial c_{\lambda\nu}}{\partial \eta^\mu} - \frac{\partial c_{\mu\nu}}{\partial \eta^\lambda} \right) \quad \forall \lambda, \mu, \nu \in [\overline{1, \Lambda}]. \quad (37)$$

Основою побудови моделі руху (36) є рівняння (23), які описують базову динаміку системи, що досліджується, а також матриця перетворення (3), яка відображає структуру її агрегату. Тому отримана модель відповідає вимогам холистичного врахування як параметричної, так і структурної організації вагона, що розглядається. Тензорний характер моделі (36) дозволяє будувати та перебудовувати модель руху у будь-якій системі координат $\eta^\lambda \forall \lambda \in [\overline{1, \Lambda}]$, тобто – для будь-якої структури розрахункової схеми вагона.

Висновки дослідження та перспективи його продовження. У дослідженні, на прикладі вагона маглев поїзда, виконано розвиток методики тензорного моделювання його динаміки. Продемонстровано переваги такої методики порівняно з іншими, загальноприйнятими. У структурі об'єкта розгляду враховано наявність основних, значущих для його рухів – механічних і електромагнітних – елементів. Методика, що розробляється, може з успіхом використовуватися при дослідженнях динаміки, зокрема, зазначеного типу поїздів. Це сприятиме спрощенню та полегшенню таких досліджень, а також підвищенню якості їх результатів.

Список використаних джерел:

- Schneider G., Schmid P., Kargl A., Liang X., Dign F. & Eberhard P. Simulation of a High-Speed Maglev Train on an Elastic Guideway of Infinite Length MAGLEV. *25th International Conference on Magnetically Levitated Systems and Linear Drives*. 2022. URL: <https://elib.uni-stuttgart.de/server/api/core/bitstreams/c26533d5-e4f3-427a-807b-4e0769eb0878/content>
- Zhiqiang L., Yun L., Xu W. On maglev train automatic operation control system based on a disturbance-rejection control algorithm. *27th Chinese Control Conference*. IEEE. 2008. P. 681–685. DOI:10.1109/CHICC.2008.4605180
- Guodong W., Lei Y., Lianchuan M. Model-based Simulation and Verification on Typical Function Scenes of High-Speed Maglev Vehicle Operation Control System. *Journal of Physics: Conference Series*. IOP Publishing. 2020. № 1654 (1). P. 12–19. URL: <https://iopscience.iop.org/issue/1742-6596/1654/1>
- Zou Z., Zheng M., Lu Q. Modeling and Simulation of Traction Power Supply System for High-Speed Maglev Train. *World Electr. Veh. J.* 2022. № 13(5). URL: <https://doi.org/10.3390/wevj13050082>
- Ying Z., Enwei J., Yongjun L., Hengxin L. Train operation simulation and capacity analysis for a high-speed maglev station Lianbo Deng. *IET Intelligent Transport Systems*. 2024. URL: <https://doi.org/10.1049/itr2.12607>
- Li X., Liu A., Yu M., Yang S., Shao Q., Lei T. Hardware-in-loop simulation of high vehicle based on Hi Gale. *Proceedings V. 136559. Second International Conference on Electronics, Electrical and Control Systems (EECS 2025)*. 1365920. 2025. URL: <https://doi.org/10.1117/12.3070680>
- Поляков В. А., Хачапуридзе Н. М. Моделирование относительного движения магнитолевитирующего поезда. *Вісник Дніпропетровського національного університету залізничного транспорту імені акад. В. Лазаряна*. 2006. № 14. С. 146-151.
- Поляков В. А., Хачапуридзе Н. М. Тензорное моделирование движения магнитолевитирующего поезда относительно пути. *Вестн. ХНТУ*. 2007. № 2 (28). С. 265–270.
- Sinje J. L. *Tensor Methods in Dynamics*. Toronto: University of Toronto Press. Department of Applied Mathematics. 1936. 43 p.
- Schilirò D. Structural change and models of structural analysis: theories, principles and methods. MPRA Paper No. 41817. URL: <https://mpra.ub.uni-muenchen.de/41817/>

-
11. Uklejewski R., Krakowski M. Electromechanical Analogies for the Theory of Consolidation. *Engineering Transactions*. 1982. V. 30(3-4). P. 317–326. URL: <https://et.ippt.pan.pl/index.php/et/article/view/1982>
 12. Діннік О. Н. Курс теоретичної механіки. Ч. 2. Динаміка. Дніропетровськ: ДВОУ Техвидав. 1931. 148 с.
 13. Кільчевський М. О., Нечипоренко Г. Д., Шальда Л. М. Основи аналітичної механіки. Київ: Наук. Думка. 1975. 220 с.
 14. Needham T. Visual differential Geometry and Forms. Princeton University Press. 2021. 531 p. URL: <https://agorism.dev/book/math/diff-geom/Visual%20Differential%20Geometry%20and%20Forms%3A%20A%20Mathematical%20Drama%20in%20Five%20Acts%20by%20Tristan%20Needham.pdf>

References:

1. Schneiderl, G., Schmidl, P., Kargl, A., Liang, X., Dign, F., & Eberhardl, P. (2022). Simulation of a High-Speed Maglev Train on an Elastic Guideway of Infinite Length MAGLEV. *25th International Conference on Magnetically Levitated Systems and Linear Drives*. Retrieved from: <https://elib.uni-stuttgart.de/server/api/core/bitstreams/c26533d5-e4f3-427a-807b-4e0769eb0878/content>
2. Zhiqiang, L., Yun, L. & Xu, W. (2008). On maglev train automatic operation control system based on a disturbance-rejection control algorithm. *27th Chinese Control Conference*. IEEE. P. 681–685. DOI:10.1109/CHICC.2008.4605180
3. Guodong, W., Lei, Y., & Lianchuan, M. (2020). Model-based Simulation and Verification on Typical Function Scenes of High-Speed Maglev Vehicle Operation Control System. *Journal of Physics: Conference Series*. IOP Publishing. № 1654 (1). P. 12–19. Retrieved from: <https://iopscience.iop.org/issue/1742-6596/1654/1>
4. Zou, Z., Zheng, M. & Lu, Q. (2022). Modeling and Simulation of Traction Power Supply System for High-Speed Maglev Train. *World Electr. Veh. J.* № 13(5). <https://doi.org/10.3390/wevj13050082>
5. Ying, Z., Enwei, J., Yongjun, L. & Hengxin, L. (2024). Train operation simulation and capacity analysis for a high-speed maglev station Lianbo Deng. *IET Intelligent Transport Systems*. <https://doi.org/10.1049/itr2.12607>
6. Li, X., Liu A., Yu, M., Yang, S., Shao, Q. & Lei, T. (2025). Hardware-in-loop simulation of high vehicle based on Hi Gale. *Proceedings V. 136559. Second Internationale Conference on Electronics, Electrical and Control Systems (EECS 2025)*. 1365920 <https://doi.org/10.1117/12.3070680>.
7. Poliakov, V. A. & Khachapurydze, N. M. (2006). Modelyrovanye odnosytnoho dvyzheniya mahnytolevytryuushcheho poezda. *Visnyk Dnipropetrovskoho natsionalnoho universytetu zaliznychnoho transportu imeni akad. V. Lazariana*. № 14. S. 146–151.
8. Poliakov, V. A. & Khachapurydze, N. M. (2007). Tenzornoe modelyrovanye dvyzheniya mahnytolevytryuushcheho poezda odnosytno puty. *Vestn. KhNTU*. № 2 (28). S. 265–270.
9. Sinje, J. L. Tensor Methods in Dynamics. (1936) Toronto: University of Toronto Press. Department of Applied Mathematics.
10. Schilirò, D. Structural change and models of structural analysis: theories, principles and methods. MPRA Paper No. 41817. Retrieved from: <https://mpra.ub.uni-muenchen.de/41817/>
11. Uklejewski, R. & Krakowski, M. (1982) Electromechanical Analogies for the Theory of Consolidation. *Engineering Transactions*. V. 30(3-4). P. 317–326. Retrieved from: <https://et.ippt.pan.pl/index.php/et/article/view/1982>
12. Dinnik, O. N. (1931) Kurs teoretychnoi mekhaniky. Ch. 2. Dynamika. Dniropetrovsk: DVOU Tekhvydav.
13. Kilchevskiy, M. O., Nechiporenko, H. D. & Shalda, L. M. (1975). Osnovy analitychnoi mekhaniky. Kyiv: Nauk. Dumka.
14. Needham, T. (2021) Visual differential Geometry and Forms. Princeton University Press. Retrieved from: <https://agorism.dev/book/math/diff-geom/Visual%20Differential%20Geometry%20and%20Forms%3A%20A%20Mathematical%20Drama%20in%20Five%20Acts%20by%20Tristan%20Needham.pdf>

Дата надходження статті: 28.10.2025

Дата прийняття статті: 17.11.2025

Опубліковано: 30.12.2025

УДК 656:658.6

DOI <https://doi.org/10.32782/2521-6643-2025-2-70.37>

Новальська Н. І., кандидат економічних наук, доцент,
доцент кафедри організації авіаційних перевезень
Державного університету «Київський авіаційний інститут»
ORCID: 0000-0002-6331-9217

Клименко В. В., кандидат економічних наук, доцент,
доцент кафедри організації авіаційних перевезень
Державного університету «Київський авіаційний інститут»
ORCID: 0000-0002-4168-3296

Єрошенко О. Р., старший викладач кафедри організації
авіаційних перевезень,
Державного університету «Київський авіаційний інститут»
ORCID: 0000-0002-8411-1116

Ситник Д. О., здобувачка другого (магістерського) рівня вищої
освіти Державного університету «Київський авіаційний інститут»
ORCID: 0009-0003-2550-1979

ПЕРСПЕКТИВИ РОЗВИТКУ ЛОГІСТИЧНОЇ ІНФРАСТРУКТУРИ ПАСАЖИРСЬКИХ ПЕРЕВЕЗЕНЬ: ОРГАНІЗАЦІЙНО-ТЕХНОЛОГІЧНІ ТА ЕКОНОМІЧНІ АСПЕКТИ

Стаття присвячена дослідженню питання подальшого розвитку логістичної інфраструктури пасажирських перевезень в Україні.

Визначено, що розвиток інфраструктури пасажирських перевезень має здійснюватися з урахуванням потреб населення в якісних, безпечних та комфортних послугах. Обґрунтовано потребу у формуванні єдиного пасажирського простору шляхом інтеграції різних видів транспорту та врахування запитів населення. Подальше вдосконалення системи пасажирських перевезень має відбуватися із застосування інноваційних та технологічних рішень, які забезпечать підвищення економічної та соціальної ефективності ринку транспортних послуг. Розглянуто приклади реалізованих проєктів взаємодії видів транспорту при організації пасажирських перевезень, що забезпечили підвищення економічної та соціальної ефективності ринку транспортних послуг в Україні.

Враховуючи вітчизняний та іноземний досвід організації перевезення пасажирів в сполученні “аеропорт-залізничний вокзал” розглянуто техніко-технологічні характеристики Міжнародного аеропорту “Одеса” як стратегічного транспортного вузла південного регіону України. Здійснено аналіз динаміки пасажиропотоку Міжнародного аеропорту “Одеса” в довоєнний період. Досліджено стан системи сполучення Міжнародний аеропорт “Одеса” – залізничний вокзал м. Одеси. Наведено результати аналізу переваг та недоліків існуючих варіантів пасажирських перевезень в аналізованій системі сполучення.

З метою забезпечення покращення транспортного обслуговування пасажирів та забезпечення зручної пересадки з одного виду транспорту на інший розглянуто проєкт забезпечення безперервного та зручного переміщення пасажирів між залізничним вокзалом та аеропортом. Реалізація запропонованих організаційно-технологічних та інфраструктурних рішень для системи пасажирських перевезень є економічно обґрунтованою та стратегічно важливою для розвитку Міжнародного аеропорту «Одеса», підвищення його конкурентоспроможності та забезпечення загальної ефективності транспортної логістики регіону. Проведено порівняння рівня комфорту для пасажирів за існуючими та пропонованими варіантами транспортування.

Ключові слова: логістична інфраструктура, пасажирські перевезення, аеропорт, транспорт, взаємодія видів транспорту, проєкт.

© Н. І. Новальська, В. В. Клименко, О. Р. Єрошенко, Д. О. Ситник, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

Novalska N. I., Klymenko V. V., Yeroshenko O. R., Sytnyk D. O. Prospects for the development of passenger transportation logistics infrastructure: organizational, technological and economic aspects

The article is devoted to the study of the issue of further development of the logistics infrastructure of passenger transportation in Ukraine.

It is determined that the development of passenger transportation infrastructure should be carried out taking into account the needs of the population in high-quality, safe and comfortable services. The need for the formation of a single passenger space by integrating different types of transport and taking into account the requests of the population is substantiated. Further improvement of the passenger transportation system should occur through the use of innovative and technological solutions that will ensure increased economic and social efficiency of the transport services market. Examples of implemented projects of interaction between modes of transport in the organization of passenger transportation, which ensured an increase in the economic and social efficiency of the transport services market in Ukraine, are considered.

Taking into account domestic and foreign experience in organizing passenger transportation in the «airport-railway station» connection, the technical and technological characteristics of the Odessa International Airport as a strategic transport hub of the southern region of Ukraine were considered. The dynamics of passenger traffic at the Odessa International Airport in the pre-war period were analyzed. The state of the connection system between the Odessa International Airport and the Odessa railway station was studied. The results of the analysis of the advantages and disadvantages of existing passenger transportation options in the analyzed connection system were represented.

In order to improve passenger transport services and ensure convenient transfers from one mode of transport to another, a project to ensure uninterrupted and convenient movement of passengers between the railway station and the airport was considered. The implementation of the proposed organizational, technological and infrastructure solutions for the passenger transportation system is economically justified and strategically important for the development of the Odessa International Airport, increasing its competitiveness and ensuring the overall efficiency of transport logistics in the region. A comparison of the level of comfort for passengers using existing and proposed transportation options was carried out.

Key words: logistics infrastructure, passenger transportation, airport, transport, interaction of transport modes, project.

Постановка проблеми. Подальший розвиток інфраструктури пасажирських перевезень в Україні пов'язаний із складним характером завдань щодо відновлення пошкоджених об'єктів та необхідністю ефективного управління логістичними процесами з урахуванням запитів ринку.

Темпи сучасного життя, зростаючі вимоги щодо надійності та якості послуг з пасажирських перевезень зумовлюють необхідність забезпечення задоволення відповідного попиту. Врахування цих вимог вимагає розробки обґрунтованих рішень щодо формування комплексу об'єктів та систем логістичної інфраструктури пасажирських перевезень. Реалізація зазначеного завдання передбачає проведення аналізу відповідних сегментів ринку пасажирських перевезень, оцінки організаційно-технологічного рівня транспортних підприємств, розробки технологічних схем пасажирських перевезень в транспортних вузлах.

Аналіз останніх досліджень і публікацій. Питання перспектив розвитку та відновлення логістичної інфраструктури пасажирських перевезень на різних видах транспорту в Україні викликають певний науковий інтерес. Зокрема, Дружиніна В.В. досліджувала питання функціонування ринку транспортних пасажирських послуг для ефективного забезпечення соціально-виробничого комплексу країни, а також визначала важливість ефективного управління логістичними процесами задля забезпечення оптимального задоволення потреб населення у якісних, безпечних та комфортних послугах [1]. В публікаціях Цвірко О.О. розглядається питання впровадження логістичних підходів у сфері пасажирських перевезень на залізничному транспорті України [2]. В публікаціях Чаркіної Т. Ю. та Задоя В. О. досліджено сучасний стан та перспективи розвитку транспортної інфраструктури в Україні, визначено необхідність застосування сучасного комплексного інноваційного підходу до відновлення транспортної інфраструктури [3].

Проте питання подальшого розвитку логістичної інфраструктури вимагає особливого підходу, оскільки має враховувати сучасні світові тенденції щодо формування технологічних схем пасажирських перевезень при взаємодії видів транспорту задля забезпечення задоволення потреб споживачів в якісних транспортних послугах на вітчизняному ринку.

Метою статті є визначення організаційно-технологічних та економічних аспектів розвитку логістичної інфраструктури пасажирських перевезень при взаємодії видів транспорту задля забезпечення максимального задоволення потреб населення.

Виклад основного матеріалу. Результативність логістичних рішень в сфері пасажирських перевезень залежить від реалізації двох напрямків в організації послуг транспорту. По-перше, забезпечення задоволення наявного та прогнозованого рівня попиту на пасажирські перевезення. По-друге, формування попиту з урахуванням наявних можливостей та перспектив розвитку транспортної інфраструктури.

Разом з тим, необхідно враховувати систему критеріїв, які визначають ефективність функціонування ринку пасажирських перевезень. Серед найбільш вживаних критеріїв визначають швидкість та вартість перевезень, безпеку, надійність, комфорт. Задля досягнення економічної, соціальної та екологічної ефективності, питання про яку в наш час стає дуже актуальним, необхідним є врахування системи інших факторів. Зокрема, макро- і мікроекономічних, науково-технічних, інформаційних, соціальних, демографічних, культурних, природних тощо [1].

Перспективи подальшого розвитку системи пасажирських перевезень потребують розробки та прийняття кардинальних рішень щодо змін принципів побудови транспортно-логістичної інфраструктури, здійснення інноваційних та технологічних рішень, які забезпечать підвищення економічної та соціальної ефективності ринку транспортних послуг. Система пасажирських перевезень має бути гнучкою, здатною швидко реагувати на мінливі умови функціонування. Реалізація вищезазначеного вимагає застосування дієвих принципів в організації та управлінні логістичною інфраструктурою пасажирських перевезень.

Формування єдиного пасажирського простору шляхом інтеграції різних видів транспорту створює зручність для пасажирів, комфортабельність проїзду, скорочення часу в дорозі, зменшення витрат. Сучасні транспортні вузли мають забезпечувати зручну пересадку між різними видами транспорту. Тісна взаємодія між перевізниками рівних видів транспорту передбачає послідовне поліпшення умов перевезення, структури та якості сервісних послуг. Щодо самих перевізників, то маємо перетворення внутрішньологістичного середовища, зміну системи внутрішньофірмових та зовнішніх комунікацій.

Серед прикладів взаємодії видів транспорту в організації пасажирських перевезень є денний рейс поїзда "Укрзалізниці" № 119/120 Київ-Хелм, що почав курсувати з вересня 2023 року. Він зручний для пасажирів, які планують встигнути на нічний та ранковий виліт із аеропорту Варшави (наприклад, до Мюнхена, Парижа, Франкфурта, Відня, Острави, Каїра). Зручність також полягає в тому, що паспортний контроль при перетині митного кордону України пасажирів проходять в день, а не вночі. На станції Хелм здійснюється протягом півгодини пересадка на польський поїзд IC21100 до Варшави, а потім міською електричкою до варшавського аеропорту імені Шопена [4].

З Південного залізничного вокзалу Києва до Міжнародного аеропорту "Бориспіль" можна доїхати на спеціальних автобусах Sky Bus (інтервал 15-45 хв. залежно від часу доби). Також було організовано залізничне сполучення експрес-поїздом Kyiv Boryspil Express (курсав тричі в день та ввечері, час в дорозі 40 хвилин).

Актуальною концепція мультимодального сполучення для організації пасажирських перевезень залишається і для інших аеропортів України. Практична реалізація таких концепцій потребує організаційно-технологічних та інфраструктурних рішень, а також визначення економічної ефективності проекту.

Важливим транспортним вузлом півдня України є міжнародний аеропорт "Одеса". Відсутність зручного транспортного сполучення між аеропортом та залізничним вокзалом Одеси створює певні незручності для пасажирів через необхідність здійснення пересадки, знижується рівень комфорту для пасажирів, збільшується загальний час в дорозі.

Враховуючи як вітчизняний так і іноземний досвід щодо організації перевезення пасажирів в сполученні "аеропорт-залізничний вокзал" доцільно розглянути питання щодо покращення транспортного обслуговування в транспортних вузлах та забезпечення зручної пересадки з одного виду транспорту на інший. Це дозволить аеропорту підвищити привабливість для пасажирів та інвесторів, зміцнивши його роль як важливого транспортного та економічного центру південного регіону України.

Одеський аеропорт як один з ключових елементів транспортної інфраструктури регіону забезпечує високий рівень логістичної взаємодії та інтегрує повітряні, морські, наземні шляхи сполучення.

Міжнародний аеропорт "Одеса" як стратегічний транспортний вузол півдня України здійснює перевезення вантажів та пасажирів у внутрішньому та міжнародних сполученнях. Аеропорт має сучасні технологічні комплекси, які відповідають міжнародним авіаційним стандартам, зокрема сучасні пасажирські термінали. В 2017 році було відкрито новий чотирьох поверховий пасажирський термінал, в якому розміщено 12 стійок реєстрації з потужністю обслуговування 1000 осіб за годину. Новий термінал здатний обслужити за рік 4 млн. пасажирів. Введення в дію в 2021 році нової злітно-посадкової смуги та супутньої інфраструктури надало можливість приймати аеропорту повітряні судна категорії «D» такі як Boeing 767 і Airbus A330, що використовуються для середньо- та далекомагістральних перельотів.

До війни Міжнародний аеропорт "Одеса" був одним з найбільш завантажених в Україні. Зокрема, в 2019 році обсяг пасажиропотоку зріс завдяки інфраструктурним змінам. Сприяли цьому технологічні оновлення, розбудова, розширення мережі маршрутів в країні Азії та Європи, активна співпраця з міжнародними авіакомпаніями. Динаміка пасажиропотоку в Міжнародному аеропорті "Одеса" свідчила про позитивні тенденції щодо його зростання (табл. 1).

В січні 2022 року було обслуговано 110575 пасажирів, з них 93556 на міжнародних рейсах та 17019 на внутрішніх. Пасажиропотік у порівнянні з аналогічним періодом минулого року зріс на 166%. Також авіакомпанія Ryanair відновила рейси до Лондона, Брюсселя, Венеції та відкрила нові рейси до Барселони, Болоньї, Катовіце. Найбільш популярні напрямки рейсів пасажирських перевезень: Стамбул, Шарм-еш-Шейх, Бориспіль, Варшава, Київ (Жуляни).

Дістатися пасажиром з Міжнародного аеропорту "Одеса" до залізничного вокзалу міста і навпаки можна, користуючись послугами маршрутного таксі № 117, міського автобусу № 137, таксі, каршерингу, особистого автомобіля. Послуги маршрутного таксі та міського автобусу є найбільш привабливим для пасажирів з невеликим бюджетом, однак в години пік цей транспорт переповнений, що створює значні незручності для пасажирів з великим багажем. Крім того автобусний маршрут довший та має більшу кількість

зупинок, в підсумку збільшується час в дорозі. Крім того в нічні години громадський транспорт не функціонує. Послуги таксі та сервіси каршерингу забезпечать більш швидке та комфортне пересування, можливість перевезення об'ємного або великого за кількістю багажу. Такий варіант проїзду, в порівнянні з попередніми, є значно дорожчим. Крім того в години пік можуть виникати проблеми з наявністю вільного для замовлення транспорту. Більш деталізовано результати порівняння можливих варіантів проїзду за досліджуваним маршрутом наведено в таблиці 2, яка складена за матеріалами [7].

Таблиця 1

Динаміка пасажиропотоку Міжнародного аеропорту «Одеса» за період 2005-2022 рр. [5, 6]

Рік	Пасажиропотік в МА «Одеса», тис. осіб	у порівнянні з аналогічним періодом минулого року, %
2005	343	
2006	465,1	35,6
2007	527,4	13,4
2008	787	49,2
2009	651	-17,3
2010	707,1	8,6
2011	824,3	16,6
2012	907,6	10,1
2013	1069,1	17,8
2014	863,9	-19,2
2015	949,1	9,8
2016	1033,5	8,9
2017	1228,102	18,3
2018	1446,521	17,7
2019	1346,2	-6,9
2020	698,706	-58
2021	1328,326	98
січень 2022	110,575	166

Також слід зазначити нерегулярність руху транспорту в нічний час. Оскільки більшість рейсів аеропорту виконуються протягом доби, а громадський транспорт працює за обмеженим графіком, пасажирів, що прибувають у нічний час, часто змушені користуватися таксі або каршерингом, що значно збільшує вартість поїздки.

Таким чином, питання швидкості, комфорту, зручності проїзду для пасажирів з Міжнародного аеропорту «Одеса» до залізничного вокзалу міста і навпаки потребує вирішення. Нинішня система сполучення між залізничним вокзалом та аеропортом є інтегрованою. Брак єдиної інформаційної платформи для бронювання квитків і визначення тривалості подорожі створює додаткові труднощі, особливо для туристів, які вперше відвідують м. Одеса.

Таблиця 2

Порівняння існуючих варіантів транспортного сполучення

Варіанти проїзду	Час у дорозі, хв.	Вартість проїзду, грн.	Частота руху	Переваги	Недоліки
Маршрутне таксі № 117	25-40	15	5-10 хв у годину пік	Дешевий варіант, доступний цілодобово	Незручність для пасажирів з багажем, низький рівень комфорту
Автобус № 137	30-45	15	15-20 хв	Дешевий варіант, простий у використанні	Тривалий час у дорозі, залежність від трафіку
Таксі	15-25	200-400	На вимогу	Зручність, швидкість	Висока вартість, можливість шахрайства
Каршеринг (Bolt, Uklon)	15-25	150-300	На вимогу	Комфорт, гнучкість маршруту	Висока вартість, залежність від наявності авто
Особистий транспорт	15-30	Витрати на паливе та паркування	Залежить від власника	Найзручніший варіант для власників авто	Проблеми з паркуванням в аеропорту

Аналіз переваг та недоліків існуючих варіантів пасажирських перевезень, а також вітчизняний та закордонний досвід засвідчує про доцільність застосування спеціального експресу або аеропортового шатлу. Розвиток власного транспортного сервісу надасть змогу аеропорту здійснювати контроль за системою перевезень пасажирів, формувати графіки руху аеропортового шатлу відповідно до розкладів авіарейсів, підвищити рівень якості обслуговування пасажирів. Такий варіант пасажирських перевезень на зазначеному маршруті руху потребує розробки організаційно-технологічних та інфраструктурних рішень, економічного обґрунтування.

З огляду на це, доцільно розглянути проект забезпечення безперебійного та зручного переміщення пасажирів між залізничним вокзалом та аеропортом Одеси. Основні етапи проектування включають вибір оптимального виду транспорту, організацію пересадок та розробку узгодженого графіку руху.

На основі аналізу варіантів визначено, що найбільш ефективним видом транспорту є експрес-автобус. Основні характеристики експрес-автобуса: сучасні низькопідлогові моделі для зручності пасажирів; місткість – 30-50 місць; частота рейсів – кожні 20-30 хвилин, відсутність зайвих зупинок для мінімізації часу у дорозі.

Для підвищення комфорту пасажирів пропонується: створення зручного пересадкового хабу біля залізничного вокзалу з виділеними платформами для експрес-автобусів; інформування пасажирів через електронні табло, мобільний додаток та оголошення в поїздах і автобусах; обладнання зон очікування з місцями для сидіння та захистом від негоди; єдиний квиток на поїзд, автобус та літак, що дозволить пасажирам уникнути зайвих черг та незручностей; організація зупинки експрес-автобуса безпосередньо біля терміналу, що мінімізує час пересадки на авіатранспорт.

Розглянемо економічні аспекти реалізації даного проекту. Варіантами джерел фінансування проекту можуть бути: власні ресурси або кредитні програми, державні або муніципальні субсидії (як частина міської транспортної програми), грантове фінансування (європейські інфраструктурні фонди можуть підтримати проект), виручка від продажу квитків, надходження від реклами на транспортних засобах та в зонах очікування, спонсорські угоди (партнерство з авіакомпаніями чи іншими бізнесами), надходження від функціонування паркінгових зон.

Основні статті витрат проекту: придбання автобусів Mercedes-Benz Citaro C2; створення інфраструктури (облаштування зупинок, електронні інформаційні табло, автоматизована система продажу квитків); операційні витрати (паливно-мастильні матеріали, технічне обслуговування, оплата праці). Джерелами грошових надходжень від реалізації запропонованого проекту є продаж квитків, реклама на транспорті.

Для оцінки економічної доцільності проекту було проведено відповідні фінансові розрахунки. Враховуючи розмір загальних капітальних витрат, величину прогнозованих грошових надходжень та поточних витрат на обслуговування, величину амортизаційних відрахувань (прямий метод розрахунку), базову ставку податку на прибуток (18%), терміну реалізації інвестиційного проекту (3 роки) визначено величину щорічного операційного грошового потоку (табл. 3).

Таблиця 3

Розрахунок планових показників реалізації інвестиційного проекту, тис. грн

№ з/п	Показники	1 рік	2 рік	3 рік
1	Чистий дохід проекту	22025	23121	24277
2	Витрати на обслуговування системи	7250	7975	8772
3	Прибуток від проекту	14775	15146	15505
4	Податок на прибуток	2650	2726	2791
5	Чистий прибуток	12115	12420	12714
6	Амортизаційні відрахування	10666	10666	10666
7	Сума грошового потоку	22781	23086	23380

Для дисконтування сум грошового потоку застосовано ставку в розмірі 25% та визначено чистий приведенний дохід інвестиційного проекту. Також з метою аналізу чутливості інвестиційного проекту проаналізовано зміну величини економічних показників за умови коливання величини ставки дисконту в сторону зменшення та збільшення на 5%. Деталізовано результати розрахунку чистого приведенного доходу інвестиційного проекту по роках реалізації наведено в (табл. 4).

Чиста теперішня вартість аналізованого інвестиційного проекту становитиме:

- 1) при ставці дисконтування 20% NPV 16544,3 тис.грн.;
- 2) при ставці дисконтування 25% NPV 12970,4 тис.грн.;
- 3) при ставці дисконтування 30% NPV 9825,7 тис.грн.

Таблиця 4

Розрахунок чистого приведенного доходу інвестиційного проекту

Рік, t	Сума грошового потоку, тис. грн.	Ставка дисконтування 20%			Ставка дисконтування 25%			Ставка дисконтування 30%		
		Дисконтний множник $1/(1+0,2)^t$	Дисконтований грошовий потік (DCF_t), тис. грн.	Накопичений DCF_t , тис. грн.	Дисконтний множник $1/(1+0,25)^t$	Дисконтований грошовий потік (DCF_t), тис. грн.	Накопичений DCF_t , тис. грн.	Дисконтний множник $1/(1+0,3)^t$	Дисконтований грошовий потік (DCF_t), тис. грн.	Накопичений DCF_t , тис. грн.
0	-32000	1	-32000		1	-32000		1	-32000	
1	22781	0,8333	18983,4	-13016,6	0,8	18224,8	-13775,2	0,7692	17523,1	-14476,9
2	23086	0,6944	16030,9	3014,3	0,64	14775	999,8	0,5917	13660	-816,9
3	23380	0,5787	13530	16544,3	0,512	11970,6	12970,4	0,4552	10642,6	9825,7

Зростання ставки дисконтування на 5% впливає на суттєве скорочення дисконтованого грошового потоку та зменшення чистої середньої вартості проекту в середньому на 22-24%. Оскільки величина NPV в усіх трьох аналізованих варіантах має позитивне значення, то інвестиційний проект є економічно вигідним за розглянутими ставками дисконтування. Визначення дисконтованого періоду окупності інвестицій засвідчило, що за ставки дисконтування 25% окупність настає наприкінці 2-го року реалізації інвестиційного проекту.

В цілому серед переваг наведеного проекту слід визначити: для аеропорту – забезпечення конкурентоспроможності, покращення зовнішнього іміджу, зростання пасажиропотоку, зростання рівня логістичного сервісу, отримання додаткового джерела доходу, підвищення ефективності господарювання; для пасажирів – комфорт, зручність, скорочення часу перебування в дорозі. В (табл. 5) показано порівняння рівня комфортності при перевезенні пасажирів за різними варіантами транспортування.

Таблиця 5

Порівняння рівня комфорту за різними варіантами транспортування

Критерій	Громадський автобус	Таксі	Експрес-автобус аеропорту
Комфортність сидінь	Середня	Висока	Висока
Клімат-контроль	Відсутній або обмежений	Так	Так
Простір для багажу	Обмежений	Обмежений	Великий
Wi-Fi	Відсутній	Іноді	Так
Відеоспостереження	Відсутнє	Відсутнє	Так
Доступність для маломобільних пасажирів	Обмежена	Обмежена	Повна

Висновки. Сучасна транспортно-логістична інфраструктура пасажирських перевезень спрямована на узгоджене використання видів транспорту задля максимального задоволення потреб споживачів транспортних послуг, раціонального використання транспортних засобів, зменшення загальнотранспортних експлуатаційних витрат.

В післявоєнний час відбудеться відновлення діяльності аеропортів в Україні. В свою чергу актуалізується питання удосконалення системи пасажирських перевезень в великих транспортних вузлах за участі авіаційного транспорту. Розробка сучасної технологічної системи пересадки пасажирів в транспортних вузлах потребуватиме організаційно-технологічного та економічного обґрунтування.

Список використаних джерел:

- Дружиніна В. В., Черніхова О. С. Управління логістичними процесами на ринку транспортних пасажирських перевезень. *Сталий розвиток економіки*. 2024. № 1(48). С. 150–158.
- Цвірко О. О. Впровадження логістичних підходів щодо пасажирських перевезень у процесі структурної реформи залізничного транспорту України. *Причорноморські економічні студії*. 2018. Випуск 33. С. 165–169.
- Чаркіна Т. Ю., Задоя В. О., Юрчик О. А. Сучасний стан та перспективи розвитку відновлення і розбудови транспортної інфраструктури в Україні. *Агросвіт*. 2024. № 6. С. 103–112.

-
4. Як дістатися до аеропорту Варшави поїздом із України: зручний маршрут. URL: <https://www.rbc.ua/rus/travel/k-distatysya-aeroportu-varshavi-poyizdom-1693817411.html>
 5. Одеса (аеропорт). *Вікіпедія*. URL: <https://surl.li/zmlybh>.
 6. Odesa Airport / Міжнародний аеропорт "Одеса". URL: https://www.facebook.com/odesa.aero/?locale=uk_UA
 7. Офіційний сайт Одеської міської ради – Розклад руху громадського транспорту в м. Одеса. URL: <https://omr.gov.ua>.

References:

1. Druzhynina, V. V., & Chernikhova, O. S. (2024). Upravlinnia lohistychnymy protsesamy na rynku transportnykh pasazhyrskykh perevezen [Management of logistics processes in the passenger transport market]. *Stalyi rozvytok ekonomiky*, (1), 150–158.
2. Tsvirko, O. O. (2018). Vprovadzhennia lohistychnykh pidkhodiv shchodo pasazhyrskykh perevezen u protsesi strukturnoi reformy zaliznychnoho transportu Ukrainy [Implementation of logistics approaches to passenger transport in the process of structural reform of Ukraine's railway transport]. *Prychornomorski ekonomichni studii*, (33), 165–169.
3. Charkina, T. Iu., Zadoia, V. O., & Yurchyk, O. A. (2024). Suchasnyi stan ta perspektyvy rozvytku vidnovlennia i rozbudovy transportnoi infrastruktury v Ukraini [Current status and prospects for the restoration and development of transport infrastructure in Ukraine]. *Ahrosvit*, (6), 103–112.
4. Yak distatysia do aeroportu Varshavy poizdom iz Ukrainy: zruchnyi marshrut [How to get to Warsaw Airport by train from Ukraine: a convenient route]. Retrieved from: <https://www.rbc.ua/rus/travel/k-distatysya-aeroportu-varshavi-poyizdom-1693817411.html>
5. Odesa (airport) [Odesa (airport)]. Wikipedia. Retrieved from: <https://surl.li/zmlybh>.
6. Odesa Airport / International Airport “Odesa”. Retrieved from: https://www.facebook.com/odesa.aero/?locale=uk_UA
7. Official website of the Odessa City Council – Public transport timetable in Odessa. Retrieved from: <https://omr.gov.ua>.

Дата надходження статті: 27.10.2025

Дата прийняття статті: 17.11.2025

Опубліковано: 30.12.2025

МОРСЬКИЙ ТА ВНУТРІШНІЙ ВОДНИЙ ТРАНСПОРТ

UDC 656.614.32:338.47

DOI <https://doi.org/10.32782/2521-6643-2025-2-70.38>

Drozhzhyn O. L., Candidate of Technical Sciences,
Associate Professor, Associate Professor at the Department
of Fleet Operation and Sea Transport Technologies
Odesa National Maritime University
ORCID: 0000-0002-9695-9296

Onyschenko S. P., Doctor of Economic Sciences, Professor,
Professor at the Department of Fleet Operation and Sea Transport
Technologies
Odesa National Maritime University
ORCID: 0000-0002-7528-4939

IMPACT ASSESSMENT OF LINER SCHEDULE DISRUPTION ON FLEET EFFICIENCY BASED ON TIME CHARTER EQUIVALENT

The article discusses the problem of liner schedule disruption (LSD) and its impact on the economic efficiency of shipping companies. It has been determined that failure to comply with container ship schedules leads to direct economic losses, which are reflected in a decrease in the time charter equivalent (TCE) indicator, which is a universal indicator of the profitability of ship operation regardless of the form of chartering. It is shown that schedule disruptions have a double negative effect: on the one hand, a decrease in efficiency due to an increase in voyage time, and on the other hand, a deterioration in the carrier's image, which in the long term reduces demand and market share.

The paper proposes a mathematical model for quantitative assessment of losses from delays in port and during the voyage. Analytical dependencies allow determining the decrease in TCE due to an increase in downtime in ports and the time a vessel spends in sea transit. The calculations showed that delays during ship movement are 2–3 times more costly than delays of equivalent duration in port, due to high bunker fuel costs and the need to recover the schedule by increasing speed.

The assessment methodology developed by the author allows analyzing the economic consequences of schedule non-compliance for both individual voyages and the entire line or fleet of the company as a whole. It takes into account the difference in the size of vessels and their contribution to the total volume of transportation, which ensures the adequacy of the assessment in strategic planning. The results obtained are of practical importance for optimizing management decisions in the field of liner shipping, in particular for improving the reliability of schedules, minimizing operational risks, and reducing financial losses.

Thus, the proposed approach transforms the abstract problem of schedule disruption into a specific financial indicator that can be used as a tool for monitoring performance and making decisions in real time, as well as for long-term strategic planning of container transport development.

Key words: liner shipping, container, Time Charter Equivalent, schedule, delay, risk, disruption.

Дрожжін О. Л., Онищенко С. П. Оцінка впливу порушення розкладу руху лінійних суден на ефективність флоту на основі тайм-чартерного еквіваленту

У статті розглянуто проблему порушення розкладу регулярних лінійних перевезень (Liner Schedule Disruption, LSD) та її вплив на економічну ефективність виробничої діяльності судноплавних компаній. Визначено, що недотримання графіків руху контейнерних суден призводить до прямих економічних втрат, які можуть бути виражені через зниження показника time charter equivalent (TCE), і розглядається виразником економічної ефективності від експлуатації суден незалежно від форми фрахтування. В роботі зазначено, що збої у розкладі мають подвійний негативний ефект: з одного боку – зменшення ефективності, з іншого – іміджеві втрати перевізника, що впливає на зниження попиту з боку вантажовласників; представлена робота присвячена вивченню тільки першої групи втрат. Було запропоновано математичну модель для кількісної оцінки втрат від затримок під час стоянкової і ходової операції. Аналітичні залежності дозволяють встановити рівні зниження TCE унаслідок збільшення часу простою під час перебування суден в портах та й на переходах між портами. Проведені розрахунки показали, що затримки під час ходової операції є у 2–3 рази більшими, ніж еквівалентні за тривалістю затримки у порту. Методика оцінки, розроблена авторами, дозволяє здійснювати аналіз економічних наслідків недотримання розкладу як для окремого рейсу, так і для всієї лінії

© O. L. Drozhzhyn, S. P. Onyschenko, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

чи флоту компанії в цілому. Вона враховує різницю у розмірах суден та їх внесок у загальний обсяг перевезень, що забезпечує адекватність оцінки при стратегічному плануванні. Отримані результати мають практичну значущість для оптимізації управлінських рішень у сфері лінійного судноплавства, зокрема для підвищення надійності графіків, мінімізації операційних ризиків та зменшення фінансових втрат.

Таким чином, запропонований підхід перетворює абстрактну проблему порушення розкладу на конкретний фінансовий показник, що може бути використаний як інструмент моніторингу ефективності та прийняття рішень у реальному часі, а також для довгострокового стратегічного планування розвитку контейнерних перевезень.

Ключові слова: лінійне судноплавство, контейнер, тайм-чартерний еквівалент, розклад, затримка, ризик, порушення.

Introduction. Risks are an integral part of the shipping business, covering both the technological aspects of maritime transport and commercial aspects. A distinctive feature of the liner shipping (LS) sector is that the source of significant costs (both for the LS- company and for the cargo owner) is non-performance, deviation, or disruption of the schedule. Like all types of commercial risks, the consequences of non-performance of the schedule are economic losses, which consist of two components. First of all, the consequences of failure to comply with the schedule, caused by delays in ports or on the route between ports, which usually leads to an increase in transit time, i.e., the voyage time for ships operating on the line, are a decrease in economic efficiency, i.e., the time charter equivalent. Second, failure to comply with the schedule negatively affects the image of the carrier company, which gradually leads to a decrease in demand in favor of competing companies. This, in turn, reduces the company's transport volumes and market share, which in monetary terms also results in certain economic losses. While these losses can only be assessed by experts by comparing the company's competitiveness with its competitors in the more or less long term, the first component of the economic consequences of schedule non-compliance can be assessed on the basis of information on voyage delays.

The significance of the problem of schedule disruptions is the focus of attention for many scientists: since a container ship operates from one port to another across the globe, a disruption by these unexpected events can cause delay, deviation, stoppage, or loss of service platform [1], considering the problem on four levels: delay, deviation, stoppage, and loss of platform service. The seriousness of the requirement to adhere to the schedule is explained by the so-called “knock-on effect” [7], which is related to the fact that a delay that occurred at the previous port of call can also cause delays at subsequent ports of call. Considering that LS is part of a broader system, intermodal delivery, it can be argued that the delay goes beyond the “sea leg” and spreads throughout the supply chain.

In scientific literature, risks that can affect the cost level of a liner operator can be described in different ways. Management articles refer to them as “business risk” or “business environment-based risk” a multidisciplinary set of economic, political, social, natural, and other risks. Based on the environment of origin, the authors distinguish between macro and micro levels.

Literature review. The authors consider delays in LS and schedule disruptions to be sources of the following risks: financial risks [1]; risks of losing customers [3]; article [4] discusses congestion risk associated with disruptions to liner schedules in the context of the pandemic; in [5], “emission risks” are considered as a consequence of disruptions to container ship schedules; the article is devoted to LS [6] classifies “Transportation Delays” as “Risk Associated with Physical Flows,” highlighting the following causes: “industrial action, port congestion, terminal productivity, weather conditions, empty container management, container shortages, planning and scheduling problems, customs inefficiency, oil (and bunkers) prices, pandemics.” In [7], schedule unreliability in LS is considered a cause of “risk of a stock-out.” T. Nottebom sees delays as a cause of negative impact on logistics costs and reputation [8].

The faults that can lead to delays in LS can be of various levels. Undoubtedly, incorrect assessment of supply and demand for transportation in the trade region [16, 18], assessment of the level of competition, the size of ships in the fleet [10, 15, 17], connection of the line to the service network [11, 12], and the level of service provided by the line [13, 14] are among the most common [16, 18], assessment of the level of competition, the size of ships on the line, fleet composition [10, 15, 17], connection of the line to the service network [11, 12, 13, 17, 18] are planning issues and lie at the strategic level [9]. Some scholars refer to shipping route design as a strategic decision-making level [19, 14].

The development of the schedule itself is considered a tactical task, while the problem of restoring the schedule is solved by regulation and lies at the operational level of decision-making [21].

Figure 1, compiled on the basis of a literature review, illustrates how complex and multifaceted the problem of schedule disruption is in terms of its impact on business. The purpose of this article is to develop an LSD assessment tool that is acceptable for fleets operating under various forms of chartering.

The purpose of the article. To assess the economic impact of delays, we propose using Time Charter Equivalent (TCE) as a generalized indicator used in the maritime economy to measure the profitability of vessel operation, regardless of the form of chartering. In LS, various forms of vessel chartering are used: voyage charter, time charter, spot agreements, etc. TCE allows these different forms to be reduced to a single measure that reflects the “equivalent rental rate for a given period” and facilitates comparison of the impact of schedule disruptions for all forms of chartered vessels involved in the liner network.

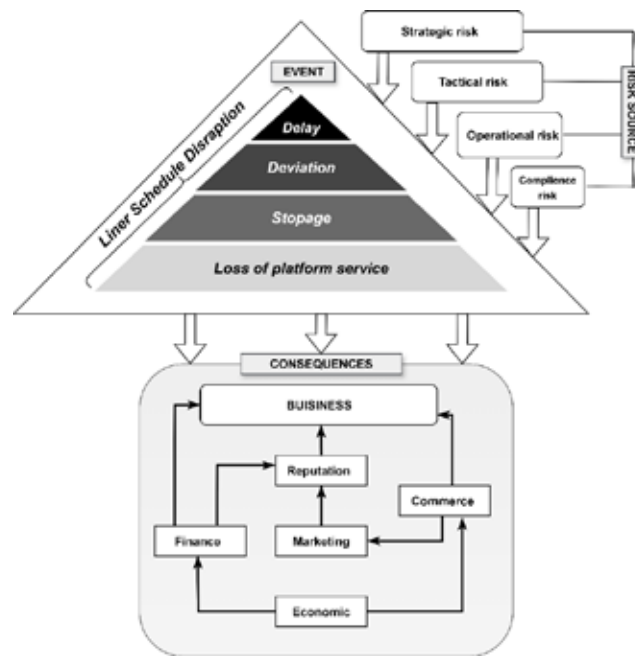


Fig. 1. LSD-problem: Risk Source-Event-Consequence (by the literature analysis basis)

It is evident that the issue of schedule disruption (LSD) can lead to significant reputational risks, threatening a loss of demand and even the closure of the regional service network. This article focuses solely on assessing the costs incurred by linear operators.

Mathematical LSD-problem representation. Therefore, the economic losses of the shipping company can be described mathematically as follows. Let us introduce the following symbols:

$i = \overline{1, n}$ – lines of the company under consideration, n – total number of lines;

$s = \overline{1, S_i}$ – ships of the company operating on the i -line S_i – total number of ships of the company on this line;

$k = \overline{1, K_s^i}$ – voyage of vessel s on line i ,

K_s^i – the number of voyages of vessel s on line i within the period under consideration (e.g., a year);

T_s^{ik} – actual transit time for vessel s on line i ;

T_s^{*ik} – scheduled transit time for vessel s on line i .

Transit time consists of two components: the time the vessel spends in the ports of the line and the vessel's underway time T_s^{p-ik} (transit between the ports of the line) T_s^{m-ik} :

$$T_s^{ik} = T_s^{p-ik} + T_s^{m-ik}. \quad (1)$$

T_s^{p-ik} is formed from the time the vessel spends in the ports of the line. If we take $l = \overline{1, L_i}, i = \overline{1, n}$ – the index of ports on line i , L_i – the total number of ports of call on the line, then:

$$T_s^{p-ik} = \sum_{l=1}^{L_i} T_s^{p-ikl}, \quad (2)$$

T_s^{p-ikl} is the ship's layover time in port l . We will also introduce indicators that reflect the planned total layover time of the ship in ports, the ship's layover time in each port of the line, and the ship's sailing time on the line, respectively, T_s^{*p-ik} , T_s^{*p-ikl} , T_s^{*m-ik} .

The basic indicator for assessing the efficiency of ships is the time charter equivalent, which allows not only to draw conclusions about efficiency, but also to compare different options for operating ships.

The general expression of the time charter equivalent for a specific vessel on a specific line is as follows:

$$TCE_s^{ik} = \frac{F_s^{ik} - (R_{sik}^{port} + R_{sik}^{can} + R_{sik}^{bunk})}{T_s^{ik}}, i = \overline{1, n}, s = \overline{1, S_i}, k = \overline{1, K_s^i}, \quad (3)$$

F_s^{ik} – freight (income) for vessel s on line i , R_{sik}^{port} port expenses, R_{sik}^{can} , R_{sik}^{bunk} costs of passing through channels/straits and bunker expenses.

It should be noted that, on the one hand, can be considered as conditionally constant and independent of time. Even with an increase in the time spent in port, these costs remain almost unchanged. In some cases, may increase

due to certain port charges R_{sik}^{port} that depend on the time the ship spends in port. For example, in Ukrainian ports, the sanitary fee is divided into two categories: “up to 10 days of stay” and “more than 10 days of stay,” with a difference of about 60 %. Therefore, when summarizing the impact of schedule non-compliance on efficiency, the possible dependence R_{sik}^{port} on the time of stay should be taken into account in some cases:

$$R_{sik}^{port} = \sum_{l=1}^{L_i} R_{sikl}^{port} (T_s^{p\_ikl}), i = \overline{1, n}, s = \overline{1, S_i}, k = \overline{1, K_s^i} \quad (4)$$

$r_{si}^{bunk\_p}, r_{si}^{bunk\_m}$ – port fees and charges.

Bunker costs are determined by the underway time and layover time, taking into account the relevant fuel consumption standards:

$$TCE_s^{ik} (T_s^{p_{ik}}, T_s^{m_{ik}}) = \frac{F_s^{ik} - (R_{sik}^{port} (T_s^{p_{ik}}) + R_{sik}^{can} + R_{sik}^{bunk} (T_s^{p_{ik}}, T_s^{m_{ik}}))}{T_s^{p_{ik}} + T_s^{m_{ik}}}, i = \overline{1, n}, s = \overline{1, S_i}, k = \overline{1, K_s^i} \quad (5)$$

the $r_{si}^{bunk\_p}, r_{si}^{bunk\_m}$, respectively, are the fuel consumption standards for the vessel's time underway and vessel's time in port.

Thus, summarizing the above regarding the dependence of individual components of the time charter equivalent on the time of the components of the voyage and the voyage as a whole, we obtain the following expression:

$$TCE_s^{ik} (T_s^{p_{ik}}, T_s^{m_{ik}}) = \frac{F_s^{ik} - (R_{sik}^{port} (T_s^{p_{ik}}) + R_{sik}^{can} + R_{sik}^{bunk} (T_s^{p_{ik}}, T_s^{m_{ik}}))}{T_s^{p_{ik}} + T_s^{m_{ik}}}, i = \overline{1, n}, s = \overline{1, S_i}, k = \overline{1, K_s^i} \quad (6)$$

Therefore, changes in the TCE under the influence of changes in transit time and compared to the base are estimated as follows:

$$\Delta TCE_s^{ik} (T_s^{p_{ik}}, T_s^{m_{ik}}) = TCE_s^{*ik} - TCE_s^{ik} (T_s^{p_{ik}}, T_s^{m_{ik}}) = \frac{F_s^{ik} - (R_{sik}^{port} (T_s^{*p_{ik}}) + R_{sik}^{can} + R_{sik}^{bunk} (T_s^{*p_{ik}}, T_s^{*m_{ik}}))}{T_s^{*p_{ik}} + T_s^{*m_{ik}}} - \frac{F_s^{ik} - (R_{sik}^{port} (T_s^{p_{ik}}) + R_{sik}^{can} + R_{sik}^{bunk} (T_s^{p_{ik}}, T_s^{m_{ik}}))}{T_s^{p_{ik}} + T_s^{m_{ik}}}, i = \overline{1, n}, s = \overline{1, S_i}, k = \overline{1, K_s^i} \quad (7)$$

In expression (7) TCE_s^{*ik} – is the planned efficiency of the vessel's operation on the line, taking into account the accepted work schedule. If the vessel did not adhere to the schedule on a given voyage, i.e., there were delays in ports or during transit between ports, then: $\Delta TCE_s^{ik} (T_s^{p_{ik}}, T_s^{m_{ik}}) > 0$.

If the schedule is met, then $\Delta TCE_s^{ik} (T_s^{p_{ik}}, T_s^{m_{ik}}) = 0$. It should be noted that a situation is theoretically possible where $\Delta TCE_s^{ik} (T_s^{p_{ik}}, T_s^{m_{ik}}) < 0$, which means that the ship also did not meet the schedule, but the flight time was less than planned. However, as a rule, this situation is purely theoretical, practically $\Delta TCE_s^{ik} (T_s^{p_{ik}}, T_s^{m_{ik}}) \geq 0$.

One more comment on (7) and (8): determining the economic consequences of not adhering to the schedule on a particular voyage involves taking into account only changes in the voyage's time parameters $T_s^{p_{ik}}, T_s^{m_{ik}}$; the other components of (7) are considered constant. Therefore, (7) reflects a decrease in efficiency only due to exceeding the flight time. In fact, there is a possibility that the freight amount will be different due to the cancellation of a certain number of containers or, conversely, the presence of additional “last minute” shipments. In any case, only changes in time are taken into account. This allows us to analyze the impact of the “time” factor.

Empirical estimates for the liner vessel. Figure 2 shows, for example, a vessel operating on a route with a transit time of 30 days, the dynamics of the time charter equivalent depending on the increase in voyage time separately from the increase in underway time ($+T_{uw}$) and vessel time in port $+T_p$.

Exploratory calculations were performed for a container shipping line on the route Suez – Jeddah – Colombo – Singapore – Shanghai (planned TT 30 days and TT 40 days).

It should be noted that these dependencies do not take into account changes in ship speed and corresponding changes in fuel consumption, but they allow us to see in general terms that an increase in voyage time significantly affects efficiency, with the impact of the underway component being significantly higher due to the cost of bunkering.

It should be noted that for this example and a transit time of 30 days, an increase in layover time by 0.5 days leads to a decrease in the time charter equivalent by an average of 3%, but an increase in sailing time by 0.5 days leads to a decrease in efficiency by 7 %-10 %.

Figure 3 for the same example of initial data shows the theoretical dependence of the time charter equivalent on both the underway time and the vessel's time in port, which allows us to trace the decrease in the time charter equivalent under the influence of the components of transit time simultaneously.

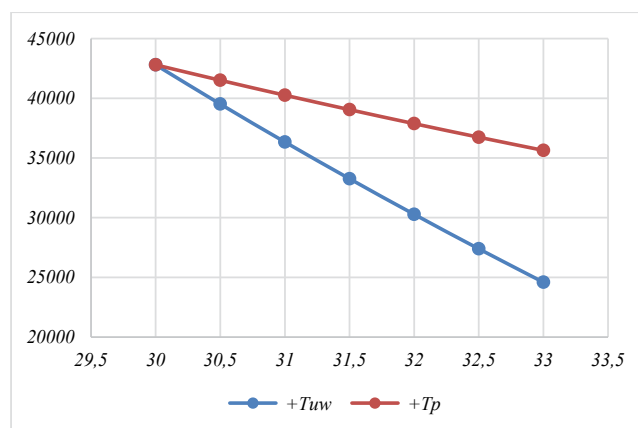


Fig. 2. Dependence of the TCE (USD/day) on the increase in TT (days) for a planned TT 30 days

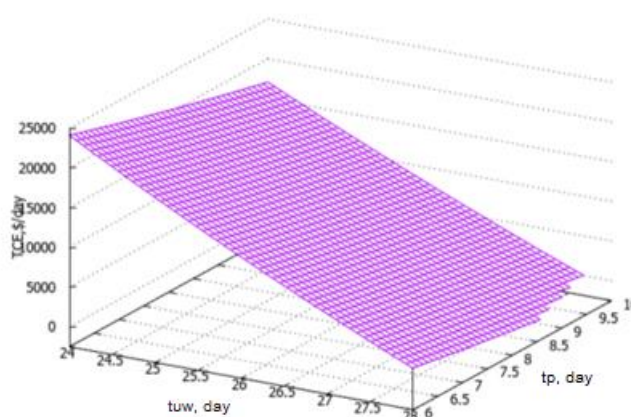


Fig. 3. Dependence of the TCE (USD/day) on underway time (days) and vessel's time in port (days) for a base TT 30 days

Similar calculations were performed for a transit time of 40 days, taking into account the corresponding changes in the sum of freight, port, and other costs.

Figures 4 and 5 indicate the dependence of the TCE on the increase in transit time and its two components. It should be noted that for longer TT, an increase in its components by 0.5 days leads to smaller changes in efficiency – 4-6 % for underway time and 1.5-1.8 % for time in port.

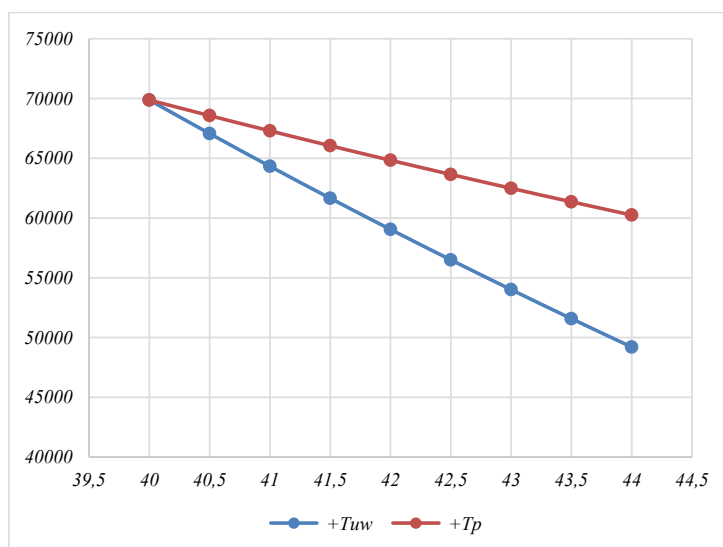


Fig. 4. Dependence of the TCE (USD/day) on the increase in TT (days) for a planned base TT of 40 days

Therefore, an increase in transit time of up to 10% can lead to a 20% or even greater reduction in the time charter equivalent.

Assessment of the LSD impact on fleet efficiency over the planned period. All of the above was based on a single voyage and a certain vessel, but if we consider the results of a vessel's operation on a route over a certain period of time (a year, for example), the average efficiency of the vessel can be expressed as follows:

$$TCE_s^i = \frac{\sum_{k=1}^{K_s^i} \left(F_s^{ik} - (R_{sik}^{port} + R_{sik}^{can} + R_{sik}^{bunk}) \right)}{\sum_{k=1}^{K_s^i} T_s^{ik}}, i = \overline{1, n}, s = \overline{1, S_i}. \quad (8)$$

$T_s^{ik} = T_s^{p-ik} + T_s^{m-ik}$ – is the time of the k -voyage of the s -vessel on line i .

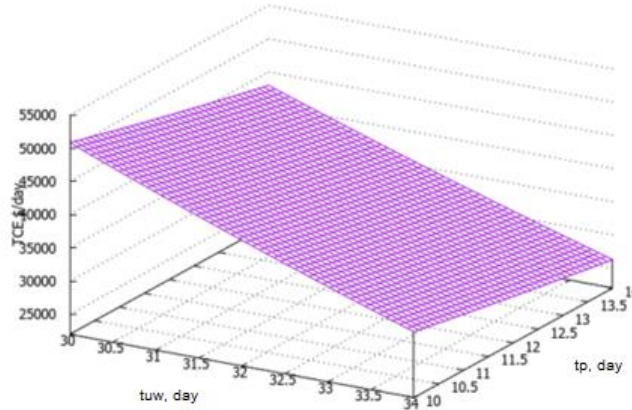


Fig. 5. Dependence of the TCE (USD/day) on underway time (days) and vessel time in port (days) for a planned TT 40 days

Therefore, by analogy with (7), the reduction in efficiency for a ship on a route over a certain period of time is estimated as:

$$\Delta TCE_s^i \left(T_s^{p-ik}, T_s^{m-ik} \mid_{k=1, K_s^i} \right) = TCE_s^{*i} - TCE_s^i \left(T_s^{p-ik}, T_s^{m-ik} \mid_{k=1, K_s^i} \right) = \frac{\sum_{k=1}^{K_s^i} \left(F_s^{ik} - (R_{sik}^{port} + R_{sik}^{can} + R_{sik}^{bunk}) \right)}{\sum_{k=1}^{K_s^i} (T_s^{p-ik} + T_s^{m-ik})}, i = \overline{1, n}, s = \overline{1, S_i}. \quad (9)$$

Average efficiency of all vessels operating on the line:

$$TCE^i = \frac{\sum_{s=1}^{S_i} \sum_{k=1}^{K_s^i} \left(F_s^{ik} - (R_{sik}^{port} + R_{sik}^{can} + R_{sik}^{bunk}) \right)}{\sum_{s=1}^{S_i} \sum_{k=1}^{K_s^i} T_s^{ik}}, i = \overline{1, n}. \quad (10)$$

It should be noted that if the vessels on the line differ significantly in size (container capacity), then for “averaging” it is necessary to take into account the share of the container capacity of the vessels in the total container capacity of the fleet operating on the line, therefore (10) is transformed as follows:

$$TCE^i = \sum_{s=1}^{S_i} \lambda_s \cdot TCE_s^i = \sum_{s=1}^{S_i} \lambda_s \left[\frac{\sum_{k=1}^{K_s^i} \left(F_s^{ik} - (R_{sik}^{port} + R_{sik}^{can} + R_{sik}^{bunk}) \right)}{\sum_{k=1}^{K_s^i} T_s^{ik}} \right], i = \overline{1, n}, \quad (11)$$

C_s – container capacity of the vessel, $\sum_{s=1}^{S_i} C_i$ – total container capacity of vessels operating on the line:

$$\lambda_s = \frac{C_s}{\sum_{s=1}^{S_i} C_i}, s = \overline{1, S_i}. \quad (12)$$

This approach allows for the unequal impact of each vessel on the average efficiency of vessels on the line.

Therefore, the average decrease in efficiency due to increased voyage time and failure to meet the schedule is determined as follows:

$$\Delta TCE^i = \sum_{s=1}^{S_i} \lambda_s \cdot \Delta TCE_s^i, i = \overline{1, n}. \quad (13)$$

This indicator characterizes the average decrease in the efficiency of ships on the line over a certain period of time, where ΔTCE_s^i – changes in the average efficiency of each ship on the line located at the base (9).

Taking into account the “daily” nature of the time charter equivalent, in absolute terms, the economic losses in efficiency amount to:

- for a specific voyage of a specific vessel on a route:

$$\Delta E_s^{ik} = \Delta TCE_s^{ik} (T_s^{p_{ik}}, T_s^{m_{ik}}) \cdot T_s^{ik} = \Delta TCE_s^{ik} (T_s^{p_{ik}}, T_s^{m_{ik}}) \cdot (T_s^{p_{ik}} + T_s^{m_{ik}}), i = \overline{1, n}, s = \overline{1, S_i}, k = \overline{1, K_s^i} \quad (14)$$

- for a specific vessel on a route during a specific period of time:

$$\Delta E_s^i = \Delta TCE_s^i (T_s^{p_{ik}}, T_s^{m_{ik}} |_{k=1, K_s^i}) \cdot \sum_{k=1}^{K_s^i} (T_s^{p_{ik}} + T_s^{m_{ik}}), i = \overline{1, n}, s = \overline{1, S_i} \quad (15)$$

- for all vessels operating on the line:

$$\Delta E^i = \Delta TCE^i \cdot T' = i = \overline{1, n} \quad (16)$$

T' – the period of time under consideration.

For example, in the case considered above, if the base flight time is 30 days, and if the durations of the five flights were 30, 30.5, 31, 31.5, and 32 days, respectively, then the average decrease in efficiency is $\Delta TCE_s^i = 2502$ USD/day, and the total decrease in efficiency for this period is $\Delta E_s^i = 387844$ USD/day.

Furthermore, if we consider the company's fleet as a whole, which operates on all routes, the total efficiency losses for the period T' are:

$$\Delta E = \sum_{i=1}^n \Delta E^i = \sum_{i=1}^n (\Delta TCE^i \cdot T') = T' \sum_{i=1}^n \Delta TCE^i \quad (17)$$

It should be noted that the consequences of increased voyage time have different significance for vessels of different sizes, taking into account the length of the route. Therefore, it is important to determine the share of a particular vessel and a particular route in the total economic losses (16-17). The following indicators can be used for this purpose:

$$I_{\Delta E}^{is} = \frac{\Delta E_s^i}{\Delta E^i}, i = \overline{1, n}, s = \overline{1, S_i}, \quad (18)$$

$$I_{\Delta E}^i = \frac{\Delta E^i}{\Delta E}, i = \overline{1, n}, \quad (19)$$

where $0 \leq I_{\Delta E}^{is} \leq 1$ – is the share of vessel s in the total losses on line i , $0 \leq I_{\Delta E}^i \leq 1$ – is the share of line i in the total losses of the carrier company. These indicators supplement the information for analysis that can be obtained on the basis of (14)-(17) for a more complete description of the situation.

Figure 6 shows a diagram that summarizes the view of the formation of economic losses of ships on the line due to failure to comply with the schedule. Thus, the diagram shows the chain of formation of absolute indicators of economic losses ΔE for each ship and line.

Results and discussions. The developed TCE framework methodology is capable not only of quantitatively determining the economic losses of a liner operator from LSD, but also, critically, provides the necessary detail, distinguishing between the economic impact of delays and deviations from announced schedules while docked in port and underway. The conclusion that delays during vessel movement (T_{mv}) are two to three times more expensive than equivalent delays in port (T_p) is key to operational strategy and should be taken into account in the context of specific issues.

The established sensitivity ratio between delays at sea and in port confirms the need for a detailed approach to managing schedule-related risks. This differentiated impact arises primarily from the fuel consumption component, which is exacerbated by the International Maritime Organization (IMO) decarbonization program. The pressure to improve energy efficiency means that any unplanned time recovery requires high-speed steaming, which directly contradicts emissions targets and incurs significant, often unplanned, bunkering costs. This conclusion is strongly supported by the analysis of Meng et al. (2023) [5], who highlight the complex trade-offs between schedule reliability, bunker consumption, and necessary speed adjustments under the Carbon Intensity Indicator (CII). Our results provide a concrete financial multiplier that justifies why minimizing unexpected speed adjustments is economically preferable to enduring minor port congestion.

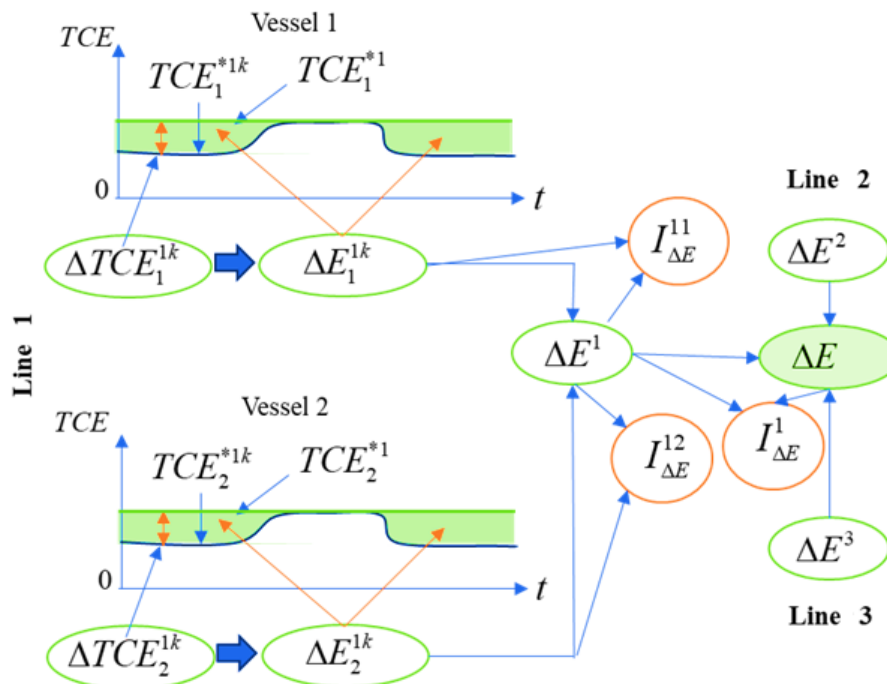


Fig. 6. Formation of economic losses of ships in a linear chain due to failure to comply with the schedule

The quantified severity of economic losses is consistent with recent studies that highlight the high systemic cost of unreliability. The most analysed articles on LSD-problem emphasize the broad systemic nature of the risk of schedule disruption, and our concept transforms this abstract risk into a concrete, measurable financial indicator. By providing daily losses, the model transforms the LSD problem from operational inefficiency to direct financial liability. This diagnostic tool promotes the idea of ensuring supply chain resilience by enabling real-time decisions based on whether the cost of restoring the schedule exceeds the projected losses in TCE.

Furthermore, the aggregated loss metric is instrumental in advancing strategic fleet planning. Traditional models for network design often rely on generalized or historical cost averages. The high-resolution operational cost data generated by the TCE framework can refine strategic planning exercises, providing more realistic inputs for optimal fleet deployment and network configuration. The consistent application of this TCE-based methodology, therefore, enables LS companies to move beyond simply acknowledging schedule unreliability to actively attributing and monetizing the financial compromises across their entire operational portfolio, ensuring strategic decisions are grounded in real-world economic impacts.

Conclusion. The difference between the planned time charter equivalent and the actual time charter equivalent forms changes in the time charter equivalent, which justifies changes in efficiency over a certain period of time, which, in general, for all lines and vessels, forms an integral indicator of economic losses.

Thus, failure to comply with the line schedule in the context of delays, which is associated with a variety of factors, leads to a decrease in the efficiency of the vessel, as measured by the time charter equivalent. In this situation, the resource of vessels – their carrying capacity over a certain period of time – decreases, which leads to a decrease in daily and overall efficiency. Monitoring changes in efficiency is the basis for taking appropriate organizational measures to ensure the necessary (planned) level of efficiency.

Bibliography:

1. Gurning R.O.S. Maritime disruptions in the Australian-Indonesian wheat supply chain: An analysis of risk assessment and mitigation strategies. University of Tasmania, 2011. Thesis. DOI: <https://doi.org/10.25959/23210093.v1>
2. Elmi Z., Singh P., Meriga V. K., Goniewicz K., Borowska-Stefańska M., Wiśniewski S., Dulebenets M.A. Uncertainties in Liner Shipping and Ship Schedule Recovery: A State-of-the-Art Review. *Journal of Marine Science and Engineering*. 2022. Vol. 10, № 5. P. 563. DOI: <https://doi.org/10.3390/jmse10050563>
3. Zhou J., Zhao Y., Yan X., Wang M. Strategy and Impact of Liner Shipping Schedule Recovery under ECA Regulation and Disruptive Events. *Journal of Marine Science and Engineering*. 2024. Vol. 12, № 8. P. 1405. DOI: <https://doi.org/10.3390/jmse12081405>
4. Gui D., Wang H., Yu M. Risk Assessment of Port Congestion Risk during the COVID-19 Pandemic. *Journal of Marine Science and Engineering*. 2022. Vol. 10. P. 150. DOI: <https://doi.org/10.3390/jmse10020150>

5. Meng L., Wang X., Jin J., Han C. Optimization Model for Container Liner Ship Scheduling Considering Disruption Risks and Carbon Emission Reduction. *Journal of Marine Science and Engineering*. 2023. Vol. 11, № 7. P. 1449. DOI: <https://doi.org/10.3390/jmse11071449>
6. Baştuğ S., Haralambides H., Akan E., Kiraci K. Risk mitigation in service industries: A research agenda on container shipping. *Transport Policy*. 2023. Vol. 141. P. 232–244. DOI: <https://doi.org/10.1016/j.tranpol.2023.07.011>
7. Vernimmen B., Dullaert W., Engelen S. Schedule unreliability in liner shipping: Origins and consequences for the hinterland supply chain. *Maritime Economics & Logistics*. 2007. Vol. 9, № 3. P. 193–213.
8. Notteboom T. E. The time factor in liner shipping services. *Maritime Economics and Logistics*. 2006. Vol. 8. P. 19–39. DOI: <http://dx.doi.org/10.1057/palgrave.mel.9100148>
9. Brouer B. D., Karsten C. V., Pisinger D. Optimization in liner shipping. *4OR*. 2017. Vol. 15, № 1. P. 1–35. DOI: <https://doi.org/10.1007/s10479-018-3023-8>
10. Mulder J., Dekker R. Methods for strategic liner shipping network design. *European Journal of Operational Research*. 2014. Vol. 235, № 2. P. 367–377. DOI: <https://doi.org/10.1016/j.ejor.2013.09.041>
11. Wilmsmeier G., Notteboom T. Determinants of liner shipping network configuration: a two-region comparison. *GeoJournal*. 2011. Vol. 76, № 3. P. 213–228. DOI: <https://doi.org/10.1007/s10708-009-9333-2>
12. Wilmsmeier G., Hoffmann J. Liner shipping connectivity and port infrastructure as determinants of freight rates in the Caribbean. *Maritime Economics and Logistics*. 2008. Vol. 10. P. 130–151. DOI: <https://doi.org/10.1186/s41072-017-0019-5>
13. Márquez Ramos L., Martínez Zarzoso I., Pérez García E., Wilmsmeier G. The interrelationship of maritime network connectivity, transport costs and maritime trade. 14th Annual Congress of the International Association of Maritime Economists (IAME), Melbourne, 3–5 July 2006.
14. Kim H. J., Son D. H., Yang W., Kim J. G. Liner ship routing with speed and fleet size optimization. *KSCE Journal of Civil Engineering*. 2019. Vol. 23, № 3. P. 1341–1350. DOI: <https://doi.org/10.1007/s12205-019-0564-6>
15. Silva L.M.R., Wang H., Soares C. G. A strategic fleet size and mix vehicle routing model to analyse the impact of demand fluctuation on river-sea liner shipping. *Ocean Engineering*. 2024. Vol. 307. P. 118096. DOI: <https://doi.org/10.1016/j.oceaneng.2024.118096>
16. Del Rosal I. Trade effects of liner shipping across world regions. *Maritime Business Review*. 2024. Vol. 9, № 1. P. 2–16. DOI: <https://doi.org/10.1016/10.1108/MABR-06-2023-0040>
17. Baird A. J., Lindsay A. J. Strategic choice in the global container shipping industry: A resource-based approach. Proceedings of the 1996 IAME Conference. Vancouver, 1996.
18. Robinson R. Liner shipping strategy, network structuring and competitive advantage: a chain systems perspective. *Research in Transportation Economics*. 2005. Vol. 12, № 1. P. 247–289. DOI: [https://doi.org/10.1016/S0739-8859\(04\)12008-8](https://doi.org/10.1016/S0739-8859(04)12008-8)
19. Tavasszy L., Minderhoud M., Perrin J.F., Notteboom T. A strategic network choice model for global container flows: specification, estimation and application. *Journal of Transport Geography*. 2011. Vol. 19, № 6. P. 1163–1172. DOI: <https://doi.org/10.1016/j.jtrangeo.2011.05.005>
20. D'agostini E., Nam H.-S., Kang S.-H. Gaining Competitive Advantage at Sea: An Overview of Shipping Lines' Strategic Decisions. *International Journal of Transportation Engineering and Technology*. 2019. Vol. 5, № 4. P. 74–81. DOI: <https://doi.org/10.11648/j.ijtet.20190504.12>
21. Meng Q., Wang S., Andersson H., Thun K. Containership routing and scheduling in liner shipping: Overview and future research directions. *Transportation Science*. 2014. Vol. 48, № 2. P. 265–280. DOI: <https://doi.org/10.1287/trsc.2013.0461>
22. Dulebenets M. A., Pasha J., Abioye O. F. et al. Vessel scheduling in liner shipping: a critical literature review and future research needs. *Flexible Services and Manufacturing Journal*. 2021. Vol. 33. P. 43–106. DOI: <https://doi.org/10.1007/s10696-019-09367-2>

References:

1. Gurning, ROS (2011). Maritime disruptions in the Australian-Indonesian wheat supply chain: An analysis of risk assessment and mitigation strategies. University of Tasmania. Thesis. <https://doi.org/10.25959/23210093.v1>
2. Elmi, Z., Singh, P., Meriga, V. K., Goniewicz, K., Borowska-Stefańska, M., Wiśniewski, S., & Dulebenets, M. A. (2022). Uncertainties in Liner Shipping and Ship Schedule Recovery: A State-of-the-Art Review. *Journal of Marine Science and Engineering*, 10(5), 563. <https://doi.org/10.3390/jmse10050563>
3. Zhou, J., Zhao, Y., Yan, X., & Wang, M. (2024). Strategy and Impact of Liner Shipping Schedule Recovery under ECA Regulation and Disruptive Events. *Journal of Marine Science and Engineering*, 12(8), 1405. <https://doi.org/10.3390/jmse12081405>
4. Gui, D.; Wang, H.; Yu, M. (2022). Risk Assessment of Port Congestion Risk during the COVID-19 Pandemic. *J. Mar. Sci. Eng.* 10, 150. <https://doi.org/10.3390/jmse10020150>
5. Meng, L., Wang, X., Jin, J., & Han, C. (2023). Optimization Model for Container Liner Ship Scheduling Considering Disruption Risks and Carbon Emission Reduction. *Journal of Marine Science and Engineering*, 11(7), 1449. <https://doi.org/10.3390/jmse11071449>

-
6. Baştuğ, S., Haralambides, H., Akan, E., Kiraci, K. (2023) Risk mitigation in service industries: A research agenda on container shipping, *Transport Policy*, Volume 141. 232–244, <https://doi.org/10.1016/j.tranpol.2023.07.011>.
 7. Vernimmen, B., Dullaert, W., & Engelen, S. (2007). Schedule unreliability in liner shipping: Origins and consequences for the hinterland supply chain. *Maritime Economics & Logistics*, 9(3), 193–213.
 8. Notteboom, T. E. (2006). The time factor in liner shipping services. *Maritime Economics and Logistics*, 8, 19–39. <http://dx.doi.org/10.1057/palgrave.mel.9100148>
 9. Brouer, B. D., Karsten, C. V., & Pisinger, D. (2017). Optimization in liner shipping. *4OR*, 15(1), 1–35. <https://doi.org/10.1007/s10479-018-3023-8>
 10. Mulder, J., & Dekker, R. (2014). Methods for strategic liner shipping network design. *European Journal of Operational Research*, 235(2), 367–377. <https://doi.org/10.1016/j.ejor.2013.09.041>
 11. Wilmsmeier, G., & Notteboom, T. (2011). Determinants of liner shipping network configuration: a two-region comparison. *GeoJournal*, 76(3), 213–228. <https://doi.org/10.1007/s10708-009-9333-2>
 12. Wilmsmeier, G., & Hoffmann, J. (2008). Liner shipping connectivity and port infrastructure as determinants of freight rates in the Caribbean. *Maritime Economics and Logistics*, 10, 130–151. <https://doi.org/10.1186/s41072-017-0019-5>
 13. Márquez Ramos, L., Martínez Zarzoso, I., Pérez Garcia, E. & Wilmsmeier, G. (2006). The interrelationship of maritime network connectivity, transport costs and maritime trade. 14th annual congress of the international association of maritime economists (IAME) in Melbourne, July 3-5 2006.
 14. Kim, H. J., Son, D. H., Yang, W., & Kim, J. G. (2019). Liner ship routing with speed and fleet size optimization. *KSCE Journal of Civil Engineering*, 23(3), 1341–1350. <https://doi.org/10.1007/s12205-019-0564-6>
 15. Silva, L. M. R., Wang, H., & Soares, C. G. (2024). A strategic fleet size and mix vehicle routing model to analyse the impact of demand fluctuation on river-sea liner shipping. *Ocean Engineering*, 307, 118096. <https://doi.org/10.1016/j.oceaneng.2024.118096>
 16. Del Rosal, I. (2024). Trade effects of liner shipping across world regions. *Maritime Business Review*, 9(1), 2–16. <https://doi.org/10.1016/10.1108/MABR-06-2023-0040>
 17. Baird, A. J., A. J. Lindsay. (1996). Strategic choice in the global container shipping industry: A resource-based approach. Proceedings of the 1996 IAME Conference, Vancouver.
 18. Robinson, R. (2005). Liner shipping strategy, network structuring and competitive advantage: a chain systems perspective. *Research in Transportation Economics*, 12(1), 247–289. [https://doi.org/10.1016/S0739-8859\(04\)12008-8](https://doi.org/10.1016/S0739-8859(04)12008-8)
 19. Tavasszy, L., Minderhoud, M., Perrin, J. F., & Notteboom, T. (2011). A strategic network choice model for global container flows: specification, estimation and application. *Journal of transport geography*, 19(6), 1163–1172. <https://doi.org/10.1016/j.jtrangeo.2011.05.005>
 20. Enrico D’agostini, Hyung-Sik Nam, Seog-Hwan Kang. Gaining Competitive Advantage at Sea: An Overview of Shipping Lines’ Strategic Decisions. *International Journal of Transportation Engineering and Technology*. Vol. 5, No. 4, 2019, pp. 74–81. <https://doi.org/10.11648/j.ijtet.20190504.12>
 21. Meng, Q., Wang, S., Andersson, H., Thun, K. (2014). Containership routing and scheduling in liner shipping: Overview and future research directions. *Transportation Science*, 48 (2). 265–280. DOI: <https://doi.org/10.1287/trsc.2013.0461>
 22. Dulebenets, M. A., Pasha, J., Abioye, O. F. et al. (2021). Vessel scheduling in liner shipping: a critical literature review and future research needs. *Flex Serv Manuf J* 33, 43–106. <https://doi.org/10.1007/s10696-019-09367-2>

Дата надходження статті: 03.10.202

Дата прийняття статті: 10.11.2025

Опубліковано: 30.12.2025

Коскіна Ю. О., доктор технічних наук, професор,
професор кафедри експлуатації флоту і
технології морських перевезень
Одеського національного морського університету
ORCID: 0000-0002-3164-6504

ВИЗНАЧЕННЯ ТРИВАЛОСТІ ПРОСТОГО РЕЙСУ СУДНА НА УМОВАХ ДОГОВОРУ ФРАХТУВАННЯ НА РЕЙС

Стаття присвячена порядку визначення елементів та загальної тривалості рейсу судна відповідно до умов договору фрахтування судна на рейс. Підставою виконання судном нерегулярного плавання рейсу з перевезення вантажу є укладений між судновласником-перевізником і фрахтувальником договір фрахтування судна на рейс, який за своєю сутністю є договором про перевезення вантажу. Умови такого договору багато у чому визначають тривалість рейсу, яка для судновласника-перевізника є певним орієнтиром зайнятості судна за укладеним договором, а фрахтувальнику дозволяють оцінити терміни перевезення вантажу на морській ділянці його доставки. До розгляду покладено простий рейс судна, який виконується на умовах «1 порт завантаження – 1 порт розвантаження». Рейс судна розглянуто у статті у вигляді окремих складових – переходу у баласті (за необхідності виконання), стоянки у порту завантаження, переходу із вантажем та стоянки у порту розвантаження. Подано порядок їх розрахунку із посиланням на умови договору фрахтування на рейс, які використовуються у відповідних калькуляціях як вихідні дані. Серед таких умов принципове значення для визначення тривалості переходів судна мають порти завантаження і розвантаження, якими визначається відстань переходу судна із вантажем, та маршрут виконання переходу. Наголошено на важливості врахування умов договору фрахтування на рейс щодо термінів та своєчасності приходу судна до порту завантаження, вказаних у договорі фрахтування на рейс датами laydays і cancelling. При розрахунках часу стоянки судна у портах під відповідними вантажними роботами важливе значення мають погоджені у договорі фрахтування на рейс норми вантажних робіт, порядок рахування вихідних та святкових днів, умови щодо початку відрахунку сталійного часу. Багатоваріантність формулювань названих умов договору фрахтування судна на рейс може вимагати внесення певних корективів, наразі принциповий порядок розрахунку тривалості елементів і загального часу рейсу залишається незмінним.

Ключові слова: простий рейс судна, тривалість рейсу, елементи часу рейсу, умови договору фрахтування на рейс, рейсове фрахтування, рейсовий чартер.

Koskina Yu. O. Determining the time of a vessel's single voyage under the terms of a voyage charter-party

The paper is devoted to the procedure for determining the elements and total voyage time in accordance with the terms of a voyage charter-party. The basis for the performance of a voyage for the carriage of goods by a tramp vessel is a voyage charter party concluded between the shipowner-carrier and the charterer, which is in essence a contract for the carriage of goods. The terms of such a contract largely determine the duration of the voyage, which for the shipowner-carrier is a certain benchmark for the employment of the vessel under the contract concluded and for the charterer allows to estimate the time of cargo transportation on the sea leg of its delivery.

The paper considers a single voyage of a vessel, which is carried out on the terms "1 port of loading – 1 port of discharging". The voyage of a vessel is considered in the paper as consisted of separate components – a ballast passage (if necessary), the vessel is at the port of loading, passage with cargo and staying at the port of discharging. The procedure for their calculation is presented with reference to the terms of the voyage charter party, which are used in the relevant calculations as initial data. Among these terms, the ports of loading and unloading are of fundamental importance for determining the duration of the vessel's passages, as far as they determine the distance of the vessel's passages with cargo and also the route of the voyage. It is emphasized the importance of considering the terms of the voyage charter party regarding the timing of the vessel's arrival at the port of loading, as specified in the voyage charter party by the laydays and cancelling dates. When calculating the time a vessel is at berth in ports for the relevant cargo operations, the cargo operations rates agreed in the voyage charter party, the procedure for counting days off and holidays, and the terms for beginning the laytime running are of great importance. The varied wording of these terms of the voyage charter party may require certain adjustments, but the basic procedure for calculating the duration of the elements and the total voyage time remains unchanged.

Key words: vessel's single voyage, voyage time, elements of voyage time, voyage charter-party terms, voyage chartering, voyage charter-party.

Постановка проблеми. Погоджені сторонами та зафіксовані у договорі фрахтування судна на рейс (рейсовому чартері), умови виконання рейсу є підставами для оцінки судновласником часу виконання рейсу, що у свою чергу, визначає не лише грошові, а і часові витрати – строк зайнятості судна на виконанні певного рейсового чартеру, а відтак – і певні можливості планування (наскільки воно взагалі є можливим у trampовому судноплаванні) подальшої роботи судна – укладання рейсового чартеру з відфрахтування судна у наступний рейс. Визначити точний час рейса судна можна лише пост-фактум, після його виконання, оскільки на тривалість його виконання впливає численна кількість факторів, врахувати які на етапі попередніх розрахунків не є можливим. Втім, виходячи з умов виконання рейсу, зафіксованих у договорі фрахтування судна на рейс, судновласник має можливість більше або менше орієнтовно визначити час зайнятості судна, яке він відфрахтовує, на перевезенні вантажу.

Аналіз останніх досліджень і публікацій. Підставами для визначення загального часу рейсу є визначення тривалості окремих його елементів, порядок визначення яких у найбільш узагальнено вигляді подано у [1-4]. Окремі автори вдаються до використання результатів розрахунків тривалості окремих елементів часу рейсу для вирішення конкретних експлуатаційних задач [5-12] або як підстави економічної оцінки рейсу судна [13-17]. Згадані публікації здебільшого ґрунтуються на основних положеннях теорії транспортних процесів і систем, наразі не враховують специфіку роботи суден на умовах рейсового фрахтування – домовленості у конкретному рейсовому чартері, саме якими власне і визначаються як окремі вихідні дані, необхідні для виконання попередніх калькуляцій, так і власне структура рейсу. Умовам договорів рейсового фрахтування присвячені доволі детальні дослідження, наукові та навчально-методичні публікації [18-21], наразі викладення матеріалу носить теоретико-пояснювальний характер, а автори обмежуються лише формулюваннями щодо загального впливу умов договору фрахтування на рейс на тривалість його виконання.

Тож **метою статті** є структуризація рейсу судна з перевезення вантажу, який виконується на умовах договору фрахтування на рейс, та формування порядку визначення тривалості окремих його складових і загального часу рейсу, виходячи із умов його виконання, зафіксованих у договорі фрахтування на рейс, який укладено судновласником і фрахтувальником.

Виклад основного матеріалу дослідження. Основний виробничий процес роботи судна – рейс – складається з декількох основних етапів, які є послідовними, взаємопов'язаними та взаємообумовленими: завантаження вантажу на судно; перехід із вантажем з порту завантаження до порту розвантаження; розвантаження вантажу у відповідному порту. За необхідності здійснення переходу у баласті з порту поточного місцезнаходження судна до порту завантаження такий баластний перехід вважається частиною того рейсу, який вимагає його виконання [1-2]. Логічну послідовність етапів виконання рейсу подано на рис. 1, де використано наступні умовні позначення:

- перехід судна у баласті p_{x_6} із часовими точками його початку $k_{x_{6n}}$ і закінчення $k_{x_{63}}$, який характеризується тривалістю t_{x_6} ;
- стоянка судна у порту завантаження p_{cm_3} із часовими точками її початку $k_{cm_{3n}}$ та закінчення $k_{cm_{33}}$, яка характеризується тривалістю t_{cm_3} ;
- перехід судна із вантажем p_{x_6} із часовими точками його початку $k_{x_{6n}}$ і закінчення $k_{x_{63}}$, який характеризується тривалістю t_{x_6} ;
- стоянка судна у порту завантаження p_{cm_p} із часовими точками її початку $k_{cm_{pn}}$ та закінчення $k_{cm_{p3}}$, яка характеризується тривалістю t_{cm_p} .

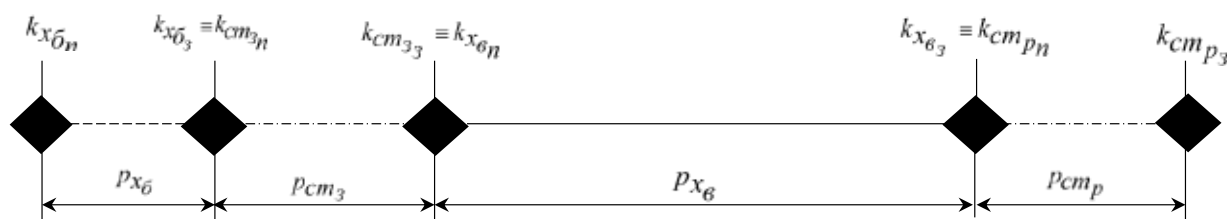


Рис. 1. Складові простого рейсу судна

Специфіка роботи суден у trampовому судноплаванні (зокрема – відсутність «прив'язки» роботи судна на перевезеннях вантажів між постійними портами) полягає у характерній практиці наявності баластних переходів, тому, досліджуючи структуру рейсу, який виконується на умовах рейсового чартеру, виходити-мемо з наявності такого переходу у структурі рейсу:

$$p_{x_6} = [k_{x_{6n}}; k_{x_{63}}], \quad (1)$$

а його тривалість визначається відношенням відстані переходу до швидкості руху судна на переході:

$$t_{x_6} = \frac{L_6}{V_{c_6}}, \quad (2)$$

де L_6 – відстань переходу судна у баласті, миль;
 V_{c_6} – швидкість руху судна у баласті, миль/добу.

Відстань переходу у баласті є відстанню між поточним місцезнаходженням судна (як правило, за трампової організації роботи суден – останній порт розвантаження у попередньому рейсі) та портом завантаження, вказаним у договорі рейсового фрахтування. До певних складнощів у розрахунках тривалості переходу у баласті можуть призвести невизначеності із точною назвою порту завантаження, адже умови договору рейсового фрахтування дають фрахтувальнику право позначити його як такий, що розташований на ренджі. Втім, навіть за такого варіанту вказівки порту завантаження, відстань переходу у баласті можна визначити більш або менш точно, поклавши у розрахунки відстань до одного (будь-якого) порту названого ренджа.

Саме цією тривалістю переходу у баласті визначається дата фактичного приходу судна до порту завантаження:

$$k_{x_{\delta_3}} \equiv k_{n_3}, \quad (3)$$

де k_{n_3} – часова точка приходу судна до порту завантаження.

Оскільки умовами укладеного договору фрахтування судна на рейс термін прибуття судна до порту завантаження фіксується lausan'ом, на переході у баласті мають враховуватися часові обмеження з прибуття судна до порту завантаження, зафіксовані у рейсовому чартері у вигляді дат laydays і cancelling:

$$k_l \leq k_{n_3} \leq k_{can}, \quad (4)$$

де k_l та k_{can} – часові точки відповідно laydays і cancelling згідно до умов договору фрахтування на рейс.

Однак виконання розрахунку тривалості баластного переходу судна за (1) є доречним, якщо очевидно є ситуація, за якої дата закінчення виконання переходу у баласті, якою і визначатиметься дата приходу до порту завантаження, знаходиться у межах lausan. Втім, на практиці можливими є ситуації, коли навіть без виконання розрахунків тривалості переходу судна у баласті очевидно, що час виконання баластного переходу є таким, що його закінчення і фактичний прихід до порту завантаження передуватимуть даті laydays. Класичні умови рейсового фрахтування не зобов'язують фрахтувальника приймати судна для завантаження до дати laydays, тому за приходу судна до вказаної у чартері дати laydays судно ризикує простоюватиме, очікуючи дату laydays. Наразі, для запобігання такого ризику судовласник має можливість зменшити швидкість руху судна на переході у баласті: фактично це означає, що швидкість руху судна на переході у баласті може визначатися «від зворотного», тобто через необхідну тривалість переходу до порту завантаження, виходячи з дати laydays. Тоді

$$p_{x_6} = [k_{x_{\delta_n}}; k_l], \quad (5)$$

і саме з урахуванням дати laydays визначається швидкість, з якою має рухатися судно на переході, щоб прибути до порту завантаження не раніше за дату laydays:

$$V'_{c_6} = \frac{L_6}{t'_{x_6}}, \quad (6)$$

де V'_{c_6} – швидкість, із якою має рухатися судно на переході у баласті задля забезпечення приходу судна у порт завантаження у дату laydays, миль/добу;

t'_{x_6} – період часу між датою звільнення судна від зобов'язань за попереднім рейсом та датою laydays, діб.

Слід зазначити, що варіювання швидкістю руху обмежено умовами технічної експлуатації суден, що має враховуватися за використання (6) при виконанні переходу у баласті:

$$V_{\min} \leq V'_{c_6} \leq V_{\max}, \quad (7)$$

де $V_{\min}$ і $V_{\max}$ – відповідно мінімальні та максимально припустимі значення швидкостей за умовами експлуатації двигуна, миль/добу.

Цілком можливим є і варіант, коли за орієнтовними оцінками часу виконання переходу у баласті судно запізнюється до дати cancelling. За такої ситуації судовласник має можливість «пришвидшити» судно,

збільшивши швидкість судна на переході у баласті, до значення, яке забезпечить приход судна до порту завантаження у дату cancelling. Тоді

$$p_{x_6} = [k_{x_{6n}}; k_{can}] \quad (8)$$

за швидкості

$$V_{c_6}'' = \frac{L_6}{t_{x_6}''}, \quad (9)$$

де V_{c_6}'' – швидкість, із якою має рухатися судно на переході у баласті задля забезпечення приходу судна у порт завантаження у дату cancelling, миль/добу;

t_{x_6}'' – період часу між датою звільнення судна від зобов'язань за попереднім рейсом та датою cancelling, діб
і дотримання

$$V_{\min} \leq V_{c_6}'' \leq V_{\max}. \quad (10)$$

Втім, умови рейсового фрахтування дозволяють судновласнику, за запізнення судна у порт завантаження до дати cancelling відправити фрахтувальнику запит щодо її переносу, а останній має дати відповідь протягом певної кількості годин щодо своє рішення щодо згоди на перенос дати cancelling або скасування чартер-партії.

Точна дата та час прибуття судна до порту завантаження у межах погодженого lausan'у поточнюється шляхом подання судном під час руху до порту завантаження відповідних повідомлень на адресу фрахтувальника – нотісів про очікуваний час прибуття судна (ETA – estimate time of arrival), кількість та періодичність яких фіксується у договорі рейсового фрахтування як результат домовленостей сторін.

Момент фактичного прибуття судна до порту завантаження і готовності до прийому вантажу знаходиться у межах lausan'у

$$k_{n_3} \in [k_l; k_{can}], \quad (11)$$

та у документальному сенсі фіксується моментом подання капітаном нотіса про готовність судна до завантаження (NOR – notice of readiness). Основний обов'язок фрахтувальника на цьому етапі полягає у забезпеченні повній підготовленості вантажу (фізична наявність у порту та документальне оформлення) до відправлення на експорт та завантаження на судно.

Безпосереднє завантаження вантажу на судно починається, як правило, зі сплином так званого пільгового періоду (grace period) після подання нотісу про готовність судна до вантажних робіт. Його тривалість частково залежить від умов договору фрахтування та, як правило, при поданні нотісу про готовність судна до завантаження у першій половині дня, рахування сталійного часу починається з обіда поточного дня; за подання вказаного нотісу у другій половині дня – з ранку наступного робочого дня [22]. Втім, можливими є і інші варіанти домовленостей щодо початку рахування сталійного часу, наразі тривалість grace period, як правило, не є значною, тому

$$k_{x_{63}} \equiv k_{n_3} \equiv k_{cm_{n_3}}. \quad (12)$$

Як зазначено раніше, на момент подання судном згаданого нотісу вантаж має бути підготовленим для початку розміщення його на судні, що означає виконання відповідних експортних формальностей та оформлення дозвільних документів. Адже рахування сталійного часу і початок виконання вантажних робіт, тривалість яких визначається чартерними домовленостями про сталійний час, мають відбутися одночасно.

Чартерні умови про сталійний час фіксують їх тривалість через добові норми завантаження, що, за загальної інформації про певну кількість вантажу, яка перевозитиметься судном та яку наведено у чартері у вигляді базисної кількості, дозволяє судновласнику із більшою або меншою точністю визначити кількість сталійних днів, призначених для завантаження вантажу:

$$t_{cmal_3} = \frac{Q}{M_3}, \quad (13)$$

де t_{cmal_3} – сталійний час завантаження вантажу, діб;

Q – кількість вантажу, вказана у рейсовому чартері, т;

M_3 – норма завантаження вантажу, вказана у чартері, т/судно-добу.

Точна кількість вантажу, яку судно може взяти на борт, залежить від чистої вантажопідйомності судна на момент знаходження у порту завантаження; порядок визначення кількості вантажу через чисту вантажопідйомність судна із врахуванням умов договору фрахтування судна на рейс викладено зокрема у [23].

Інтенсивність вантажних робіт з завантаження, яка визначається добовими нормами завантаження вантажу, має оцінюватиметься судновласником саме виходячи з тієї норми завантаження, яка вказана у рейсовому чартері, незалежно від його інформованості щодо фактичних норм завантаження вантажу, який перевозитиметься, у порту завантаження. Це пояснюється тим, що відповідно до умов рейсового фрахтування, звичайною практикою є організація та виконання вантажних робіт фрахтувальником, який наразі і відповідає перед судновласником за дотримання погоджених у договорі рейсового фрахтування термінів виконання завантаження вантажу. Така відповідальність має грошовий характер – у будь-якому рейсовому чартері є умови щодо контрсталійного часу і демереджа, саме тому фрахтувальник і має забезпечити одночасність моментів початку рахування сталійного часу та початку фактичного завантаження вантажу, як зазначено раніше.

Кількість сталійних днів, визначена відповідно до умов договору фрахтування за (3), є для судновласника лише орієнтиром у визначенні тривалості стояночного часу, який наразі є більшим, ніж тривалість сталійного часу, оскільки враховує, окрім тривалості безпосереднього виконання вантажних робіт, тривалість grace period, час виконання процедур і формальностей, пов'язаних із приходом судна та його відходом з порту, а також період часу, коли вантажні роботи не виконуються. До останніх за умовами договору рейсового фрахтування можуть відноситися домовленості про виключення вихідних та святкових днів з рахування сталійного часу за чартерної домовленості SHEx (Sundays and Holidayas Excepted). Фактичний час знаходження судна у порту під вантажними роботами, який власне і цікавить судновласника для визначення тривалості відповідної складової рейсу у більшому ступені, ніж тривалість сталійного часу (хоча наразі саме остання і є основою формування стояночного часу), значно залежить також і від погодженого порядку рахування сталійного часу – рахування чи нерахування вихідних та святкових днів як сталійних. Це може подовжити час знаходження судна у порту порівняно із калькуляціями сталійного часу. Наразі на етапі попередніх часових оцінок тривалості рейсу саме умови про сталійний час є єдиними вихідними даними, якими може оперувати судновласник.

Процес безпосереднього виконання вантажних робіт $p_{стал_3}$ тривалістю $t_{стал_3}$ характеризується точками його початку $k_{стал_3_n}$ та закінчення $k_{стал_3_3}$, при чому

$$\left[k_{стал_3_n}; k_{стал_3_3} \right] \in \left[k_{см_3_n}; k_{см_3_3} \right], \quad (14)$$

і відповідно

$$t_{стал_3} < t_{см_3}. \quad (15)$$

Із закінченням розміщення вантажу на судні капітан судна має підписати коносамент, функція якого у цьому сенсі полягає у підтвердженні прийому судном певної кількості визначеного вантажу. Фактично підписання коносаменту фіксує закінчення завантаження вантажу, і після оформлення документів та проходження процедур і формальностей, пов'язаних із відходом судна, судно залишає порт завантаження. Аналогічно до тривалості відповідних процедур на прибутті судна до порту завантаження, їх тривалість на відході є незначною.

Стояночний час завантаження як складова частина рейсу судна, враховуючи викладене вище, може бути сформульований наступний чином:

$$p_{см_3} = \left[k_{см_3_n}; k_{см_3_3} \right], \quad (16)$$

при чому

$$k_{п_3} \equiv k_{см_3_n}, \quad (17)$$

і

$$k_{см_3_3} \equiv k_{\theta_3}, \quad (18)$$

де k_{θ_3} – часова точка виходу судна з порту завантаження, а структура стояночного часу у порту завантаження із ключовими точками окремих складових подана на рис. 2.

Загалом тривалість часу знаходження судна у порту завантаження (стояночного часу) може бути визначена як:

$$t_{см_3} = t_{gp} + \frac{Q}{M_3} + t_{SHEx} + t_{\phi_3}, \quad (19)$$

де t_{gp} – тривалість періоду часу між приходом судна до порту завантаження та початком виконання вантажних робіт, діб;

t_{SHEx} – тривалість неробочих періодів часу, які, відповідно до чартерної домовленості SHEx не рахуються як сталійний час, діб; за протилежної чартерної домовленості SHInc (Sundays Holidays Included)

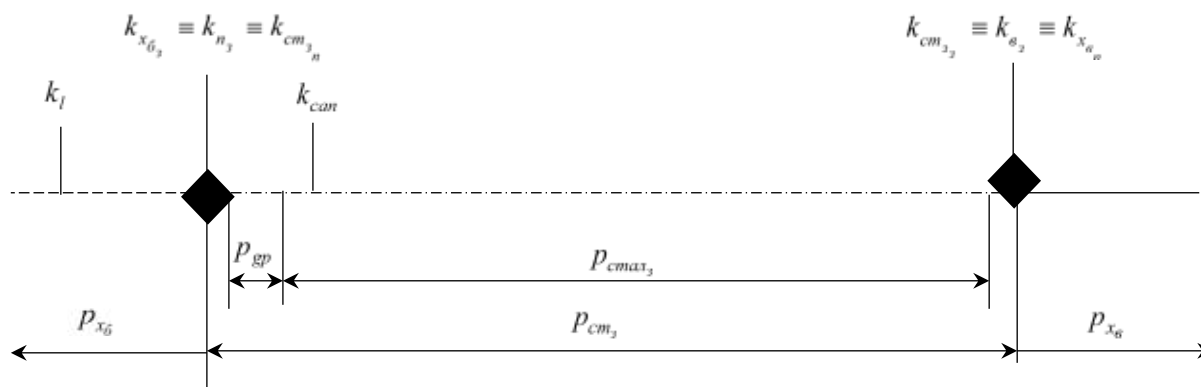


Рис. 2. Структура стояночного часу у порту завантаження

$$t_{SHEx} = 0 \quad (20)$$

t_{ϕ_3} – тривалість оформлення документів та виконання процедур і формальностей, пов’язаних із відходом судна з порту завантаження, діб.

Очевидно, що визначення стояночного часу рейсу у порту завантаження за (19), окрім чартерних домовленостей, багато у чому визначається датою та часом приходу судна k_{n_3} , на чому наголошено зокрема у [24-25] і точно встановити яку на етапів попередніх калькуляцій, виходячи лише з умов договору фрахтування на рейс, є вкрай складним. Викладені у [25] наразі дозволяють із певною часткою умовності врахувати тривалість окремих елементів стояночного часу рейсу судна. Тим не менш, калькуляції стояночного часу рейсу лише відповідно до умов договору фрахтування на рейс носять характер орієнтовних – таких, що дозволяють надати лише приблизну оцінку часу стоянки судна у порту завантаження.

Закінчення виконання формальностей, пов’язаних із відходом судна з порту завантаження, фактично є початком наступного етапу виконання рейсу – переходу із вантажем до порту розвантаження:

$$p_{x_6} = [k_{x_{en}}; k_{x_{e3}}] \quad (21)$$

$$k_{\theta_n} \equiv k_{\theta_3} \quad (22)$$

тривалість якого можна визначити, аналогічно визначення тривалості переходу судна у баласті (2) із використанням відповідних вихідних даних – відношенням відстані переходу (відстані між портом завантаження і портом розвантаження) та швидкості руху судна на переході із вантажем.

Терміни виконання судном такого переходу у рейсовому чартері не фіксуються жодними датами або часовими інтервалами. У договорах рейсового фрахтування є лише застереження загального характеру щодо необхідності руху судна до порту розвантаження найкоротшим навігаційно рекомендованим маршрутом із розумною швидкістю, а відхилення від курсу дозволяється лише за виключних ситуацій/обставин. Відсутність жорстких обмежень щодо швидкості руху судна дає судовласникам певні можливості варіювати швидкістю з метою зменшення витрат на паливо – зокрема, у періоди високих цін на бункер вони активно вдаються до так званого slow steaming, суттєво заощаджуючи на паливних витратах [3]. Втім, варіювання швидкості руху судна на переходах із вантажем має враховувати мінімально і максимально припустимі величини швидкостей судна, про що згадано раніше.

Відстань переходу судна із вантажем визначається відстанню між портами завантаження і розвантаження, і, аналогічно до визначення відстані переходу судна у баласті, перед судовласником може постати проблема із визначенням точної відстані між портами, якщо вони вказані у чартері як такі, що розташовані на певному ренджі. За такої ситуації у попередні калькуляції тривалості переходу судна із вантажем відстань між портами можна покласти, прийнявши до розгляду певний порт, розташований на названому ренджі.

Аналогічно до переходу судна до порту завантаження, рухаючись до порту розвантаження, судно має подавати на адресу зацікавлених осіб повідомлення із вказівкою очікуваної дати та часу прибуття судна (ETA – estimate time of arrival). Так само, із прибуттям судна під розвантаження воно має подати відповідний нотіс про готовність судна до вантажних робіт та пройти необхідні портові прикордонні, санітарні та митні формальності. Оскільки тривалість таких процедур, як правило, доволі незначна, можна покласти

$$k_{\theta_3} \equiv k_{n_p} \equiv k_{cm_{p_n}} \quad (23)$$

де k_{n_p} – часова точка приходу судна до порту розвантаження.

Знаходження судна у порту розвантаження характеризується наразі тими самими складовими, що і знаходження судна у порту завантаження: із прибуттям судно має подати відповідний нотіс, після подання якого починається grace period; аналогічно до знаходження судна у порту завантаження є необхідність проходження формальностей, пов'язаних із оформленням приходу судна до порту і початком розвантажувальних робіт; у розрахунках стояночного часу розвантаження слід враховувати порядок рахування сталієвих днів, зазначений у договорі рейсового фрахтування, який наразі може відрізнятися від тих самих умов по порту завантаження. Після закінчення розвантаження судно має пройти дозвільні портові формальності, пов'язані із закінченням розвантаження та виходом судна з порту, що власне і є кінцем рейсу судна. Розрахунок тривалості сталієвого часу у порту розвантаження визначається аналогічно до (13) із використанням відповідних вихідних даних щодо зокрема норм вантажних робіт та порядку рахування сталієвих днів. Тими самими залишаються також і складнощі, пов'язані із визначенням дати та часу приходу судна до порту розвантаження, притаманні аналогічним розрахункам стояночного часу по порту завантаження, на чому наголошено раніше.

Наразі загальний часу простого рейсу судна фактично є проміжком часу між вказаними на рис. 1 точками його початку і закінчення відповідно $k_{cm_{p3}}$ і $k_{x_{6n}}$, його загальна тривалість визначається сумуванням тривалостей окремих його складових:

$$t_p = t_{x_6} + t_{cm_3} + t_{x_6} + t_{cm_p}, \quad (24)$$

де t_p – тривалість часу простого рейсу судна, діб.

Висновки. Тривалість часу виконання судном рейсу на підставі договору фрахтування на рейс у великому ступені визначається окремими умовами договору. На етапі попередніх калькуляцій, зокрема – судновласником для оцінки орієнтовної зайнятості судна на роботі за певним договором це наразі єдині вихідні дані, які можуть ним використовуватися. Тривалість ходового часу рейса визначається чартерними домовленостями про порти завантаження і розвантаження та, за внесення їх чітких назв, дозволяє визначити орієнтовну тривалість виконання переходу судна. умови договору рейсового фрахтування щодо кількості вантажу та сталієвого часу по портах завантаження і розвантаження є підставами для оцінки тривалості виконання вантажних робіт, якою наразі формується стояночний час рейсу. Визначити точний час рейсу, звісно, можна лише за фактом його виконання, адже на його тривалість мають вплив певні фактори, передбачити які заздалегідь не є можливими (наприклад, погодні умови, черги у портах тощо). Втім, використання погоджених умов договору рейсового фрахтування є наразі єдиним чітким орієнтиром при попередньому визначенні тривалості рейсу судна.

Список використаних джерел:

1. Шибасєв О. Г., Боровік С. С., Тихоніна І. І. Основи теорії транспортних процесів і систем: навчальний посібник. Одеса : ОНМУ, 2023. 125 с.
2. Шибасєв О. Г., Кравченко О. А., Тихоніна І. І. Управління роботою флоту: навчальний посібник. Одеса : ОНМУ, 2023. 74 с.
3. Кириллова О.В. Теоретичні основи управління роботою флоту у транспортно-технологічних системах: дис. ... доктора техн. наук: 05.22.01. Одеса, 2016. 470 с.
4. Мельник О. Сучасні методи організації роботи флоту і портів. *Збірник наукових праць Державного економіко-технологічного університету транспорту. Сер. : Економіка і управління.* 2015. Вип. 33. С. 195–206.
5. Zhongbin Zhao, Xifu Wang, Hao Wang, Suxin Cheng, Wei Lui. Fleet Deployment Optimization for LNG Shipping Vessels Considering the Influence of Mixed Factors. *Journal of Marine Science and Engineering.* 2022. 10(12). 2034. DOI:10.3390/jmse10122034
6. Wiśniewski B., Korwin-Piotrowski T., Wielgosz M., Siódma S. Navigational aspects of ship's voyage planning taking into account calculations of ETA (Estimated Time of Arrival). *Scientific Journals Maritime University of Szczecin.* 2012. 29 (101) pp. 182–187.
7. Шибасєв О. Г. Формалізація рішень щодо управління морськими перевезеннями у транспортній мережі. *Вісник Херсонського національного технічного університету.* 2025. 1(1(92)) С. 291–295. DOI:10.35546/kntu2078-4481.2025.1.1.39
8. Покотілов І. П., Кошкін К. В. Вибір судна для реалізації проектів ефективних морських перевезень. *Управління розвитком складних систем.* 2011. № 7. С. 48–52.
9. Just in Time Arrival Guide. URL: <https://greenvoyage2050.imo.org/wp-content/uploads/2021/01/GIA-just-in-time-hires.pdf>
10. Tomita F., Kido M. An approximate travel time calculation and a robust GNSS-acoustic positioning method using an MCMC technique. *Earth, Planets and Space.* 2022. volume 74, Article number: 176. <https://doi.org/10.1186/s40623-022-01740-0>
11. Коскіна Ю. О. Планування складових часу рейса з урахуванням умов чартер-партії щодо сталієвого часу. *Вісник ОНМУ,* 2019. Вип. 2 (59). С. 166–182. DOI 10.33082/2226-1915 -2-2019-166-182

-
12. Щербина О. В., Дрожжин О. Л., Тихоніна І. О. Методика обґрунтування порту бункерування судна в рейсі. *Науковий технологічний журнал*. 2019. № 3(43). С. 377–384. DOI: 10.18372/2310-5461.43.13989
 13. Мельник О. М., Малаксіано М. О. Обґрунтування вибору суден враховуючи можливість їх використання для перевезення негабаритних вантажів. *Вчені записки ТНУ ім. В. І. Вернадського*. 2020. 31(70)2 № 1. С. 135–140. DOI:10.32838/2663-5941/2020.1-2/25
 14. Малаксіано М. О., Мельник О. М. Обґрунтування вибору судна для фрахтування на умовах тайм-чартеру з урахуванням можливості його використання для перевезень негабаритних вантажів. *Вісник Вінницького політехнічного інституту*. 2020. № 1(148). С. 90–96. DOI: <https://doi.org/10.31649/1997-9266-2020-148-1-90-96>
 15. Гловацьких В. І. Ефективність роботи суден залежно від швидкісного режиму. *Розвиток транспорту*. 2021. № 4 (11). С. 104–115. DOI: <https://doi.org/10.33082/td.2021.4-11.10>
 16. Vyshnevsky O., Vishnevsky D. Investigation of the effect of ship size and marine transportation distance on the possible decrease of voyage efficiency. *Technology Audit and Production Reserves*. 2017. № 1(2(33)). pp. 38–44. DOI:10.15587/2312-8372.2017.92965
 17. Стовба Т. А. Методичні аспекти експлуатаційно-економічного аналізу роботи судна в рейсі. *Таврійський науковий вісник*. 2011. № 77. С. 327–332.
 18. Лапкін О. І., Онищенко С. П., Коскіна Ю. О. Теорія і практика фрахтових операцій: монографія. Одеса, 2017. 151 с.
 19. Cooke J. et al. Voyage Charters. London: Informa Law, 2014. 1463 p.
 20. Gorton L., et al. Shipbrokering and Chartering Practice. London: Informa Law, 2009. 400 p.
 21. Plomaritou E., Papadopoulos A. Shipbrokering and Chartering Practice. London: Informa Law. 2017. 771 p.
 22. Commencement of Laytime. URL: <https://sc.bimco.org/TrialContract/14029>
 23. Коскіна Ю. О., Дрожжин О. Л., Тихоніна І. І. Визначення завантаження судна «важким» вантажем у рейсі на умовах договору рейсового фрахтування. *Науковий технологічний журнал*. 2024. Том 64. № 4. С. 528–538. DOI: <https://doi.org/10.18372/2310-5461.63.19763>
 24. Онищенко С. П., Коскіна Ю. О. Оцінка стояночного часу рейса за різних формулювань умов рейсової чартер-партії. *Збірник наукових праць ДУІТ. Серія «Транспортні системи і технології»*. 2018. Вип. 32. Т.2. С. 146–155. doi.org/10.32703/2617-9040-2018-32-2-146-155
 25. Онищенко С. П., Коскіна Ю. О. Формування стояночного часу рейса з урахуванням умов чартер-партії. *Науковий вісник Херсонської державної морської академії*. 2017. № 2 (17). С. 69–78.

References:

1. Shybaiev, O. H., Borovik, S. S., Tykhonina, I. I. (2023). Osnovy teorii transportnykh protsesiv i system [Fundamentals of Transport Processes and Systems Theory]. Odesa : ONMU. 125 p.
2. Shybaiev, O. H., Kravchenko, O. A., Tykhonina, I. I. (2023). Upravlinnia robotoiu flotu [Fleet management]. Odesa : ONMU. 74 p.
3. Kyrylova, O. V. (2017). Teoretychni osnovy upravlinnia robotoiu flotu u transportno tekhnolohichnykh systemakh [Theoretical foundations of fleet management in transport and technological systems] (Doctoral dissertation, Odesa National Maritime University). 470 p.
4. Melnyk, O. (2015). Suchasni metody orhanizatsii roboty flotu i portiv [Modern methods of organising the work of fleets and ports]. *Collection of scientific papers of the State University of Economics and Technology of Transport: Economics and Management*, (33), 195–206.
5. Zhongbin Zhao, Xifu Wang, Hao Wang, Suxin Cheng, Wei Lui. (2022). Fleet Deployment Optimization for LNG Shipping Vessels Considering the Influence of Mixed Factors. *Journal of Marine Science and Engineering*, (10(12)), 2034. DOI:10.3390/jmse10122034
6. Wiśniewski, B., Korwin-Piotrowski, T., Wielgosz, M., Siódmak, S. (2012). Navigational aspects of ship's voyage planning taking into account calculations of ETA (Estimated Time of Arrival). *Scientific Journals Maritime University of Szczecin*, (29 (101)), 182–187.
7. Shybaiev, O. H. (2025). Formalization of decisions regarding maritime transportation management in transport network. *Visnyk of Kherson National Technical University*, (1(92)), 291–295. DOI:10.35546/kntu2078-4481.2025.1.1.39
8. Pokotilov, I. P., Koshkin, K. V. (2011). Selecting vessels for implementing efficient maritime transport projects. *Managing the development of complex systems*, (7), 48–52.
9. Just in Time Arrival Guide. Retrieved from: <https://greenvoyage2050.imo.org/wp-content/uploads/2021/01/GIA-just-in-time-hires.pdf>
10. Tomita, F., Kido, M. (2022). An approximate travel time calculation and a robust GNSS-acoustic positioning method using an MCMC technique. *Earth, Planets and Space*, (74), Article number: 176. <https://doi.org/10.1186/s40623-022-01740-0>

-
11. Koskina, Yu. O. (2019). Planning the components of voyage time considering the terms of charter-party on laytime. *Herald of the Odessa National Maritime University*, (2 (59)), 166–182. DOI 10.33082/2226-1915-2-2019-166-182
12. Shcherbyna, O. V., Drozhzhyn, O. L., Tykhonina, I. O. (2019). Metodyka obhruntuvan-nia portu bunkeruvannia sudna v reisi [Substantiation method of bunker port for the ship during the voyage]. *Science-Based Technologies*, (3(43)), 377–384. DOI: 10.18372/2310-5461.43.13989
13. Melnyk, O. M., Malaksiano, M. O. (2020). Obgruntuvannia vyboru suden vrakho-vuiuchy mozhlyvist yikh vykorystannia dlia perevezennia nehabarynykh vantazhiv [Vessel selection prospects and suitability assessment for oversized cargo transportation]. *Scientific notes of Taurida National V.I. Vernadsky University. Series: Technical Sciences*, (31(70) 2 № 1), 135–140. DOI:10.32838/2663-5941/2020.1-2/25
14. Malaksiano, M. O., Melnyk, O. M. (2020). Obgruntuvannia vyboru sudna dlia frakhtuvannia na umovakh taim-charteru z urakhuvanniam mozhlyvosti yoho vyko-rystannia dlia perevezen nehabarynykh vantazhiv [Vessel choice justification for chartering on a time-charter basis, considering feasibility for oversized cargo shipment]. *Visnyk of Vinnytsia Polytechnical Institute*, (1 (148)), 90–96. DOI: <https://doi.org/10.31649/1997-9266-2020-148-1-90-96>
15. Hlovatskykh, V. I. (2021). Efektyvnist roboty suden zalezhno vid shvydkisnoho rezhymu [Efficiency of vessels depending on speed mode]. *Transport Development*, (4 (11)), 104–115. DOI <https://doi.org/10.33082/td.2021.4-11.10>
16. Vyshnevskaya, O., Vishnevsky, D. (2017). Investigation of the effect of ship size and marine transportation distance on the possible decrease of voyage efficiency. *Technology Audit and Production Reserves*, (1(2(33))), 38–44. DOI:10.15587/2312-8372.2017.92965
17. Stovba, T. A. (2011). Metodichni aspekty ekspluatatsiino-ekonomichnoho analizu roboty sudna v reisi [Methodological aspects of operational and economic analysis of ship performance during voyages]. *Tavriya Scientific Bulletin*, (77), 327–332.
18. Lapkin, O. I., Onyshchenko, S. P., Koskina, Yu. O. (2017). Teoriia i praktyka frakhtovykh operatsii [Theory and practice of chartering operations]. Odesa, 151.
19. Cooke, J. et al. (2014). Voyage Charters. London: Informa Law, 1463 p.
20. Gorton, L., et al. (2009). Shipbroking and Chartering Practice. London: Informa Law, 400 p.
21. Plomaritou, E., Papadopoulos A. (2017). Shipbroking and Chartering Practice. London: Informa Law. 771 p.
22. Commencement of Laytime. Retrieved from: <https://sc.bimco.org/TrialContract/14029>
23. Koskina, Yu. O., Drozhzhyn, O. L., Tykhonina, I. I. (2024). Vyznachennia zavantazhennia sudna «vazhkym» vantazhem u reisi na umovakh dohovora reisovoho frakhtuvannia [Determination the loading of a vessel with “heavy” cargo on a voyage under the terms of a voyage charter party]. *Science-Based Technologies*, (64 № 4), 528–538. DOI: <https://doi.org/10.18372/2310-5461.63.19763>
24. Onyshchenko, S. P., Koskina, Yu. O. (2018). Otsinka stoianochnoho chasu reisa za riznykh formuliuvan umov reisovoi charter-partii [Estimation of vessel's time in port at different wordings of the voyage charter-party's terms]. *Collection of Scientific Papers of the State University of Infrastructure and Technologies of the Ministry of Education and Science of Ukraine: Series "Transport Systems and Technologies"*. (32/2), 146–155. doi.org/10.32703/2617-9040- 2018-32-2-146-155
25. Onyshchenko, S. P., Koskina, Yu. O. (2017). Formuvannia stoianochnoho chasu reisa z urakhuvanniam umov charter-partii [Basics of forming the vessel's time in port considering the charter-party's terms]. *Scientific Bulletin of the Kherson State Maritime Academy*, (2 (17)), 69–78.

Дата надходження статті: 25.10.2025

Дата прийняття статті: 10.10.2025

Опубліковано: 30.12.2025

ЕЛЕКТРОНІКА, ЕЛЕКТРОННІ КОМУНІКАЦІЇ, ПРИЛАДОБУДУВАННЯ ТА РАДІОТЕХНІКА

УДК 528.8:528.7:528.4:528.06

DOI <https://doi.org/10.32782/2521-6643-2025-2-70.40>

Альперт С. І., кандидат технічних наук,
доцент кафедри аерокосмічної геодезії та землеустрою
факультету архітектури, будівництва та дизайну
Державного університету «Київського авіаційного інституту»
ORCID: 0000-0002-7284-6502

ОБРОБКА СУПУТНИКОВИХ ЗНІМКІВ З ВИКОРИСТАННЯМ АПАРАТУ ТЕОРІЇ НЕЧІТКИХ МНОЖИН

Теорія нечітких множин – це розділ прикладної математики, присвячений методам аналізу та обробки невизначених та неточних даних. Основним поняттям даної теорії є функція приналежності. Функція приналежності визначає ступінь приналежності деякого елемента чи об'єкта до нечіткої множини. Теорія нечітких множин може точно описувати математичні моделі складних природних об'єктів і процесів та широко застосовується у різних сферах, таких як: інформатика, штучний інтелект, розпізнавання образів та інші. Ця теорія використовується для моделювання, аналізу, вивчення різних складних систем із невизначеними даними. Теорія нечітких множин є важливим математичним інструментом у дистанційному зондуванні Землі. Поняття «нечіткої множини» застосовуються у таких задачах класифікування, де звичайні методи класифікування дають неточні результати. Дана теорія описує класи земного покриття зі ступенями приналежності об'єкта кожному із цих класів. Цей математичний підхід застосовується при роботі із неточними та невизначеними границями між класами. Застосування нечітких множин надає можливість будувати більш реалістичні моделі для вирішення задач дистанційного зондування. Зазначено, що запропонована теорія відіграє важливу роль у задачах обробки супутникових знімків при вирішенні задач дистанційного зондування Землі, оскільки дана методика дає змогу точно охопити структуру супутникового зображення. Описано деякі логічні операції з нечіткими числами із застосуванням методу альфа-зрізів. Розглянуто на числових прикладах логічні операції з нечіткими числами. Детально проаналізовано нечітку кластеризацію методом С-середніх. Зазначено, що даний підхід широко застосовується для проведення неконтрольованого класифікування супутникових зображень. Він дозволяє пікселям належати до декількох класів із різним ступенем приналежності, на відміну від інших відомих методів кластеризації. Нечітка кластеризація методом С-середніх використовується для обробки, аналізу та класифікування змішаних пікселів. Апарат теорії нечітких множин може бути застосований для вирішення таких актуальних та важливих задач дистанційного зондування Землі, як: сільськогосподарські задачі, моніторинг водних ресурсів, класифікування рослинності, розв'язок екологічних та геологічних задач.

Ключові слова: обробка даних ДЗЗ, класифікування зображень, теорія нечітких множин, нечітке число, нечітка кластеризація, задачі ДЗЗ.

Alpert S. I. Processing of satellite images using the apparatus of fuzzy set theory

Fuzzy set theory is a branch of applied mathematics devoted to methods of analysis and processing of vague and inaccurate data. Main concept of this theory is membership function. Membership function defines the degree to which a certain element or object belongs to a fuzzy set. Fuzzy set theory accurately describes mathematical models of complex natural objects and processes, it is widely applied in various fields, such as: computer science, artificial intelligence, pattern recognition, and others. This theory is used to model, analyze and study various complex systems with uncertain data. Fuzzy Set Theory is a considerable mathematical tool in remote sensing. The concept of «fuzzy set» is used in classification tasks when common classification methods give inaccurate results. This theory describes the classes of land cover with the degrees of belonging of object to each of these classes. This mathematical approach is useful when dealing with imprecise and ambiguous boundaries between classes. The use of fuzzy sets makes it possible to build more realistic models for solving remote sensing problems. It was noted that the proposed theory plays a serious role in the processing of satellite images for solution of remote sensing tasks, because this technique allows to accurately describe the structure of the satellite image. Some logical operations with fuzzy numbers applying the alpha-cut method were described. These logical operations with fuzzy numbers were considered using numerical

© С. І. Альперт, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

examples. Fuzzy clustering using the C-means method was analyzed in detail. It was noted that this approach is widely applied for unsupervised classification of satellite images. It allows, that pixels belong to multiple classes with different degrees of membership, unlike other known clustering methods. Fuzzy C-Means clustering is useful for processing, analyzing and classification mixed pixels. Apparatus of Fuzzy Set Theory can be applied for solution of such actual and important remote sensing tasks, such as: agricultural tasks, water resources monitoring, vegetation classification, various ecological and geological problems.

Key words: remote sensing data processing, image classification, Fuzzy Set Theory, fuzzy number, fuzzy clustering, remote sensing tasks.

Вступ. Як відомо, теорія нечітких множин була заснована у 1965 р. відомим американським математиком та логіком Заде Л. для розробки та вдосконалення математичних методів, які здатні обробляти невизначену і неточну інформацію. Поняття нечіткої множини є певним узагальненням відомого поняття класичної множини, але, на відміну від поняття класичної множини, воно дає можливість елементам належати до множини частково, тобто вводиться певний ступінь належності елемента до множини, який, в свою чергу, приймає значення від «0» до «1».

Основними складовими апарату теорії нечітких множин є функції приналежності, нечіткі правила та алгоритми для виведення нечіткого висновку.

У роботах Заде Л. зазначалося, що складні системи в умовах неточності та невизначеності не можуть бути описані класичними математичними теоріями. Теорія нечітких множин знайшла широке застосування у різноманітних сферах, таких як: штучний інтелект, задачі розпізнавання образів, нейронні мережі, задачі управління ресурсами, керування безпілотними літальними апаратами, автоматика, системи керування, тощо.

Слід зазначити, що класифікування невизначених, неповних та неточних супутникових даних є однією із найскладніших процедур при розв'язанні задач у сфері дистанційного зондування Землі (ДЗЗ).

Як правило спектральні характеристики різних об'єктів часто мають розмиті та нечіткі межі. Більшість природних об'єктів, а саме, ліси, водойми, ґрунтовий покрив не можна однозначно класифікувати, оскільки їх спектральні ознаки не мають чітких меж. У таких випадках традиційне бінарне класифікування дає не досить точні результати, оскільки воно не здатне враховувати усі особливості спектральних характеристик природних об'єктів.

Тому у наш час актуальною постає задача саме розробки нових методів обробки та класифікування супутникових зображень за умов наявності неточних та невизначених вхідних даних. Саме у такому випадку рекомендується застосовувати методи обробки аерокосмічних даних, що ґрунтуються на теорії нечітких множин, оскільки бінарне класифікування може давати не досить точні результати.

Актуальність даної роботи полягає у необхідності розробки рекомендацій щодо застосування теорії нечітких множин, її основних понять та методик для підвищення точності процедури класифікування супутникових знімків.

Аналіз останніх джерел та публікацій. У даній статті було розглянуто основні можливості застосування математичного апарату теорії нечітких множин для підвищення точності класифікування аерокосмічних знімків в умовах неточних та невизначених вхідних даних.

У наукових роботах Кауфмана А., Гупта М., Дюбуа Д., Праде Х. були наведені та проаналізовані фундаментальні поняття та означення теорії нечітких множин, логічні (арифметичні) операції з нечіткими числами та основні сфери застосування [7, 11].

У статтях Бінагі Е. та Вана Е. були описані різні методики класифікування мультиспектральних зображень із використанням апарату теорії нечітких множин для вирішення численних задач ДЗЗ [2, 14]. Наголошувалося на тому, що методи обробки та класифікування супутникових знімків із застосуванням теорії нечітких множин дають більш точні результати ніж методи класифікування, що ґрунтуються на класичній теорії ймовірностей.

У роботах Канона Р., Гопала С. та Вудкока К. було проаналізовано методи сегментації супутникових знімків з метою створення тематичних карт та методи оцінки точності тематичних карт із використанням методик теорії нечітких множин [4, 10].

Отже, **метою дослідження** є аналіз фундаментальних понять теорії нечітких множин та сучасної методики кластеризації супутникових зображень на основі нечіткої логіки для вирішення актуальних задач ДЗЗ, таких як: класифікування лісів та урбанізованих територій, вирішення сільськогосподарських та геологічних задач, проведення екологічного моніторингу.

Виклад основного матеріалу. Застосування нечітких множин дає можливість описувати, обробляти, аналізувати та відображати взаємозв'язки між неточними даними.

У даному пункті розглянемо фундаментальні поняття та базові означення теорії нечітких множин [8, 12, 13].

Розглянемо універсальну множину X .

Задаємо нечітку підмножину $A \subset X$ на основі характеристичної функції, а саме:

$$\mu_A : X \rightarrow [0,1]. \quad (1)$$

Будь-якому елементу множини $x \in A$ ставиться у відповідність дійсне значення, що приймає характеристична функція $\mu_A(x)$ із інтервалу $[0,1]$.

Значення $\mu_A(x)$ характеризує вагу елемента (об'єкта) x у заданій множині A .

Розглянемо нечітку множину $A \subset X$ та довільне дійсне число $\alpha \in [0,1]$.

α – рівень чи α – зріз, який відповідає нечіткій множині A , є чіткою множиною та позначається через ${}^\alpha A$, а саме:

$${}^\alpha A = \{x \in X : \mu_A(x) \geq \alpha\}. \quad (2)$$

Центральний елемент виражається із використанням α -зрізу:

$${}^\alpha A_i = \{x : \mu_A(x) \geq \alpha_i\} = [{}^\alpha A_{i_{lower}}, {}^\alpha A_{i_{upper}}], \quad (i = 1, 2, 3, \dots, n), \quad (3)$$

де $\alpha_0 = 1 > \alpha_1 > \alpha_2 > \dots > \alpha_n = 0$, $\alpha \in [0,1]$.

Далі наведемо означення “трикутного числа”.

Трикутне нечітке число A будується на основі чисел $[a, b, c]$ та характеристичної функції, яка розраховується наступним чином:

$$\mu_A(x) = \begin{cases} \frac{x-a}{b-a}, & a \leq x \leq b \\ \frac{c-x}{c-b}, & b \leq x \leq c \end{cases}, \quad (4)$$

де a – мінімальне (найменше) значення;

b – найвірогідніше значення;

c – максимальне (найбільше) значення.

Далі побудуємо α – зріз множини A . Для цього покладемо, що $\alpha \in [0,1]$.

Звідси маємо:

$$\alpha = \frac{x-a}{b-a}; \quad \alpha = \frac{c-x}{c-b}. \quad (5)$$

Із отриманого виразу (5) виражаємо першу та другу границі інтервалу.

Виражаємо x через α :

$$\begin{aligned} x-a &= \alpha(b-a), \\ x &= (b-a)\alpha + a; \\ c-x &= \alpha(c-b), \\ x &= c - (c-b)\alpha. \end{aligned} \quad (6)$$

Ми отримали першу та другу границю інтервалу.

Застосовуючи вище описані рівності, отримуємо вираз для α -зрізу множини A :

$${}^\alpha A = [(b-a)\alpha + a, c - (c-b)\alpha]. \quad (7)$$

Логічні операції над нечіткими числами із використанням методу α -зрізів

Додавання нечітких чисел

Припустимо, що ми маємо наступні вхідні дані:

$X = [a, b, c]$ та $Y = [p, q, r]$ – нечіткі числа, чії характеристичні функції розраховуються наступним чином:

$$\mu_X(x) = \begin{cases} \frac{x-a}{b-a}, & a \leq x \leq b \\ \frac{c-x}{c-b}, & b \leq x \leq c \end{cases}, \quad \mu_Y(x) = \begin{cases} \frac{x-p}{q-p}, & p \leq x \leq q \\ \frac{r-x}{r-q}, & q \leq x \leq r \end{cases}. \quad (8)$$

Далі із формули (8) виражаємо α – зрізи нечітких чисел X та Y :

$$\begin{aligned} {}^\alpha X &= [(b-a)\alpha + a, c - (c-b)\alpha], \\ {}^\alpha Y &= [(q-p)\alpha + p, r - (r-q)\alpha]. \end{aligned} \quad (9)$$

Для знаходження суми нечітких чисел X та Y треба спочатку знайти суму α -зрізів нечітких чисел X та Y , спираючись на інтервальну арифметику, а саме:

$$\begin{aligned} {}^{\alpha}X + {}^{\alpha}Y &= [(b-a)\alpha + a, c - (c-b)\alpha] + \\ &+ [(q-p)\alpha + p, r - (r-q)\alpha] = \\ &= [a + p + (b-a+q-p)\alpha, c + r - (c-b+r-q)\alpha]. \end{aligned} \quad (10)$$

Для розрахунку характеристичної функції $\mu_{X+Y}(x)$ треба ліву та праву границі інтервалу із попереднього виразу прирівняти до x .

Маємо:

$$\begin{aligned} x &= a + p + (b-a+q-p)\alpha, \\ x &= c + r - (c-b+r-q)\alpha. \end{aligned} \quad (11)$$

Далі виражаємо α через x , покладаючи, що $\alpha=0$ та $\alpha=1$ у виразі (10):

$$\begin{aligned} \alpha &= \frac{x - (a+p)}{(b+q) - (a+p)}, \quad (a+p) \leq x \leq (b+q), \\ \alpha &= \frac{(c+r) - x}{(c+r) - (b+q)}, \quad (b+q) \leq x \leq (c+r). \end{aligned} \quad (12)$$

Із виразу (12) виражаємо характеристичну функцію для суми 2 нечітких чисел X та Y :

$$\mu_{X+Y}(x) = \begin{cases} \frac{x - (a+p)}{(b+q) - (a+p)}, & (a+p) \leq x \leq (b+q), \\ \frac{(c+r) - x}{(c+r) - (b+q)}, & (b+q) \leq x \leq (c+r). \end{cases} \quad (13)$$

Віднімання нечітких чисел

За умовою ми маємо $X=[a,b,c]$ та $Y=[p,q,r]$ – нечіткі числа, чиї характеристичні функції знаходяться, використовуючи формулу (8), а їх α -зрізи розраховуються по формулам (9).

Для обчислення різниці нечітких чисел X та Y нам слід знайти різницю α -зрізів цих нечітких чисел X та Y , застосовуючи інтервальну арифметику, так само як і при додаванні нечітких чисел:

$$\begin{aligned} {}^{\alpha}X - {}^{\alpha}Y &= [(b-a)\alpha + a, c - (c-b)\alpha] - [(q-p)\alpha + p, r - (r-q)\alpha] = \\ &= [(b-a)\alpha + a - (r - (r-q)\alpha), c - (c-b)\alpha - ((q-p)\alpha + p)] = \\ &= [(a-r) + (b-a+r-q)\alpha, (c-p) - (c-b+q-p)\alpha]. \end{aligned} \quad (14)$$

Для обчислення характеристичної функції $\mu_{X-Y}(x)$ слід прирівняти ліву та праву границі інтервалу із виразу (14) до x .

Тоді отримуємо:

$$\begin{aligned} x &= (a-r) + (b-a+r-q)\alpha, \\ x &= (c-p) - (c-b+q-p)\alpha. \end{aligned} \quad (15)$$

Виражаємо α через x , покладаючи, що $\alpha=0$ та $\alpha=1$ у виразі (14):

$$\begin{aligned} \alpha &= \frac{x - (a-r)}{(b-q) - (a-r)}, \quad (a-r) \leq x \leq (b-q), \\ \alpha &= \frac{(c-p) - x}{(c-p) - (b-q)}, \quad (b-q) \leq x \leq (c-p). \end{aligned} \quad (16)$$

Далі із виразу (16) виражаємо характеристичну функцію різниці 2 нечітких чисел X та Y :

$$\mu_{X-Y}(x) = \begin{cases} \frac{x - (a-r)}{(b-q) - (a-r)}, & (a-r) \leq x \leq (b-q) \\ \frac{(c-p) - x}{(c-p) - (b-q)}, & (b-q) \leq x \leq (c-p) \end{cases}. \quad (17)$$

Числові приклади логічних операцій над нечіткими числами із використанням методу α -зрізів

Нехай за умовою задачі маємо наступні вхідні дані:

$X = [2, 3, 5]$ та $Y = [4, 6, 8]$ – 2 нечітких числа, чий характеристичні функції, обчислюються за формулою (4):

$$\mu_X(x) = \begin{cases} \frac{x-2}{3-2}, & 2 \leq x \leq 3 \\ \frac{5-x}{5-3}, & 3 \leq x \leq 5 \end{cases}, \quad \mu_Y(x) = \begin{cases} \frac{x-4}{6-4}, & 4 \leq x \leq 6 \\ \frac{8-x}{8-6}, & 6 \leq x \leq 8 \end{cases},$$

$$\text{або } \mu_X(x) = \begin{cases} \frac{x-2}{1}, & 2 \leq x \leq 3 \\ \frac{5-x}{2}, & 3 \leq x \leq 5 \end{cases}, \quad \mu_Y(x) = \begin{cases} \frac{x-4}{2}, & 4 \leq x \leq 6 \\ \frac{8-x}{2}, & 6 \leq x \leq 8 \end{cases}.$$

Знаходимо α -зрізи нечітких чисел X та Y , використовуючи формулу (9):

$${}^{\alpha}X = [2 + \alpha, 5 - 2\alpha],$$

$${}^{\alpha}Y = [2\alpha + 4, 8 - 2\alpha].$$

1) Розрахунок суми двох нечітких чисел

Розраховуємо суму α -зрізів для нечітких чисел X та Y :

$${}^{\alpha}X + {}^{\alpha}Y = [2 + \alpha, 5 - 2\alpha] + [2\alpha + 4, 8 - 2\alpha] = [3\alpha + 6, 13 - 4\alpha].$$

Для обчислення характеристичної функції $\mu_{X+Y}(x)$ ми маємо прирівняти ліву та праву границі інтервалу із виразу до x .

Звідси виражаємо:

$$x = 3\alpha + 6,$$

$$x = 13 - 4\alpha.$$

Потім виражаємо α через x , вважаючи, що $\alpha = 0$ та $\alpha = 1$:

$$\alpha = \frac{x-6}{3}, \quad 2+4 \leq x \leq 3+6,$$

$$\alpha = \frac{13-x}{4}, \quad 3+6 \leq x \leq 5+8.$$

Далі знаходимо характеристичну функцію суми 2 нечітких чисел X та Y :

$$\mu_{X+Y}(x) = \begin{cases} \frac{x-6}{3}, & 6 \leq x \leq 9, \\ \frac{13-x}{4}, & 9 \leq x \leq 13. \end{cases}$$

2) Розрахунок різниці двох нечітких чисел

Будуємо різницю α -зрізів нечітких чисел X та Y :

$${}^{\alpha}X - {}^{\alpha}Y = [2 + \alpha, 5 - 2\alpha] - [2\alpha + 4, 8 - 2\alpha] = [3\alpha - 6, 1 - 4\alpha].$$

Для розрахунку характеристичної функції $\mu_{X-Y}(x)$ прирівнюємо ліву та праву границі інтервалу з даного виразу до x .

Звідси маємо:

$$x = 3\alpha - 6,$$

$$x = 1 - 4\alpha.$$

Виражаємо α через x , покладаючи, що у даному виразі $\alpha = 0$ та $\alpha = 1$:

$$\alpha = \frac{x+6}{3}, \quad 2-8 \leq x \leq 3-6,$$

$$\alpha = \frac{1-x}{4}, \quad 3-6 \leq x \leq 5-4$$

$$\begin{aligned} \alpha &= \frac{x+6}{3}, & -6 \leq x \leq -3, \\ \text{чи} \\ \alpha &= \frac{1-x}{4}, & -3 \leq x \leq 1. \end{aligned}$$

Отже, отримуємо характеристичну функцію різниці 2 нечітких чисел X та Y :

$$\mu_{X-Y}(x) = \begin{cases} \frac{x+6}{3}, & -6 \leq x \leq -3; \\ \frac{1-x}{4}, & -3 \leq x \leq 1. \end{cases}$$

Застосування нечіткої кластеризації для обробки супутникових знімків. Нечітка кластеризація для класифікування супутникових зображень є ефективним методом для проведення аналізу та інтерпретації даних ДЗЗ. На відміну від традиційної або чіткої кластеризації, де кожен об'єкт відноситься тільки до одного кластера, нечітка кластеризація допускає приналежність об'єкта одночасно до декількох кластерів із різним ступенем приналежності. Це є важливим для вирішення задач ДЗЗ, оскільки пікселі супутникових зображень часто характеризуються змішаними спектральними ознаками, що відповідають різним типам земного покриття, тобто: рослинність, вода, ґрунт можуть частково змішуватися у одному пікселі на супутниковому знімку [1, 3].

Зазначимо, що для проведення найбільш ефективного аналізу та інтерпретації даних ДЗЗ слід використовувати метод нечіткої кластеризації С-середніх (Fuzzy C-Means) [4, 5].

Алгоритм методу нечіткої кластеризації С-середніх

1) Спочатку задаємо вхідні дані:

$X = \{x_1, x_2, \dots, x_n\}$ – множина спектральних векторів пікселів,

де x_i – повний сигнал пікселя, а саме, n -вимірний вектор, компонентами якого є значення сигналу пікселя на осях спектрального простору (спектральний вектор пікселя);

c – число кластерів, що обирається попередньо користувачем на основі аналізу даних чи емпіричним шляхом;

m – коефіцієнт (параметр) нечіткості, який відповідає за ступінь розмитості кластерів. Зазвичай, $m = 2$;

v_j – центри кластерів, що випадковим чином ініціалізуються,

де $j = 1, \dots, c$.

2) На даному кроці задається формула критерію мінімізації (функціонал):

$$J(U, V) = \sum_{i=1}^n \sum_{j=1}^c (\mu_{ij})^m \|x_i - v_j\|^2, \quad 1 \leq m < \infty, \quad (18)$$

де x_i – спектральний вектор i -го пікселя;

μ_{ij} – ступінь приналежності i -го пікселя до j -го кластера, що обчислюється за наступною формулою:

$$\mu_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{\|x_i - v_j\|}{\|x_i - v_k\|} \right)^{\frac{2}{m-1}}}. \quad (19)$$

З даної формули випливає, що чим ближче піксель до центру конкретного кластера, тим його ступінь приналежності є вищим.

При цьому має виконуватися умова нормування, а саме:

$$\sum_{j=1}^c \mu_{ij} = 1, \quad \forall i = 1, \dots, n.$$

m – коефіцієнт нечіткості.

v_j – центр j -го кластера;

Після перерахунку нових ступенів приналежності слід перерахувати (оновити) центри кластерів v_j за наступною формулою:

$$v_j = \frac{\sum_{i=1}^n \mu_{ij}^m x_i}{\sum_{i=1}^n \mu_{ij}^m}. \quad (20)$$

3) Перевіряємо умови зупинки даного алгоритму, а саме продовжуємо перераховувати ступені приналежності та оновлювати центри кластерів за вище наведеними формулами, поки не буде досягнута умова зупинки (наприклад, зміна центрів кластерів стане менше заданого порогового значення ϵ , чи кількість ітерацій перевищить нами задане максимальне значення).

Метод нечіткої кластеризації широко може застосовуватися у ДЗЗ, де межі класів розмиті, а одним пікселям чи областям знімка властиві ознаки відразу кількох типів об'єктів. Даний алгоритм враховує неоднозначність приналежності пікселя до того чи іншого класу та ефективно розділяє змішані зони.

Метод Fuzzy C-Means застосовується для вирішення численних практичних завдань в ДЗЗ, таких як: виділення різних типів земного покриття (ліс, вода, сільськогосподарські угіддя, урбанізовані території) на багатоспектральних супутникових зображеннях, сегментація гіперспектральних зображень, виділення водних об'єктів (річок, озер), моніторинг водних ресурсів, резервуарів на оптичних знімках [4, 6, 9]. Також слід зазначити, що гібридні методи, що використовують у комплексі апарат теорії нечітких множин та водні індекси дають більш точні результати при розпізнаванні на супутникових знімках берегової лінії та мілководних ділянок.

Висновки. У статті було розглянуто та проаналізовано фундаментальні поняття та положення теорії нечітких множин. При цьому акцентувалося на тому, що за останні десятиліття апарат теорії нечітких множин здобув широке застосування в автоматизації, керуванні безпілотними літальними апаратами, нейронних мережах, штучному інтелекті та при вирішенні задач ДЗЗ. Так, дана теорія застосовується для обробки та розпізнавання аерокосмічних знімків в умовах неточних та неповних вхідних даних ДЗЗ. Застосування нечітких методів дає змогу підвищити точність і стійкість класифікування супутникових зображень за рахунок врахування неоднозначності спектральних ознак.

У статті були наведені основні логічні операції над нечіткими числами із використанням методу альфа-зрізів, зокрема, операції додавання та віднімання із числовими прикладами.

Було зазначено, що більшість задач ДЗЗ оперують саме нечіткими поняттями, оскільки різні класи різних природних об'єктів є нечіткими за своєю природою, тобто мають розмиті границі. Перехід від належності до неналежності об'єктів до даних класів є поступовим.

Тому при вирішенні задач класифікування доцільно застосовувати саме теорію нечітких множин, оскільки чіткі класи представляють спрощену картину, що не відповідає реальності та може призвести до невірних результатів класифікування.

Було запропоновано використовувати метод нечіткої кластеризації C-середніх для проведення обробки супутникових знімків та детально описано алгоритм даного методу.

Наголошувалося на тому, що теорія нечітких множин та нечіткі числа можуть бути застосовані при вирішенні численних задач ДЗЗ, зокрема, при вирішенні різноманітних природно-ресурсних задач, а саме: класифікування земного покриття (визначення ґрунтів та рослинності), аналіз урбанізованих територій (розпізнавання об'єктів інфраструктури), моніторинг надзвичайних ситуацій та природних катастроф, оцінка стану лісів (моніторинг деградації лісових масивів) та вирішення задач сільського господарства (прогноз врожайності).

Список використаних джерел:

1. Alpert S. I. Data combination method in Remote Sensing tasks in case of conflicting information sources. *Ukrainian Journal of Remote Sensing*. 2021. Vol. 8(3). P. 44–48.
2. Binaghi E., Madella P., Montesano M. G., Rampini A. Fuzzy contextual classification of multiresource remote sensed images. *IEEE Transactions on Geoscience and Remote Sensing*. 1997. Vol. 35(2). P. 326–340.
3. Blonda P. N., Pasquariello G., Losito S., Mori A., Posa F., Ragno, D. An experiment for the interpretation of multitemporal remotely sensed images based on a fuzzy logic approach. *Int. J. Remote Sensing*. 1991. Vol. 12(3). P. 463–476.
4. Cannon R., Dave J., Bezdek J. C., Trivedi M. Segmentation of a thematic mapper image using the fuzzy c-means clustering algorithm. *IEEE Trans. Geo. and Remote Sensing*. 1986. Vol. 24(3). P. 400–408.
5. Canters F. Evaluating the uncertainty of area estimates derived from fuzzy land-cover classification. *Photogrammetric Engineering and Remote Sensing*. 1997. Vol. 63(4). P. 403–414.
6. Chang C. I. Hyperspectral data processing. Algorithm design and analysis. *Hoboken, N J: John Wiley and Sons*. 2013. 1164 p.
7. Dubois D., Prade H. Fundamentals of Fuzzy sets. *Boston: Kluwer Academic Publishers*. 2000. P. 125–193.
8. Dutta P., Boruah H., Ali T. Fuzzy arithmetic with and without using α -cut method: a comparative study. *International Journal of Latest trends in Computing*. 2011. Vol. 2 (1). P. 99–108.
9. Fisher P. F., Pathirana S. The evaluation of fuzzy membership of land cover classes in the suburban zone. *Remote Sensing of the Environment*. 1990. Vol. 34. P. 121–132.
10. Gopal S., Woodcock C. Theory and methods for accuracy estimation of thematic maps using fuzzy sets. *Photogrammetric Engineering and Remote Sensing*. 1994. Vol. 60. P. 181–188.

-
11. Kaufman A., Gupta M. M. Introduction to Fuzzy Arithmetic, Theory and applications. *Van Nostrand Reinhold Co. Inc., Workingham, Berkshire*. 2003. 361 p.
 12. Kent J. T., Mardia K. V. Spatial Classification Using Fuzzy Membership Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1988. Vol. 10(5). P.659–671.
 13. Lu D., Weng Q. A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*. 2007. Vol. 28(5). P. 823–870.
 14. Wang E. Improving Remote Sensing Image Analysis through Fuzzy Information Representation. *Photogrammetric Engineering and Remote Sensing*. 1990. Vol. 56(8). P.1163–1169.

References:

1. Alpert, S. I. (2021). Data combination method in Remote Sensing tasks in case of conflicting information sources. *Ukrainian Journal of Remote Sensing*, 8(3), 44–48.
2. Binaghi, E., Madella, P., Montesano, M. G., Rampini, A. (1997). Fuzzy contextual classification of multiresource remote sensed images. *IEEE Transactions on Geoscience and Remote Sensing*, 35(2), 326–340.
3. Blonda, P. N., Pasquariello, G., Losito, S., Mori, A., Posa, F., Ragno, D. (1991). An experiment for the interpretation of multitemporal remotely sensed images based on a fuzzy logic approach. *Int. J. Remote Sensing*, 12(3), 463–476.
4. Cannon, R., Dave, J., Bezdek, J. C., Trivedi, M. (1986). Segmentation of a thematic mapper image using the fuzzy c-means clustering algorithm. *IEEE Trans. Geo. and Remote Sensing*, 24(3), 400–408.
5. Canters, F. (1997). Evaluating the uncertainty of area estimates derived from fuzzy land-cover classification. *Photogrammetric Engineering and Remote Sensing*, 63(4), 403–414.
6. Chang, C. I. (2013). Hyperspectral data processing: Algorithm design and analysis. *Hoboken, N J: John Willey and Sons*, 1164 p.
7. Dubois, D., Prade, H. (2000). Fundamentals of Fuzzy sets. *Boston: Kluwer Academic Publishers*, 125–193.
8. Dutta, P., Boruah, H., Ali, T. (2011). Fuzzy arithmetic with and without using α -cut method: a comparative study. *International Journal of Latest trends in Computing*, 2 (1), 99–108.
9. Fisher, P. F., Pathirana, S. (1990). The evaluation of fuzzy membership of land cover classes in the suburban zone. *Remote Sensing of the Environment*, 34, 121–132.
10. Gopal, S., Woodcock, C. (1994). Theory and methods for accuracy estimation of thematic maps using fuzzy sets, *Photogrammetric Engineering and Remote Sensing*, 60, 181–188.
11. Kaufman, A., Gupta, M. M. (2003). Introduction to Fuzzy Arithmetic, Theory and applications. *Van Nostrand Reinhold Co. Inc., Workingham, Berkshire*, 361 p.
12. Kent, J. T., Mardia K. V. (1988). Spatial Classification Using Fuzzy Membership Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(5), 659–671.
13. Lu, D., Weng, Q. (2007). A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*, 28(5), 823–870.
14. Wang, E. (1990). Improving Remote Sensing Image Analysis through Fuzzy Information Representation. *Photogrammetric Engineering and Remote Sensing*, 56(8), 1163–1169.

Дата надходження статті: 08.08.202

Дата прийняття статті: 10.09.2025

Опубліковано: 30.12.2025

ГОТЕЛЬНО-РЕСТОРАННА СПРАВА ТА КЕЙТЕРИНГ

УДК 338.48:379.8

DOI <https://doi.org/10.32782/2521-6643-2025-2-70.41>

Горб К. М., кандидат географічних наук, доцент
Університету митної справи та фінансів
ORCID: 0000-0002-3779-484X

Корнєєв М. В., доктор економічних наук,
професор Університету митної справи та фінансів
ORCID: 0000-0002-4005-5335

Павлов В. В., здобувач вищої освіти
Університету митної справи та фінансів
ORCID: 0009-0005-2888-5167

ТЕХНОЛОГІЇ ГОТЕЛЬНОГО ОБСЛУГОВУВАННЯ ТУРИСТИЧНИХ ГРУП

Статтю присвячено аналізу підходів, узагальненню досвіду та систематизації технологій готельного обслуговування туристичних груп на різних етапах гостьового циклу, за різних форматів туристичних подорожей у різних засобах розміщення туристів, включно з рухомими (туристичні потяги, круїзні судна, автобуси та літаки зі спальними місцями тощо), з метою подальшої стандартизації та удосконалення групового сервісного процесу в туризмі та гостинності. Технологічні процеси обслуговування туристичних груп у готелях мають низку суттєвих відмінностей від надання готельних послуг індивідуальним постояльцям. На етапі бронювання готельних послуг для групи туристів важливо замовнику надати готелю список туристів з їх паспортними даними, узгодити розподіл туристів по номерах різних категорій, з пріоритетом компактного розташування номерів, де проживатимуть туристи групи, а також питання трансферу групи до готелю при заїзді та зворотного трансферу при виїзді, подбати про стоянку при готелі групового транспортного засобу. На етапі розміщення важливо оперативно заселити учасників групи до номерів готелю через узгоджені дії чергового адміністратора готелю та супроводжувача туристичної групи. Під час проживання групи в готелі важливим є забезпечення сумісності мети перебування групи із метою перебування решти гостей. На останньому етапі остаточних розрахунків, виселення та виїзду групи з готелю важливим є забезпечення оперативності взаємоузгоджених дій щодо перевірки номерів, здачі ключів, складання особистих речей та завантаження їх на транспортний засіб тощо.

При груповому готельному обслуговуванні туристів на кожному етапі гостьового циклу ключовим завданням є уникнення черг та мінімізація часу на будь-які очікування більшою частиною групи решти туристів. Усі зазначені положення готельного обслуговування туристичних груп справедливі як для стаціонарних засобів та закладів розміщення, так і для рухомих, а саме для тих туристичних транспортних засобів, де передбачається надання туристам спальних місць для ночівель. У перспективі доречно приділити увагу більш детальному обґрунтуванню концепції групового гостьового циклу закладу розміщення, сутність якого полягатиме у послідовності та безперервності забезпечення готельного сервісу групового характеру.

Ключові слова: груповий гостьовий цикл закладу розміщення, груповий туризм, готельне обслуговування туристичних груп, стаціонарні та рухомі засоби розміщення туристів, груповий готельний сервіс.

Horb K. M., Korneyev M. V., Pavlov V.V. Hotel service technologies of tourist groups

The article is devoted to the analysis of approaches, generalization of experience and systematization of technologies of hotel service of tourist groups at different stages of the guest cycle, in different formats of tourist trips in different means of accommodation of tourists, including mobile ones (tourist trains, cruise ships, buses and airplanes with sleeping places, etc.), with the aim of further standardization and improvement of the group service process in tourism and hospitality. Technological processes of service of tourist groups in hotels have a number of significant differences from the provision of hotel services to individual guests. At the stage of booking hotel services for a group of tourists, it is important for the customer to provide the hotel with a list of tourists with their passport data, to agree on the distribution of tourists by rooms of different categories, with the priority of compact mutual arrangement of rooms where the tourists of the group will live, as well as the issue of group trans-

© К. М. Горб, М. В. Корнєєв, В. В. Павлов, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

fer to the hotel upon arrival and return transfer upon departure, to take care of parking at the hotel of the group vehicle. At the accommodation stage, it is important to promptly settle the group members in the hotel rooms through the coordinated actions of the hotel administrator on duty and the tour group guide. During the group's stay at the hotel, it is important to ensure the compatibility of the purpose of the group's stay with the purpose of the rest of the guests. At the last stage of final settlements, eviction and departure of the group from the hotel, it is important to ensure the efficiency of mutually agreed actions regarding checking rooms, handing over keys, packing personal belongings and loading them into a vehicle, etc. When providing group hotel services to tourists at each stage of the guest cycle, the key task is to avoid queues and minimize the time for any waiting by the majority of the group of the rest of the tourists. All of the above provisions for hotel services for tourist groups are valid for both stationary facilities and accommodation establishments and for mobile ones, namely for those tourist vehicles where it is intended to provide tourists with sleeping places for overnight stays. In the future, it is appropriate to pay attention to a more detailed justification of the concept of a group guest cycle of an accommodation facility, the essence of which will be the consistency and continuity of providing group hotel service.

Key words: group guest cycle of an accommodation facility, group tourism, hotel service for tourist groups, stationary and mobile tourist accommodation facilities, group hotel service.

Постановка проблеми. Питання організації та технологій надання як туристичних, так і готельних послуг є широко опрацьованими та достатньо глибоко розробленими в сучасній науковій та навчально-методичній літературі та активно використовуються на практиці, модернізуючись та оновлюючись із часом відповідно до поступового й послідовного опанування досягнень науково-технічного прогресу та впровадження інноваційних технологій. Проте доволі слабо, з нашої точки зору, простежуються обґрунтування різних форматів туристичного обслуговування в готелях, зокрема, надання комплексу готельних послуг туристичним групам, сформованим як туроператорськими та турагентськими компаніями, так і «самостійно» на корпоративній основі. Процес саме групового обслуговування у засобах розміщення туристів має низку специфічних особливостей, що потребує окремого розгляду та обґрунтування. Тому можна вважати актуальним аналіз та систематизацію процесу готельного обслуговування туристичних груп.

Аналіз останніх досліджень і публікацій. Інноваційні та традиційні технології та схеми обслуговування гостей, туристів у закладах розміщення є предметом вивчення багатьох українських та зарубіжних авторів як у науковій літературі, так і в навчальних та навчально-методичних працях. Їх обґрунтування можна зустріти як у роботах, що охоплюють предметне поле організації туристичної діяльності, так і у працях, присвячених сервісним процесам виключно у царині готельного господарства. Певні роботи присвячені схемам співпраці готелів з туроператорами, але здебільшого вони стосуються організаційно-економічних та правових питань, не розглядаючи детально техніки обслуговування гостей, запрошених туроператором як посередником головним чином між готелем та туристами, у т. ч. і у груповому форматі.

Серед науковців, що присвятили свої розробки вказаним питанням готельного сервісу для туристів, першочергово слід назвати О. Куракіна [6], Д. Кучер [7], О. Нодь [10], І. Тіщенко [17], А. Козлову [4], Т. Лисюк, О. Терещук, М. Пасічника [9], О. Романуху [12], Г. Круль, О. Заячук [5], І. Нестерчук, С. Тищенко, Т. Чернишова, Б. Шевчука, А. Осіпчук [15], О. Сущенко, Ю. Лолу, Н. Козубову [14], Р. Ладиженську [8], Т. Ткаченко, Є. Козловського [18], К. Трухачову, Н. Власенко, О. Воскресенську, А. Софієнко, В. Шукліну [20], І. Чуєву, Л. Ніколаєву, Н. Косарєву, [21], Т. Скутар [16], В. Джинджояна [3], О. Юдіну, Н. Небабу, Є. Сайгак, Д. Татаренко [1] тощо. Вказані та деякі інші праці лягли в основу авторських розробок та пропозицій щодо технологій готельного обслуговування туристичних груп у цій статті.

Метою статті є аналіз підходів, узагальнення досвіду та систематизація технологій готельного обслуговування туристичних груп на різних етапах гостьового циклу, за різних форматів туристичних подорожей у різних засобах розміщення туристів, включно з рухомими (туристичні потяги, круїзні судна, автобуси та літаки зі спальними місцями тощо), з метою подальшої стандартизації та удосконалення групового сервісного процесу в туризмі та гостинності.

Виклад основного матеріалу. В основу систематизації технологій обслуговування гостей у закладах готельного господарства в теорії та на практиці покладено, у першу чергу, циклічність процесу обслуговування, що відповідає концепції гостьового циклу закладу розміщення як системи послідовних етапів єдиного сервісного процесу для гостя, починаючи з бронювання місця або номеру в готелі та закінчуючи остаточним розрахунком за усі отримані протягом перебування послуги та товари і виїздом гостя з готелю ([5], [15] та ін.). Відповідно до даної усталеної концепції, гостьовий технологічний цикл складається із чотирьох узгоджених послідовних етапів, на кожному з яких застосовуються окремі, специфічні за технологією надання послуги, різні за генезисом та характером, але однакові за рівнем сервісу.

У більшості робіт з обґрунтування гостьового циклу кожен його етап розглядається здебільшого виключно з позиції індивідуального обслуговування від одного до декількох гостей, що придбали даний цикл готельних послуг за роздрібною ціною. В окремих роботах зазначається можливість наявності посередника на першому етапі циклу (бронювання та замовлення) [15], але при цьому не акцентується на груповому характері надання послуг. Може бути представлено та проаналізовано специфіку фінансово-економічного або юридичного оформлення таких заявок, але з технологічної точки зору готельний сервіс при групових заїздах, перебуваннях та виїздах туристів фактично не розглядається.

У зв'язку з вищевикладеним ми здійснили спробу провести порівняльний аналіз специфіки готельного обслуговування «індивідуальних» гостей, з одного боку, та корпоративних і туристичних груп, з іншого, що в узагальненому вигляді представлено у табл. 1. Варто проаналізувати специфіку цих відмін з акцентом на технологічних особливостях групового сервісу на кожному етапі гостьового циклу.

Перший етап готельного обслуговування у межах зазначеного циклу передбачає обрання готельного закладу для споживання послуг та замовлення і бронювання номерів, зазвичай у дистанційному форматі. За умови бронювання номерів для групи гостей (як правило, від 6 до 50 осіб) в туropolерейтингу прийнято, крім домовленостей між замовником (туроператором) та виконавцем (готелем) щодо характеру оплати та розподілу прав і відповідальності, обов'язково узгодити кількість туристів у групі та розподілити їх по категоріях номерів готелю (можливо, і по конкретних номерах) згідно попередньо ретельно з'ясованих побажань самих туристів. Характерною рисою таких замовлень може бути і прагнення туроператора до більш компактного розміщення групи по готельних номерах, у міру можливостей. На цьому етапі туроператор передає готелю й ексклюзивну інформацію з паспортними даними кожного туриста. З'ясовуються також питання часу прибуття і заселення групи, терміну перебування в готелі та часу виселення і виїзду групи. За умови автобусного туру визначається характер паркування і перебування автобусу в режимі відстою протягом проживання групи в готельному закладі.

Таблиця 1

Відмінності між індивідуальним та груповим готельним обслуговуванням туристів протягом гостьового циклу

Етап гостьового циклу	Специфіка готельного обслуговування	
	індивідуального	групового
Бронювання (Reservation)	Консультації клієнта здебільшого в дистанційно-заочному форматі, оформлення індивідуальної заявки	Переговори з посередником (туроператором, турагентом, івент-компанією) або представником корпоративного замовлення, оформлення групової заявки із зазначенням категорії номеру для кожного туриста
Прибуття (Arrival), трансфер до готелю, реєстрація (Check in Procedure) та розміщення клієнта (Accommodation)	Сприяння зустрічі гостя у транспортному хабі (на вокзалі, в аеропорті тощо) та його трансферу до готелю; паркування (за потреби) індивідуального транспортного засобу, оформлення проживання за індивідуальним готельним тарифом та заселення в номер	Сприяння зустрічі групи в хабі, безпосередньо зустріч поблизу готелю, паркування групового транспортного засобу (за необхідності), швидке поселення групи під контролем старшого групи (супроводжувача) згідно з попередніми домовленостями з організатором подорожі
Проживання (Staying) та надання основних і додаткових послуг протягом терміну перебування	Надання основних послуг згідно готельному тарифу та додаткових послуг за бажанням гостя	Стандартне обслуговування кожного туриста згідно домовленостей з організатором подорожі та надання додаткових послуг туристам в індивідуальному форматі згідно їх бажань, здебільшого з відома супроводжувача групи
Остаточний розрахунок за усі отримані під час перебування послуги та товари, виїзд (Departure) та зворотній трансфер	Узгодження та кінцеві розрахунки гостя за увесь наданий у готелі сервіс, виселення, сприяння трансферу до транспортного хабу (за потреби), провадження гостя	Оперативне виселення групи з готелю, остаточний розрахунок за надані основні послуги супроводжувачем групи згідно домовленостей з організатором подорожі, за додаткові – індивідуально з кожним туристом, сприяння трансферу до хабу (за потреби), провадження групи

Джерело: складено авторами

Починаючи з другого етапу та до кінця циклу весь організаційно-технологічний сервісний процес під'їзду, заселення, перебування і виселення туристичної групи з готелю відбувається виключно через супроводжувача – офіційного представника туристичної компанії (ширше – організатора подорожі), який стає головним посередником між адміністрацією готельного закладу та туристичною групою і певною мірою долучається до «команди сервісного забезпечення» групи на час її проживання в готелі, діє у тісній співпраці з представниками усіх служб закладу розміщення, що обслуговують туристів групи, від адміністратора до покоївки.

На другому етапі гостьового циклу в умовах обслуговування туристичної групи приблизно за годину до прибуття в готель супроводжувач зв'язується з черговим адміністратором та уточнює спосіб і час прибуття групи до готелю. Питання трансферу в цьому випадку зазвичай вже узгоджені попередньо, у випадку автобусного туру туристичний автобус поєднує функції основного («далекого») та трансферного транспорту. При прибутті до готелю супроводжувач оперативно вирішує питання заселення групи до номерів та паркування або влаштування транспортного засобу «на відстій».

Сам процес заселення до номерів згідно попередньо узгодженої схеми з метою уникнення черг та тривалих очікувань туристів групи відбувається спрощено. Власний паспорт адміністратору готелю пред'являє та залишає на певний час лише супроводжувач групи. Адміністратор послідовно видає супроводжувачу ключі від кожного номеру згідно зі сформованою ще перед початком подорожі схемою розселення, після чого відразу супроводжувач передає ключ тим туристам, що заселятимуться до цього номера, відмічаючи у своєму списку, хто з туристів групи у якому номері проживає.

Після отримання ключів туристи поступово заселяються до номерів. При цьому черговий адміністратор та його помічники, а також носії активно проявляють гостинність, всіляко сприяючи комфортному та оперативному заселенню: проводять або роз'яснюють короткий шлях до номеру, за потреби допомагають доправити важкі валізи.

Інформацію щодо правил перебування в готелі, часу та місця сеансів харчування, можливостей надання додаткових послуг тощо може бути надано туристам групи як під час під'їзду до готелю ще в транспортному засобі, так і перед видачею ключів або вже після неї. У будь-якому разі відразу після закінчення процедури заселення групи супроводжувач прописує у груповому чаті всі актуальні відомості щодо перебування в готелі та орієнтування у його просторі.

Після заселення групи у деяких туристів можуть виникнути питання (або навіть претензії чи скарги) щодо сервісного забезпечення проживання в номері: деє може не вистачати ковдр або рушників, хтось із туристів не може самостійно з'ясувати спосіб вмикання телевізора або кондиціонера, комусь може терміново знадобитися пароль від готельного Wi-Fi, і т. ін. Усі ці питання вирішуються з представниками адміністрації готелю виключно через супроводжувача групи.

Тільки після завершення процедури заселення групи і вирішення всіх вищезгаданих поточних питань супроводжувач сам заселяється до відповідного номеру на проживання та повідомляє групі (зазвичай через чат), в якому номері він проживає. Як і протягом усієї подорожі, під час перебування групи в готелі супроводжувач цілодобово знаходиться «в режимі чергового» для забезпечення максимально комфортного відпочинку кожного туриста групи.

Третій етап гостьового циклу з точки зору варіативності сценаріїв перебування гостей у готелі та різноманіття послуг, що надаються гостям, є найскладнішим та залежить як від тривалості, так і від мети такого перебування, і відповідно від категорії та профілю готельного закладу. Це стосується й «індивідуальних» гостей, і туристичних груп.

Найпростішим варіантом проживання групи в готелі є її поселення виключно на одну ночівлю, з вечора до ранку, що характерно здебільшого для екскурсійно-пізнавальних турів. У цьому випадку різноманіття послуг є мінімальним, перебування зазвичай включає лише користування зручностями готельного номеру для відновлення фізичних і частково моральних сил, та сніданок в ресторанній залі наступного дня. Але і в таких умовах дехто з туристів може скористатися і додатковими послугами при вечірньому відвідуванні ресторану при готелі, спа-комплексу та басейну тощо.

У випадку відносно тривалого перебування туристів у готелі, у т. ч. і у складі груп, що є типовим передусім для відпочинкових турів та курортних готелів, різноманіття рекреаційних занять, а відповідно і наданих послуг, суттєво розширюється. У якості як основних (що входять і в готельний тариф, і в туристичний пакет), так і додаткових (за окрему плату) послуг за таких умов можуть виступати оздоровчі та лікувальні, екскурсійні та похідно-спортивні, розважально-анімаційні та видовищні, додаткове харчування, побутові (перукарські, чищення та прання тощо), торгівельні послуги та ін.

Обидва з вищенаведених прикладів крайнього прояву широти спектру готельних послуг на даному етапі гостьового циклу зазвичай не передбачають суттєвих відмін між обслуговуванням «індивідуальних» гостей та туристичних груп, крім можливого варіанту компактної розсадки учасників групи у залі ресторану під час сеансів харчування. В усьому іншому учасники туристичної групи (для великих готелів, можливо, одночасно і не однієї) на час проживання зазвичай ніби «розчиняються» серед інших гостей і відпочивають відповідно до своїх уподобань та можливостей готельного сервісу.

Проте система відносин та стосунків всередині туристичної або корпоративної групи, контрастність відміни внутрішньогрупової корпоративної культури від загальноприйнятої в готелі атмосфери перебування гостей згідно його профілю та концепції, в окремих випадках може викликати конфлікти, що негативно позначиться на репутації і готелю, і туристичної компанії – організатора подорожі і замовника корпоративних готельних послуг. Міра такої сумісності і загальної «вписуваності» групи у загальний «гостьовий процес» готелю, на наш погляд, має своєчасно, вже на етапі проєктування туру та попередніх домовленостей організатора подорожі з готелем, відслідковуватись та аналізуватись обома сторонами з метою уникнення

зазначеного роду конфліктів. Наприклад, замовлення проведення весілля, конференції, святкування ювілею фірми та інших корпоративних заходів у готельно-ресторанному комплексі завжди вимагає оцінки сумісності поведінки корпоративної спільноти з умовами перебування решти гостей.

Окремо у світлі вищевикладеного варто зазначити реалізацію пакету послуг конференц-сервісу готелю на третьому етапі гостьового циклу. Він може бути замовлений туристичною (зазвичай туроператорською) або івент-компанією, що працює, в тому числі, й у сфері MICE-туризму, так і безпосередньо компанією – організатором конференц-заходу. У цьому випадку формується своєрідний проміжний варіант готельного обслуговування, що включає елементи і групового, й індивідуального сервісу. Гості, що є учасниками заходу, можуть з'їжджатися до готелю та після завершення роз'їжджатися у різний час з (до) різних місць, що потребує для кожного з них окремих послуг зустрічі, трансферу та поселення, виселення та проводжання. Але програма перебування в готелі для них усіх є єдиною, і питання її сумісності із загальним відпочинковим форматом проживання решти гостей готелю повинні, з нашої точки зору, бути предметом підвищеної уваги і замовників, і реалізаторів пакету MICE-послуг у готелі.

Останній (четвертий) етап гостьового циклу передбачає серію організаційних та технологічних процедур щодо остаточних розрахунків гостей за всі отримані додаткові послуги (в окремих ситуаціях – і за товари), їх виселення, проводжання та зворотного трансферу. У випадку виїзду з готелю туристичної групи, особливо за умов дефіциту часу (наприклад, сніданок о 7.00, виїзд групи о 8.00), надзвичайно важливим є забезпечення оперативності оформлення виїзду. Це стосується і «здачі» та перевірки стану номерів, і остаточних розрахунків кожного туриста групи за усі отримані та неоплачені (або недоплачені) додаткові готельні сервіси, і збирання особистих речей та їх завантаження у транспортний засіб, і т. ін. Контрольні функції забезпечення такої оперативності покладаються на супроводжувача групи, який і підбадьорює своїх туристів, і в міру можливостей та потреб прискорює роботу співробітників готелю.

Окремо слід згадати і технології надання послуг готельного генезису на транспортних засобах, де передбачено ночівлю на спальних місцях (та частково харчування) під час їх руху в процесі туристичних подорожей. Традиційно такі послуги відносять до транспортних, а не до послуг розміщення, хоча за своєю технологічною специфікою вони відрізняються від класичних стаціонарних закладів розміщення в основному лише рухомістю місць ночівлі. Особливо, на наш погляд, це важливо для чартерних рейсів на засобах усіх видів транспорту, де обслуговування ночівель на спальних місцях та харчувань (як у відкритому просторі в салоні (плацкарт), так і в окремих приміщеннях (каюти, купе тощо)) має виключно груповий формат.

Відповідні туристичні засоби транспорту досить часто називають «готелями на колесах», «плавучими» та «літаючими» готелями, особливо це характерно для круїзних суден та висококласних туристичних потягів. Технологічна сутність забезпечення готельного сервісу як для груп, так і для індивідуальних туристів у цих випадках фактично не відрізняється від його стаціонарного варіанту, за окремими винятками, пов'язаними з логістичними питаннями графіку поновлення відповідних ресурсів (води, джерел енергії, продуктів харчування та напоїв, засобів чищення або нових комплектів постільної білизни тощо) не лише в часі, а також і у просторі. Відміни між індивідуальним та груповим форматом готельного обслуговування на транспортних засобах зі спальними місцями фактично не відрізняються від стаціонарних засобів розміщення.

Отже, в цілому можна стверджувати, що готельне обслуговування туристичних груп на окремих етапах гостьового циклу в технологічному вимірі суттєво відрізняється від забезпечення готельного сервісу для індивідуальних туристів та гостей, незалежно від того, придбали вони даний сервісний пакет особисто в готелі чи через посередника, у складі більшого пакету послуг (наприклад, туристичної подорожі або програми конференції), чи як окремий готельний продукт. Головні технологічні відміни групового сервісного формату від індивідуального в цьому контексті спостерігаються насамперед у процедурах поселення групи у готель та виселення з нього, суттєве значення має «вписуваність» групи у загальний антураж та атмосферу життя готелю, сумісність мети перебування групи з метою перебування решти гостей. Усе вищевикладене дає підстави запропонувати окрему концепцію «групового гостьового циклу закладу розміщення», сутність якого полягатиме у послідовності забезпечення готельного сервісу групового формату, з урахуванням його специфіки на кожному етапі, від бронювання компактного блоку номерів до виїзду групи з готелю.

Висновки. Технологічні процеси обслуговування туристичних груп у готелях мають низку суттєвих відмінностей від надання готельних послуг індивідуальним постояльцям незалежно від того, придбали вони дані послуги прямо у готелю чи через посередника. На етапі бронювання готельних послуг для групи туристів важливо замовнику надати готелю список туристів з їх паспортними даними, узгодити розподіл туристів по номерах різних категорій, з пріоритетом компактного взаєморозташування номерів, де проживатимуть туристи групи, а також питання трансферу групи до готелю при заїзді та зворотного трансферу при виїзді, подбати про стоянку при готелі групового транспортного засобу. На етапі розміщення важливо оперативно заселити учасників групи до номерів готелю через узгоджені дії чергового адміністратора готелю та супроводжувача туристичної групи. Під час проживання групи в готелі важливим є забезпечення сумісності мети перебування групи із метою перебування решти гостей. На останньому етапі остаточних розрахунків, виселення та виїзду групи з готелю важливим є забезпечення оперативності взаємоузгоджених дій щодо перевірки номерів, здачі ключів, складання особистих речей та завантаження їх на транспортний засіб тощо.

При груповому готельному обслуговуванні туристів на кожному етапі гостьового циклу ключовим завданням є уникнення черг та мінімізація часу на будь-які очікування більшою частиною групи решти туристів. Усі зазначені положення готельного обслуговування туристичних груп справедливі як для стаціонарних засобів та закладів розміщення, так і для рухомих, а саме для тих туристичних транспортних засобів, де передбачається надання туристам спальних місць для ночівель. У перспективі доречно приділити увагу більш детальному обґрунтуванню концепції групового гостьового циклу закладу розміщення, сутність якого полягатиме у послідовності та безперервності забезпечення готельного сервісу групового формату, головним чином з метою подальшої стандартизації та удосконалення групового сервісного процесу в туризмі та гостинності.

Список використаних джерел:

1. Yudina O. I., Nebaba N. O., Saihak Ye. L., Tatarenko D. K. Peculiarities and modern trends in the organization of hotel services. *Системи та технології*. Т. 69. № 1. 2025. С. 300–306.
2. Горб К. М., Корнеєв М. В., Кучер М. М. Гостинність в готельно-ресторанному та туристичному бізнесі: загальні засади досліджень та особливості підготовки фахівців. *Науковий погляд: економіка та управління*. 2024. № 1 (85). С. 111–119.
3. Джинджоян В. В. Сучасні інноваційні технології в менеджменті туризму та гостинності. *Ефективна економіка*. 2021. № 6. URL: <http://www.economy.nauka.com.ua/?op=1&z=8951>
4. Козлова А. О. Технологія обслуговування в готелях та туркомплексах: конспект лекцій для студентів денної і заочної форм навчання освітнього рівня «бакалавр» за спеціальністю 073 – Менеджмент. Харків : ХНУМГ ім. О. М. Бекетова, 2019. 114 с.
5. Круль Г., Заячук О. Організація і технологія надання послуг гостинності : навч. посіб. Чернівці : Чернівець. нац. ун-т ім. Ю. Федьковича, 2022. 488 с.
6. Куракін О. Б. Дослідження методологічних підходів до оцінки якості готельних послуг. *Innovations and technologies in the service sphere and food industry*. № 1 (7). 2023. С. 5–9.
7. Кучер Д. Б. Підвищення якості готельних послуг через внутрішню оптимізацію процесів надання послуг на вітчизняних підприємствах готельного господарства. *Вісник Чернівецького торговельно-економічного інституту. Економічні науки*. 2012. Вип. 1. С. 230–237.
8. Ладиженська Р. С. Конспект лекцій з дисципліни «Технологія обслуговування в готелях і туркомплексах» (для студентів 4 курсу усіх форм навчання за напрямом підготовки «Менеджмент»). Харків : ХНАМГ, 2010. 254 с.
9. Лисюк Т. В., Терещук О. С., Пасічник М. П. Інноваційні технології у готельно-ресторанному господарстві. *Економіка та суспільство*. 2022. Вип. 40. URL: <https://economyandsociety.in.ua/index.php/journal/article/view/1450/1395>
10. Нодь О. Інноваційні тенденції в сфері готельного обслуговування в умовах сучасних викликів. *Modeling the development of the economic systems*. 2024. № 2. С. 289–294.
11. Опорний конспект лекцій з дисципліни «Організація сфери туристичних послуг» для студентів спеціальності 242 «Туризм» денної та заочної форм навчання / Укл.: С. В. Гресь-Євреїнова. Запоріжжя: НУ «Запорізька політехніка», 2021. 43 с.
12. Романуха О. М. Методичні рекомендації з вивчення дисципліни «Технологія готельних послуг», ступінь «бакалавр». Кривий Ріг : ДонНУЕТ, 2021. 62 с.
13. Співпраця з туристичними агентствами: як залучати групових та корпоративних клієнтів. URL: <https://hotel-solution.com.ua/news/cooperation-with-travel-agencies-how-to-attract-group-and-corporate-clients>
14. Сущенко О. А., Лола Ю. Ю., Козубова Н. В. Готельний бізнес : навчальний посібник для студентів спеціальності 242 «Туризм» першого (бакалаврського) рівня. Харків : ХНЕУ ім. С. Кузнеця, 2018. 167 с.
15. Технологія готельної справи: навчальний посібник / Укл.: І. К. Нестерчук, С. В. Тищенко, Т. М. Чернишова, Б. Л. Шевчук, А. С. Осіпчук. Житомир : 2019. 223 с.
16. Технологія туристичної діяльності : лабораторний практикум / укл. Т. Д. Скутар. Чернівці : Чернівець. нац. ун-т ім. Ю. Федьковича, 2020. 64 с.
17. Тищенко І. В. Сучасні тенденції інновацій у сфері готельного господарства. *Трансформаційна економіка*. 2025. № 2 (11). С. 118–122.
18. Ткаченко Т. І., Козловський Є. В. Особливості встановлення договірних відносин туристичного оператора з постачальниками послуг, агентами та партнерами. *Економіка та держава*. 2018. № 5. С. 85–88.
19. Топові технології в готельному господарстві: як покращити сервіс та зручність гостей. URL: <https://hotel-solution.com.ua/news/top-technologies-in-the-hotel-industry-how-to-improve-service-and-convenience-of-guests>
20. Трухачова К. В., Власенко Н. А., Воскресенська О. Є., Софієнко А. В., Шукліна В. В. Особливості технологій обслуговування туризму в event-менеджменті: екополітика, логістика, маркетинг у інфраструктурно-ресурсному та анімаційному плануванні події. *Вісник Херсонського національного технічного університету*. 2021. № 4. С. 231–239.

21. Чуєва, І., Ніколаєва, Л., Косарева, Н. М. Інноваційні технології туристичної діяльності на світовому ринку туристичних послуг. *Економіка та суспільство*. 2021. № 30. URL: <https://economyandsociety.in.ua/index.php/journal/article/view/622/596>

References:

1. Yudina, O. I., Nebaba, N. O., Saihak, Ye. L., & Tatarenko, D. K. (2025). Peculiarities and modern trends in the organization of hotel services. *Systemy ta tekhnolohii*. T. 69. № 1. P. 300–306.
2. Horb, K. M., Kornieiev, M. V., & Kucher, M. M. (2024). Hostynnist v hotelno-restorannomu ta turystychnomu biznesi : zahalni zasady doslidzhen ta osoblyvosti pidhotovky fakhivtsiv [Hospitality in the hotel, restaurant and tourism business: general principles of research and features of specialist training]. *Naukovyi pohliad: ekonomika ta upravlinnia*. № 1 (85), pp. 111–119.
3. Dzhyndzhoian, V. V. (2021). Suchasni innovatsiini tekhnolohii v menedzhmenti turyzmu ta hostynnosti [Modern innovative technologies in tourism and hospitality management]. *Efektivna ekonomika*. № 6. Retrieved from: <http://www.economy.nayka.com.ua/?op=1&z=8951>
4. Kozlova, A. O. (2019). Tekhnolohiia obsluhovuvannia v hoteliakh ta turkompleksakh: konspekt lektsii dlia studentiv dennoi i zaochnoi form navchannia osvithnoho rivnia «bakalavr» za spetsialnistiu 073 – Menedzhment [Service technology in hotels and tourist complexes: lecture notes for full-time and part-time students of the «bachelor» educational level in the specialty 073 – Management]. Kharkiv : KhNUMH im. O. M. Beketova. 114 p.
5. Krul, H., & Zaiachuk, O. (2022). Orhanizatsiia i tekhnolohiia nadannia posluh hostynnosti : navch. posib. [Organization and technology of providing hospitality services: a textbook]. Chernivtsi : Chernivets. nats. un-t im. Yu. Fedkovycha. 488 p.
6. Kurakin, O. B. (2023). Doslidzhennia metodolohichnykh pidkhodiv do otsinky yakosti hotelnykh posluh [Research on methodological approaches to assessing the quality of hotel services]. *Innovations and technologies in the service sphere and food industry*. № 1 (7). P. 5–9.
7. Kucher, D. B. (2012). Pidvyshchennia yakosti hotelnykh posluh cherez vnutrishniu optymizatsiiu protsesiv nadannia posluh na vitchyzniannykh pidpriemstvakh hotelnoho hospodarstva [Improving the quality of hotel services through internal optimization of service provision processes at domestic hotel enterprises]. *Visnyk Chernivetskoho torhovelno-ekonomichnoho instytutu. Ekonomichni nauky*. Vyp. 1. P. 230–237.
8. Ladyzhenska, R. S. (2010). Konspekt lektsii z dystsypliny «Tekhnolohiia obsluhovuvannia v hoteliakh i turkompleksakh» (dlia studentiv 4 kursu usikh form navchannia za napriamom pidhotovky «Menedzhment») [Lecture notes on the discipline «Service Technology in Hotels and Tourist Complexes» (for 4th year students of all forms of study in the field of «Management»)]. Kharkiv : KhNAMH. 254 p.
9. Lysiuk, T. V., Tereshchuk, O. S., & Pasichnyk, M. P. (2022). Innovatsiini tekhnolohii u hotelno-restorannomu hospodarstvi [Innovative technologies in the hotel and restaurant industry]. *Ekonomika ta suspilstvo*. Vypusk 40. Retrieved from: <https://economyandsociety.in.ua/index.php/journal/article/view/1450/1395>
10. Nod, O. Innovatsiini tendentsii v sferi hotelnoho obsluhovuvannia v umovakh suchasnykh vyklykiv (2024). [Innovative trends in the field of hotel services in the face of modern challenges]. *Modeling the development of the economic systems*. № 2. P. 289–294.
11. Opornyi konspekt lektsii z dystsypliny «Orhanizatsiia sfery turystychnykh posluh» dlia studentiv spetsialnosti 242 «Turyzm» dennoi ta zaochnoi form navchannia (2021). [Reference lecture notes on the discipline «Organization of the tourism services sector» for students of specialty 242 «Tourism» of full-time and part-time forms of study] / Ukl.: S. V. Hres-Yevreinova. Zaporizhzhia: NU «Zaporizka politekhnika». 43 p.
12. Romanukha, O. M. (2021). Metodychni rekomendatsii z vyvchennia dystsypliny «Tekhnolohiia hotelnykh posluh», stupin «bakalavr» [Methodological recommendations for studying the discipline «Hotel Services Technology», Bachelor's degree]. Kryvyi Rih : DonUET. 62 p.
13. Spivpratsia z turystychnymy ahentstvamy: yak zaluchaty hrupovykh ta korporatyvnykh kliientiv [Cooperation with travel agencies: how to attract group and corporate clients]. Retrieved from: <https://hotel-solution.com.ua/news/cooperation-with-travel-agencies-how-to-attract-group-and-corporate-clients>
14. Sushchenko, O. A., Lola, Yu. Yu., & Kozubova, N. V. (2018). Hotelnyi biznes : navchalnyi posibnyk dlia studentiv spetsialnosti 242 «Turyzm» pershoho (bakalavrskoho) rivnia [Hotel business: a study guide for students of specialty 242 «Tourism» of the first (bachelor's) level]. Kharkiv : KhNEU im. S. Kuznetsia. 167 p.
15. Tekhnolohiia hotelnoi spravy: navchalnyi posibnyk (2019) [Hotel Technology: A Study Guide] / Ukl.: I. K. Nesterchuk, S. V. Tyshchenko, T. M. Chernyshova, B. L. Shevchuk., A. S. Osipchuk. Zhytomyr :. 223 p.
16. Tekhnolohiia turystychnoi diialnosti : laboratornyi praktykum (2020) [Technology of tourism activities: laboratory workshop] / ukl. T. D. Ckutar. Chernivtsi : Chernivets. nats. un-t im. Yu. Fedkovycha., 64 p.
17. Tishchenko, I. V. (2025). Suchasni tendentsii innovatsii u sferi hotelnoho hospodarstva [Current trends in innovation in the hospitality industry]. *Transformatsiina ekonomika*. № 2 (11). P. 118–122.
18. Tkachenko, T. I., & Kozlovskiy, Ye. V. (2018). Osoblyvosti vstanovlennia dohovirnykh vidnosyn turystychnoho operatora z postachalnykamy posluh, ahentamy ta partneramy [Features of establishing contractual relations between a tour operator and service providers, agents and partners]. *Ekonomika ta derzhava*. № 5. P. 85–88.

19. Topovi tekhnolohii v hotelnomu hospodarstvi: yak pokrashchyty servis ta zruchnist hostei [Top technologies in the hotel industry: how to improve service and guest comfort]. Retrieved from: <https://hotel-solution.com.ua/news/top-technologies-in-the-hotel-industry-how-to-improve-service-and-convenience-of-guests>

20. Trukhachova, K. V., Vlasenko, N. A., Voskresenska, O. Ye., & Sofienko, A. V., Shuklina, V. V. (2021). Osoblyvosti tekhnolohii obsluhovuvannia turyzmu v event-menedzhmenti: ekopolityka, lohistyka, marketynh u infrastruktorno-resursnomu ta animatsiinomu planuvanni podii [Features of tourism service technologies in event management: environmental policy, logistics, marketing in infrastructure, resource and animation planning of an event]. *Visnyk Khersonskoho natsionalnoho tekhnichnoho universytetu*. № 4. P. 231–239.

21. Chuieva, I., Nikolaieva, L., & Kosarieva, N. M. (2021). Innovatsiini tekhnolohii turystychnoi diialnosti na svitovomu rynku turystychnykh posluh [Innovative technologies of tourism activities in the global market of tourist services]. *Ekonomika ta suspilstvo*. № 30. Retrieved from: <https://economyandsociety.in.ua/index.php/journal/article/view/622/596>

Дата надходження статті: 30.10.2025

Дата прийняття статті: 17.11.2025

Опубліковано: 30.12.2025

UDC 658.381

DOI <https://doi.org/10.32782/2521-6643-2025-2-70.42>

Horozhankina N. A., PhD in Geography, Associate Professor
at the Department of International Tourism and Hotel and Restaurant
Business University of Customs and Finance
ORCID: 0000-0003-2773-8630

Saihak Ye. L., PhD in Economics, Department of International
Tourism and Hotel and Restaurant Business
University of Customs and Finance
ORCID: 0000-0002-6406-9272

Grytsenko A. V., PhD in Economics, Department of International
Tourism and Hotel and Restaurant Business
University of Customs and Finance
ORCID: 0000-0002-8986-0714

Topol D. S., University of Customs and Finance
ORCID: 0009-0004-7910-0692

OPTIMISATION OF THE STAFF MOTIVATION SYSTEM IN THE HOTEL BUSINESS

The personnel motivation system is considered on the example of LLC "Hotel World (the hotel "Dnipropetrovsk")". Based on the questionnaire method, a check was made of the compliance of the existing personnel motivation system with the requirements of the company's employees.

The compliance of the existing personnel motivation system with the requirements of the company's employees was checked by means of a questionnaire using a specially developed questionnaire. All hotel employees (100%) participated in the questionnaire.

Factors of employee dissatisfaction with their work were identified, and the need to improve the personnel policy for employees was assessed. 38% of respondents are completely satisfied with their work in the institution, 43% are more satisfied than dissatisfied, 14% are more dissatisfied and 5% are completely dissatisfied. It was found that 60% of employees are satisfied with the organisation of their workplace, 40% are completely or partially dissatisfied.

An assessment of the motivation system was conducted to assess the motivation system of the staff of the hotel "Dnipropetrovsk". It was found that many motivational elements of the hotel "Dnipropetrovsk" require improvement. The main problems associated with the motivation system are: lack of assessment of well-organised professional development and training activities for staff; lack of an adaptation process for employees; lack of career planning for employees; lack of awareness of employees; low level of financial incentives for employees.

As a result of the audit of the personnel management system of the hotel "Dnipropetrovsk", it was found that the staff motivation system at the hotel needs improvement, which affects labour productivity.

The following measures are recommended to improve the staff motivation system at the hotel "Dnipropetrovsk":
1. Implement an effective staff motivation system that increases labour productivity. When using schemes and bonuses in the salary structure, there should be a clear connection between the contribution of a particular employee and the amount of compensation. 2. Improve the system of material and non-material incentives for employees, including the introduction of additional salary payments at the end of the year, the issuance of certificates of the best employees, etc.

A SWOT analysis of the staff motivation system of the hotel "Dnipropetrovsk" was conducted, based on which the shortcomings of the staff motivation system were identified: the staff motivation system is represented only by wages in the form of direct compensation and is practically limited; failure to use the entire motivational potential of employees to improve existing processes and results of the enterprise.

The main opportunity of the hotel "Dnipropetrovsk" is to gain time during the recovery period after the economic crisis, reconstruction after Ukraine's victory in the war with Russia and use it to develop a motivation system that meets modern market conditions.

Key words: staff motivation, personnel management, staff turnover, hotel "Dnipropetrovsk", business, optimization, system.

© N. A. Horozhankina, Ye. L. Saihak, A. V. Grytsenko, D. S. Topol, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

Горожанкіна Н. А., Сайгак Є. Л., Гриценко А. В., Тополь Д. С. Оптимізація системи мотивації персоналу в готельному бізнесі

Розглянуто систему мотивації персоналу на прикладі ТОВ «Готельний світ (готель «Дніпропетровськ»)). На основі методу анкетування проведено перевірку відповідності існуючої системи мотивації персоналу вимогам співробітників компанії.

Виявлено фактори незадоволеності працівників роботою та оцінено необхідність покращення кадрової політики щодо працівників. 38% респондентів повністю задоволені своєю роботою в установі, 43% більш задоволені, ніж незадоволені, 14% більш не задоволені і 5% повністю не задоволені. Встановлено, що 60% співробітників задоволені організацією свого робочого місця, 40% повністю або частково незадоволені. Аналіз бажаного методу винагороди показав, що 50% респондентів отримували бонуси. Організація колективних заходів показує бажання людей об'єднатися. 9% співробітників хочуть продемонструвати свої таланти, а 11% зацікавлені в більш значущій роботі.

Проведено оцінку системи мотивації персоналу готелю «Дніпропетровськ». Встановлено, що багато мотиваційних елементів готелю «Дніпропетровськ» вимагають поліпшення. Основними проблемами, пов'язаними з системою мотивації персоналу, є: відсутність оцінки добре організованих заходів з професійного розвитку та навчання персоналу; відсутність процесу адаптації для співробітників; відсутність планування кар'єри для співробітників; відсутність обізнаності співробітників; низький рівень фінансового стимулювання для працівників.

Запропоновано заходи для поліпшення системи мотивації персоналу в готелі «Дніпропетровськ»: впровадити ефективну систему мотивації персоналу для підвищення його продуктивності праці; вдосконалити систему матеріального та нематеріального стимулювання працівників, що включають введення додаткових виплат заробітної плати в кінці року, видачу сертифікатів найкращих співробітників.

Ключові слова: мотивація персоналу, управління персоналом, плинність кадрів, готель «Дніпропетровськ», бізнес, оптимізація, система.

Problem statement. In hotel businesses, the main task is to create an optimal management system in which staff motivation is an important organisational component. In the current realities of a market economy, the functioning of any structure is impossible without regular, systematic study and determination of the effectiveness of the use of available resources, which includes assessing the productivity of personnel management mechanisms.

In every enterprise that uses a modern approach to staff motivation, the human resources of the enterprise (including hotel businesses) are the most important resource, and the development of its potential is considered one of the most important goals.

Analysis of recent research and publications. The main methodological works can be considered the works of the authors: O. Havrylova, L. Kovalenko, V. Nepochatenko, I. Chernysh, M. Kozyk, D. Gayova, M. Malska, I. Pandyak and others.

The goal of the article: to study the optimisation of the personnel motivation system in the hotel business using the example of the activities of the enterprise LLC "Hotel World (the hotel "Dnipropetrovsk")".

Presentation of the main material. Company-wide HR policy is the most important part of HR management. To achieve current goals, it is important to apply certain strategies, such as employee motivation, when creating.

The goal of motivation management in personnel management is to realise a person's abilities to perform the tasks assigned to them. Motivating and encouraging employees means changing the structure of values and interests in accordance with the specified parameters, developing a core of appropriate motivation, and providing long-term influence on employees in order to develop the potential of employees on this basis [2;5]. In personnel policy in the field of incentives and motivation, it is necessary to set goals for each employee, the achievement of which becomes a guarantee of receiving guaranteed benefits [1].

Incentive is the driving force that not only influences a person's character, but also influences the external state of their work through the profit distribution mechanism, since there is an interest in obtaining certain guaranteed benefits that the employee actively motivates [4].

Stimulating labour actively to achieve the goals of the enterprise stimulates employees to work, increases productivity and quality of work through material and moral influence [3].

Let's move on to considering the personnel motivation system using the example of LLC "Hotel World (the hotel "Dnipropetrovsk")", which is located in the city of Dnipro, Sicheslavsk Naberezhna Street, building 33.

The organisational and legal form of the company "Hotel World" LLC is a limited liability company.

The location of the hotel is very advantageous, opposite is the Dnipro River and a refined embankment with a large number of different cafes, restaurants, coffee shops, the building of the State Circus, the city centre is a 10-minute walk, Taras Shevchenko Park is located 1.5 km away.

The hotel has a three-star status. Each room has cable TV, a balcony and a refrigerator. The private bathroom has a bathtub and free toiletries.

Guests can enjoy European and Ukrainian cuisine in the hotel's restaurants or relax in the bar.

The bus stop "Naberezhna Sicheslavsk Street" is 20 metres from the hotel. The city's train station is 3 km away.

Many employees participate in the personnel management system of the hotel "Dnipropetrovsk". The general director is the head of the enterprise. All department heads carry out his orders.

There is a clearly defined scheme of interaction between managers and subordinates in the hotel "Dnipropetrovsk". The heads of each department ensure a high-quality solution to the assigned tasks. At the same time, the hotel manager and his deputy rely on the work of their subordinates and have more time to solve other tasks of the organisation. Thus, the overall productivity of the hotel is at a high level.

A hotel manager is responsible for performing human resource management functions such as recruiting, hiring, evaluating, and terminating staff. In addition, his responsibilities include the following functions: planning the development of the organisational structure and moral climate of the hotel.

Verification of the compliance of the existing personnel motivation system with the requirements of the company's employees was carried out by means of a questionnaire using a specially developed questionnaire. All hotel employees (100%) participated in the questionnaire.

The study was able to identify factors of employee dissatisfaction with work and assess the need to improve employee personnel policy.

38% of respondents are completely satisfied with their work in the institution, 43% are more satisfied than dissatisfied, 14% are more dissatisfied and 5% are completely dissatisfied (Fig. 1).

The survey found that 60% of employees are satisfied with the organisation of their workplace, while 40% are completely or partially dissatisfied.

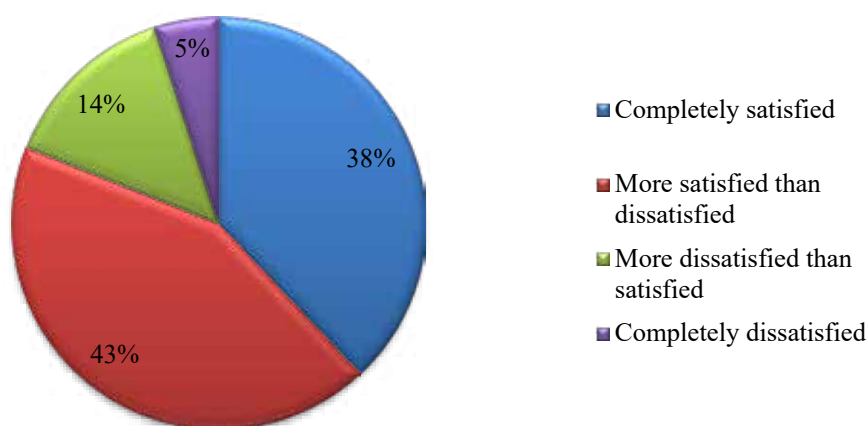


Fig. 1. The hotel "Dnipropetrovsk" staff satisfaction

Source: compiled by the authors

When analysing opportunities for training and improving staff skills, we took into account employee benefits, the effects of individual training and on-the-job training:

- 1) professional development and on-the-job training – 40%;
- 2) training and seminars held outside of working hours on weekends – 10%;
- 3) providing paid time for personal training of personnel – 50%.

An analysis of the preferred reward method showed that 50% of respondents received bonuses. Organising collective events revealed people's desire to unite. 9% of employees want to showcase their talents, and 11% are interested in more meaningful work.

An assessment of the staff motivation system of the hotel "Dnipropetrovsk" was carried out using a five-point scale (table 1).

Analysing the data obtained, we can conclude that many motivational elements of the hotel "Dnipropetrovsk" require improvement. The main problems associated with the motivation system are:

- 1) lack of evaluation of well-organised professional development and training activities for staff;
- 2) lack of an adaptation process for employees;
- 3) lack of career planning for employees;
- 4) lack of awareness of employees;
- 5) low level of financial incentives for employees.

In addition to the undeniable achievements in the field of personnel management of the hotel "Dnipropetrovsk", several important issues were identified during the process of familiarisation and analysis of the company under study.

As a result of an audit of the personnel management system of the hotel "Dnipropetrovsk", it became clear that the hotel's staff motivation system needs improvement, which affects labour productivity.

Table 1

Evaluation of the motivation system at the hotel "Dnipropetrovsk"

Components of motivation	Rating score
1. Organisational culture. A system of values and norms shared by all employees of an enterprise.	3
2. Identification with the legal entity. Visible image of the company from the staff and the outside world.	3
3. Attendance system. Participation of employees in the distribution of general economic results, participation in the capital of the enterprise, development of cooperation.	2
4. Personnel care. All types of social assistance, services provided for recreation, sports, and care for employees who need assistance.	4
5. Workplace organisation. Providing the workplace with ergonomic and organisational support.	4
6. HR policy. Planning and selection of professional development activities and mobile devices takes into account the needs, desires, and professional abilities of employees.	3
7. Organisation of working hours. Convenient setting of working hours according to the needs of employees and companies.	3,5
8. Communicating information to employees. Providing employees with information about the company's work.	3

Source: compiled by the authors

To improve the staff motivation system at the hotel "Dnipropetrovsk", we recommend the following measures:

1. Implement an effective staff motivation system that increases labour productivity. When using schemes and bonus plans in the salary structure, there should be a clear connection between the contribution of a particular employee and the amount of compensation.

2. Improve the system of material and non-material incentives for employees, including the introduction of additional salary payments at the end of the year, issuing certificates to the best employees, etc.

A SWOT analysis of the staff motivation system of the hotel "Dnipropetrovsk" was conducted (table 2).

Analysis of the staff motivation system of the hotel "Dnipropetrovsk" revealed the following shortcomings:

1. The staff motivation system is represented only by wages in the form of direct compensation and is practically limited. The organisation has developed rules regarding bonuses and established a possible salary increase of 5-10%.

Table 2

SWOT-analysis of the staff motivation system of the hotel "Dnipropetrovsk"

External environment	Opportunities	Threats
Internal environment	1. The opportunity to gain time during the recovery period after the economic crisis, reconstruction after Ukraine's victory in the war with Russia and use it to develop a motivation system that meets modern market conditions. 2. Maintaining a market position in a competitive environment.	1. Decrease in sales volumes compared to planned indicators.
Strengths	Field of strengths and opportunities 1. Management does not allow the deterioration of the socio-psychological climate in the organisation, a decrease in the effectiveness of teamwork, or a decrease in the effectiveness of individual employees. 2. Awareness of the problems faced by the company and the availability of the necessary financial resources to eliminate them in sufficient time will allow the development of projects that will improve the company's motivation system.	Field of strengths and threats 1. Using all the strengths of the hotel enterprise in order to overcome the decline in staff productivity and focus on achieving the goals facing the organisation.
Weaknesses	Field of weaknesses and opportunities 1. Development of projects aimed at improving the motivation system (including the development of an effective reward system and elements of non-material and non-monetary incentives).	Field of weaknesses and threats 1. The company needs to develop and implement a new staff motivation system that best meets modern conditions, thereby increasing employees' interest in their work and increasing labour productivity.

Source: compiled by the authors

Those who work in such a system have a financial interest in achieving high overall productivity for the entire team to improve the quality of their work results.

2. Failure to use the entire motivational potential of employees to improve existing processes and results of the enterprise.

The main opportunity of the hotel "Dnipropetrovsk" is to gain time during the recovery period after the economic crisis, reconstruction after Ukraine's victory in the war with Russia, and use it to develop a motivation system that meets modern market conditions.

Based on the SWOT analysis of the personnel motivation system, in order to improve the personnel motivation system of the hotel "Dnipropetrovsk", we can suggest the following measures:

- use of all your strengths to maximise the opportunities arising in the external environment;
- description of the possibility of trying to overcome the weaknesses of the organisation;
- use of your strengths to eliminate the threat.

The procedure for providing additional payments should be reflected in the rules of the internal labour regulations and the regulation "Rules for the remuneration of employees of the hotel "Dnipropetrovsk". Such assistance is justified in accordance with the Labour Code, since it does not worsen the employee's situation. However, such assistance should be established for all employees. In order to consolidate the company's personnel and avoid staff turnover, an allowance for work experience is introduced.

To increase incentives and benefits for company employees, we propose to introduce a correction coefficient depending on the category of employees and the amount of additional payments allocated by the company.

Conclusions from this study and prospects for further exploration in this direction. In summary, the main recommendations for achieving the goal of our study are as follows. In order to improve the staff motivation system at the hotel "Dnipropetrovsk", it is recommended to implement the following measures:

1. Implement an effective staff motivation system that increases labour productivity. When using schemes and bonus plans in the salary structure, there should be a clear connection between the contribution of a particular employee and the amount of compensation.

2. Improve the system of material and non-material incentives for employees, including the introduction of additional salary payments at the end of the year, issuing certificates to the best employees, etc.

Bibliography:

1. Бойко З., Горожанкіна Н., Грушка В. Модернізація систем управління підприємств сфери обслуговування. *Development Service Industry Management*. 2025. № 1. С. 29–34.
2. Гаврилова О. В., Осійчук В. В. Мотивація персоналу як один із методів управління у готельних підприємствах. *Modern Economics*. 2023. № 38(2023). С. 22–27.
3. Коваленко Л. Г. Мотивація персоналу підприємств готельно-ресторанного господарства. *Глобальні та національні проблеми економіки*. 2017. № 16. С. 323–326.
4. Непочатенко В. Мотивація персоналу в готельному господарстві. *Економіка та суспільство*. 2021. Вип.31. URL: <https://doi.org/10.32782/2524-0072/2021-31-42>
5. Черниш І., Козик М. Сучасні форми і методи мотивації персоналу підприємств сфери послуг. *Економіка і регіон*. Полтава: ПНТУ, 2021. Т. 1(80). С. 87–91.

References:

1. Boiko, Z., Horozhankina, N., & Hrushka, V. (2025). Modernizatsiia system upravlinnia pidpriemstv sfery obsluhovuvannia [Modernisation of management systems of service enterprises]. *Development Service Industry Management*. № 1, 29–34.
2. Havrylova, O., & Osiychuk, V. (2023). Motyvatsiia personalu yak odyz iz metodiv upravlinnia u hotelnykh pidpriemstvakh [Staff motivation as one of the management methods in hotels]. *Modern Economics*. № 38, 22–27.
3. Kovalenko, L. H. (2017). Motyvatsiia personalu pidpriemstv hotelno-restorannoho hospodarstva [Motivation of the personnel of the enterprises of hotel and restaurant economy]. *Hlobalni ta natsionalni problemy ekonomiky [Global and national economic problems]*. 16, 323–326.
4. Nepochatenko, V. (2021). Motyvatsiia personalu v hotelnomu hospodarstvi [Staff motivation in the hotel industry]. *Ekonomika ta suspilstvo [Economy and society]*. (31). Retrieved from: URL: <https://doi.org/10.32782/2524-0072/2021-31-42>
5. Chernysh, I., & Kozyk, M. (2021). Suchasni formy i metody motyvatsii personalu pidpriemstv sfery posluh [Modern forms and methods of motivating personnel of service enterprises]. *Ekonomika i rehion [Economy and region]*. Poltava: PNTU. Vol. 1(80), 87–91.

Дата надходження статті: 30.10.2025

Дата прийняття статті: 17.11.2025

Опубліковано: 30.12.2025

Kucher M. M., PhD in Economics, Associate Professor
at the Department of International Tourism and Hotel and Restaurant
Business University of Customs and Finance
ORCID: 0000-0002-3868-5311

Korneyev M. V., Doctor of Economics, Professor,
Dean of the Faculty of Innovative Technologies
University of Customs and Finance
ORCID: 0000-0002-4005-5335

Shcholakova H. V., PhD in Political Sciences, Associate Professor
at the Department of International Tourism and Hotel and Restaurant
Business
University of Customs and Finance
ORCID: 0000-0001-9396-9184

Zakharenko O. O., University of Customs and Finance
ORCID: 0009-0006-0284-2748

CURRENT ASPECTS OF PERSONNEL MOTIVATION IN THE HOTEL AND RESTAURANT SECTOR OF UKRAINE

The article explores the current aspects of employee motivation in Ukraine's hotel and restaurant sector under conditions of social, economic, and technological uncertainty caused by martial law, large-scale population migration, a shortage of qualified personnel, and rapid digitalisation. Employee motivation in the hospitality industry serves as a key factor in ensuring productivity, service quality, workforce stability, and the competitiveness of enterprises. The features of modern approaches to staff motivation in the hotel and restaurant sector of Ukraine were studied, effective methods of stimulation were identified, and innovative solutions adapted to today's challenges were substantiated.

Within the framework of the research, motivational methods have been classified into five categories: financial incentives, non-financial incentives, indirect methods, elimination of demotivating factors, and disciplinary measures. Their effectiveness has been analysed taking into account corporate culture, the age structure of personnel, and the specific features of labour processes. The study applies classical motivational theories (A. Maslow, V. Vroom, F. Herzberg, J. Adams, M. Seligman) as well as generational theory (X, Y, Z), which makes it possible to consider the social and value-based differences among employees.

Particular attention is paid to innovative approaches to motivation, including flexible work models, digital gamification, personalised development programs, mental health support, transparent bonus systems, and employee participation in decision-making. The SWOT assessment of motivation strategies within hospitality enterprises highlighted key tendencies, including the enhanced significance of non-financial motivators, the prioritisation of emotional stability, intensified attention to professional development and career progression, the institutionalisation of KPI-linked compensation mechanisms, the integration of new social groups, and the increasing utilisation of digital human resource tools.

The study substantiates that wages remain a fundamental motivator, especially under inflationary conditions; however, an effective motivational system should combine both financial and non-financial stimuli, be flexible, transparent, and socially sensitive. Effective management of employee motivation in the hotel and restaurant business represents a strategic resource that ensures resilience, operational efficiency, and sustainable development of hospitality enterprises.

Key words: employee motivation, HR analytics, management under martial law, non-financial incentives, hotel and restaurant sector of Ukraine.

Кучер М. М., Корнеев М. В., Щолокова Г. В., Захаренко О. О. Актуальні аспекти мотивації персоналу в готельно-ресторанній сфері України

Розглянуто питання мотивації персоналу в готельно-ресторанній сфері України в умовах соціальної, економічної та технологічної невизначеності, спричиненої воєнним станом, масштабною міграцією населення, нестачею кваліфікованих кадрів та діджиталізацією. Мотивація працівників у сфері гостинності є ключовим чинником забезпечення продуктивності, якості обслуговування, стабільності трудового колективу та конкурентоспроможності підприємств.

© М. М. Kucher, M. V. Korneyev, H. V. Shcholakova, O. O. Zakharenko, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

Досліджено особливості сучасних підходів до мотивації персоналу в готельно-ресторанній сфері України, визначено ефективні методи стимулювання та обґрунтування інноваційних рішень, адаптованих до викликів сьогодення.

У межах дослідження класифіковано мотиваційні методи за п'ятьма напрямками: матеріальні стимули, нематеріальні стимули, непрямі підходи, усунення демотивуючих чинників та дисциплінарні заходи. Проведено аналіз їх ефективності з урахуванням корпоративної культури, вікової структури персоналу та специфіки трудових процесів. Застосовано теоретичні моделі мотивації (А. Маслоу, В. Врум, Ф. Герцберг, Дж. Адамс, М. Селігман), а також теорію поколінь (X, Y, Z), що дозволяє врахувати соціальні та ціннісні відмінності між працівниками.

Особливу увагу приділено інноваційним підходам до мотивації: гнучким моделям роботи, цифровій гейміфікації, персоналізованим програмам розвитку, підтримці ментального здоров'я, прозорій системі бонусів та залученню працівників до прийняття рішень. На основі SWOT-аналізу мотиваційної політики підприємств готельно-ресторанного бізнесу виявлено ключові тренди: зростання ролі нематеріальних стимулів, потреба в емоційній стабільності, орієнтація на навчання та кар'єрне зростання, впровадження KPI-орієнтованих систем преміювання, інтеграція нових соціальних груп, активне використання цифрових HR-рішень.

У статті обґрунтовано, що заробітна плата залишається базовим мотиватором, особливо в умовах інфляції, проте ефективна мотиваційна система повинна поєднувати матеріальні та нематеріальні стимули, бути гнучкою, прозорою та соціально чутливою. Ефективне управління мотивацією персоналу в готельно-ресторанному бізнесі є стратегічним ресурсом, що забезпечує стійкість, ефективність та розвиток підприємств у сфері гостинності.

Ключові слова: мотивація персоналу, HR-аналітика, управління в умовах воєнного стану, нематеріальні стимули, готельно-ресторанна сфера України.

Problem statement. The hotel and restaurant sector represents one of the most dynamic and socially responsive segments of Ukraine's economy. Its overall productivity, quality of services, employee engagement, and consequently, the financial sustainability of enterprises, are largely determined by the effectiveness of personnel motivation. In the context of martial law and ongoing socio-economic turbulence, hospitality enterprises face deep structural transformations caused by massive internal migration, a shortage of skilled professionals, and accelerated digital integration. These factors redefine employees' motivational orientations, reshape employers' management priorities, and necessitate a revision of classical motivation models traditionally applied in the hospitality industry.

Under such extraordinary conditions, the human resource system must balance between financial stability, psychological well-being, and adaptive forms of professional engagement. Therefore, the rethinking of motivation in the hospitality sector should not only respond to present challenges but also form a resilient basis for workforce sustainability and long-term competitiveness.

Analysis of recent research and publications. The problem of personnel motivation continues to attract significant scholarly attention in management and hospitality studies. Researchers, including M. Hakova [1], I. Kantsur, A. Pastukh [2], L. Kovalenko [3], A. Kotlyk, D. Tereshchenko [4], V. Nepochatenko [5], L. Chepurda, I. Herman [6] and others, have examined various aspects of this topic, ranging from material incentive systems and monitoring of motivational needs to behavioural factors influencing employee engagement. Their findings emphasise the multidimensional character of motivation and its impact on enterprise efficiency.

However, despite the substantial academic groundwork, certain questions remain open. Contemporary studies rarely consider the motivational behaviour of hospitality personnel under martial law, nor do they adequately reflect the psychological and social effects of wartime migration on human capital. In addition, the increasing average age of the workforce and the integration of new demographic groups require differentiated incentive mechanisms and flexible management solutions. Consequently, the need for scientifically justified, context-specific motivation strategies for hotel and restaurant enterprises in Ukraine remains pressing and of practical relevance.

The goal of the article is to analyse and systematise the current directions of personnel motivation in hotel and restaurant enterprises in Ukraine and to identify effective approaches to employee stimulation.

Presentation of the main material. A fundamental managerial function in the hospitality industry is to establish a motivational environment that encourages creativity, initiative, and professional commitment among employees. Effective motivation connects organisational objectives with the personal aspirations of staff, ensuring higher productivity, service quality, and team cohesion. Understanding how motivation operates as both a psychological and managerial mechanism is crucial for constructing systems that transform external incentives into sustained internal engagement.

The need-driven nature of motivation defines it as a multidimensional behavioural process. Individual differences in needs, values, and priorities lead to significant variation in motivational responses; therefore, a single universal model cannot guarantee lasting effectiveness. Successful hospitality enterprises tend to integrate economic, professional, and socio-psychological components, combining individual and collective incentive strategies to maintain a balance between tangible and intangible rewards.

Based on this approach, motivation methods in the hotel and restaurant business can be systematised into several key categories. These categories encompass both direct and indirect influences on employee behaviour, ranging from monetary incentives to the creation of a favourable psychological climate. The classification summarised below (table 1) outlines the principal types of motivation, their mechanisms, and associated benefits and limitations within the hospitality context.

Table 1

Methods of motivating personnel

Motivation category	Specific methods	Advantages	Risks / limitations
Material	Salary, bonuses, profit share, extra payments	Direct stimulation, simple evaluation of results	Limited long-term effect, saturation effect
Non-material	Praise, recognition, respect, sense of belonging	Increases emotional engagement, no financial costs	Requires an individual approach, subjectivity
Indirect efforts	Training, development, career advancement	Long-term effect, improved qualification	Requires investment, delayed results
Elimination of negative factors	Removing injustice, conflicts, toxic atmosphere	Improved microclimate, reduced staff turnover	Requires changes in corporate culture
Disciplinary methods	Warnings, fines, dismissal	Control, discipline maintenance	Demotivation, fear, loss of initiative

Source: compiled by the authors

The empirical analysis carried out across hotel enterprises was based on the following motivational theories: V. Vroom's Expectancy Theory, B. Skinner's Reinforcement Theory, A. Maslow's Self-Actualisation Theory, F. Herzberg's Two-Factor Theory, M. Seligman's Well-Being Theory, J. Adams's Equity Theory, and the Participation Theory (table 2).

Table 2

Possibilities of applying motivational theories in the hotel and restaurant sector

Theory	Author	Application in the hotel and restaurant sector
Expectancy theory	V. Vroom	Identification of individual employee expectations regarding rewards
Reinforcement theory	B. Skinner	Use of bonuses, incentives, and gamification
Self-actualization theory	A. Maslow	Creating conditions for professional development
Two-factor theory	F. Herzberg	Differentiation between hygiene and motivational factors
Well-being theory	M. Seligman	Support for mental health and maintenance of a positive workplace climate
Equity theory	J. Adams	Transparent system of rewards and recognition
Participation theory	Modern HR approaches	Involvement of employees in decision-making processes

Source: compiled based on [7–8]

Employee motivation in the hotel and restaurant industry has undergone a profound conceptual evolution – from the rigid, authoritarian management paradigms of the nineteenth century to the contemporary philosophy of service-oriented leadership. During the latter half of the twentieth century, material remuneration mechanisms dominated as the central motivational driver. However, beginning in the 1990s, attention progressively shifted toward the cultivation of corporate culture, continuous professional development, and the enhancement of emotional intelligence as critical determinants of workforce engagement. In the post-COVID-19 era, motivation within the hospitality sector has further evolved toward flexibility, personalisation, and a holistic focus on employees' psychological well-being and life satisfaction, reflecting a transition to human-centred management models.

Classical motivational models based on A. Maslow's hierarchy of needs, V. Vroom's expectancy theory, and F. Herzberg's two-factor model establish a linear relationship between employee needs and productivity. However, under the conditions of high market turbulence, labour market instability, generational shifts, and changing social orientations, these theories, while remaining conceptually valid, have demonstrated a decline in practical effectiveness.

The generational theory (X, Y, Z) provides an analytical framework that enables the consideration of social, technological, and value-based differences among various groups of employees [9]. It is important to emphasise that in the hotel and restaurant business, where representatives of all three generations work simultaneously, the motivational system must be multi-level and adaptive (table 3). For instance, Generation X tends to prioritise stability and mentorship; Generation Y is motivated by opportunities for professional development and workplace flexibility; whereas Generation Z responds best to digital platforms, rapid feedback, and active social engagement.

The results of the conducted research demonstrate that an efficient motivation system within the hospitality industry must balance tangible and intangible incentives, taking into consideration the distinctive characteristics of hospitality labour – such as emotional demands, extensive client communication, seasonal fluctuations, and shift-based employment structures.

Table 3

Differences in the age structure of enterprise personnel

Generation	Characteristics	Effective motivators
X (1965–1980)	Stability, experience, loyalty	Years of service recognition, participation in decision-making, mentorship
Y (1981–1996)	Flexibility, development, balance	Career growth, flexible schedule, professional training
Z (1997–2012)	Technological orientation, speed, self-expression	Recognition, digital tools, teamwork

Source: compiled by the authors

Material incentives (such as salaries, bonuses, and additional payments) remain a fundamental component of motivation; however, their stimulating power tends to diminish over time due to the so-called “saturation effect.” Non-material incentives (such as recognition, belonging, and self-realisation) have a longer-lasting impact on employees’ emotional engagement and overall job satisfaction.

Employee development is the cornerstone of long-term motivation. Indirect motivation methods, including training programs, workshops, and the development of creative potential, are viewed as investments in human capital that enhance an enterprise’s competitiveness. Employees who are provided with opportunities for professional growth tend to demonstrate higher levels of loyalty and productivity. Encouraging learning fosters a culture of development, which constitutes a strategic advantage for hotels and restaurants.

The elimination of demotivating factors is an essential element of motivational policy. The presence of injustice, conflicts, favouritism, or toxic interpersonal relationships reduces the effectiveness of any motivational initiatives. Therefore, the creation of a healthy psychological climate is a prerequisite for the successful implementation of a motivation strategy.

Disciplinary methods should be applied in a balanced manner. Although disciplinary measures are necessary to maintain order, their excessive use can lead to loss of initiative and overall demotivation. It is recommended to combine control mechanisms with positive reinforcement techniques [3, 10, 11, 12].

The examination of contemporary motivational practices in the hotel and restaurant industry reveals a distinct shift toward a comprehensive, personalised, and analytically substantiated approach that integrates material, non-material, and developmental incentives. Central to this system is an individualised strategy that recognises each employee as a unique contributor, taking into account their personal values, needs, and professional aspirations.

Figure 1 illustrates a conceptual model of motivational methods tailored to the personnel structure of enterprises within the hospitality sector, reflecting the interrelation between managerial tools and the psychological characteristics of staff engagement.

The existence of a variety of approaches and numerous methods that implement them complicates the managers’ task of selecting the one that best corresponds to the goals of a particular enterprise and current conditions. Therefore, the choice of methods and the evaluation of motivational policies should be based on HR analytics. As key indicators for assessing the effectiveness of motivational measures, which can be used to adjust and improve motivational programs, both quantitative and qualitative parameters are commonly applied: staff turnover – changes in personnel numbers expressed as a percentage before and after the implementation of a motivational program; employee satisfaction index – derived from surveys or interviews; labor productivity – quantitative performance indicators; economic effect – profit growth, reduction in recruitment costs, and minimization of downtime.

Personnel motivation in the hotel and restaurant business is evolving from a simple incentive mechanism into a complex, dynamic, and personalised process that integrates psychological, managerial, and economic aspects. Effective motivation management is not limited to salary; it is about creating an environment where employees are willing to work, develop, and remain within the organisation.

An analysis of practical case studies presented on the websites of leading companies in the hotel and restaurant industry [10–11] has revealed innovative directions in motivation that contribute to improving business efficiency and resilience (table 4).

Among the effective innovative approaches to personnel motivation in the hotel and restaurant business under conditions of social, economic, and technological uncertainty are the following:

- flexible work models – the introduction of variable schedules, part-time employment, and opportunities for self-scheduling contributes to reducing burnout and staff turnover. Expected results: staff turnover decrease by 15–25%;
- digital gamification – the use of mobile applications to reward achievements, encourage participation in team competitions, and distribute bonuses. Expected results: efficiency increase by 10–20%;
- continuous learning and development – the implementation of online courses, mentorship programs, and internal academies promotes professional growth. Expected results: improvement in service quality by 15–25%;

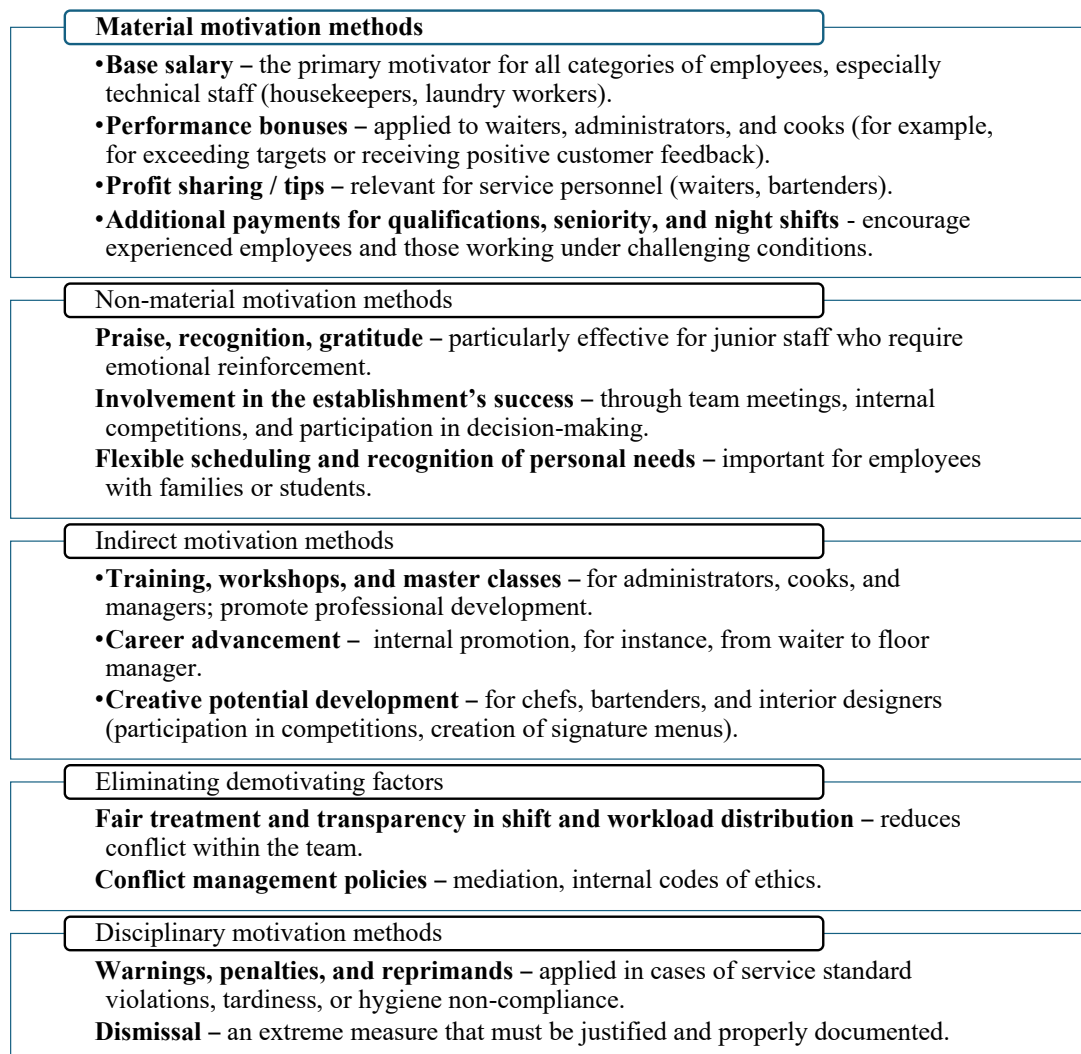


Fig. 1. Methods of employee motivation considering the specific structure of personnel in the hotel and restaurant sector
Source: compiled by the authors

Table 4

Motivational programs in the global hospitality industry

Company	Practice of using motivational programs
Marriott International	<i>Spirit to Serve</i> program: training, recognition, and employee involvement
Reikartz	Emotional support programs for staff
McDonald's Ukraine	Gamification through mobile applications
Premier Hotels	Flexible shift planning and self-managed scheduling

Source: compiled by the authors

- personalised motivation programs – individual development plans, regular meetings with HR managers, and consideration of employees' personal goals. Expected results: loyalty increase by 25–40%;
- mental health support – emotional well-being programs, access to psychologists, and anti-stress initiatives. Expected results: stress reduction by 30–50%;
- recognition and transparent bonus systems – public recognition of achievements, regular bonuses, and quality-based reward programs. Expected results: productivity increase by 15–30%;
- employee involvement in decision-making – surveys, brainstorming sessions, and participation in the development of new services. Expected results: increase in employee initiative by 50–70%.

Innovative trajectories in personnel motivation within the hotel and restaurant sector provide enterprises with greater institutional adaptability, enabling them to address contemporary socio-economic challenges while maintaining organisational resilience. The integration of flexible work arrangements, digital transformation tools, and

comprehensive well-being programs constitutes a decisive foundation for sustaining competitiveness in the hospitality domain.

Building upon prior scholarly contributions [2, 6, 12] and incorporating the outcomes of motivational monitoring alongside a SWOT assessment of incentive strategies implemented by hospitality enterprises in Dnipro (table 5), this study identifies the principal trends shaping personnel motivation in Ukraine's hotel and restaurant industry amid the conditions of martial law. These trends reveal an evolving paradigm of human capital engagement that emphasises both psychological sustainability and data-driven human resource management.

Table 5

SWOT-analysis of motivational policy in hotel and restaurant enterprises in Ukraine under martial law

Strengths	Weaknesses
– Flexibility of motivational policy – Increase in average wages – Growth of team cohesion	– Lack of psychological support – Shortage of qualified personnel – Imperfect performance evaluation system
Opportunities	Threats
– Involvement of internally displaced persons and veterans – Expansion of training programs – Implementation of HRIS (Human Resource Information Systems)	– Prolonged instability – Emotional burnout – Competition with foreign employers

Source: developed by the authors

The main trends in personnel motivation in Ukraine under martial law include:

- the growing role of non-material incentives such as recognition, involvement, and flexibility;
- the increasing need for emotional stability and psychological support;
- a focus on professional development, including training, mentorship, and career advancement;
- the transition toward KPI-based reward systems;
- the inclusion of new social groups;
- the implementation of digital HR solutions, including HRIS and motivation analytics.

In the Ukrainian hotel and restaurant sector, wages remain the fundamental motivator, especially in conditions of inflation and economic instability. Considering inflationary pressures, enterprises must ensure not only competitive pay levels but also dynamic wage indexation. Motivational programs should include an annual salary review based on the results of employee certification, competency assessments, and the achievement of key performance indicators (KPIs).

Under current conditions of instability, particularly during martial law, the incentive system must be adaptive, transparent, and socially sensitive.

Conclusions from this study and prospects for further research in this area. Under the current operating conditions of Ukraine's hotel and restaurant industry – marked by martial law, instability in the labour market, rapid digital transformation, and pronounced generational change – personnel motivation emerges as a decisive determinant of organisational efficiency, resilience, and competitiveness. The findings of this research affirm that the motivational architecture within the hospitality sector must be multidimensional, strategically integrated, and responsive to the individual needs and behavioural patterns of employees.

The structured systematisation of motivational approaches has made it possible to delineate five principal categories: material and non-material incentives, indirect stimulation methods, the removal of demotivating factors, and disciplinary interventions. The overall effectiveness of these mechanisms is contingent upon contextual variables such as corporate culture, demographic composition of the workforce, and the specific characteristics of labour processes within hospitality enterprises.

The analysis of motivational theories (by V. Vroom, A. Maslow, F. Herzberg, J. Adams, M. Seligman, and others) demonstrated the need to combine classical approaches with modern concepts of service leadership, well-being, and participatory management. Particular attention should be paid to Generation Z (Zoomers), which forms new expectations regarding motivation – digital platforms, quick feedback, and social engagement.

An effective motivational policy must be based on HR analytics, which allows evaluating the efficiency of motivational measures using both quantitative and qualitative indicators such as staff turnover, satisfaction index, labour productivity, and economic impact.

Thus, personnel motivation in the hotel and restaurant sector is evolving from a basic system of incentives into a strategic component of management, necessitating a flexible, individualised, and evidence-based approach. Future research should concentrate on developing adaptive motivation models that reflect contemporary socio-economic challenges and ensure the sustainable engagement of human capital.

Bibliography:

1. Гакова М. В. Мотиваційний моніторинг потреб персоналу готельно-ресторанних підприємств. *Сталый розвиток економіки*. 2024. № 1 (48). С. 344–350.
2. Канцур І.Г., Пастух А.О. Сучасні тренди матеріального стимулювання персоналу в готельно-ресторанному бізнесі. *Актуальні проблеми економіки*. 2024. № 11(281). С. 226–234.
3. Коваленко Л. Г. Мотивація персоналу підприємств готельно-ресторанного господарства. *Глобальні та національні проблеми економіки*. 2017. Вип. 16. С. 323–326.
4. Котлик А.В., Терещенко Д.А. Впровадження системи гейміфікації для мотивації персоналу відділу продажів. *Український журнал прикладної економіки та техніки*. 2025. Том 10. № 2. С. 185–188.
5. Непочатенко В. О. Мотивація персоналу в готельному господарстві. *Економіка та суспільство*. 2021. № 31. URL: <https://economyandsociety.in.ua/index.php/journal/article/view/709>
6. Чепурда Л.М., Герман І.В. Удосконалення методів управління матеріальним стимулюванням персоналу готельного бізнесу. *Економіка та суспільство*. 2025. Вип. 73. С. 112–117. URL: <https://economyandsociety.in.ua/index.php/journal/article/view/5857>
7. Літвін О.Г., Турло Н.П. Сучасні теорії мотивації праці та їх застосування в Україні. *Вчені записки ТНУ імені В. І. Вернадського. Серія: Економіка і управління*. 2025. Том 36 (75). № 1. С. 28–33.
8. Теорії мотивації та їх значення для управління персоналом. Команда HURMA. 2019. URL: <https://hurma.work/blog/teoriyi-motyvacziyi-ta-yih-znachennya-dlya-upravlinnya-personalom/>
9. Денисенко М.П., Давиденко Н.В. Теорія поколінь та який її вплив на сучасний бізнес. *Економіка та держава*. 2020. № 1. С. 46–49.
10. Marriott International. Spirit to Serve Program. URL: <https://news.marriott.com/news/2018/05/31/marriott-international-takes-spirit-to-serve-into-the-communities>
11. McDonald's. URL: <https://jobs.mchire.com/>
12. Левицька І.В., Климчук А.О. Управління персоналом готельно-ресторанної сфери в умовах воєнного стану. *Economics and Business Management*. 2022. Vol. 14, № 3. С. 40–51.

References:

1. Hakova, M. V. (2024). Motyvatsiinyi monitorynh potreb personalu hotelno-restorannykh pidpriemstv [Motivational monitoring of the needs of personnel in hotel and restaurant enterprises]. *Stalyi rozvytok ekonomiky [Sustainable economic development]*, 1(48), 344–350.
2. Kantsur, I. H., & Pastukh, A. O. (2024). Suchasni trendy materialnoho stymuliuvannia personalu v hotelno-restorannomu biznesi [Modern trends in material incentives for personnel in the hotel and restaurant business]. *Actual Problems of Economics*, 11(281), 226–234.
3. Kovalenko, L. H. (2017). Motyvatsiia personalu pidpriemstv hotelno-restorannoho hospodarstva [Motivation of personnel in hotel and restaurant enterprises]. *Hlobalni ta natsionalni problemy ekonomiky [Global and National Problems of the Economy]*, (16), 323–326.
4. Kotlyk, A. V., & Tereshchenko, D. A. (2025). Vprovadzhennia systemy heimifikatsii dlia motyvatsii personalu viddilu prodazhiv [Implementation of gamification systems for motivating sales department personnel]. *Ukrainskyi zhurnal prykladnoi ekonomiky ta tekhniky [Ukrainian Journal of Applied Economics and Technology]*, 10(2), 185–188.
5. Nepochatenko, V. O. (2021). Motyvatsiia personalu v hotelnomu hospodarstvi [Personnel motivation in the hotel industry]. *Ekonomika ta suspilstvo [Economy and Society]*, (31). Retrieved from: <https://economyandsociety.in.ua/index.php/journal/article/view/709>
6. Chepurda, L. M., & Herman, I. V. (2025). Udoskonalennia metodiv upravlinnia materialnym stymuliuvanniam personalu hotelnoho biznesu [Improvement of material incentive management methods in the hotel business]. *Ekonomika ta suspilstvo [Economy and Society]*, (73), 112–117. Retrieved from: <https://economyandsociety.in.ua/index.php/journal/article/view/5857>
7. Litvin, O. H., & Turlo, N. P. (2025). Suchasni teorii motyvatsii pratsi ta yikh zastosuvannia v Ukraini [Modern theories of work motivation and their application in Ukraine]. *Vcheni zapysky TNU imeni V. I. Vernadskoho. Serii: Ekonomika i upravlinnia [Scientific Notes of V. I. Vernadsky Taurida National University. Series: Economics and Management]*, 36 (75) (1), 28–33.
8. Hurma Team. (2019). Teorii motyvatsii ta yikh znachennia dlia upravlinnia personalom [Motivation theories and their importance for personnel management]. Retrieved from: <https://hurma.work/blog/teoriyi-motyvacziyi-ta-yih-znachennya-dlya-upravlinnya-personalom/>
9. Denysenko, M. P., & Davydenko, N. V. (2020). Teoriia pokolin ta yii vplyv na suchasnyi biznes [The theory of generations and its impact on modern business]. *Ekonomika ta derzhava [Economics and the State]*, (1), 46–49.
10. Marriott International. (2018). Spirit to Serve Program. Retrieved from: <https://news.marriott.com/news/2018/05/31/marriott-international-takes-spirit-to-serve-into-the-communities>

-
11. McDonald's (2025). Retrieved from: <https://jobs.mchire.com/>
12. Levytska, I. V., & Klymchuk, A. O. (2022). Upravlinnia personalom hotelno-restorannoi sfery v umovakh voiennoho stanu [Personnel management in the hotel and restaurant sector under martial law]. *Economics and Business Management*, 14(3), 40–51.

Дата надходження статті: 30.10.2025

Дата прийняття статті: 17.11.2025

Опубліковано: 30.12.2025

UDC 334.01

DOI <https://doi.org/10.32782/2521-6643-2025-2-70.44>

Stebliuk N. F., PhD in Economics, Associate Professor
at the Department of International Tourism and Hotel and Restaurant
Business University of Customs and Finance
ORCID: 0000-0002-4488-769X

Nebaba N. O., Doctor of Economics, Professor at the Department
of Economic Modeling, Accounting and Statistics
Oles Honchar Dnipro National University
ORCID: 0000-0003-1264-106X

Razinkova M. Yu., Doctor of Economics, Professor at the Department
of International Tourism and Hotel and Restaurant Business
University of Customs and Finance
ORCID: 0000-0002-1590-7428

Hryshchenko S. V., University of Customs and Finance
ORCID: 0009-0005-0209-8819

IMPLEMENTING CUSTOMER-CENTRIC APPROACHES IN THE RESTAURANT INDUSTRY: ANALYTICS, MOTIVATION, EFFICIENCY

The article explores current trends in the development of the restaurant business, which are associated with the need to respond promptly to changes in market conditions and increase the level of consumer expectations, and actualize the implementation of customer-oriented approaches in the service system of business models of establishments. The impact of a customer-oriented approach, staff motivation and key performance indicators (KPI) on the quality of service and the effectiveness of the activities of restaurant business establishments is determined. Key performance indicators are determined, which allow not only to determine the level of customer satisfaction, but also to identify weaknesses in the service process, and therefore, to improve business processes based on the obtained analytics. Specialized training programs are proposed, which are aimed at increasing the efficiency of personnel management in the restaurant business and improving the quality of customer service. It has been established that the focus on service quality, personalized interaction with consumers and the creation of a comfortable communication environment contributes to an increase in customer satisfaction, reducing their churn and stimulating repeated consumption. It has been determined that systematic monitoring of key performance indicators allows for timely identification of problem areas, optimization of business processes and effective use of marketing resources. The feasibility of implementing specialized motivational programs that positively affect staff productivity, reduce staff turnover and create a favorable internal climate has been substantiated.

The practical significance of the results obtained lies in the possibility of their direct application in the activities of restaurant business establishments to improve the customer service system based on a customer-oriented approach. The results obtained can also be used to increase the economic efficiency and sustainability of restaurant business establishments in conditions of high competition.

Prospects for further research are the development of digital KPI monitoring tools, the study of the impact of staff emotional intelligence on customer experience, as well as a comparative analysis of the effectiveness of customer-oriented strategies in different formats of restaurant business establishments.

Key words: customer orientation, entrepreneurship, restaurant, service quality, key performance indicators, consumer market, consumer experience, motivational programs, organizational and economic support, business analytics.

Стеблюк Н. Ф., Небаба Н. О., Разінькова М. Ю., Грищенко С. В. Впровадження клієнтоорієнтованих підходів у ресторанному бізнесі: аналітика, мотивація, ефективність

У статті проведено дослідження сучасних тенденцій розвитку ресторанного бізнесу, які пов'язані з необхідністю оперативного реагування на зміни ринкових умов та підвищення рівня споживчих очікувань, актуалізують впровадження клієнтоорієнтованих підходів в системі обслуговування бізнес-моделей закладів. Визначено вплив клієнтоорієнтованого підходу, мотивації персоналу та ключових показників ефективності (КПІ) на якість обслуговування та результативність діяльності закладів ресторанного бізнесу. Визначено ключові показники ефективності, які дозво-

© N. F. Stebliuk, N. O. Nebaba, M. Yu. Razinkova, S. V. Hryshchenko, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

ляють не лише визначити рівень задоволення клієнтів, але й виявити слабкі місця у процесі обслуговування, а отже, вдосконалити бізнес-процеси на основі отриманої аналітики. Запропоновано спеціалізовані навчальні програми, які спрямовані на підвищення ефективності управління персоналом у сфері ресторанного бізнесу та покращення якості обслуговування клієнтів. Встановлено, що фокус на якості обслуговування, персоналізованій взаємодії зі споживачами та створенні комфортного комунікаційного середовища сприяє зростанню рівня задоволеності клієнтів, зменшенню їх відтоку та стимулюванню повторного споживання. Визначено, що систематичний моніторинг ключових показників ефективності дозволяє своєчасно виявляти проблемні зони, оптимізувати бізнес-процеси та ефективно використовувати маркетингові ресурси. Обґрунтовано доцільність впровадження спеціалізованих мотиваційних програм, які позитивно впливають на продуктивність персоналу, знижують плинність кадрів та формують сприятливий внутрішній клімат.

Практичне значення одержаних результатів полягає у можливості їх безпосереднього застосування в діяльності закладів ресторанного бізнесу для вдосконалення системи обслуговування споживачів на основі клієнтоорієнтованого підходу. Отримані результати також можуть бути використані для підвищення економічної результативності та стійкості закладів ресторанного бізнесу в умовах високої конкуренції.

Перспективи подальших досліджень полягають у розробленні цифрових інструментів моніторингу KPI, вивченні впливу емоційного інтелекту персоналу на клієнтський досвід, а також у порівняльному аналізі ефективності клієнтоорієнтованих стратегій у різних форматах закладів ресторанного бізнесу.

Ключові слова: клієнтоорієнтованість, підприємницька діяльність, ресторан, якість обслуговування, ключові показники ефективності, споживчий ринок, споживчий досвід, мотиваційні програми, організаційно-економічне забезпечення, бізнес-аналітика.

Problem statement. In the context of globalization and intensifying competition within the restaurant industry, the quality of service acts as one of the key determinants of an enterprise's sustainable competitive position. In a saturated market environment, consumers demonstrate increasingly high demands for the level of service, giving preference to establishments that combine high product quality with a proper service culture. This, in turn, shapes a positive consumer experience and increases the likelihood of repeat visits [1, 2].

The relevance of this research topic is determined by the growing competition in the restaurant business, where service quality plays a crucial role in fostering customer loyalty, maintaining a positive image of the establishment, and ensuring its long-term development. Under conditions of continuous market transformation, rising consumer expectations, and increasing service requirements, the customer-oriented approach serves as a key element of the modern business management system.

Analysis of recent research and publications. Modern academic research increasingly emphasizes that the customer-oriented approach is not merely a management philosophy but a prerequisite for achieving sustainable competitive advantages in business. N. Savitska, H. Chmil, and N. Dzhgutashvili [3] propose a systemic approach to customer-oriented management, focusing on the diagnosis and evaluation of customer experience. Scholars [3, 4] highlight the lack of systematization in the implementation of customer-oriented approaches, particularly due to the absence of comprehensive digital solutions and the weak integration of service and marketing processes. Yu. Tvardovskyi and O. Popko [4] argue that in a saturated market, customer orientation becomes a strategic asset. They emphasize the problem of the absence of a systematic approach in Ukrainian enterprises, where customer orientation often remains an initiative of individual employees rather than a formalized business process.

The study [5] outlines a system of indicators for assessing the customer orientation of hotel enterprises within each factor of customer-oriented activity. In the scientific discourse, the essence of customer orientation is revealed as a systemic approach encompassing the development of long-term relationships with consumers, ensuring high levels of satisfaction, personalizing services, and continuously improving the quality of service [5, 6, 7]. Such an approach enables the creation of a unique customer experience, which becomes a source of competitive advantage for restaurant enterprises.

Despite significant scientific progress, contemporary studies have not sufficiently examined the internal factors of staff motivation and their relationship with the level of customer orientation. Moreover, there remains a limitation in empirical data, particularly concerning the effectiveness of digital tools' implementation. Therefore, further research should focus on developing comprehensive adaptive models that integrate staff motivation, customer experience diagnostics, and the implementation of modern technologies aimed at constructing an efficient customer-oriented service system.

The purpose of this study is to determine the impact of the customer-oriented approach, staff motivation, and key performance indicators (KPIs) on service quality and the performance of restaurant enterprises.

Presentation of the main material. According to data from Kantar DX (Digital Analytics) based on the analysis of global trends, the Top-4 key consumer needs for the coming years have been identified [8]:

- 1) protection – consumers want to feel confident that they are both physically and informationally secure;
- 2) well-being – consumers aim to control their stress levels, maintain a balanced diet, and look fit and attractive;
- 3) flow – consumers seek to save time and have access to products and services whenever and wherever they desire;

4) experience – consumers continuously look for new impressions and sensations that can enrich their daily lives.

For restaurant enterprises, these needs define the strategic guidelines for shaping service quality and business models. Ensuring protection implies compliance with sanitary and hygienic standards, safe food storage, transparency in meal preparation, and protection of customers' personal data. The focus on well-being requires including healthy and balanced dishes in the menu, creating a comfortable atmosphere, and providing attentive, caring service. The need for flow is addressed through the implementation of digital solutions such as online ordering, mobile applications, fast service, and flexible working hours. Finally, the desire for experience stimulates the development of unique gastronomic concepts, creative food presentation, and thematic events.

Ensuring service quality has acquired the status of a strategic management priority, as reflected in international standards where customer orientation is viewed as a key component of quality management systems. Consequently, quality in its modern interpretation is considered a fundamental prerequisite for achieving competitiveness, strengthening business reputation, and ensuring the long-term success of an enterprise [8]. The corporate culture of an enterprise also plays a crucial role in ensuring service quality. Restaurants that invest in the professional development of their employees, create comfortable working conditions, and implement motivational programs typically demonstrate higher levels of service. In addition, modern technologies such as restaurant automation systems (CRM, POS systems) contribute to increasing service efficiency and reducing the number of operational errors. Table 1 presents the main structural components of service quality in the restaurant industry.

Table 1

Structural components of service quality in the restaurant industry

Service quality component	Description	Evaluation methods
Staff professionalism	Knowledge of service standards, communication skills, ability to resolve conflicts, ethical behavior, appearance, and emotional stability.	Certification, mystery shopper evaluations, customer surveys, observation, managerial assessment.
Physical environment	Cleanliness of premises, interior aesthetics, comfort of furniture, noise level, music accompaniment, temperature, lighting, and accessibility for people with limited mobility.	Visual assessment, checklists, photo documentation, customer feedback, sanitary audits.
Service quality	Freshness and quality of ingredients, cooking technology, food presentation aesthetics, menu compliance, taste consistency, and kitchen hygiene.	Tasting sessions, laboratory testing, customer questionnaires, control purchases, internal quality standards.
Service efficiency	Speed of service, order accuracy, response to complaints, clarity of communication, and coordination among staff.	Time tracking, receipt analysis, CRM reports, customer feedback, CES (Customer Effort Score) evaluation.
Emotional interaction	Friendliness, attentiveness, empathy, ability to create a positive atmosphere, capacity to uplift the customer's mood, and non-verbal communication.	Customer reviews, observation, interviews, NPS evaluation, emotional analysis of comments.

Source: [2;8]

Such characteristics as assortment, taste properties of dishes, price level, visual design of the interior, reputation, or location play a significant role in shaping the perception of a restaurant; however, they cannot compensate for a lack of attentiveness or professionalism on the part of the service staff. At the same time, an enterprise that ensures a high level of service obtains a sustainable competitive advantage, which contributes to the formation of a customer-oriented approach. The latter implies that the business focuses on identifying, understanding, and satisfying the individual needs of its clients, as well as building long-term relationships based on trust, quality, and mutual benefit [9, 10].

A customer-oriented approach includes:

- a deep understanding of the target audience's needs;
- personalized interaction with customers;
- flexibility in service delivery;
- proactive response to guests' feedback and suggestions;
- consistency in maintaining high-quality service.

Modern studies demonstrate that customer orientation directly influences the key performance indicators of hospitality enterprises, transforming ordinary visits into meaningful experiences that foster long-term relationships with guests. The correlation between customer orientation and the economic efficiency of hospitality establishments is presented in table 2.

Effective measurement of customer orientation in the restaurant business requires a comprehensive approach based on an integrated system of metrics. The Customer Satisfaction Score (CSAT) allows assessing the level of satisfaction immediately after a specific interaction; the Net Promoter Score (NPS) measures overall loyalty and the client's willingness to recommend the establishment; while the Customer Effort Score (CES) evaluates the degree of effort a client must exert to receive a service. The synergistic application of these indicators provides a multidimensional understanding of the customer experience, allowing managers to identify critical service points and respond promptly to changes in brand perception [11]. An important element of a customer-oriented strategy is the implementation of multichannel communication and the encouragement of customers to provide feedback.

Table 2

Key aspects of the impact of customer orientation on enterprise development

Impact criterion	Main effects	Performance indicators
Customer loyalty growth	Formation of a stable customer base, increased visit frequency, reduced customer churn, and strengthened brand trust.	NPS (Net Promoter Score) – customer willingness to recommend; Customer Retention Rate – customer retention level; CLV (Customer Lifetime Value) – long-term customer value.
Competitive advantage	Creation of unique market positioning, enhanced reputation, and increased brand recognition.	Market share growth by 15–25%; Brand Awareness Index; USP (Unique Selling Proposition).
Economic efficiency	Growth of average check, increased profit margins, cost optimization, improved profitability, higher ROMI, lower CAC, and growth in AOV and CR.	Increase in average check by 15–20%; ROI (Return on Investment) – 25–35%; ROMI (Return on Marketing Investment); CAC (Customer Acquisition Cost); AOV (Average Order Value); CR (Conversion Rate); profit growth.
Customer experience improvement	Positive emotions, personalized service, reduced customer effort, emotional connection with the brand, quick response to inquiries, comfortable interaction, consideration of individual needs, and increased satisfaction.	Higher CSAT (Customer Satisfaction Score), CES (Customer Effort Score), and NPS (Net Promoter Score); reduced FRT (First Response Time).
Corporate culture development	Employee motivation, reduced staff turnover, enhanced professionalism, teamwork building, and improved internal climate.	Higher employee satisfaction and lower staff turnover rate.

Source: [11;12]

Negative feedback can also serve as a valuable source of information for improving service quality and adapting business processes to current consumer demands. The analysis of digital marketing reports is a key managerial task, as it directly affects the effectiveness of decision-making processes. The structure of such reports varies depending on the complexity of marketing funnels, which necessitates their systematic digitalization and the identification of new “growth points” (parameters) whose optimization can significantly enhance economic outcomes. Examples of such indicators include the CTR (Click-Through Rate) – a clickability coefficient that can be increased through the creation of optimized and creative advertising materials, and the CR (Conversion Rate) – a conversion coefficient that can be improved through A/B testing and continuous optimization of the sales funnel. Key performance indicators (KPIs) make it possible not only to determine customer satisfaction levels but also to identify weaknesses in the service process, thereby improving business operations based on the obtained analytics.

Every improvement in customer service – for example, reducing service time or increasing the quality of dishes – raises customer satisfaction (CSAT) and loyalty (NPS). Collecting and analyzing NPS survey results across specific segments is a relatively simple process that enables the company to clearly determine which customer groups report positive interaction experiences and which segments require improved service and greater attention. Reducing the First Response Time (FRT) or the time required to resolve an issue contributes to a higher Customer Effort Score (CES), reflecting improved ease of interaction. Enhancing service or food quality can lead to an increase in the Average Order Value (AOV), Customer Lifetime Value (CLV), and Customer Retention Rate, thereby ensuring greater profitability and brand stability.

The main growth points that companies should pay attention to remain CTR (clickability) and CR (conversion). Improving CTR requires investment in dynamic creative development – the creation and testing of new, more effective advertising materials tailored to different platforms (table 3).

Table 3

Digital marketing performance results of the restaurant

Indicator	Facebook	Instagram	Google Search	Youtube	Total	Value for CEO
Budget (UAH)	1000	1500	800	900	4200	Important for assessing resource allocation and investment potential.
Reach	243902	306122	210526	236842	997393	Number of unique users who viewed the advertisement. Indicates the potential for audience expansion.
CPM	4,1	4,9	3,8	3,8	-	Cost per 1,000 impressions. A lower CPM implies more efficient budget utilization for audience reach.
CTR (%)	1,20%	1,40%	1,80%	0,90%	7,10%	Click-through rate. One of the key “growth levers.” A high CTR reflects the attractiveness of creatives.
Traffic	2927	4286	3789	2132	13134	Total number of website visits – the foundation for subsequent conversions.
CPC (UAH)	0,34	0,35	0,21	0,42	0,32	Cost per click. A lower CPC allows attracting more traffic within the same budget.
CR Leads (%)	6%	7%	7%	9%	6%	Conversion rate to leads. Demonstrates how effectively traffic is converted into potential customers.
CPL (UAH)	5,69	5	3,02	4,69	4,5	Desirable to have the lowest possible cost of acquiring a single potential client.
CR Orders (%)	25%	23%	30%	15%	24%	Conversion rate for orders. The second key “growth lever.” A high CR indicates an efficient sales funnel.
Sales (units)	44	69	80	29	222	Total number of successful transactions. A direct indicator of performance.
AOV (UAH)	150	148	152	147	150	Calculation of total revenue and profitability.
CAC (UAH)	22,73	21,74	10	31,03	18,92	Customer acquisition cost. A critical metric for evaluating spending efficiency.
AOV / CAC Ratio	6,6	6,81	15,2	4,74	7,91	Ratio of average order value to customer acquisition cost. Demonstrates profitability of each acquired customer – the higher, the more efficient the channel. Google Search Ads shows leadership in this indicator.
Sales volume (UAH)	6600	10212	12160	4263	33235	Total revenue from sales.
ROMI (%)	560%	581%	1420%	374%	691%	Return on marketing investment. A key indicator for the CEO, showing how much revenue the company earns per unit of currency invested. Google Search Ads demonstrates the highest profitability, indicating the need to reallocate the budget toward the most efficient channels (Google Search Ads).

Source: [11;12]

It should be noted that there is a significant gap between the high value consumers place on quality service (socio-psychological factors) and the low level of operational implementation, which is manifested in the indifference and lack of precision in staff performance. Employee qualifications and professional recruitment are crucial factors, as professionalism, service culture, responsiveness, and understanding of guests' needs directly affect the level of service. However, the identified indifference among employees is a direct indicator of the inefficiency of both behavioral and operational customer orientation.

Therefore, the competence of team members is a critically important factor that determines the success of any business. An effective personnel motivation system in the restaurant industry should be comprehensive, consistent, and take into account the specific features of different staff categories, such as cooks, waiters, administrators, and others. Each of these professional groups has its own unique tasks, working conditions, and motivational needs, which requires management to design customized incentive programs that reflect their responsibilities and roles within the organization [13, 14].

An essential aspect is the alignment of motivational initiatives with the corporate culture of the enterprise, which ensures their organic integration and acceptance among employees. The development of motivational programs should be based on a systematic analysis of the actual needs and expectations of the staff, as well as consideration of external factors influencing the work environment, such as working conditions, wage levels, opportunities for professional growth, and social protection (table 4). This document serves as a sample of motivational programs structured by day, week, month, and quarter, which can be adapted by an enterprise to its business standards.

Table 4

**Types of motivational programs in the restaurant business
(*the specific reward is determined by the establishment)**

1	Daily		
	Evaluation criteria	Reward	Comments
	Highest average check	Priority shift choice	Based on revenue
	Highest total sales	Movie ticket	At least 7 menu items
	Check fullness	Bonus hour	
	Dish from the “go-list”	Complimentary “go-list” dish	Develop dish cards
	Dish lottery	0 UAH*	
	“Wish game”	The lowest performer fulfills the wish of the best one	
	Sales tracking every 3 hours (12:00, 15:00, 18:00, 21:00, 24:00). Reports by waiter and total sales, point distribution.	Maximum score – 3000	Smile, appearance, sales technique, tray handling – evaluation criteria
2	Evaluation checklist (KLO)	0 UAH deposit*	Based on revenue
	Weekly		
	Evaluation criteria	Reward	Comments
	Highest check	Deposit of 0 UAH*	
	Highest total sales	Beauty salon certificate	
	Check fullness	Exemption from general cleaning	
	“Who sold the most?”	Bonus hour	Desserts, steaks, wine, etc.
	Mystery guest	Motivation retention	
	Best waiter	Badge + hookah in chain venue	Create a “Best” badge
3	Hostess	Make-up certificate	Internal Evaluation Sheet (KLO)
	Bartender	0 UAH* meal deposit	Evaluation sheet
	Hookah master	0 UAH* meal deposit	Evaluation sheet
	Monthly		
	Evaluation criteria	Reward	Comments
	Highest total sales	0 UAH* bonus to salary	
	Highest average check	Certificate for 0 UAH* in network venue	
	Check fullness	SPA certificate	
	Most charming smile	Professional photoshoot	
4	Best manager	0 UAH* bonus to salary	Which shift achieved higher turnover
	Hostess	Tanning subscription	
	Bartender	Personalized bottle of alcohol	
	Hookah master, Hookah record of the month	Gym membership, Bowling or cinema tickets	
	Waiter’s new record (without banquet)	0 UAH* deposit in network venue	
	Quarterly		
	Evaluation criteria	Reward	Comments
	Point system	Trip	
	Situational board		Poster paper and markers
4	Barrier table		Implement sticker board
	Suggestion box		Box with notes (open once a month)
	Kindness box	Clothing store certificate	Reward for the most mentioned employee (“kindness joke”)
	Good deeds		Labor therapy (prepare list – regulations)
	“Bring a friend”	0 UAH*	If a friend passes certification and works for two months
	Mentorship	0 UAH* per participant	Internship
	“Smoking token”		Create a token (manager’s responsibility)

Conclusions. The implementation of the customer-oriented concept in the restaurant business serves as a strategic tool for enhancing the competitiveness of establishments. The study revealed that focusing on service quality, personalized interaction with consumers, and the creation of a comfortable communication environment contributes to increased customer satisfaction, reduced churn, and stimulation of repeat consumption. It has been determined that systematic monitoring of key performance indicators (KPIs) allows for the timely identification of problem areas, optimization of business processes, and efficient use of marketing resources. The feasibility of implementing specialized motivational programs has been substantiated, as they positively affect employee productivity, reduce staff turnover, and foster a favorable internal climate. The obtained results have practical significance and can be applied to improve the service system in restaurant establishments, increase their economic performance, and ensure sustainability under high competition. Future research prospects involve a deeper analysis of the impact of digital technologies on customer experience and the development of effective personnel motivation models in the restaurant business. An important direction also lies in assessing the correlation between the level of customer orientation and the financial performance of enterprises.

Bibliography:

1. Однолько В.О. Теоретичні основи сервісу на підприємстві готельно-ресторанної справи. *Розвиток методів управління та господарювання на транспорті*. 2022. 4(81). С. 112–123.
2. Жегус О. В., Михайлова М. В., Афанасьєва О.П. Передумови формування моделі розумного споживання продуктів харчування в Україні. *Бізнес Інформ*. 2024. № 6. С. 389–398.
3. Савицька Н. Л., Чміль Г. Л., Джугташвілі Н. М. Клієнт-орієнтоване управління в готельно-ресторанному бізнесі : монографія. Харків: вид. Іванченко І. С. 2021. 209 с.
4. Твардовський Ю. В., Попко О. В. Проблема та перспективи формування конкурентних переваг на основі застосування клієнтоорієнтованого підходу. *Економіка та суспільство*. 2024. Вип. 70. URL: <https://economyandsociety.in.ua/index.php/journal/article/view/5354/5298>
5. Петлін І. В., Вихлюк М. І. Клієнтоорієнтований підхід до формування маркетингової стратегії підприємств готельного бізнесу. *Підприємництво та інновації*. 2024. Вип.32. С. 14–20.
6. Гаркавенко В. О., Стець О. В. Економіко-математична модель управління клієнтською базою підприємства. *Стратегія економічного розвитку України*. 2022. Вип.50. С. 177–196.
7. Мельник І. М., Антошкова Н. А. Потенціал підприємства ресторанного господарства в забезпеченні клієнтоорієнтованого сервісу. *Інтелект XXI*. 2019. № 6(1). С. 133–138.
8. Rau. Основні маркетингові тренди 2025 року – дослідження Kantar. URL: <https://rau.ua/dosvid/osnovni-marketingovi-trendi-2025/>
9. Illiashenko S., Shypulina Y., Illiashenko N., Golysheva I. Prospects for the Development of the Digital Labour Market in Ukraine: The National and Regional Aspects. *Studia Regionalne i Lokalne*. 2023. № 2(92). P.50–66.
10. Stebliuk N., Nebaba N., Ozierova S. Competitiveness of restaurant business enterprises: evaluation and ways to improve. *Науковий погляд: економіка та управління*. 2024. № 2 (86). С. 83–89.
11. Бондарчук К. П., Кифяк О. В., Горішевський П. А. Аналіз взаємодії маркетингу й обслуговування клієнтів у готельно-ресторанному бізнесі України. *Академічні візії*. 2024. Вип. 30. URL: <https://archer.chnu.edu.ua/xmlui/handle/123456789/10205>
12. Легенчук С. Ф., Завалій Т. О. Big Data в маркетинговій аналітиці: можливості та проблеми використання. *Проблеми теорії та методології бухгалтерського обліку, контролю і аналізу*. 2023. Вип. 1 (54). С. 52–58.
13. Чаркіна Т. Ю., Марценюк Л. В., Задоя В. О., Пікуліна О. В. Стратегічні напрями управління готельно-ресторанним бізнесом в умовах кризи. *Економіка та держава*. 2021. № 2. С. 19–23.
14. Кирчата І. М., Приходько Д.О. Клієнтоорієнтованість як стратегічний орієнтир формування моделі купівельної поведінки споживача та розвитку бізнесу. *Проблеми і перспективи розвитку підприємництва*. 2023. (30). С. 182–193.

References:

1. Odnolko, V. O. (2022). Teoretychni osnovy servisu na pidpriemstvi hotelno-restorannoi spravy [Theoretical foundations of service at hotel and restaurant enterprises]. *Development of Management and Economy Methods in Transport*, 4(81), 112–123.
2. Zhehus, O. V., Mykhailova, M. V., & Afanasieva, O. (2024). Peredumovy formuvannia modeli rozumnoho spozhyvannia produktiv kharchuvannia v Ukraini [Preconditions for the formation of a smart food consumption model in Ukraine]. *Business Inform*, (6), 389–398.
3. Savitska, N. L., Chmil, H. L., & Dzhgutashvili, N. M. (2021). Kliient-oriientovane upravlinnia v hotelno-restorannomu biznesi : monohrafiia [Client-oriented management in the hotel and restaurant business: Monograph]. Kharkiv: Ivanchenko I. S. Publishing House.

-
4. Tvardovskyi, Yu. V., & Popko, O. V. (2024). Problema ta perspektyvy formuvannia konkurentnykh perevah na osnovi zastosuvannia kliientoorientovanoho pidkhodu [Problems and prospects of forming competitive advantages based on the application of a client-oriented approach]. *Economics and Society*, (70). Retrieved from: <https://economyandsociety.in.ua/index.php/journal/article/view/5354/5298>
 5. Petlin, I. V., & Vykliuk, M. I. (2024). Kliientoorientovanyi pidkhid do formuvannia marketynhovoï stratehii pidpriemstv hotelnogo biznesu [Client-oriented approach to forming the marketing strategy of hotel enterprises]. *Entrepreneurship and Innovation*, (32), 14–20.
 6. Harkavenko, V. O., & Stets, O. V. (2022). Ekonomiko-matematychna model upravlinnia kliientskoïu bazoiu pidpriemstva [Economic-mathematical model of customer base management in enterprises]. *Strategy of Economic Development of Ukraine*, (50), 177–196.
 7. Melnyk, I. M., & Antoshkova, N. A. (2019). Potentsial pidpriemstva restorannoho hospodarstva v zabezpechenni kliientoorientovanoho servisu [The potential of a restaurant enterprise in providing customer-oriented service]. *Intellect XXI*, 6(1), 133–138.
 8. Rau. Kantar. (2025). [Key marketing trends of 2025 – research]. Retrieved from: <https://rau.ua/dosvid/osnovni-marketingovi-trendi-2025/>
 9. Illiashenko, S., Shypulina, Y., Illiashenko, N., & Golysheva, I. (2023). Prospects for the development of the digital labour market in Ukraine: The national and regional aspects. *Studia Regionalne i Lokalne*, 2(92), 50–66.
 10. Stebliuk, N. F., Nebaba, N. O., & Ozerova, S. O. (2024). Competitiveness of restaurant business enterprises: Evaluation and ways to improve. *Scientific View: Economics and Management*, 2(86), 83–89.
 11. Bondarchuk, K. P., Kifyak, O. V., & Gorishevsky, P. A. (2024). Analiz vzaiemodii marketynhu y obsluhovuvannia kliientiv u hotelno-restorannomu biznesi Ukrainy [Analysis of the interaction between marketing and customer service in the hotel and restaurant business of Ukraine]. *Academic Visions*, (30). Retrieved from: <https://archer.chnu.edu.ua/xmlui/handle/123456789/10205>
 12. Lehentshuk, S. F., & Zavaliy, T. O. (2023). Big Data v marketynhovii analitytsi: mozhlyvosti ta problemy vykorystannia [Big Data in marketing analytics: Opportunities and challenges]. *Problems of Theory and Methodology of Accounting, Control and Analysis*, 1(54), 52–58.
 13. Charkina, T. Yu., Marcenyuk, L. V., Zadoya, V. O., & Pikulina, O. V. (2021). Stratehichni napriamy upravlinnia hotelno-restorannym biznesom v umovakh kryzy [Strategic directions of hotel and restaurant business management in times of crisis]. *Economics and State*, (2), 19–23.
 14. Kyrchata, I. M., & Prykhodko, D. O. (2023). Kliientoorientovanist yak stratehichnyi oriientyr formuvannia modeli kupivelnoi povedinky spozhyvacha ta rozvytku biznesu [Customer orientation as a strategic guideline for shaping consumer behavior and business development]. *Problems and Prospects of Entrepreneurship Development*, (30), 182–193.

Дата надходження статті: 30.10.2025

Дата прийняття статті: 10.11.2025

Опубліковано: 30.12.2025

UDC 338.48-640.4

DOI <https://doi.org/10.32782/2521-6643-2025-2-70.45>

Yudina O. I., Doctor of Economics, Professor at the
Department of International Tourism and Hotel and Restaurant
Business University of Customs and Finance
ORCID: 0000-0003-3699-5321

Kucher M. M., PhD in Economics, Associate Professor
at the Department of International Tourism and Hotel and Restaurant
Business University of Customs and Finance
ORCID: 0000-0002-3868-5311

Sabirov O. V., PhD, Associate Professor at the Department
of International Tourism and Hotel and Restaurant Business
University of Customs and Finance
ORCID: 0000-0001-9436-0477

Boholiubova A. V.,
University of Customs and Finance
ORCID: 0009-0001-8850-3950

CURRENT TRENDS AND CREATIVE APPROACHES TO INNOVATIVE DEVELOPMENT OF HOSPITALITY INDUSTRY ENTERPRISES IN WORLDWIDE

The article considers current trends in innovative development of hospitality industry enterprises in Europe and worldwide, emphasising creative approaches that shape modern operational strategies within hotel and restaurant sectors. The features of innovative development of hospitality industry enterprises were investigated. Trends and creative approaches to the implementation of innovations in the "La Générale" restaurant were identified.

The conducted analysis highlights several major trends defining contemporary transformation in the restaurant industry: the growing role of delivery and takeaway services, the promotion of healthy and balanced menus, the proliferation of casual dining concepts combining refined cuisine with quick service, and the accelerated digitalisation and automation of business processes through intelligent technologies. Additionally, the research identifies crucial directions in hotel innovations, such as the use of smart technologies, mobile client interaction, AI-based recognition systems, and environmental certification as markers of sustainable development.

There has been a positive trend in investment activity of hotel companies in the global market and a downward trend in investment in new technologies in European restaurant establishments. The example of "La Générale" demonstrates how the integration of digital, environmental, managerial, and cultural innovations fosters operational efficiency, brand differentiation, and competitiveness.

Overall, the hospitality sector today evolves under the influence of digitalisation, automation, ecological responsibility, and creativity in service design. Within the restaurant domain, innovations concentrate on health-oriented nutrition and digital service channels, whereas the hotel industry introduces smart and contactless solutions that enhance sustainability and comfort. The outcomes of the study confirm that modern approaches to innovation at "La Générale" ensure the enterprise's resilience, economic adaptability, and distinct market position.

Key words: innovative development, hospitality industry, enterprises, hotel and restaurant business, trends, creative approaches, digitalisation, technologies.

Юдіна О. І., Кучер М. М., Сабіров О. В., Боголюбів А. В. Сучасні тренди та креативні підходи інноваційного розвитку підприємств індустрії гостинності в світі

У статті розглядаються сучасні тренди інноваційного розвитку підприємств індустрії гостинності, визначаються основні заходи та креативні підходи, що впливають на формування моделей прогресивної діяльності закладів готельної та ресторанної сфери. Досліджено особливості інноваційного розвитку підприємств індустрії гостинності. Визначено тренди та креативні підходи щодо впровадження нововведень в ресторані "La Générale".

У роботі проведено аналіз сучасних інноваційних трендів у ресторанному бізнесі, до яких відносяться: послуги служби доставки та пропозиція страв «на винос»; орієнтація на корисне меню та принципи здорового харчування;

© O. I. Yudina, M. M. Kucher, O. V. Sabirov, A. V. Boholiubova, 2025

Стаття поширюється на умовах ліцензії CC BY 4.0

поширення закладів харчування формату «casual», що характеризуються високою якістю страв і швидкістю обслуговування; цифровізація та автоматизація на основі інтелектуальних технологій; впровадження QR-кодів у процес обслуговування; багатоканальність, пов'язана з поширенням ресторанів нового формату, сфокусованих на приготуванні та доставці їжі. Встановлено креативні тренди інноваційного розвитку в сфері готельного бізнесу, зокрема: інтелектуальні номери, віртуальні тури, мобільні сервіси для споживачів, сертифікація екологічної відповідальності, технології розпізнавання на основі ШІ тощо. Також визначено актуальні напрями інноваційної діяльності готельних підприємств світової індустрії гостинності, що включають: використання мобільних додатків, розумних технологій, ШІ та машинного навчання, реалізацію ініціатив сталого розвитку.

Виявлено позитивну динаміку інвестиційної активності готельних підприємств на глобальному ринку та спадаючий тренд вкладень у нові технології у закладах ресторанного господарства Європи. Досліджено особливості інноваційного розвитку швейцарського ресторану "La Générale" внаслідок нововведень та проведення креативних гастрономічних, культурно-розважальних, управлінських та екологічних заходів.

Визначено, що індустрія гостинності формується під впливом сучасних інноваційних трендів: екологізації, цифровізації, автоматизації, появи креативних сервісних підходів, активного застосування штучного інтелекту. У ресторанному секторі нововведення фокусуються на здоровому харчуванні, доставці їжі, QR-сервісах. Креативність та інноваційна активність у готельному бізнесі проявляється в інтелектуальних номерах, екосертифікації, автоматизованому бронюванні, безконтактному сервісі тощо. Встановлені сучасні напрями інноваційного розвитку ресторану «La Générale» визначають його ефективність, унікальність, економічну стійкість та конкурентоспроможність.

Ключові слова: інноваційний розвиток, індустрія гостинності, підприємства, готельний та ресторанний бізнес, тренди, креативні підходи, цифровізація, технології.

Problem statement. The hospitality industry functions as a complex and multifaceted system of services that provide clients with accommodation, food, recreation, and various travel-related experiences. Its distinctive feature lies in the combination of material and non-material components that together form value for the consumer. In recent years, this sector has demonstrated a pronounced dynamic of continuous modernisation influenced by technological progress, the growing sophistication of consumer expectations, and global imperatives of sustainable development. Currently, hospitality has transcended national borders, becoming an essential part of the global service economy. Its rapid transformation is largely driven by digitalisation, the introduction of advanced information and management technologies, and the reorientation toward innovation-centred business models. As a result, the sector increasingly acts as a catalyst of economic growth, generating significant multiplier effects and stimulating the development of related industries such as transport, catering, and entertainment. The inflow of investments, the intensification of digital transformations, and the growing innovation potential of enterprises in this field facilitate their adaptation to volatile market conditions. These processes ensure the resilience of hospitality enterprises, support the principles of sustainable development, and enhance the efficiency of national and international economies through improved competitiveness and productivity.

Analysis of recent research and publications. The issue of innovative development of enterprises has been extensively discussed by both domestic and foreign scholars, reflecting the growing importance of innovation as a strategic factor in ensuring competitiveness and long-term sustainability. I. M. Pidkaminnyi and V. S. Tsipurinda define innovative development as a purposeful and systematic process that integrates the interests of employees and the enterprise as a unified system, emphasising the coordination of all managerial subsystems [1]. T. M. Pilia-voz interprets innovative development as a sequential, controlled chain of transformations that lead the enterprise toward a balanced and stable innovative condition [2].

However, despite the significant theoretical groundwork, many aspects related to the specific characteristics of innovative progress in hospitality enterprises remain insufficiently explored. The hospitality sector possesses distinctive organisational, technological, and emotional features that differentiate it from other service industries. Thus, identifying the peculiarities of innovation under modern conditions – particularly in the context of creativity, sustainability, and digital transformation – remains a pressing academic and practical task. Special attention is also required to analyse how creative and technological approaches manifest within specific enterprises, such as the Swiss restaurant “La Générale”, which exemplifies the integration of innovative technologies with environmental, managerial, and cultural initiatives aimed at enhancing efficiency and customer satisfaction.

The goal of the article is to study the features and directions of innovative development of hospitality industry enterprises, identify trends and creative approaches to the implementation of innovations in the restaurant "La Générale".

Presentation of the main material. To define the distinctive features and differentiate the nature of innovations within the hospitality industry, it is advisable to analyse hotel and restaurant operations separately, as they constitute the core components of this sector. Although both industries strive to deliver the highest quality of service and to maximise customer satisfaction, their technological and operational processes, as well as their approaches to innovation, differ considerably. Understanding these differences enables the formulation of development strategies tailored to each type of enterprise, thereby improving the performance outcomes of economic entities through the implementation of innovative solutions. Such an approach allows the identification of key areas that contribute to

the formation of an innovative environment within a given segment of the hospitality industry and, consequently, the determination of priority pathways for its progressive advancement.

Innovative development relies on creative approaches aimed at generating non-standard solutions that enhance service quality, foster the creation of new products and services, and increase both the efficiency and adaptability of enterprises within a rapidly changing environment.

The results of the analysis of current innovation trends and creative development directions in the restaurant and hotel sectors are presented in table 1 and figure 1.

Table 1

Modern innovative trends in the restaurant business

Innovative trend	Essential characteristics
Multichannel operations	New-format restaurants that operate exclusively for food preparation and delivery without serving customers on-site. The focus is on using modern delivery applications to reduce costs and increase convenience for consumers. Integration with kitchen equipment and expansion of ordering channels are ensured.
Takeaway dishes and delivery services	Integration of digital platforms with traditional sales points allows customers to place orders online and receive them offline, thus ensuring service flexibility.
Trend toward healthy eating and casual concepts	Orientation toward nutritious menus comparable in quality to premium-segment offerings but more affordable in price. Expansion of concepts promoting fast yet healthy dining.
Digitalisation and automation	Implementation of intelligent technologies to accelerate service, reduce staff workload, and increase operational efficiency.
QR codes in service	Use of digital solutions for payment and menu access via smartphones, supporting environmental standards and meeting the demand for contactless service—particularly relevant in the post-COVID-19 context.

Source: compiled by the authors based on [1, 2, 5]

Modern innovative solutions in the restaurant business are aimed primarily at automating service processes to improve their quality and optimise operational performance. The COVID-19 pandemic had a substantial impact on the transformation of innovations in the hospitality industry. During this period, QR codes were actively introduced to provide access to digital menus, food delivery services expanded significantly, and takeaway options emerged, allowing restaurant establishments to adapt to crisis conditions and maintain profitability.

At present, these innovative models have evolved into a sustainable trend characterised by creativity, high consumer demand for remote services, and the emergence of a broad segment of online restaurants.

During periods of instability and economic crisis, the need arises to form creative approaches to process organisation, the renewal of production and service technologies, and the modernisation of operational and management activities of hospitality enterprises. These measures aim to enhance competitive advantages and ensure effective lead generation.

The automation of hotel services, which has now become a global trend, makes it possible to optimise operational activities and enterprise expenditures while improving managerial efficiency. The creative trends of innovative development in the hotel business are presented in figure 1.

Modern hotel enterprises are actively implementing innovative solutions aimed at increasing the level of service automation, optimising expenditures, and creating a comfortable and safe environment for their guests. The primary focus is placed on the application of technologies that allow for improved service quality while simultaneously reducing resource consumption and minimising dependence on human factors.

At the same time, a significant rise in the popularity of a new direction, the environmentalisation of the hotel business, has been observed [3;4]. In light of the growing environmental awareness among consumers, an increasing number of guests now prefer hotels that introduce green technologies, perceiving them as a more modern, safer, and socially responsible choice (figure 2).

Thus, hotel enterprises in the global hospitality industry increasingly focus on implementing innovations related to sustainable development, AI, and machine learning, as well as on the development of contactless services, the need for which became particularly urgent during the COVID-19 pandemic. At present, sustainable development, based on environmental responsibility, economic efficiency, technological advancement, and the flexibility of enterprises under crisis conditions, is viewed as a key strategic direction in maintaining leading competitive positions within the hospitality market. The adoption of environmentally friendly approaches and digital technologies that reduce resource consumption contributes to the optimisation of operational outcomes, improvement in service quality, and enhancement of the overall competitiveness of hospitality establishments. Among the crucial directions

for the development of the restaurant industry is the growing emphasis on offering healthy nutrition options, which reflect contemporary consumer values and global trends in responsible consumption.

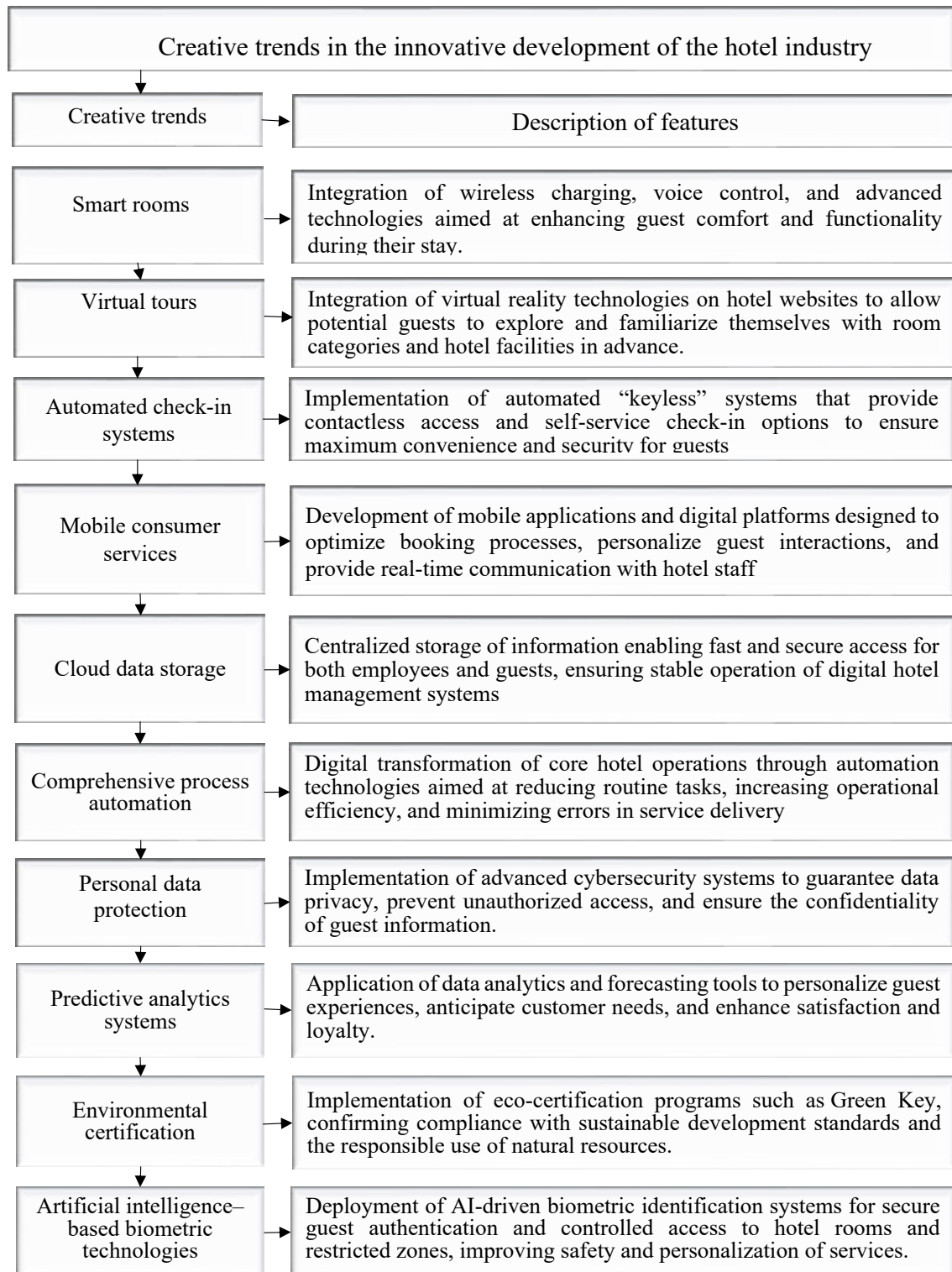


Fig. 1. Creative trends in the innovative development of the hotel industry

Source: compiled by the authors based on [1, 2, 3, 4]

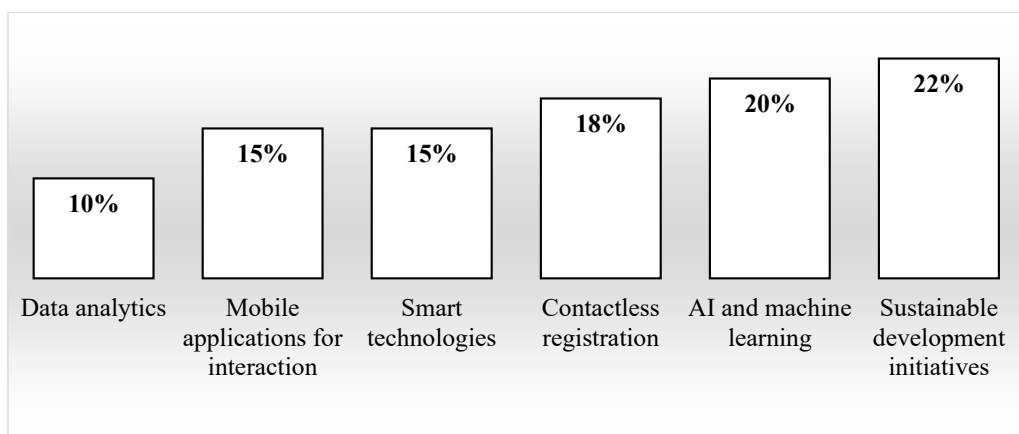


Fig. 2. Directions of innovative activity of hotel enterprises in the global hospitality industry during 2023–2024

Source: compiled by the authors based on [3, 4]

Innovative activity in the hospitality industry is directly dependent on the scale of investment inflows. The dynamics of global investment activity of hotel enterprises from 2018 to 2023 are presented in figure 3.

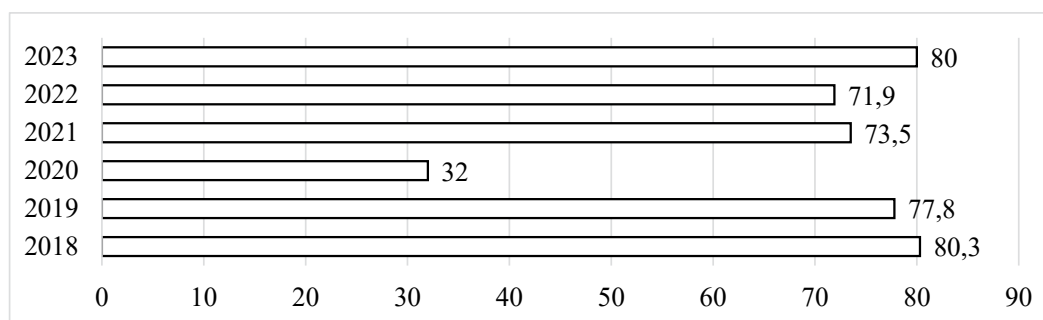


Fig. 3. Dynamics of investment activity of hotel enterprises at the global level during the period 2018–2023, billion USD

Source: compiled by the authors based on [3, 4]

The highest volume of investments in the hotel industry was recorded in 2018. With the onset of the COVID-19 pandemic, the amount of capital investment decreased sharply; however, beginning in 2021, a gradual increase in investment activity was observed. By 2023, the investment volume had approached the pre-pandemic level. This indicates a progressive recovery of the hotel business from the crisis and its transition toward stable development.

Furthermore, based on the conducted analysis, the dynamics of investments in the implementation of advanced technologies by food service enterprises were also identified (fig. 4).

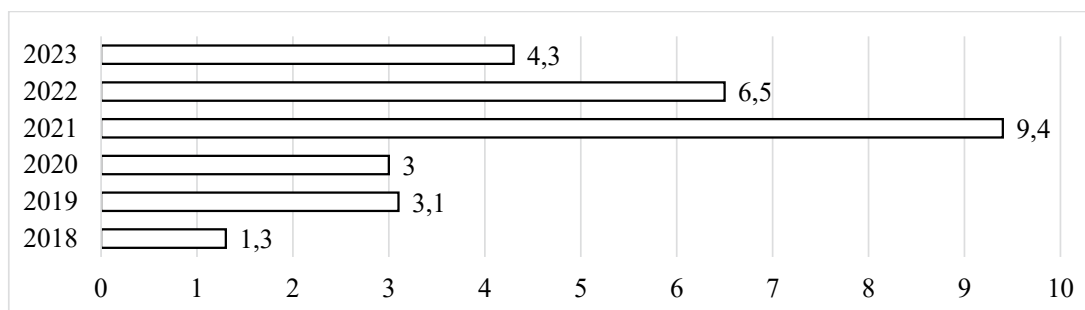


Fig. 4. Dynamics of investments in new technologies in the restaurant industry of Europe during the period 2018–2023, billion USD

Source: compiled by the authors based on [4, 5]

The graph presented in figure 4 demonstrates that, in the post-pandemic period, investments in the restaurant industry reached their highest levels, exceeding the pre-crisis indicators. However, during the last two years, the

total volume of investments has declined. Beginning in 2022, a substantial number of new restaurant establishments entered the market, which led to intensified competition, market oversaturation, and greater challenges in implementing innovative concepts. As a result, investors have redirected their capital flows toward more profitable sectors, perceiving the restaurant industry as unstable and relatively high-risk.

The identification of innovative trends and approaches in the development of hospitality enterprises was carried out using the example of the Swiss restaurant “La Générale”. The innovative activity of “La Générale” is aimed at transforming traditional business processes, developing creative approaches to management, and implementing innovations that contribute to the progressive development and strengthening of the restaurant’s competitiveness. The innovative trends and creative approaches applied in the development of “La Générale” are illustrated in figure 5.

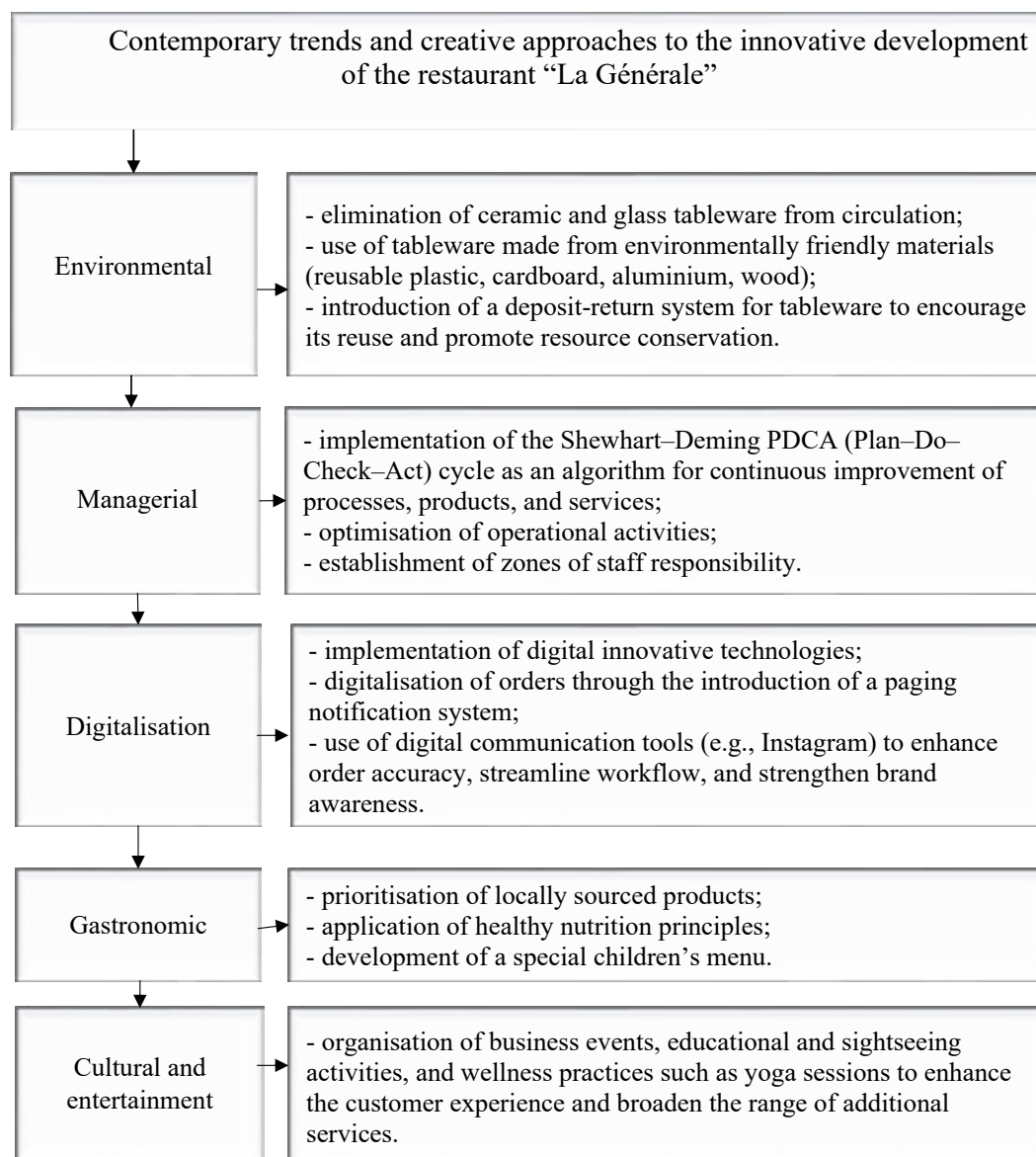


Fig. 5. Modern trends and creative approaches to the innovative development of the restaurant “La Générale”

Source: compiled by the authors based on [1, 2, 5]

The study of the operational aspects of the restaurant “La Générale”, located in Pully, Switzerland, on the shore of Lake Geneva, revealed several key innovative trends and creative approaches that have a positive impact on the development of this establishment:

– an environmental trend is aimed at the withdrawal of ceramic and glass tableware from circulation and the use of tableware made from environmentally friendly materials, including reusable plastic, cardboard, aluminium, and wood. A deposit-return system for tableware has been introduced to stimulate its repeated use. According to this system, customers pay a deposit for the tableware and receive this money back upon its return. Such a non-standard solution allows for a significant reduction of waste and contributes to resource conservation;

– managerial approaches based on the development of an algorithm for continuous improvement of processes, products, and services in accordance with the PDCA (Plan–Do–Check–Act) cycle, the optimisation of operational activities, and the establishment of zones of staff responsibility;

– gastronomic trend is characterised by the use of local products, the application of the principles of healthy nutrition, and the introduction of a special menu for children. The application of these approaches in the restaurant's activities ensures an individual approach and care for the health of guests, promotes the development of the local economy, and reduces the negative impact on the environment;

– digitalisation of the restaurant's operational processes, including the implementation of digital innovative technologies, the digitalisation of the ordering process through the introduction of a pager notification system, and the use of a digital communication tool (Instagram). The use of pagers and digital communication tools such as Instagram accelerates the processing of orders, increases their accuracy, and optimises work processes. It also contributes to the formation and dissemination of the brand, the expansion of a wide target audience, and the increase of customer loyalty to the establishment;

– cultural and entertainment approaches. Within these initiatives, business and entertainment events are organised, including film screenings, concerts, festivals, and sports broadcasts, as well as educational and cognitive excursions and wellness practices such as yoga sessions. These measures add attractiveness and value to the establishment, expand the range of additional services, and attract new consumers.

The conducted assessment demonstrates that after implementing innovative and creative initiatives – alongside the standardisation of managerial and service procedures – the restaurant “La Générale” achieved a substantial improvement in its operational efficiency. The average receipt grew by 15%, the number of tables served per hour rose by 31%, the number of orders processed per employee increased by 24%, and overall turnover expanded by 42%. This progress confirms that the systematic integration of innovative tools contributes directly to productivity and profitability.

A distinctive synergy of ecological practices, technological upgrades, and cultural projects has enabled the restaurant to establish a lasting competitive advantage that is challenging for other market participants to replicate. This combination of sustainability, digitalisation, and creativity forms a unique strategic identity for the enterprise and strengthens its position within the European hospitality sector.

Conclusions from this study and prospects for further research in this area. The study confirms that the hospitality industry today evolves under the influence of key innovation drivers such as environmental sustainability, digital transformation, automation, creative service design, and the growing application of artificial intelligence. These processes are stimulated by the acceleration of scientific and technological progress, as well as by the adaptive responses of enterprises to post-pandemic challenges and persistent economic turbulence.

In the restaurant segment, innovative activity primarily focuses on promoting health-oriented menus, improving delivery and takeaway systems, and implementing digital and QR-based tools that streamline interaction between the client and the service provider. Meanwhile, in the hotel business, innovation manifests through the adoption of smart room systems, environmental certification programs, automated booking technologies, and other digital solutions that enhance operational efficiency and sustainability.

Based on the conducted analysis of the restaurant “La Générale”, this research identified several innovation clusters determining its effectiveness and unique market positioning: environmental, gastronomic, managerial, technological, and cultural and entertainment. These areas collectively reinforce the enterprise's economic stability, foster a positive brand image, stimulate local community development, and contribute to forming a customer-centred service environment.

The revealed tendencies and creative mechanisms driving innovation in hospitality are not only practical tools for improving service quality but also fundamental prerequisites for sustainable economic growth in the sector. In future studies, it would be appropriate to focus on measuring the economic impact of such innovations, developing quantitative models for assessing their efficiency, and comparing creative approaches in different regional and institutional contexts of the hospitality industry.

Bibliography:

1. Підкамінний І. М., Ціпурида В. С. Системні фактори впливу на інноваційний розвиток підприємства. *Ефективна економіка*, 2011. № 3. URL: <http://www.economy.nayka.com.ua/?op=1&z=480>
2. Пілявоз Т. М. Інноваційний розвиток підприємства як важливий аспект розвитку економіки. *Інноваційна економіка*, 2012. № 3. С. 185–190.
3. Міжнародна статистика готелів світу. Statista. URL: <https://www.statista.com/topics/1102/hotels/#topicOverview>
4. Статистична аналітика «Alied market research». URL: <https://www.alliedmarketresearch.com/hotel-market-intelligence-software-market-A203252>
5. Міжнародний статистично-дослідницький ресурс: «DigitalFoodLab». URL: <https://www.digitalfoodlab.com/>

References:

1. Pidkaminnyy, I. M., & Tsipurynda, V. S. (2011). Systemni faktory vplyvu na innovatsiynyy rozvytok pidpryyemstva [System factors of influence on the innovative development of the enterprise]. *Efektivna ekonomika [Efficient economy]*, No. 3. Retrieved from: <http://www.economy.nayka.com.ua/?op=1&z=480>
2. Pilyavoz, T. M. (2012). Innovatsiynyy rozvytok pidpryyemstva yak vazhlyvyi aspekt rozvytku ekonomiky [Innovative development of the enterprise as an important aspect of economic development]. *Innovatsiyna ekonomika [Innovative economy]*, No. 3, pp. 185–190.
3. Mizhnarodna statystyka hoteliv svitu [*International statistics on hotels worldwide*]. Statista. Retrieved from: <https://www.statista.com/topics/1102/hotels/#topicOverview>
4. Statystychna analityka «Alled market research» [Statistical analytics “Alled market research”]. Retrieved from: <https://www.alliedmarketresearch.com/hotel-market-intelligence-software-market-A203252>
5. Mizhnarodnyy statystychno-doslidnyts'kyi resurs: «DigitalFoodLab» [International statistical research resource “DigitalFoodLab”]. Retrieved from: <https://www.digitalfoodlab.com/>

Дата надходження статті: 30.10.2025

Дата прийняття статті: 17.11.2025

Опубліковано: 30.12.2025

НОТАТКИ